British Mycological Society promoting fungal science

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

The genome of *Xylona heveae* provides a window into fungal endophytism

Romina GAZIS[a,*], *Alan KUO*[b], *Robert RILEY*[b], *Kurt LABUTTI*[b], *Anna LIPZEN*[b], *Junyan LIN*[b], *Mojgan AMIREBRAHIMI*[b], *Cedar N. HESSE*[c,d], *Joseph W. SPATAFORA*[c], *Bernard HENRISSAT*[e,f,g], *Matthieu HAINAUT*[e], *Igor V. GRIGORIEV*[b], *David S. HIBBETT*[a]

[a]*Clark University, Biology Department, 950 Main Street, Worcester, MA 01610, USA*
[b]*US Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA*
[c]*Oregon State University, Department of Botany and Plant Pathology, Corvallis, OR 97331, USA*
[d]*Los Alamos National Laboratory, Bioscience Division, Los Alamos, NM, USA*
[e]*Aix-Marseille Université, CNRS, UMR 7257, Marseille, France*
[f]*Aix-Marseille Université, Architecture et Fonction des Macromolécules Biologiques, 13288 Marseille cedex 9, France*
[g]*King Abdulaziz University, Department of Biological Sciences, Jeddah 21589, Saudi Arabia*

ARTICLE INFO

Article history:
Received 12 August 2015
Received in revised form
18 September 2015
Accepted 5 October 2015
Available online 22 October 2015
Corresponding Editor:
Teun Boekhout

Keywords:
CAZymes
Glycoside hydrolases
Horizontally transmitted endophytes
Sapwood endophytes
Symbiotaphrina
Trinosporium

ABSTRACT

Xylona heveae has only been isolated as an endophyte of rubber trees. In an effort to understand the genetic basis of endophytism, we compared the genome contents of *X. heveae* and 36 other Ascomycota with diverse lifestyles and nutritional modes. We focused on genes that are known to be important in the host–fungus interaction interface and that presumably have a role in determining the lifestyle of a fungus. We used phylogenomic data to infer the higher-level phylogenetic position of the Xylonomycetes, and mined ITS sequences to explore its taxonomic and ecological diversity. The *X. heveae* genome contains a low number of enzymes needed for plant cell wall degradation, suggesting that *Xylona* is a highly adapted specialist and likely dependent on its host for survival. The reduced repertoire of carbohydrate active enzymes could reflect an adaptation to intercellulary growth and to the avoidance of the host's immune system, suggesting that *Xylona* has a strictly endophytic lifestyle. Phylogenomic data resolved the position of Xylonomycetes as sister to Lecanoromycetes and Eurotiomycetes and placed the beetle-endosymbiont *Symbiotaphrina* as a member of this class. ITS data revealed that *Trinosporium* is also part of the Xylonomycetes, extending the taxonomic and ecological diversity of this group.

Introduction

Horizontally transmitted endophytes (HTE) are known for their ubiquitous presence, broad taxonomic host associations and high alpha and beta diversity (Arnold & Lutzoni 2007; Rodriguez *et al.* 2009). HTE have also been the target of several bioprospecting studies, which have reported their potential as sources of novel antibiotics, enzymes with applications in the bio-fuel industry and drugs (e.g., Strobel & Daisy 2003; Suryanarayanan *et al.* 2009; Aly *et al.* 2013). Despite the

* *Corresponding author*. Tel.: +1 (508) 793 7332; fax: +1 (508) 793 7174.
E-mail address: rgazis@clarku.edu (R. Gazis).
http://dx.doi.org/10.1016/j.funbio.2015.10.002
1878-6146/

research interest towards this ecologically and taxonomically diverse group, much of their biology remains to be understood. For instance, it is not known if all HTE are metabolically active while living intercellularly within their hosts or what is the extent of their colonization. The degree of specificity that HTE have with their host is also unclear, with some studies suggesting evidence of host specificity (Chaverri *et al.* 2011; Quilliam & Jones 2012; Unterseher *et al.* 2013) and others supporting a host generalist habit (Cannon & Simmons 2002; Higgins *et al.* 2011).

The mechanisms that gave rise to the evolution of endophytism and the genetic basis of the endophytic habit are also poorly understood. It has been speculated that endophytic lineages have evolved several times from plant pathogenic ancestors, but it has also been suggested that plant pathogenic species have arisen from endophytic lineages (Delaye *et al.* 2013; Xu *et al.* 2014). While the mechanisms underlying evolutionary changes between endophytic and parasitic lifestyles are expected to be lineage-specific, the genetic background of the fungal lineage should influence the transition potential.

The main reason why the ecological role of HTE cannot be generalized is that they are polyphyletic (Rodriguez *et al.* 2009). Most classes of Ascomycota have been reported as having endophytic representatives (Arnold & Lutzoni 2007), suggesting that different species of endophytes probably interact with their substrate or host in a distinct manner. The 'endophytic continuum' (Schulz & Boyle 2005) model suggests that the outcome of the plant–fungus interaction, which can range from mutualism to parasitism, depends on the fungal species, the host genetic background and the environment (Kogel *et al.* 2006).

Gazis *et al.* (2012) described the Ascomycota class Xylonomycetes based on the species *Xylona heveae*. Isolates of *X. heveae* were found mainly as sapwood endophytes of *Hevea brasiliensis*, the main source of natural rubber (Chengalroyen & Dabbs 2013). DNA sequences from *X. heveae* isolates (ITS, Internal Transcribed Spacer and LSU, Large Subunit of the ribosomal RNA) were compared to an extensive database comprising saprotrophs, plant pathogens, endophytes and sequences obtained from environmental surveys. No close relatives were found, suggesting that *X. heveae* might represent a strict endophyte, in the sense that it cannot be associated with any other lifestyle. The same study also reported that, in contrast to what has been reported for many leaf endophytes (Oses *et al.* 2006; Promputtha *et al.* 2010), *X. heveae* was not able to degrade cellulose or lignin under laboratory conditions.

In this study, we analysed the genome of *X. heveae* in a comparative framework to address the genetic basis of the endophytic habit and shifts along the 'endophytic continuum'. We focused on genes that are thought to play a role in the host–fungus interaction, with emphasis on carbohydrate active enzymes (CAZymes). Recent comparative studies have demonstrated a strong relationship between the suite of CAZymes in fungal genomes and their lifestyle and nutritional mode (Parrent *et al.* 2009; Eastwood *et al.* 2011; Ohm *et al.* 2012; Kohler *et al.* 2015). We also performed a phylogenomic analysis to infer the placement of the Xylonomycetes within the Ascomycota and a phylogenetic analysis of an expanded ITS–LSU dataset that addresses the taxonomic composition of the Xylonomycetes.

Materials and methods

Fungal strains

Xylona heveae strain TC161 (type strain) was selected for whole-genome sequencing. TC161 was isolated as a sapwood endophyte from cultivated *Hevea brasiliensis* in San Martin, Peru. The strain is deposited in the Centraalbureau voor Schimmelcultures (CBS), Utrecht, The Netherlands (CBS 132557).

Genomic DNA and RNA extraction

Xylona heveae was grown in 500 ml Potato Dextrose Broth media (PDB, Difco) at 25 °C for one month. Fungal mass was collected, dried and frozen in liquid nitrogen. DNA was extracted from ~10 g of frozen mycelium using the Qiagen Blood and Cell Culture DNA Kit Maxi (Catalog # 13362) according to the manufacturer's instructions. A total concentration of 72 ng ul^{-1} was used in the genome sequencing. Total RNA was extracted from a 2 week-old PDB culture mycelial mat using the Qiagen RNeasy Midi Kit (Catalog # 75144), following manufacturer's instructions.

Genome sequencing assembly and annotation

The genome of *Xylona heveae* was sequenced using two Illumina libraries, standard with 270 bp insert size and 4 Kbp long insert paired end (LFPE) and supplemented with a single transcriptomics Illumina library, procedures were similar as the ones described in Floudas *et al.* (2015). Reads were QC filtered for artifact/process contamination and subsequently assembled together with AllPathsLG version R46652 (Gnerre *et al.* 2011) and *de novo* transcript contigs were assembled using Rnnotator (Martin *et al.* 2010). General aspects of library construction and sequencing can be found at the JGI website http://www.jgi.doe.gov/.

The genome of *X. heveae* was annotated using the JGI annotation pipeline, which combines multiple tools for gene prediction, annotation and analysis (Grigoriev *et al.* 2014; Kuo *et al.* 2014). The assembled genomic scaffolds were masked using RepeatMasker ver. 3 (Smit *et al.* 2004). Multiple sets of gene models were predicted using different gene predictors and automatically filtered based on similarity to proteins from other organisms and gene expression support to produce a final non-redundant catalog of genes representing the best gene model found at each locus site. Multi-gene families were predicted using the Markov clustering algorithm (MCL, Enright *et al.* 2002), a part of the JGI annotation pipeline, and annotated using PFAM domains detected in cluster member sequences.

All other genomes included in the analysis were annotated as above, except the *Epichlöe festucae* genome, which was annotated by the working group based at the University of Kentucky (http://csbio-l.csr.uky.edu/ef2011/). To assess the genome coverage and the reliability of protein-coding gene prediction, a search for homologues of a set of highly

conserved fungal core genes (Table S1), was performed for all included genomes, using the BLAST tool implemented in the MycoCosm and the *E. festucae* portals. The presence or absence of a gene was determined based on an e-value <1e−5.

Genome assembly and annotations used in this study are available via the JGI fungal genome portal MycoCosm (http://jgi.doe.gov/fungi). The genome of *X. heveae* can be accessed at http://genome.jgi.doe.gov/Xylhe1, and has been deposited in DDBJ/EMBL/GenBank under the accession JXCS00000000, BioSample SAMN02744854. Species tree has been deposit in TreeBase under accession S18263.

Datasets

We assembled a dataset of 37 fungal genomes that were used for phylogenomic analysis, as well as comparative analyses of genes encoding enzymes with functions putatively related to endophytism, saprotrophy and pathogenicity. We also assembled a 48-taxon dataset of ITS and LSU sequences that was used for phylogenetic analysis. We selected genomes to represent diverse classes, nutritional modes, and lifestyles, including animal-associated fungi, plant pathogens, mutualists and saprotrophs (Table S2). We also included strains that were isolated as endophytes, some of which belong to genera known for their saprotrophic lifestyle (i.e., *Daldinia eschscholzii*, *Hypoxylon* sp. CI-4A, *Hypoxylon* CO27-5 and *Hypoxylon* EC38) and the vertically transmitted endophyte and known mutualist *Epichlöe festucae*. Genomic data from *Xanthoria parietina* were included in the species tree reconstruction but not in the comparative analysis, and *E. festucae* was not included in the species tree reconstruction (because *E. festucae* is not part of the MycoCosm genome cluster) but was part of the comparative analysis. The rRNA dataset was composed of taxa representing most of the known Ascomycota classes (Arthoniomycetes, Dothideomycetes, Eurotiomycetes, Lecanoromycetes, Leotiomycetes, Lichinomycetes, Sordariomycetes and Xylonomycetes) in addition to data from four strains of Xylonomycetes, which were gathered from GenBank (Table S3). Two of these strains were isolated as endophytes and classified as *Symbiotaphrina kochii*. A third strain was isolated from an *Amauroderma* sp. sporocarp and labelled as *Trinosporium guianense* (Dothideomycetes, *incertae sedis*), but ITS and LSU sequences showed it to be closely related to *Xylona heveae*. A fourth strain, *Hyphozyma lignicola*, was included because its phylogenetic position was resolved as sister to *Symbiotaphrina* in a previous phylogenetic study (Schoch *et al.* 2009). Finally, well-curated sequences of *Symbiotaphrina buchneri*, *S. kochii*, and *X. heveae* were included.

Species tree reconstruction and extended taxonomic diversity analysis

To reconstruct the genome-based phylogeny, 174 orthologous groups of genes (present in one copy for each taxa included in the dataset) were derived from MCL clusters and their amino acid sequences were concatenated. These sequences were aligned using MAFFT ver. 7.123b (Katoh & Standley 2013) and well-aligned regions were extracted using Gblocks ver. 0.91b with less stringent settings (Castresana 2000). Gene trees were estimated in RAxML 7.5.4 (Stamatakis 2006), using the standard algorithm and the PROTGAMMAWAG model of sequence evolution, with partitions and 100 rapid bootstraps. For the ITS + LSU analysis (rRNA dataset, see Datasets section), sequences were aligned in MAFFT ver. 7 (Katoh & Standley 2013) and phylogenetic relationships and node robustness were estimated using maximum likelihood (ML) as implemented in raxmlGUI ver. 1.3 (Silvestro & Michalak 2012) with 1000 replicates using the general time-reversible evolutionary model with a Gamma distribution to account for rate heterogeneity among sites (GTRGAMMA).

Comparative genomic analyses

Unique clusters and small secreted proteins (SSP)

Gene clusters present only in *Xylona heveae* were determined using the cluster produced by JGI (inflation parameter 2.0, 'Select Ascomycota 04/24/14.419'). From this set of unique clusters, potential small secreted proteins (SSP) were predicted. Following Ohm *et al.* (2012), SSP are defined as proteins that contain less than 200 amino acids (aa), have no transmembrane domain (TMM) as determined by TMHMM 2.0 (Krogh *et al.* 2001) and have a secretion signal as determined by SignalP ver. 3.0 (Bendtsen *et al.* 2004). We also calculated the percentage of cysteine residues in each of the potential SSP, which were labelled 'cysteine rich' when the percentage of cysteine residues in the SSP was at least twice as high as the average percentage of cysteine residues in all predicted proteins of that organism.

Gene mining and hierarchical clustering

The detection, module composition and family assignment of all CAZymes were performed as described previously (Cantarel *et al.* 2009; Levasseur *et al.* 2013; Lombard *et al.* 2014). Briefly, the method combines BLAST and HMMER searches conducted against sequence libraries and HMM profiles made of the individual functional modules featured in the CAZy database (http://www.cazy.org). Proteins that had more than 50 % identity over the entire domain length of an entry in the CAZy database were directly assigned to the same family. Proteins with less than 50 % identity to a protein in CAZy were manually inspected (searching for catalytic residues when possible). All positive hits were manually examined for final validation. All CAZymes classes were taken into account, including Glycoside Hydrolases (GHs), Carbohydrate Esterases (CEs), Glycoside Transferases (GTs), Polysaccharide Lyases (PLs) and Carbohydrate-Binding Modules (CBMs).

Lignocellulose catabolism potential was evaluated by determining the copy number for a group of CAZymes now classified as Auxiliary Activities (AA), due to their potential to help GHs, PLs and CEs enzymes gain access to the carbohydrates embedded within the plant cell wall. Among others, this group includes the lytic polysaccharide monooxygenases (LPMO = AA9), cellobiose dehydrogenases (CDH = AA3_1) and laccases (Levasseur *et al.* 2013), known to be involved, but not exclusively, in lignin breakdown. Finally, we also assessed the number of genes for Cytochrome P450 monooxygenases and Cytochrome b_{562}, which have also been associated, but not exclusively, with lignocellulose degradation (Floudas *et al.* 2012).

Following Amselem *et al.* (2011), CAZymes were grouped based on their putative substrates. Three groups were delimited, enzymes that are able to degrade components of Plant Cell Walls (PCW), Fungal Cell Walls (FCW) and PCW or FCW. CAZymes involved in PCW degradation were further divided into families targeting cellulose, hemicellulose, hemicellulose-pectin and pectin. Genes related to the transport of carbohydrates from the plant into the fungal cells were also mined, among them the ATP-binding cassette (ABC) transporters and the major facilitator superfamily (MFS).

An additional set of comparisons were performed, involving genes associated with pathogenicity and virulence, other than CAZymes, such as appressorium related genes, Necrosis Inducing Protein (NP1), phytotoxin Cerato-platanin (CP) protein and hydrophobins (Tudzynski & Sharon 2003; Linder *et al.* 2005; Motteram *et al.* 2009; Amselem *et al.* 2011; Bayry *et al.* 2012; Chen *et al.* 2013). Because the genome of *Epichlöe festucae* is not included in the MycoCosm portal, we mined for genes of interests (non-CAZymes) through the BLAST tool implemented in www.endophyte.uky.edu/. The presence or absence of a gene was determined based on an e-value <1e−5. The rest of the genomes were mined for non-CAZymes genes by searching for PFAM domain signatures (http://pfam.xfam.org).

To eliminate the influence of the genome size (total number of predicted genes) on the copy number of the targeted gene families, only Pezizomycotina were considered in the statistical analyses and comparisons were also conducted based on percentages. Copy number, for the genes of interest, was compared across the 32 Pezizomycotina genomes (again, *Xanthoria parietina* was excluded) and basic parametric comparisons were conducted to determine if *Xylona heveae* contained significantly lower or higher copy numbers of the genes of interest. The average in copy number and its 95 % confidence intervals were calculated and significant results were assumed when the copy number in X. *heveae* fell outside this confidence interval ($p \leq 0.05$). Hierarchical clustering was used to identify groups of species with similar hydrolytic profiles, using MeV ver. 4.8 (Multi experiment Viewer; Saeed *et al.* 2003). Following Ohm *et al.* (2012), Euclidean distance was used as the distance metric and complete lineage clustering as the linkage method, with ordering optimization. Heat-maps were also built using MeV.

Secondary metabolite production was explored for the endophytic strains using antiSMASH (antibiotics and Secondary Metabolite Analysis SHell) ver.2 pipeline (Blin *et al.* 2013). Further BLAST searches of the biosynthetic genes against NCBI protein database were performed for the clusters present in X. *heveae*. Fungal sex-related genes that have been functionally verified in *Aspergillus nidulans* were used for blastp search against the genome of X. *heveae* to retrieve the respective homologues (Dyer and O'Gorman 2012). To explain the possibility that X. *heveae* could use components of natural rubber as nutrient source, we searched for homologues of latex-clearing proteins (Lcp) and rubber oxygenase A (RoxA), which have been characterized in *Streptomyces* and *Xanthomonas* (Rose & Steinbüchel 2005; Chengalroyen & Dabbs 2013).

Likelihood analysis of gene gain and loss

To study the evolution of CAZyme family size variation (expansion or contraction), we used CAFE (Computational Analysis of gene Family Evolution) ver. 3 (Han *et al.* 2013). As input data we used the observed copy numbers for the specific homologues of each individual CAZyme family and an ultrametric version of the Ascomycota species tree with integer branch lengths, from which CAFE inferred the most likely gene family size at internal nodes. CAFE analyses were also performed using the copy number for CAZymes grouped by enzymes targeting PCW, including cellulose, hemicellulose, hemicellulose-pectin and pectin. CAFE was also used to estimate the ML value(s) of the birth-and-death rate parameter (λ) that is defined as the change in gene number per million years and assumes that all genes have equal λ of loss or gain. The estimated λ values were then used to calculate the probable ancestral states of family size for each node in the species tree and to run Monte Carlo sampling (1000) to estimate the *p*-values for each gene family describing the likelihood of the observed sizes given random gain or loss by employing the Viterbi test. A *p*-value of 0.05 was used as threshold.

Results

Phylogenetic position of Xylonomycetes and taxonomic diversity

The maximum likelihood phylogeny inferred from the alignment of the 174 single-copy-orthologs indicates that Xylonomycetes (represented by *Xylona heveae*) is sister to *Symbiotaphrina kochii*, which is currently classified as *incertae sedis* within Pezizomycotina (Kirk *et al.* 2008). The clade formed by X. *heveae* and S. *kochii* was resolved as sister to the clade formed by Eurotiomycetes and Lecanoromycetes (Fig 1A). Phylogenetic analysis using the ITS and LSU regions revealed that *Trinosporium guianense*, originally described as a member of Dothideomycetes (Crous *et al.* 2012), represents a third genus in Xylonomycetes (Fig 2). This analysis, in agreement with the phylogenomic results, placed species from the genus *Symbiotaphrina* (S. *kochii* and S. *buchneri*) within Xylonomycetes. In addition, it confirms the placement of the two S. *kochii* endophytic strains mined from NCBI nucleotide database. A third taxon, *Hyphozyma lignicola*, was recovered as part of the Symbiotaphrinales.

General genome features

The genome of *Xylona heveae* was assembled into 27 scaffolds (25 > 2 kb) containing 56 contigs with an average read coverage of 124.5× and a genome size of 24.34 Mbp. Using diverse gene prediction methods which included the use of RNA-Seq data, we predicted 8205 protein-coding genes, 97.5 % of which were supported by EST data. General properties of the X. *heveae* genome can be found in Table 1. In comparison to the other Pezizomycotina included in the analysis, X. *heveae* presented a smaller genome and lower number of predicted proteins (Fig 1B and C; Table 2). All of the 69 fungal core genes used to evaluate genome coverage were found in the X. *heveae* draft genome (Table S1). RepeatMasker identified 1.93 % of the genome

as being repetitive or corresponding to transposable elements, of which 78 (out of 94) were LTR retrotransposons (Long Terminal direct Repeats). The latter is less than the mean found in most Ascomycota (969) but comparable to the number found in species of *Aspergillus* (Muszewska *et al.* 2011). The GC content was estimated to be 48.1 %, which falls within the average for fungi (47.96 %, Li & Du 2014). Based on the SignalP analysis and using a hidden Markov model probability cutoff of 0.8 (Gaulin *et al.* 2008), a total of 504 predicted proteins (6.14 %) presented a secretion signal. From the total number of predicted proteins, 5635 (68.7 %) contained an identifiable domain. The largest gene families include transporters and signalling domains (MFS, WD40, *Ras* and protein kinases; Table S4).

Using as reference the genome of *Aspergillus nidulans*, we found that *X. heveae* contains homologues of genes predicted to encode proteins involved in sexual recombination and production of fruiting bodies (Dyer and O'Gorman 2012). In addition, the *X. heveae* TC161 genome has only one MAT locus, MAT 1-2 containing a high-mobility group mating-type (MAT-HMG, protein Xylhe1_285192), suggesting that it may be heterothallic (Table S5). No homologous regions of the genes known to be involved in the degradation of latex were found in *X. heveae*.

Comparative genomic analyses

Unique clusters and small secreted proteins

The 8205 *Xylona heveae* predicted proteins were grouped into 6922 clusters. From these, 1088 clusters (= 1655 predicted proteins) were shared among the included Ascomycota and 973 clusters (= 1009 predicted proteins) were unique to *X. heveae*. The average length (aa) of these 'orphan' predicted proteins was 182.61 ± 9.71 and 66 of them had no RNA coverage. Out of the total unique clusters, 416 (43 %) had a secretion signal, 960 (99 %) had no predicted conserved domain and 784 (81 %) had no BLAST hits. Most of these unique clusters (92 % = 891 clusters = 918 predicted proteins) had no homologues even when compared to all the Ascomycota genomes included in MycoCosm (cluster size = 208; http://genome.jgi.doe.gov/clm/run/ascomycota.43) and 113 (12 %) had the highest top BLAST hits to members of other groups such as animals, bacteria or plants (Table S6). Based on our definition of SSP (see Materials and methods), 165 (17 %) of the unique *Xylona*-clusters could be considered as potential SSP. The percentage of cysteine residues in the potential SSPs was ca. three times higher than in the other proteins (3.18 % vs 1.08 %).

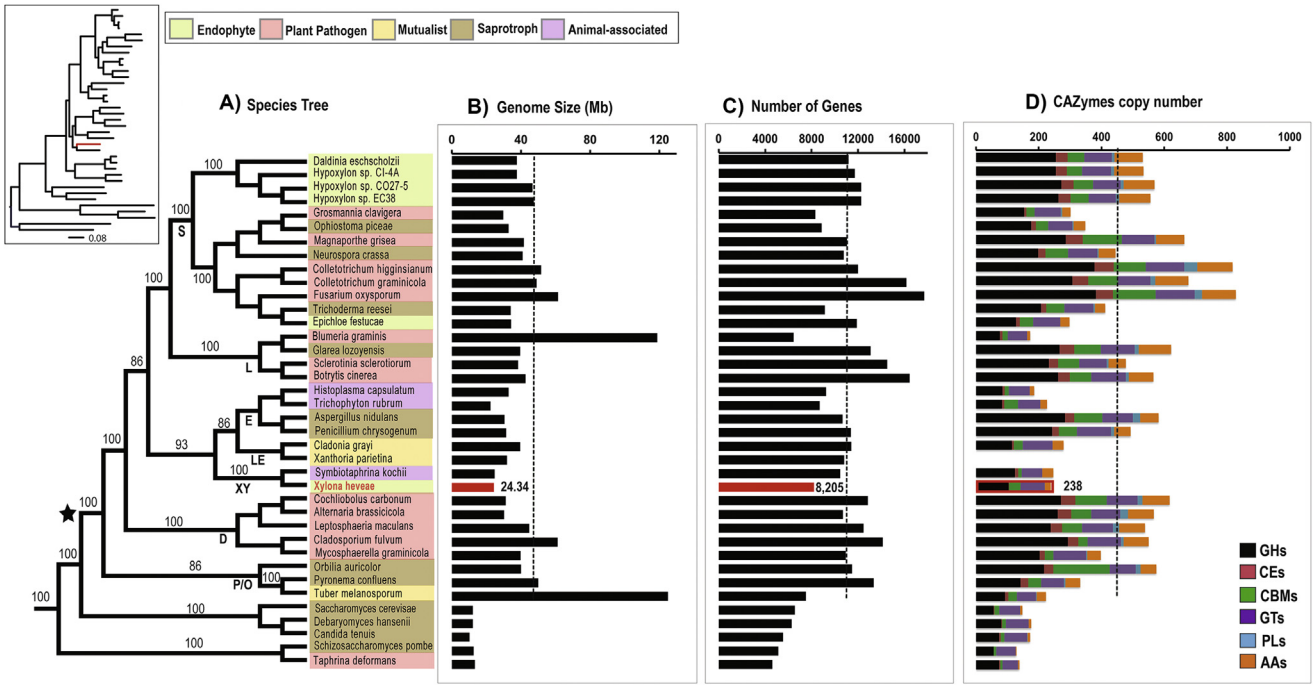

Fig 1 – Phylogeny and distribution of genome size, number of genes and CAZymes. (A) The phylogeny was inferred by the maximum likelihood method using RAxML from a concatenated alignment of 174 single-copy-orthologs conserved across species. Numbers above the branches indicate bootstrap support and letters represent Ascomycota classes (D = Dothideomycetes, E = Eurotiomycetes, L = Leotiomycetes, LE = Lecanoromycetes, PO = Pezizomycetes + Orbiliomycetes, S = Sordariomycetes, XY = Xylonomycetes and ★ = Pezizomycotina); inset depicts the same tree but reflecting branch lengths. (B), (C) Distribution of genome size and number of genes among the strains included in the dataset, dashed line represents the average for Pezizomycotina (43.77 ± 8.00 and 11 481 ± 921.03, respectively). (D) Distribution of the different classes of CAZymes and associated modules among the strains included in the dataset (GHs = Glycoside Hydrolases, CEs = Carbohydrate Esterases, CBMs = Carbohydrate-Binding Modules, GTs = GlycosylTransferases, PLs = Polysaccharide Lyases and Auxiliary Activities (AAs)). Dashed line represents the average number of CAZymes for Pezizomycotina = 459.88 ± 60.56. CAZymes data for *X. parietina* is not shown and the phylogenetic position of *E. festucae* in the species tree was based on data from six loci (Spatafora *et al.* 2007).

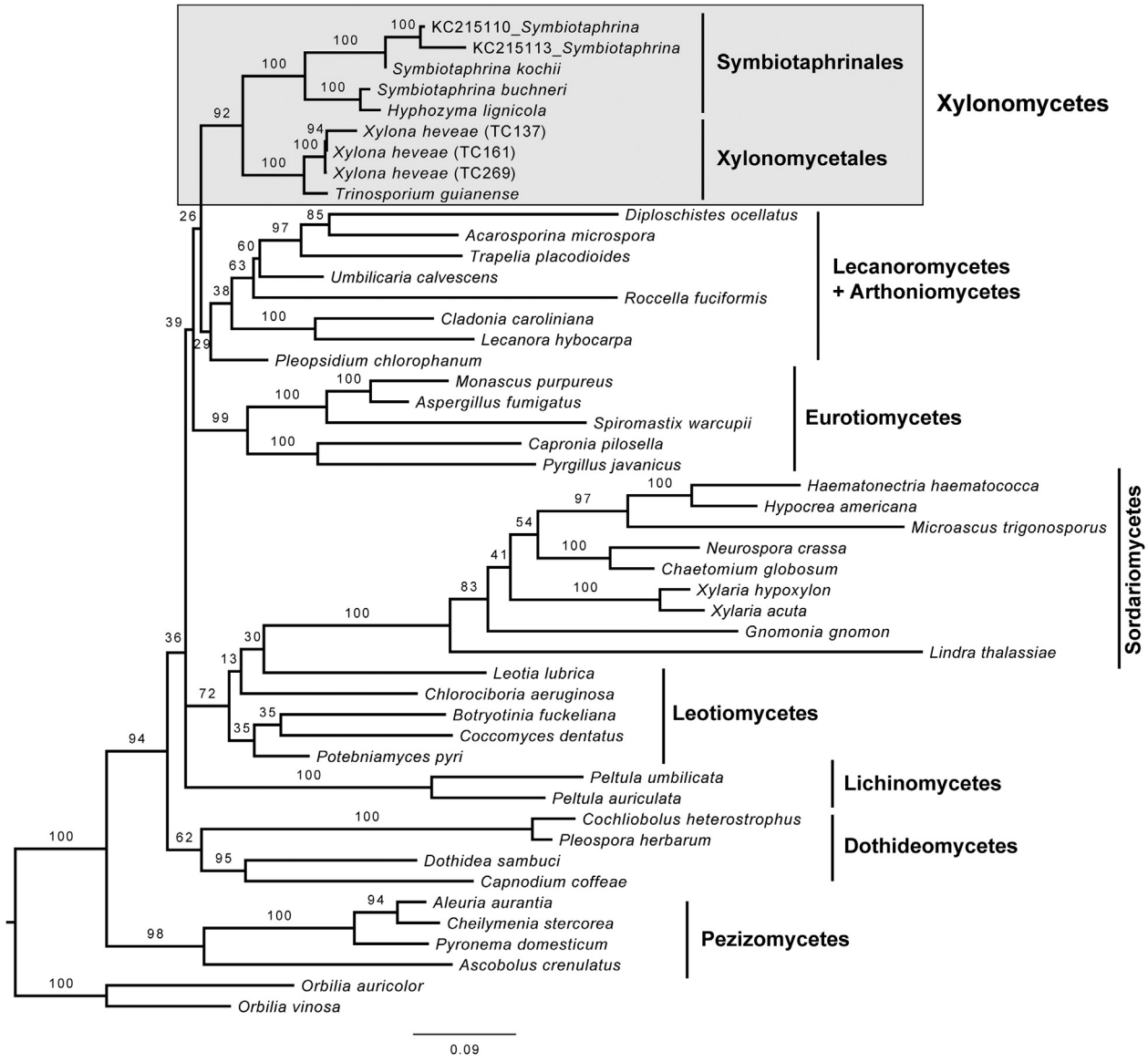

Fig 2 – Phylogenetic tree showing the placement of three strains of *Xylona heveae* and two other genera, *Trinosporium* and *Symbiotaphrina*, within the Xylonomycetes. A third genus is shown, *Hyphozyma*, clustering within the *Symbiotaphrina* clade (see Discussion). The phylogeny was inferred by the maximum likelihood method using RAxML from a concatenated alignment of the ITS and LSU ribosomal regions.

Only one cluster was exclusively present in the endophytic strains (*Daldinia eschscholzii*, *Epichlöe festucae*, *Hypoxylon* sp. CI-4A, *Hypoxylon* sp. EC38, *Hypoxylon* sp. CO27-5 and *X. heveae*) but no domain was recognized. Thirteen clusters were found exclusively shared by *X. heveae* and its closest relative *Symbiotaphrina kochii*, eleven of which had no predicted conserved domain, while two clusters were predicted as a transcription factor (Helix-loop-helix DNA-binding domain, Xylhe1_266133) and a fungal hydrophobin (Xylhe1_271306).

Carbohydrate active enzymes (CAZYmes)

We found a positive correlation between the number of predicted genes and the number of CAZymes and PCW degrading enzymes ($R^2 = 0.680$ and $R^2 = 0.609$, respectively) (Fig S1). The genome of *Xylona heveae* contained 238 genes encoding putative CAZymes, which is lower than the average in Pezizomycotina and less than expected for its genome size (Fig S1; Table 2; Tables S7–S11; protein ID for *X. heveae* CAZymes Table S12). No genes containing the domain associated with PLs were found in *X. heveae* and AAs were present in fewer numbers (Table 2; Tables S9 and S13). Only four copies of the glucose-methanol-choline oxidoreductase superfamily gene (GMC, AA3) were found in *X. heveae*, from which none of them could be considered as a potential CDH (=AA3_1, regions containing GMC linked to a cytochrome b_{562} region). In contrast to most plant-associated filamentous fungi (except for *Tuber melanosporum*), some specific CAZymes were missing from *X. heveae*. For instance, no homologues of C5 (cutinase)

Table 1 – Main features of the *Xylona heveae* strain TC161.

Feature	*X. heveae*
Genome Assembly size (Mbp)	24.34
Sequencing read coverage depth	124.5×
G+C content (%)	48.1
Total transposable elements (%)	1.93
Number of contigs	56
Number of scaffolds	27
Number of scaffolds ≥2 Kbp	25
% of scaffold length in gaps	0.1 %
Number of ESTs	23 510
Number of ESTs mapped to genome	22 911 (97.5 %)
Number of gene models	8205
Average gene length (bp)	1955
Average protein length (aa)	489
Predicted proteins represented by RNA (%)	98.7 %
Average exons per gene	3.41
Orphan proteins[a]	1009 (12.30 %)
Number of predicted secreted proteins	504 (6.14 %)

a Based on all the strains included in the comparative analysis.

Table 2 – Summary of the comparative analyses for general features of *X. heveae* genome and for a key set of genes related to carbohydrate metabolism and transportation.

	X. heveae	Average ± CI in Pezizomycotina
Genome size (Mb)	**24.34**	43.77 ± 8.00
Number of genes	**8205**	11 481 ± 921.03
Total number of CAZYs	**238**	459.88 ± 60.56
CAZYs genome %	**2.90**	3.94 ± 0.37
Total number of GHs	**104**	213 ± 28.68
Total number of CEs	**2**	27.24 ± 5.70
Total number of GTs	**78**	91.09 ± 5.53
Total number of PLs	**0**	8.79 ± 3.15
Total number of CBMs	**36**	60.58 ± 12.78
Total number AAs	**18**	60.88 ± 10.78
Total number of laccases	**6**	9.22 ± 1.68
Total number of peroxidases class II	**1**	4.78 ± 0.96
Cytochrome P450	**20**	88 ± 18.69
Cytochrome b_{562}	19	21.22 ± 1.98
The Major Facilitator Superfamily (MFS)	158	177.38 ± 35.82
ABC transporters	**31**	41.69 ± 4.87

Total CAZYs = GHs + GTs + CBM + PL + CE + AAs.
GHs = Glycoside Hydrolases.
CEs = Carbohydrate Esterases.
GTs = Glycoside Transferases.
PLs = Polysaccharide Lyases.
CBMs = Carbohydrate-Binding Module.
AAs = Auxiliary Activities (AA), includes redox enzymes that act in conjunction with CAZymes.
Significant lower abundance is denoted in bold.

genes or for pectin-degrading CAZymes (GH28, GH78, GH88, GH95, GH105, GH115, PL1, PL3, PL4, PL9, PL11 and CE8) were found.

The most abundant GHs were GH16 (endoglucanases) and GH18 (chitinases). In comparison to the other species included in the analysis, *X. heveae* genome appears to be enriched in GH15, GH32 and GH49, which are known to be involved in the degradation of starch, sucrose and dextran, respectively (Table S8). The heuristic clustering revealed that the repertoire of CAZymes present in *X. heveae* is similar to that of mutualists such as *Cladonia grayi*, *T. melanosporum* and *Epichlöe festucae* (Fig 3).

Xylona heveae presents fewer genes associated with PCW degradation, than the other Pezizomycotina analysed (Table 3; Table S14). The same pattern was found for enzymes related to FCW degradation and for enzymes that can degrade both PCW and FCW. The heuristic clustering, based on the abundance and composition of CAZymes associated with different substrates, revealed that the repertoire of enzymes present in *X. heveae* is most similar to the one present in animal-associated and mutualists filamentous fungi (Fig 4).

CAFE revealed a significant expansion of the GH32 (+3 gains, $p \leq 0.005$), GH71 (+4 gains, $p \leq 0.004$), GT1 (+4 gains, $p \leq 0.004$) and CBM50 (+6 gains, $p \leq 0.01$) families in *X. heveae*. Significant contractions were only detected for the CE4 (−3 losses, $p \leq 0.03$) family (Fig S2). Significant contractions were also found when CAZymes were grouped by substrate, such as for enzymes associated with the degradation of hemicellulose (7 losses, $p \leq 0.03$) and pectin (5 losses, $p \leq 0.002$; Fig 5A and B). No significant contractions or expansions were predicted for *Symbiotaphrina kochii* when CAZymes were grouped by substrate, but expansions in CE4 (4 gains, $p \leq 0.02$) and contractions in CBM50 (6 losses, $p \leq 0.003$) and GT1 (3 losses, $p \leq 0.01$) were detected (Fig S2).

Transporters

The genome of *Xylona heveae* presented fewer ABC-domain copies in comparison to other Pezizomycotina (Table 2; Table S15) and included one copy of the Abc1p domain (Xylhe1_246960), known to be involved in the detoxification of plant defence compounds by plant-associated fungi (Del Sorbo *et al.* 2000). On the other hand, the copy number of the MFS domain fell within the average (Table 2; Table S15) and included three copies of the SpSUT1 (Xylhe1_279241, Xylhe1_210589 and Xylhe1_239008) known for its high affinity to sucrose (Doidy *et al.* 2012).

Pathogenicity related genes

Some specific genes coding for enzymes involved in the plant infection and colonization process were absent from the genome of *Xylona heveae* but present in most of the species analysed, including the endophytic strains (Table 4). Some of these genes code for enzymes such as cutinases and pectinases, and for essential proteins involved in the development of appressoria. No homologues of NPP1 or CP genes were found in *X. heveae*, but they were present in all the other endophytic strains and all of the plant-associated taxa (except *Blumeria graminis* which lacks NPP1). Additional genes suggested to be involved in host-pathogen interactions as well as fungal virulence, such as those encoding hydrophobins, were present in lower number in *X. heveae* in comparison to the other endophytic strains (Table 4; Table S16).

Secondary metabolites

We detected 25 secondary metabolite gene clusters in *Xylona heveae*, containing 26 biosynthetic genes of which eight belonged to the nonribosomal peptide synthases

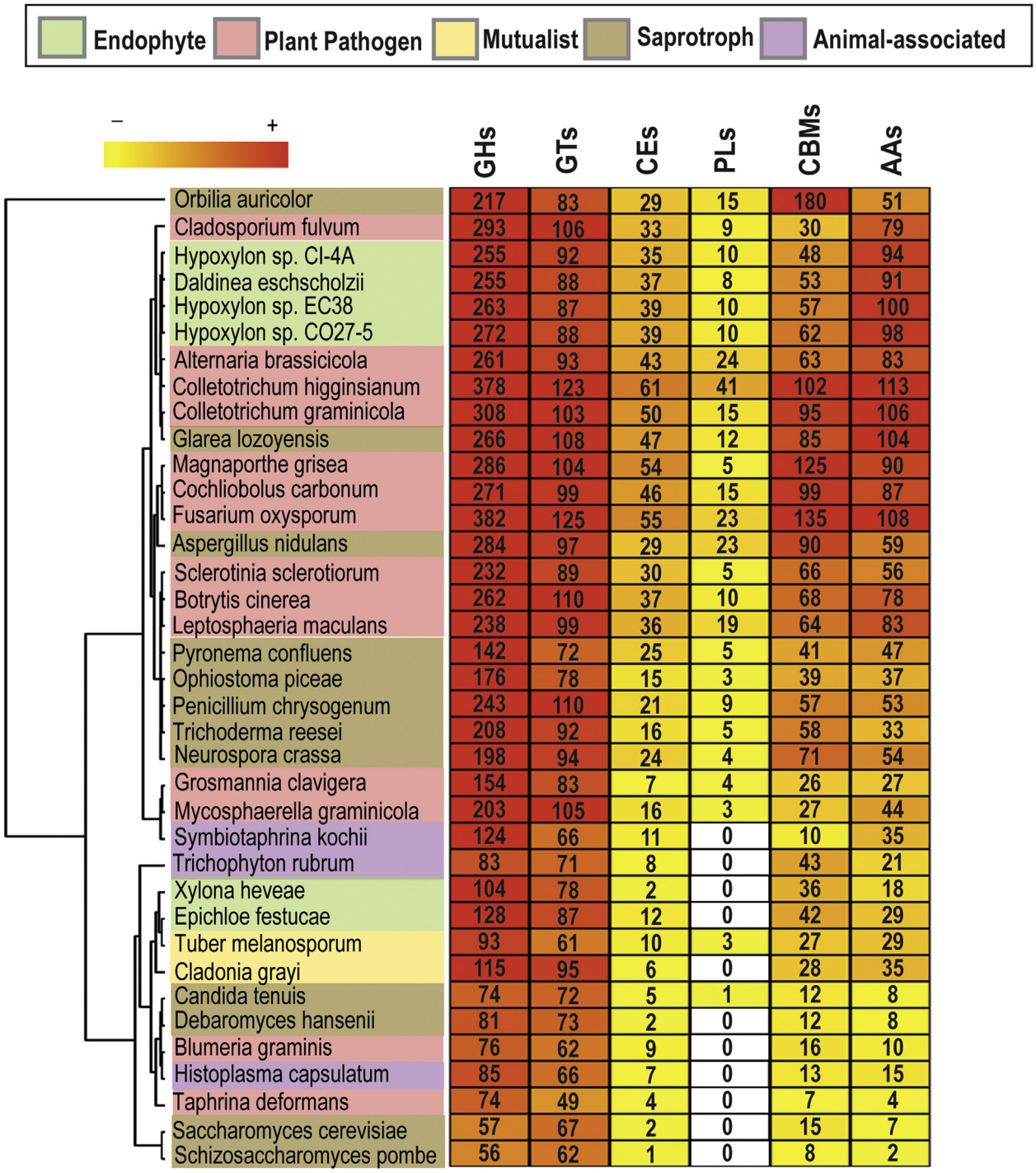

Fig 3 – Hierarchical clustering of *Xylona heveae* and 36 other Ascomycota genomes based on the abundance and composition of CAZymes classes. The number of copies for each CAZyme class is indicated and the heatmap indicates relative abundance. CAZymes categories include Glycoside Hydrolases (GHs), GlycosylTransferases (GTs), Carbohydrate Esterases (CEs), Polysaccharide Lyases (PL), Carbohydrate-Binding Modules (CBMs) and the Auxiliary Activities enzymes (AAs).

Table 3 – Summary of the comparative analyses for number of CAZYs in relation to plant and fungal cell wall degradation. Significant lower abundance is denoted in bold. Relative abundance based on number of genes is indicated within parenthesis.

	X. heveae	Average ± CI in Pezizomycotina
Plant cell wall degrading enzymes	**9 (0.11)**	92.25 ± 20.08 (0.75 ± 0.15)
Cellulose	**1 (0.01)**	21.59 ± 5.75 (0.18 ± 0.05)
Hemicellulose	**8 (0.01)**	36.72 ± 7.24 (0.30 ± 0.05)
Hemicellulose-pectin	**0 (0.0)**	8.44 ± 1.71 (0.07 ± 0.01)
Pectin	**0 (0.0)**	24.50 ± 7.22 (0.20 ± 0.05)
Fungal cell wall degrading enzymes	**34 (0.41)**	66.94 ± 6.98 (0.59 ± 0.06)
Fungal or plant cell wall degrading enzymes	**16 (0.18)**	46.47 ± 7.25 (0.39 ± 0.05)

Plant cell wall degrading (PCW) enzymes include CAZYs families involve in cellulose, hemicellulose, hemicellulose or pectin side chains and pectin (PCW = C + H + HP + P). Cellulose (C) = GH6, GH7, GH12, GH45, GH74, GH94, AA9 (ex GH61). Hemicellulose (H) = GH10, GH11, GH26, GH27, GH29, GH31, GH35, GH36, GH39, GH67, CE1, CE2, CE3, CE5, CE15, CE16. Hemicellulose or pectin side chains (HP) = GH43, GH51, GH53, GH54, GH62, GH93, CE12. Pectin (P) = GH28, GH78, GH88, GH95, GH105, GH115, PL1, PL3, PL4, PL9, CE8.
Fungal cell wall degrading CAZYs: GH16, GH17, GH18, GH20, GH55, GH64, GH71, GH72, GH75, GH76, GH81, GH85, GH92, CE4.
Fungal or plant cell wall degrading CAZYs: GH1, GH2, GH3, GH5.

(NRSP) group, nine to the polyketide synthases (PKS) group, two to the dimethylallytrytophan synthase (DMATS) group, five to the terpene synthase group, two to the NRPS/terpene hybrid group and one to the NRPS/PKS hybrid group. When compared to the other endophytic strains included in the analysis, we found that *X. heveae* presented fewer clusters (Table 5). The majority of the predicted proteins were most similar to proteins present in species within Eurotiomycetes although in some cases top hits were to proteins found in Dothideomycetes or Sordariomycetes taxa (Table S17).

Discussion

Xylonomycetes form a strongly supported clade that is sister group of the Eurotiomycetes plus Lecanoromycetes and that contains two orders, the Xylonomycetales and Symbiotaphrinales. The latter order was resolved by Schoch *et al.* (2009) but not formally named. The phylogenetic analysis using the rRNA dataset clarified the position of *Trinosporium guianense*, placing it within the Xylonomycetales. This analysis also showed that Symbiotaphrinales contains *Symbiotaphrina buchneri*, *Symbiotaphrina kochii* and a third taxon *Hyphozyma lignicola*. *Symbiotaphrina* is paraphyletic so future nomenclatural revisions will be necessary. The mining of ITS sequences from GenBank revealed two unpublished sequences that are closely related to *Symbiotaphrina*. These isolates were collected as endophytes of the monocot *Dracaena cambodiana* ('Dragon's blood') in China (KC215110 and KC215113), unveiling a previously unknown part of the biology of these beetle-associated fungi. The two endophytic *Symbiotaphrina* strains were the only sequences found to be closely related to the Xylonomycetes, even when environmental and uncultured sequences were included in the ITS mining.

Comparative genomic analyses

The genome of *Xylona heveae* is smaller than the average Pezizomycotina genome (Kelkar & Ochman 2012). Symbiotic organisms that have a tight relationship with their host often have reduced genomes (McCutcheon & Moran 2012), although this is not always the case. For instance, the obligate biotroph *Blumeria graminis* and the mycorrhizal species *Tuber melanosporum* have large genomes, but as a reflection of a high incidence of transposable elements and not due to an expanded gene content. Although it is species-dependent, the proliferation of transposable elements has been linked to a mutualistic habit (Hess *et al.* 2014; Xu *et al.* 2014). The genome of *X. heveae* presents low incidence of transposable elements, and has fewer genes than the other Pezizomycotina taxa, including its closest relative *Symbiotaphrina kochii*. A reduction of the number of genes has also been considered a sign of a close symbiosis between fungi and their host (Spanu *et al.* 2010).

We found a depauperate repertoire of CAZymes, especially those implicated in the degradation of PCW components, including cellulose, hemicellulose, pectin and lignin. The CAZymes profile found in *X. heveae* agrees with the lack of *in vitro* lignin and cellulose degradation found by Gazis *et al.* 2012. All CAZyme classes in *X. heveae* were found in lower numbers than the average Pezizomycotina. The highest number of GHs was found in species known to be aggressive plant pathogens such as *Fusarium oxysporum*, *Colletotrichum higginsianum* and *Colletotrichum graminicola*. The smallest number of GHs was found in organisms that have no significant interaction with plants (i.e., *Histoplasma capsulatum*, *Trichophyton rubrum*), with some exceptions. For instance, a low GH copy number was found in the plant pathogen *B. graminis* although this may be a reflection of its nutrition mode as a haustoria-producing obligate biotroph. In contrast to the plant-associated taxa analysed here, *X. heveae* presented no predicted genes involved in the degradation of cutin or pectin (Table 3).

Xylona heveae has a hydrolytic profile similar to mutualists and to animal-associated fungi. On the other hand, the xylariaceous strains isolated as endophytes (*Daldinia eschscholzii*, *Hypoxylon* sp. CI-4A, *Hypoxylon* CO27-5 and *Hypoxylon* EC38) have a profile resembling plant pathogens and saprotrophs (Figs 3 and 4). The oxidative potential for cellulose and lignin catabolism in *X. heveae* is also lower than that predicted for other filamentous ascomycetes. Oxidative enzymes implicated in lignocellulose degradation such as LPMO (AA9), CDH

(AA3_1) and class II peroxidases (AA2) were absent or present in low copy number. *Symbiotaphrina kochii*, the closest relative of *X. heveae*, has a slightly greater repertoire of CAZymes and a greater potential for PCW degradation, including the presence of pectinases and cutinases-encoding genes. However, its CAZyme genome content was also appreciably less than the average for Pezizomycotina.

When CAZymes were grouped by substrate, CAFE detected significant contractions for enzymes related to the degradation of pectin and hemicellulose. Larger numbers of genes encoding pectin-degrading enzymes have generally been related to a necrotrophic rather than biotrophic lifestyle in fungi (Sprockett *et al.* 2011), enabling the invasion of pectin-rich plant tissues (van Kan 2006; Amselem *et al.* 2011). However, some biotrophs such as *Cladosporium fulvum* also have high number of pectin-degrading enzymes, which are used to locally modify the primary cell walls of mesophyll cells and allow the fungus to grow in the apoplast of tomato leaves

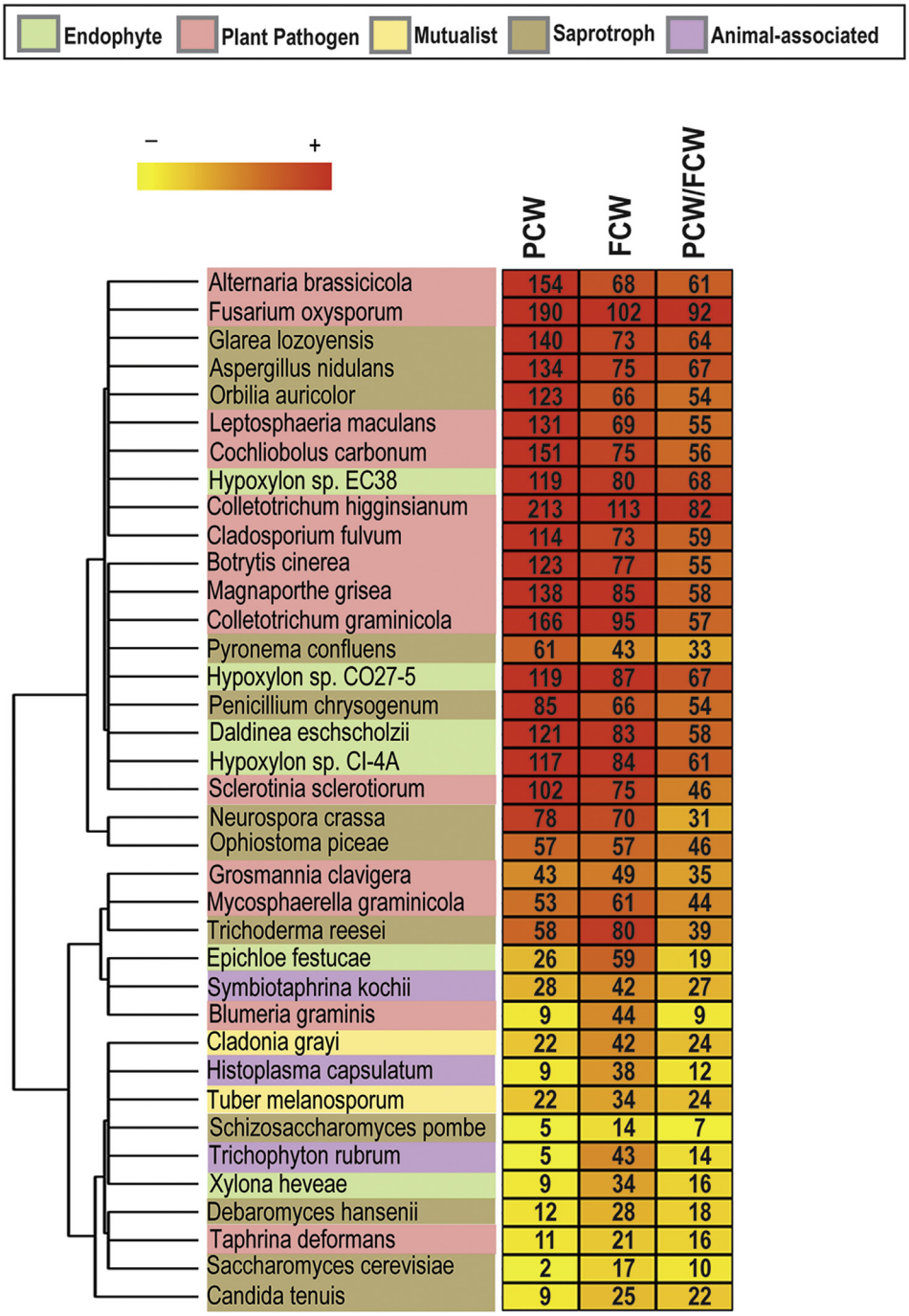

Fig 4 – Hierarchical clustering of *Xylona heveae* and 36 other Ascomycota genomes based on the abundance and composition of CAZymes associated with different substrates (PCW = Plant Cell Wall, FCW = Fungal Cell Wall, PCW/FCW = Plant Cell Wall or Fungal Cell Wall). The number of copies for the enzymes associated with each substrate is indicated and the heatmap indicates relative abundance. CAZymes were grouped in substrates following Amselem *et al.* 2011.

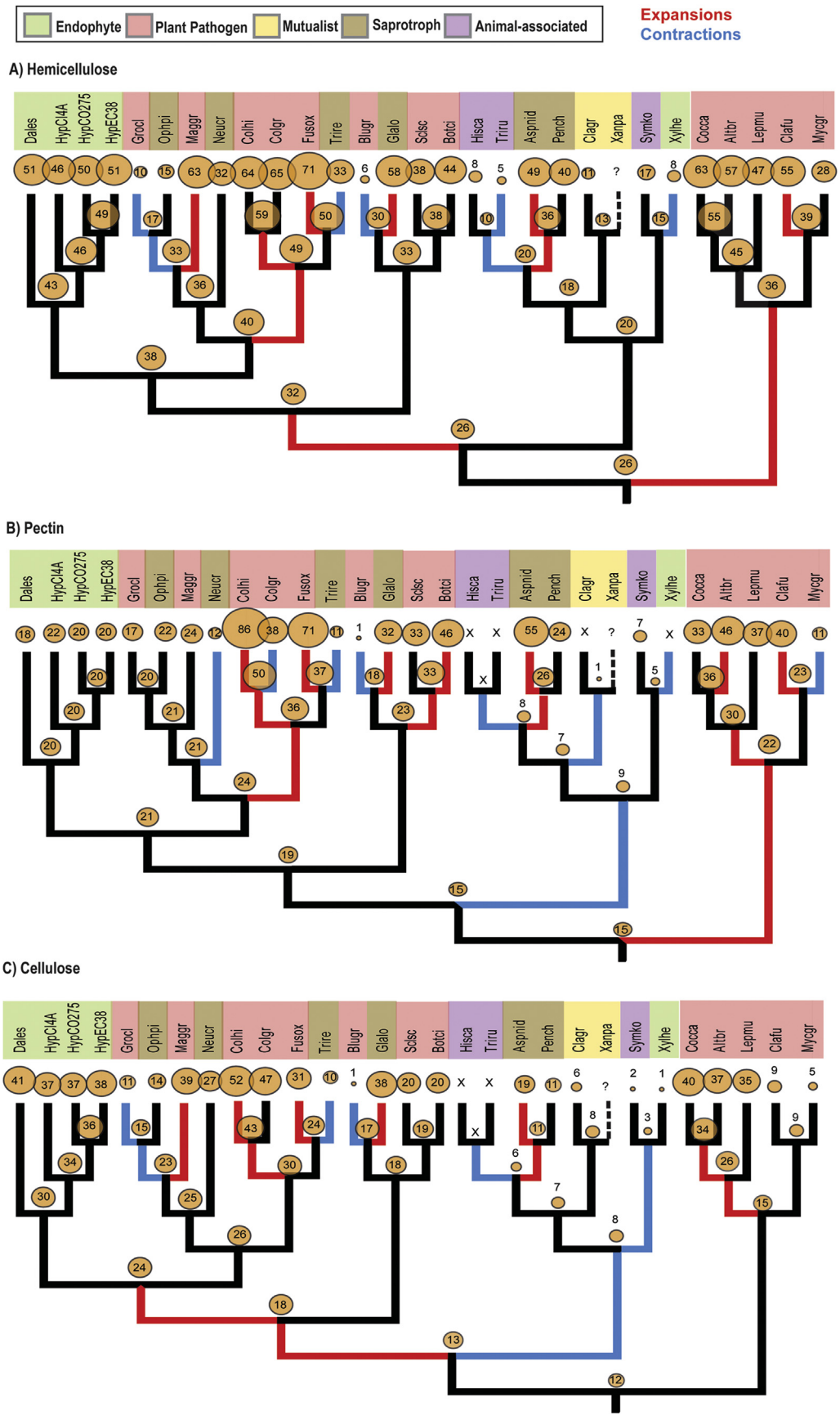

Fig 5 – Results from CAFE analysis. Distribution of CAZyme gain and loss associated with degradation of (A) Hemicellulose, (B) Pectin and (C) Cellulose. Numbers at the tips represent the number of copies for the corresponding categories in the

Table 4 – Comparative analysis among the endophytic strains, based on copy number of homologous genes associated with fungal pathogenicity. Gene count for all strains included in the analysis is shown in Table S16.

	Function	Xh	Ef	De	H1	H2	H3
Necrosis inducing protein (NPP1)	Induce transcript accumulation of pathogenesis-related (PR) genes, production of ROS and ethylene, callose apposition, and HR-like cell death	0	1	1	1	1	1
Cerato-platanin (CT)	Plays a role in the host-pathogen interaction, inducing cell necrosis and act as elicitor of the plant's immune system	0	1	1	1	1	1
LysM domain (=CBM50)	Present in enzymes involve in the degradation of bacterial cell walls. In plant it prevents chitin-trigger immunity by sequestrating chitin molecular product of fungal cell wall cleavage.	13	16	16	14	20	16
Hydrophobins	Important in the water–air interface, helps fungal pathogens evade plant immune system by masking fungal colonization	1	2	5	3	4	4
Appressorium – related genes (*Magnaporthe grisea* proteins)							
Psl1 (MGG_12594.6)	Appressorium penetration; tetraspanin	0	1	1	1	1	1
mpg1 (MGG_10315.6)	Adhesion; hydrophobin	0	0	0	0	0	0
pth11 (MGG_05871.6)	Appressorium differentiation; integral membrane protein	6	13	84	92	107	100
pde1 (MGG_00111.6)	No penetration hyphae; aminophospholipid translocase (APT) family of P-type ATPase	1	7	2	2	2	2
emp1 (MGG_00527.6)	Appressorium formation; extracellular matrix protein	1	7	2	2	2	2
atg8 (MGG_01062.6)	Penetration hyphae; autophagy	1	1	1	1	1	1
pth2 (MGG_01721.6)	Penetration hyphae; carnitine O-acetyl transferase	1	3	1	1	1	1
mst12 (MGG_12958.6)	Appressorium maturation; steA transcription factor	1	1	1	1	1	1
chm1 (MGG_06320.6)	Appressorium formation and penetration; PAK protein kinase	2	1	2	2	2	2
mas1 (asg1; gas1) (MGG_12337.6)	Appressorium penetration	0	5	4	5	3	3
mas2# (MGG_04202.6)	Appressorium penetration	0	5	4	5	3	3
mas3# (MGG_11595.6)	Uncharacterized	0	4	4	5	3	3
mas3# (MGG_00703.6)	Uncharacterized	0	5	4	5	3	3
cas1 (MGG_09875.6)	Uncharacterized	0	5	4	5	3	3
mas-related (MGG_00992.6)	Uncharacterized	0	2	4	5	3	3
mas-related (MGG_02253.6)	Uncharacterized	0	4	4	5	3	3
KP4 protein gene family							
Kp4-like (*Xylona heveae* 251361)	Secreted toxin with antifungal properties	1	1	0	1	2	2

H1: Hypoxylon sp. CI-4A.
H2: Hypoxylon sp. CO27-5.
H3: Hypoxylon sp. EC38.

Table 5 – Number of genes encoding secondary metabolites (SM) enzymes and number of SM clusters in *Xylona heveae* and five other endophytes fungal genomes.

	Total clusters	Total genes	Pks	Terpene	Nrps	Nrps-pks	Nrps-terpene	Pks-pks	DMAT	Unknown
Epichlöe festucae	41	42	7	6	21	2	2	1	3	0
Daldinia eschscholzii	52	53	22	10	14	2	2	0	3	0
Hypoxylon sp. CI-4A	48	43	22	11	9	1	0	0	0	5
Hypoxylon sp. CO27-5	48	45	22	11	8	1	0	0	3	3
Hypoxylon sp. EC38	63	65	29	12	16	2	3	0	3	0
Xylona heveae	25	27	9	5	8	1	2	0	2	0

PKS (polyketide synthase), Nrps (nonribosomal peptide-synthase), Pks-nrps (hybrid PKS/NRPS), Terpene (terpene synthase), DMATS (dimethylallyltryptophan synthase). Number of SM clusters predicted by antismash 2.0.

(de Wit *et al.* 2012). Similarly, significant expansions of enzymes involved in the cleavage of hemicellulose residues have been found in saprotrophic and plant pathogenic fungi, especially in trunk pathogens (Morales-Cruz *et al.* 2015). A contraction of hemicellulolytic and pectin depolymerizing enzymes in X. *heveae* could be a reflection of its intercellular growth, expected from a true endophyte. CAFE also detected significant contraction in the CE4 (xylan esterases) CAZyme family. Carbohydrate esterases facilitate the action of GHs on complex polysaccharides and therefore the reduction of CE4 might be related to the low GH-potential of X. *heveae* to degrade large carbohydrates. No significant contractions in

genome of each species. Numbers at the internal nodes represent the predicted number of copies. The size of the circles is proportional to these numbers. CAZymes were grouped in substrates following Amselem *et al.* 2011. Taxa in the tips follow JGI abbreviation for the strains (Table S2). Coloured branches indicate a significant (*p*-values ≤ 0.05) expansion (red), or contraction (blue) of the targeted gene family. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

terms of substrate were found in *S. kochii*, suggesting that the losses of CAZymes in *X. heveae* could be adaptations to the endophytic habit.

The overall decrease in CAZymes in *X. heveae* is accompanied by an enrichment of few CAZymes families, such as GH15, GH32 and GH49, which are involved in the catabolism of low molecular weight carbohydrates such as fructose and sucrose, as well as larger molecules such as dextran and starch (Akeboshi *et al.* 2004; Parrent *et al.* 2009; Wan *et al.* 2012). CAFE analysis detected an expansion of GH32 but not for GH15 or GH49. In addition, CAFE detected the expansion of the GH71 family (also called mutanases), a group of glucanases likely involved in α-1,3-glucan degradation. In *Trichoderma harzianum*, α-1,3-glucanases have been shown to degrade the cell wall of plant pathogenic fungi (Ait-Lahsen *et al.* 2001; Sanz *et al.* 2005) and extracellular polysaccharides produced by bacteria (Fuglsang *et al.* 2000). Whether *X. heveae* uses GH71 for the degradation of their own cell walls, allowing remodelling, or the cell walls of antagonistic bacteria or fungi remains unknown. Another expansion detected by CAFE was of CBM50 (= LysM) family, which is a binding domain that can sequester chitin molecules released as a consequence of hyphal remodelling. The latter would allow *X. heveae* to grow and colonize the intercellular space without being detected by its host.

Our results are similar to those of Gianoulis *et al.* (2012) who found a lower number of GHs in the genome of the endophytic *Ascocoryne sarcoides* (Leotiomycetes) in comparison to other non-endophytic filamentous Ascomycota. In contrast, Wang *et al.* (2015) reported a high content of CAZymes in the genome of the endophytic *Pestalotiopsis fici* (Sordariomycetes, Ascomycota). Furthermore, their clustering analysis based on the composition and abundance of CAZymes, grouped *P. fici* with aggressive plant pathogens such as *F. oxysporum* and *Fusarium verticilloides*. Similarly, Xu *et al.* (2014) found that the mutualistic endophyte *Harpophora oryzae* (Sordariomycetes, Ascomycota) contained a higher number of CAZymes compared to its plant pathogenic relatives, including the aggressive rice pathogen *Magnaporthe oryzae*, as well as *Gaeumannomyces graminis* and *Magnaporthe poae*. The diversity of CAZymes in these endophytes underscores their polyphyletic nature.

Secondary metabolites are important compounds for the colonization of hosts by fungi and important players in the host-pathogen interaction interface (Keller *et al.* 2005; Yu & Keller 2005; Tsuge *et al.* 2013). Many endophytes have been reported to protect their hosts through the production of secondary metabolites that act against insects and other fungi. Examples of this are species of the vertically transmitted endophyte genus *Epichlöe/Neotyphodium*, which produce a range of compounds that are insect deterrents such as indole-diterpenes and loline alkaloids (Schardl *et al.* 2012; 2013). Horizontally transmitted endophytes have also been found to produce alkaloids that have a role in protecting their host against insect herbivory (Sumarah *et al.* 2010) and fungal pathogens (Sumarah *et al.* 2011).

The genome of *X. heveae* presents significantly fewer secondary metabolite clusters than the other five endophytes included in the analysis and to endophytic genomes reported elsewhere (Xu *et al.* 2014; Wang *et al.* 2015). AntiSMASH predicted five terpene clusters and two Nrps-terpene hybrid clusters form which only one contained a secretion signal (Xylhe285165). The presence of terpene-coding biosynthetic genes does not necessarily imply the production of antifungal or insect deterrents since this class of secondary metabolites are also involved in primary metabolic pathways (e.g., ergosterol biosynthesis). *Xylona heveae* represents an evolutionary distinct lineage, and therefore we did not expect that if it were able to produce and secrete secondary metabolites involved in host protection (insect deterrent or antibiotics), these would be closely related to the ones produced by endophytic clavicipitaceous fungi. All of the biosynthetic genes had homologues in the genomes of other Pezizomycotina (cluster 'RunPezizomycotina UM 2015'; http://genome.jgi.doe.gov/clm/run/pezizomycotina-um-2015) and most of them had highest similarities to proteins present in Eurotiomycetes (results from BLAST searches against NCBI protein database).

The number of ABC transporter genes in *X. heveae* is lower than the average for Pezizomycotina. Since ABC transporters are involved in the protection against plant defence compounds (Del Sorbo *et al.* 2000; Schoonbeek *et al.* 2002) a lower copy number could suggest that *X. heveae*, in comparison to other plant-associated Ascomycota, does not require much protection from plant defence compounds, perhaps due to its low eliciting potential. The copy number of MFS, which could be involved in the secretion of SSP, was comparable to that in other fungal taxa. MFS transporters are also known to have specificity to certain carbohydrates and we found three copies of the SpSUT1 known for its high sucrose affinity (Talbot 2010; Doidy *et al.* 2012).

Ecological role of Xylona heveae *and its potential relationship with its host*

Based on its genome content, *X. heveae* could represent a true endophyte (= a commensal that does not decrease the fitness of its host and cannot switch to a different lifestyle). Experimental studies have shown that pathogenic species present a greater hydrolysis profile than non-pathogenic species and that most of these hydrolytic enzymes give the fungus the ability to access to the carbohydrates stored within plant cells (Amselem *et al.* 2011; King *et al.* 2011; Islam *et al.* 2012). The genome of *X. heveae* has few CAZymes involved in host colonization and plant cell wall degradation. Therefore, in contrast to the other endophytes included in this analysis, we hypothesize that *X. heveae* cannot become pathogenic. Also, based on its CAZymes content, and in contrast with the other endophytes included in this study, a potential to switch to a saprotrophic lifestyle after host senescence seems unlikely.

Given the reduced CAZyme profile of *X. heveae*, the lack of pectinases, and the low oxidative potential to degrade complex carbohydrates, we question how *X. heveae* is able to access nutrients while growing within the functional sapwood of its host. We hypothesize that *X. heveae* acts as a biotroph, feeding from the apoplastic washing fluid (AWF), which contains simple carbohydrates, such as glucose, fructose and sucrose (Sattelmacher 2001). This view is supported by the apparent expansion of GHs from families associated with the cleavage, transport and uptake of sucrose (i.e., GH32). Sucrose molecules are simple low weight sugars that can be

easily depolymerized and transported into the fungal cell. In most plants, sucrose is the major transportable product of photosynthesis that flows from the source organs to the sink organs. In this process, sucrose has to exit the mesophyll cells, pass through the apoplasm and enter the phloem (Kühn & Grof 2010; Reinders *et al.* 2012). The direct uptake of sucrose by *X. heveae* can be a strategy to avoid elicitation of plant defence mechanisms induced by glucose utilization (Rolland *et al.* 2006; Doidy *et al.* 2012).

The endophytic habit requires avoidance of the host's immune system, which can be accomplished in several ways and might involve a diverse array of proteins and signalling peptides. The plant host's immune system can be triggered by the release of CAZymes or by the components of the substrates attacked by these enzymes (i.e., products of the cleavage of cellulose or hemicellulose; Misas-Villamil & van der Hoorn 2008). Therefore, a reduction of CAZymes in *X. heveae* could be a result of the adaptation to an endophytic lifestyle. In addition, the genome of *X. heveae* presents an expanded copy number of enzymes that could play a role in allowing its growth and survival in the plant's intercellular spaces, avoiding the host's immune system. For instance, we found three copies of GH13_GT5 hybrids (= GH13-22 group) within its genome, which are enzymes that have a α-1,3-glucan synthase activity (Stam *et al.* 2006). Certain fungi have been shown to accumulate α-1,3-glucan in their cell surface as a way of masking recognizable pathogenicity factors (Fujikawa *et al.* 2012). Furthermore, CAFE detected an expansion of the LysM domain, which is a binding domain that can sequester the chitin molecules that are released as a consequence of hyphal remodelling. Even though the patterns of expansions and contractions of CAZymes found in *X. heveae* correspond to what we would expect in the genome of a true endophyte, these patterns were also found in other fungal lineages with diverse lifestyles. Therefore, we cannot propose specific changes in CAZymes content that could be related to the rise of a true endophytic lineage, but it appears to be a combination of reductions and expansions in CAZymes related to the degradation of plant cell wall and to the avoidance of the host's immune system, respectively.

The lack of cutinases, pectinases and the low copy number of enzymes needed for PCW degradation, suggest that *X. heveae* is unable to invade its host through direct penetration. Alternative pathways of host colonization include natural openings (i.e., stomata or lenticels) or through the root's vascular system. Nevertheless, based on the low potential of *X. heveae* to live as a saprotroph in the soil ecosystem and its close phylogenetic relationship with the beetle endosymbionts *Symbiotaphrina kochii* and *Symbiotaphrina buchneri*, we speculate that *X. heveae* could have an insect associated transmission mode.

Acknowledgements

We are grateful to Betsy Arnold and Kayla Arendt for providing a pure culture of *X. heveae*. We thank Daniele Armaleo, Gohn Glanden, Paul Dyer and Christopher Schardl for granting us permission to use genomic data from *C. grayi*, xylariaceous endophytes, *X. parietina* and *E. festucae*, respectively. We also thank Pedro Crous and Cony Decock for providing a culture of *T. guianense* in addition to collection data and to Dimitrios Floudas and László Nagy for their very helpful comments on this article. R. Gazis worked on this article while a postdoctoral fellow in the Open Tree of Life project, supported by the NSF (Grant # DEB-12008809). 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.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.funbio.2015.10.002.

REFERENCES

Ait-Lahsen H, Soler A, Rey M, de la Cruz J, Monte E, Llobell A, 2001. An antifungal exo-α-1,3-glucanase (AGN13.1) from the biocontrol fungus *Trichoderma harzianum*. *Applied and Environmental Microbiology* **67**: 5833–5839.

Akeboshi H, Tonozuka T, Furukawa T, Ichikawa K, Aoki H, Shimonishi A, Nishikawa A, Sakano Y, 2004. Insights into the reaction mechanism of glycosyl hydrolase family 49. *European Journal of Biochemistry* **271**: 4420–4427.

Aly AH, Debbab A, Proksch P, 2013. Fungal endophytes–secret producers of bioactive plant metabolites. *Die Pharmazie–An International Journal of Pharmaceutical Sciences* **68**: 499–505.

Amselem J, Cuomo CA, van Kan JA, Viaud M, Benito EP, Couloux A, Coutinho PM, de Vries RP, Dyer PS, Fillinger S, *et al.*, 2011. Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and Botrytis cinerea. *PLoS Genetics* **7**: e1002230.

Arnold AE, Lutzoni F, 2007. Diversity and host range of foliar fungal endophytes: are tropical leaves biodiversity hotspots? *Ecology* **88**: 541–549.

Bayry J, Aimanianda V, Guijarro JI, Sunde M, Latge JP, 2012. Hydrophobins—unique fungal proteins. *PLoS Pathogens* **8**: e1002700.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S, 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **340**: 783–795.

Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T, 2013. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Research* **41**: W204–W212.

Cannon PF, Simmons CM, 2002. Diversity and host preference of leaf endophytic fungi in the Iwokrama Forest Reserve, Guyana. *Mycologia* **94**: 210–220.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B, 2009. The carbohydrate-active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Research* **37**: D233–D238.

Castresana J, 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**: 540.

Chaverri P, Gazis RO, Samuels GJ, 2011. *Trichoderma amazonicum*, a new endophytic species on *Hevea brasiliensis* and *H. guianensis* from the Amazon basin. *Mycologia* **103**: 139–151.

Chen H, Kovalchuk A, Keriö S, Asiegbu FO, 2013. Distribution and bioinformatic analysis of the cerato-platanin protein family in Dikarya. *Mycologia* **105**: 1479–1488.
Chengalroyen MD, Dabbs ER, 2013. The biodegradation of latex rubber: a minireview. *Journal of Polymers and the Environment* **21**: 874–880.
Crous PW, Summerell BA, Shivas RG, Burgess TI, Decock CA, Dreyer L, Granke LL, Guest DI, Hardy GE, Hausbeck MK, Hüberli D, Jung T, Koukol O, Lennox CL, *et al.*, 2012. Fungal Planet description sheets: 107–127. *Persoonia: Molecular Phylogeny and Evolution of Fungi* **28**: 138.
de Wit PJ, van der Burgt A, Ökmen B, Stergiopoulos I, Abd-Elsalam KA, Aerts A, Bahkali AH, Beenen HG, Chettri P, Cox MP, *et al.*, 2012. The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PLoS Genetics* **8**: e1003088.
Del Sorbo G, Schoonbeek HJ, De Waard MA, 2000. Fungal transporters involved in efflux of natural toxic compounds and fungicides. *Fungal Genetics and Biology* **30**: 1–15.
Delaye L, García-Guzmán G, Heil M, 2013. Endophytes *versus* biotrophic and necrotrophic pathogens—are fungal lifestyles evolutionarily stable traits? *Fungal Diversity* **60**: 125–135.
Doidy J, Grace E, Kühn C, Simon-Plas F, Casieri L, Wipf D, 2012. Sugar transporters in plants and in their interactions with fungi. *Trends in Plant Science* **17**: 413–422.
Dyer PS, O'Gorman CM, 2012. Sexual development and cryptic sexuality in fungi: insights from *Aspergillus* species. *FEMS Microbiology Reviews* **36**: 165–192.
Eastwood DC, Floudas D, Binder M, Majcherczyk A, Schneider P, Aerts A, Asiegbu FO, Baker SE, Barry K, *et al.*, 2011. The plant cell wall–decomposing machinery underlies the functional diversity of forest fungi. *Science* **333**: 762–765.
Enright AJ, Van Dongen S, Ouzounis CA, 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research* **30**: 1575–1584.
Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS, *et al.*, 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **336**: 1715–1719.
Floudas D, Held BW, Riley R, Nagy LG, Koehler G, Ransdell AS, Younus H, Chow J, Chiniquy J, Lipzen A, *et al.*, 2015. Evolution of novel wood decay mechanisms in Agaricales revealed by the genome sequences of *Fistulina hepatica* and *Cylindrobasidium torrendii*. *Fungal Genetics and Biology* **76**: 78–92.
Fuglsang CC, Berka RM, Wahleithner JA, Kauppinen S, Shuster JR, Rasmussen G, Halkier T, Dalbøge H, Henrissat B, 2000. Biochemical analysis of recombinant fungal mutanases a new family of α1, 3-glucanases with novel carbohydrate-binding domains. *Journal of Biological Chemistry* **275**: 2009–2018.
Fujikawa T, Sakaguchi A, Nishizawa Y, Kouzai Y, Minami E, Yano S, Koga H, Meshi T, Nishimura M, 2012. Surface α-1,3-glucan facilitates fungal stealth infection by interfering with innate immunity in plants. *PLoS Pathogens* **8**: e1002882.
Gaulin E, Madoui MA, Bottin A, Jacquet C, Mathé C, Couloux A, Wincker P, Dumas B, 2008. Transcriptome of *Aphanomyces euteiches*: new oomycete putative pathogenicity factors and metabolic pathways. *PLoS One* **3**: e1723.
Gazis R, Miadlikowska J, Lutzoni F, Arnold AE, Chaverri P, 2012. Culture-based study of endophytes associated with rubber trees in Peru reveals a new class of Pezizomycotina: Xylonomycetes. *Molecular Phylogenetics and Evolution* **65**: 294–304.
Gianoulis TA, Griffin MA, Spakowicz DJ, Dunican BF, Sboner A, Sismour AM, Kodira C, Egholm M, Church GM, *et al.*, 2012. Genomic analysis of the hydrocarbon-producing, cellulolytic, endophytic fungus *Ascocoryne sarcoides*. *PLoS Genetics* **8**: e1002558.
Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Halla G, Shea TP, Sykes S, *et al.*, 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 1513–1518.
Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, *et al.*, 2014. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research* **42**: D699–D704.
Han MV, Thomas GW, Lugo-Martinez J, Hahn MW, 2013. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Molecular Biology and Evolution* **30**: 1987–1997.
Hess J, Skrede I, Wolfe BE, LaButti K, Ohm RA, Grigoriev IV, Pringle A, 2014. Transposable element dynamics among asymbiotic and ectomycorrhizal *Amanita* fungi. *Genome Biology and Evolution* **6**: 1564–1578.
Higgins KL, Coley PD, Kursar TA, Arnold AE, 2011. Culturing and direct PCR suggest prevalent host generalism among diverse fungal endophytes of tropical forest grasses. *Mycologia* **103**: 247–260.
Islam MS, Haque MS, Islam MM, Emdad EM, Halim A, Hossen QMM, Hossain MZ, Ahmed B, Rahim S, Rahman MS, *et al.*, 2012. Tools to kill: genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina*. *BMC Genomics* **13**: 493.
Katoh K, Standley DM, 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.
Kelkar YD, Ochman H, 2012. Causes and consequences of genome expansion in fungi. *Genome Biology and Evolution* **4**: 13–23.
Keller NP, Turner G, Bennett JW, 2005. Fungal secondary metabolism—from biochemistry to genomics. *Nature Reviews Microbiology* **3**: 937–947.
King BC, Waxman KD, Nenni NV, Walker LP, Bergstrom GC, Gibson DM, 2011. Arsenal of plant cell wall degrading enzymes reflects host preference among plant pathogenic fungi. *Biotechnology for Biofuels* **4**: 1–14.
Kirk PM, Cannon PF, Minter DW, Stalpers JA, 2008. *Dictionary of the Fungi*. CABI International, Wallingford, UK.
Kogel KH, Franken P, Hückelhoven R, 2006. Endophyte or parasite–what decides? *Current Opinion in Plant Biology* **9**: 358–363.
Kohler Annegret, Kuo Alan, Nagy Laszlo G, Morin Emmanuelle, Barry Kerrie W, Buscot Francois, Canbäck Björn, *et al.*, 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nature Genetics* **47**: 410–415.
Krogh A, Larsson B, Von Heijne G, Sonnhammer EL, 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* **305**: 567–580.
Kühn C, Grof CP, 2010. Sucrose transporters of higher plants. *Current Opinion in Plant Biology* **13**: 287–297.
Kuo A, Bushnell B, Grigoriev IV, 2014. Fungal genomics: sequencing and annotation. *Advances in Botanical Research* **70**: 1–52.
Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B, 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for Biofuels* **6**: 41.
Li XQ, Du D, 2014. Variation, evolution, and correlation analysis of C+G content and genome or chromosome size in different kingdoms and phyla. *PloS One* **9**: e88339.
Linder MB, Szilvay GR, Nakari-Setälä T, Penttilä ME, 2005. Hydrophobins: the protein-amphiphiles of filamentous fungi. *FEMS Microbiology Reviews* **29**: 877–896.

Lombard V, Golaconda RH, Drula E, Coutinho PM, Henrissat B, 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research* **42**: D490–D495.
Martin J, Bruno VM, Fang Z, Meng X, Blow M, Zhang T, Sherlock G, Snyder M, Wang Z, 2010. Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. *BMC Genomics* **11**: 663.
McCutcheon JP, Moran NA, 2012. Extreme genome reduction in symbiotic bacteria. *Nature Reviews Microbiology* **10**: 13–26.
Misas-Villamil JC, Van der Hoorn RA, 2008. Enzyme–inhibitor interactions at the plant–pathogen interface. *Current Opinion in Plant Biology* **11**: 380–388.
Morales-Cruz A, Amrine KCH, Blanco-Ulate B, Lawrence DP, Travadon R, Rolshausen PE, Baumgartner K, Cantu D, 2015. Distinctive expansion of gene families associated with plant cell wall degradation, secondary metabolism, and nutrient uptake in the genomes of grapevine trunk pathogens. *BMC Genomics* **16**: 469.
Motteram J, Küfner I, Deller S, Brunner F, Hammond-Kosack KE, Nürnberger T, Rudd JJ, 2009. Molecular characterization and functional analysis of MgNLP, the sole NPP1 domain-containing protein, from the fungal wheat leaf pathogen *Mycosphaerella graminicola*. *Molecular Plant-Microbe Interactions* **22**: 790–799.
Muszewska A, Hoffman-Sommer M, Grynberg M, 2011. LTR retrotransposons in fungi. *PLoS One* **6**: e29425.
Ohm RA, Feau N, Henrissat B, Schoch CL, Horwitz BA, Barry KW, Condon BJ, Copeland AC, Dhillon B, Glaser F, *et al.*, 2012. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLoS Pathogens* **8**: e1003037.
Oses R, Valenzuela S, Freer J, Baeza J, Rodríguez J, 2006. Evaluation of fungal endophytes for lignocellulolytic enzyme production and wood biodegradation. *International Biodeterioration and Biodegradation* **57**: 129–135.
Parrent JL, James TY, Vasaitis R, Taylor AF, 2009. Friend or foe? Evolutionary history of glycoside hydrolase family 32 genes encoding for sucrolytic activity in fungi and its implications for plant-fungal symbioses. *BMC Evolutionary Biology* **9**: 148.
Promputtha I, Hyde KD, McKenzie EH, Peberdy JF, Lumyong S, 2010. Can leaf degrading enzymes provide evidence that endophytic fungi becoming saprobes? *Fungal Diversity* **41**: 89–99.
Quilliam RS, Jones DL, 2012. Evidence for host-specificity of culturable fungal root endophytes from the carnivorous plant *Pinguicula vulgaris* (common butterwort). *Mycological Progress* **11**: 583–585.
Reinders A, Sivitz AB, Ward JM, 2012. Evolution of plant sucrose uptake transporters. *Frontiers in Plant Science* **3**: 22.
Rodriguez RJ, White Jr JF, Arnold AE, Redman RS, 2009. Fungal endophytes: diversity and functional roles. *New Phytologist* **182**: 314–330.
Rolland F, Baena-Gonzalez E, Sheen J, 2006. Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annual Review of Plant Biology* **57**: 675–709.
Rose K, Steinbüchel A, 2005. Biodegradation of natural rubber and related compounds: recent insights into a hardly understood catabolic capability of microorganisms. *Applied and Environmental Microbiology* **71**: 2803–2812.
Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, *et al.*, 2003. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**: 374–378.
Sanz L, Montero M, Redondo J, Llobell A, Monte E, 2005. Expression of an α-1,3-glucanase during mycoparasitic interaction of *Trichoderma asperellum*. *FEBS Journal* **272**: 493–499.
Sattelmacher B, 2001. The apoplast and its significance for plant mineral nutrition. *New Phytologist* **149**: 167–192.
Schardl CL, Young CA, Faulkner JR, Florea S, Pan J, 2012. Chemotypic diversity of epichloae, fungal symbionts of grasses. *Fungal Ecology* **5**: 331–344.
Schardl CL, Young CA, Hesse U, Amyotte SG, Andreeva K, Calie PJ, Fleetwood DJ, Haws DC, Moore N, Oeser B, *et al.*, 2013. Plant-symbiotic fungi as chemical engineers: multi-genome analysis of the clavicipitaceae reveals dynamics of alkaloid loci. *PLoS Genetics* **9**: e1003323.
Schoch CL, Sung GH, López-Giráldez F, Townsend JP, Miadlikowska J, Hofstetter V, Robbertse B, Matheny PB, Kauff F, Wang Z, *et al.*, 2009. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. *Systematic Biology* **58**: 224–239.
Schoonbeek HJ, Raaijmakers JM, De Waard MA, 2002. Fungal ABC transporters and microbial interactions in natural environments. *Molecular Plant-Microbe Interactions* **15**: 1165–1172.
Schulz B, Boyle C, 2005. The endophytic continuum. *Mycological Research* **109**: 661–686.
Silvestro D, Michalak I, 2012. RaxmlGUI: a graphical front-end for RAxML. *Organisms Diversity and Evolution* **12**: 335–337.
Smit AFA, Hubley R, Green P, 2004. *RepeatMasker Open-3.0*. http://www.repeatmasker.org.
Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, van Themaat EVL, Brown JKM, Butcher SA, Gurr SJ, *et al.*, 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330**: 1543–1546.
Spatafora JW, Sung GH, Sung JM, Hywel-Jones NL, White JF, 2007. Phylogenetic evidence for an animal pathogen origin of ergot and the grass endophytes. *Molecular Ecology* **16**: 1701–1711.
Sprockett DD, Piontkivska H, Blackwood CB, 2011. Evolutionary analysis of glycosyl hydrolase family 28 (GH28) suggests lineage-specific expansions in necrotrophic fungal pathogens. *Gene* **479**: 29–36.
Stam MR, Danchin EG, Rancurel C, Coutinho PM, Henrissat B, 2006. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of α-amylase-related proteins. *Protein Engineering Design and Selection* **19**: 555–562.
Stamatakis A, 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688.
Strobel G, Daisy B, 2003. Bioprospecting for microbial endophytes and their natural products. *Microbiology and Molecular Biology Reviews* **67**: 491–502.
Sumarah MW, Kesting JR, Sørensen D, Miller JD, 2011. Antifungal metabolites from fungal endophytes of Pinus strobus. *Phytochemistry* **72**: 1833–1837.
Sumarah MW, Puniani E, Sørensen D, Blackwell BA, Miller JD, 2010. Secondary metabolites from anti-insect extracts of endophytic fungi isolated from *Picea rubens*. *Phytochemistry* **71**: 760–765.
Suryanarayanan TS, Thirunavukkarasu N, Govindarajulu MB, Sasse F, Jansen R, Murali TS, 2009. Fungal endophytes and bioprospecting. *Fungal Biology Reviews* **23**: 9–19.
Talbot NJ, 2010. Living the sweet life: how does a plant pathogenic fungus acquire sugar from plants? *PLoS Biology* **8**: e1000308.
Tsuge T, Harimoto Y, Akimitsu K, Ohtani K, Kodama M, Akagi Y, Egusa M, Yamamoto M, Otani H, 2013. Host-selective toxins produced by the plant pathogenic fungus *Alternaria alternata*. *FEMS Microbiology Reviews* **37**: 44–66.
Tudzynski P, Sharon A, 2003. 9-Fungal pathogenicity genes. *Applied Mycology and Biotechnology* **3**: 187–212.
Unterseher M, Gazis R, Chaverri P, Guarniz CFG, Tenorio DHZ, 2013. Endophytic fungi from Peruvian highland and lowland habitats form distinctive and host plant-specific assemblages. *Biodiversity and Conservation* **22**: 999–1016.

van Kan JA, 2006. Licensed to kill: the lifestyle of a necrotrophic plant pathogen. *Trends in Plant Science* **11**: 247–253.

Wan J, Li Y, Yi R, Mongkolthanaruk W, Kinjo Y, Terashita T, Yamanaka K, Shimomura N, Yamaguchi T, Aimi T, *et al.*, 2012. Characterization of the glycoside hydrolase family 15 glucoamylase gene from the ectomycorrhizal basidiomycete *Tricholoma matsutake*. *Mycoscience* **53**: 194–202.

Wang X, Zhang X, Liu L, Xiang M, Wang W, Sun X, Che Y, Guo L, Liu G, Guo L, *et al.*, 2015. Genomic and transcriptomic analysis of the endophytic fungus *Pestalotiopsis fici* reveals its lifestyle and high potential for synthesis of natural products. *BMC Genomics* **16**: 28.

Xu XH, Su ZZ, Wang C, Kubicek CP, Feng XX, Mao LJ, Wang JY, Chen C, Lin FC, *et al.*, 2014. The rice endophyte *Harpophora oryzae* genome reveals evolution from a pathogen to a mutualistic endophyte. *Scientific Reports* **4**.

Yu JH, Keller N, 2005. Regulation of secondary metabolism in filamentous fungi. *Annual Review of Phytopathology* **43**: 437–458.