The fungus that came in from the cold: Dry rot’s pre-adapted ability to invade buildings

S.V. Balasundaram¹, J. Hess¹, M. B. Durling², S. C. Moody³, L. Thorbek¹, C. Progida¹, K. LaButti⁴, A. Aerts⁴, K. Barry⁴, I. V. Grigoriev⁴, L. Boddy⁵, N. Höglberg², H. Kauserud¹, D. C. Eastwood³, I. Skrede¹*

¹Department of Biosciences, University of Oslo, Oslo, Norway; ²Department of Forest Mycology, Swedish Agricultural University, Uppsala, Sweden; ³Department of Biosciences, Swansea University, Swansea, UK; ⁴United States Department of Energy Joint Genome Institute, Walnut Creek, CA, USA; ⁵School of Biosciences, Cardiff University, Cardiff, UK;

Correspondence and request for materials should be addressed to I.S. (inger.skrede@ibv.uio.no)
Abstract

Many organisms benefit from being pre-adapted to niches shaped by human activity, and have successfully invaded man-made habitats. One such species is the dry-rot fungus *Serpula lacrymans*, which has a wide distribution in buildings in temperate and boreal regions, where it decomposes coniferous construction wood. Comparative genomic analyses and growth experiments using this species and its wild relatives revealed that *S. lacrymans* evolved a very effective brown rot decay compared to its wild relatives, enabling an extremely rapid decay in buildings under suitable conditions. Adaptations in intracellular transport machineries promoting hyphal growth, and nutrient and water transport may explain why it is has become a successful invader of timber in houses. Further, we demonstrate that *S. lacrymans* has poor combative ability in our experimental set-up, compared to other brown rot fungi. In sheltered indoor conditions, the dry rot fungus may have limited encounters with other wood decay fungi compared to its wild relatives. Overall, our analyses indicate that the dry rot fungus is an ecological specialist with poor combative ability against other fungi.
Species worldwide are negatively affected by anthropogenic habitat destruction. However, for those few species originally living in natural habitats that resemble the man-made ecosphere, the opposite is also the case. Animals like the Norwegian rat (*Rattus norvegicus*) and the German cockroach (*Blatella germanica*) have extended their distribution dramatically (Robinson, 1965; Nentwig, 2008). Likewise, many plant pathogenic fungi have become extremely widespread as monotypic crop cultivation creates large habitats, and the trade and transport of these crops aid their dispersal (Anderson *et al.*, 2004; Grunwald *et al.*, 2008; Stukenbrock *et al.*, 2011). A similar pattern is seen with the few wood decay fungi that have expanded their realized niche into the human built environment.

Probably the best-known example of a successful fungal invader of the built environment is the dry rot fungus *Serpula lacrymans* var. *lacrymans* (subsequently referred to as var. *lacrymans*), which is distributed in houses in temperate and boreal regions worldwide causing brown rot decay. It spreads with human transport of timber over long distances and colonizes new buildings in its vicinity by air-borne spores (Kauserud *et al.*, 2007; 2012). Colonization of construction timber in buildings is characterised by rapid vegetative mycelial growth and formation of thick (up to 2 cm diameter, Figure 1) mycelial cords that mediate the transport of nutrition and water to new wood substrates (Jennings and Bravery, 1991). This allows quick growth and reallocation of resources via the transport of nutrition and water to the new wood substrates (Jennings and Bravery, 1991; Boddy *et al.*, 2007).

Comparative genomic approaches have shown that var. *lacrymans* and other brown rot fungi have a reduced set of plant cell wall hydrolysing enzymes to decompose wood compared to the ancestral white-rot fungi (Eastwood *et al.*, 2011; Arantes and Goodell,
A recent study has suggested that the set of secreted enzymes responsible for decomposition of var. *lacrymans* is even smaller than that of some other brown rot fungi (Presley and Schilling, 2017). The loss of enzymes by brown rot fungi is correlated with a strategy in which the initial attack of the wood is mediated by hydroxyl radicals produced by chelator-mediated Fenton (CMF) chemistry (Eastwood *et al*., 2011; Floudas *et al*., 2012; Riley *et al*., 2014). These initial attacks have been suggested to be controlled by differential gene expression of the fungi (Zhang *et al*., 2016; Presley and Schilling, 2017). The attacked wood structure is then further depolymerised by oxidising and hydrolysing enzymes that target cellulose and hemicellulose elements in the wood, while leaving modified lignin behind.

Var. *lacrymans* has a scattered natural range in high altitude mountain regions of North-East Asia, thriving in moraine-dominated habitats around the treeline where woody resources are heterogeneously distributed (Kauserud *et al*., 2012). Human transport of infected wood appears to have facilitated the colonization in the human domain in temperate regions world-wide. It is widespread in buildings in Europe and Japan, and it is also found in buildings in temperate parts of North and South America (Chile), Australia and New Zealand, but with less abundance (White *et al*., 2001; Kauserud *et al*., 2007). The large European house-colonizing population of var. *lacrymans* has low genetic variation (Kauserud *et al*., 2007; Skrede *et al*., 2013), suggesting a severe population bottleneck during the colonisation of the European built environment (Kauserud *et al*., 2012; 2007).

*Serpula lacrymans* var. *shastensis* (subsequently referred to as var. *shastensis*) is a close relative of var. *lacrymans*, from high altitude mountain regions in the Cascade mountain range (North America), but has not been reported in the built environment (Harmsen, 1960; Palfreyman *et al*., 2003). Although genetically well-separated, the two sub-species are able to form a dikaryotic mycelium when paired *in vitro* (Palfreyman *et al*., 2003; Kauserud et
In the habitat close to the treeline in the Cascades (Figure 1), var. *shastensis* colonizes and decays large logs of *Abies magnifica* (Kauserud *et al.*, 2004; 2012). Both varieties of *S. lacrymans* appear to be ecological specialists, thriving in exposed mountainous habitats with patchy resource distribution.

In contrast to the confined niches of *S. lacrymans*, its sister species *Serpula himantioides* has a widespread circumboreal distribution in natural habitats in temperate and boreal regions (Carlsen *et al.*, 2011). As with *S. lacrymans*, *S. himantioides* causes brown rot of conifers, but decomposes wood more slowly, as shown on spruce (Harmsen, 1960), and produces smaller fruit bodies and smaller cords. *Serpula himantioides* is rarely found in buildings, and when it is, it decomposes wood more slowly than var. *lacrymans*. Unlike var. *lacrymans*, indoor colonization by *S. himantioides*, as with the majority of other wood-decay fungi, represent random, and repeated colonisations from nature (Kauserud *et al.*, 2012).

It is not evident which characteristics have made var. *lacrymans* such a successful invader of the built environment compared to its wild relatives. Pinpointing contrasting genomic differences among the lineages is a first step towards detecting the genetic basis of var. *lacrymans* invasiveness and persistence. In this study we, therefore, set out to reveal which genomic features separate var. *lacrymans* from its predominantly wild relatives. We analysed which genes have undergone shifts in selective pressure and then, which gene families have expanded or contracted during divergence between variants or species. This was achieved by sequencing and *de novo* genome assembly of var. *lacrymans* and var. *shastensis* strains and comparing these to the genome of the sister species *S. himantioides*.

Genomic analyses were complemented by two growth experiments investigating differences in decomposition ability and interspecific competition, to provide more direct evidence for how each of these factors may contribute to var. *lacrymans*’ success in the built environment.
Materials and Methods

Strains. Three strains were used for physiological experiments and genome comparisons in this study. The *S. himantioides* strain (MUCL38935) was cultured from soil in the UK in 1994, the var. *shastensis* strain (SHA17-1) was collected in California, US on *Abies* in 2004 and the var. *lacrymans* (SL200) was collected from a house in Poland in 1953. Since these strains have been maintained in culture for extended periods of time, caution should be used when interpreting the results as the strains may have changed their behaviour through these years.

DNA extraction, sequencing assembly and gene predictions. More details of the DNA extraction, library preparation, sequencing procedure, and gene prediction pipeline can be found in the Supplementary text. DNA of all three strains was extracted by a modified phenol-chloroform protocol available at the JGI webpage ([http://jgi.doe.gov/collaborate-with-jgi/pmo-overview/protocols-sample-preparation-information/](http://jgi.doe.gov/collaborate-with-jgi/pmo-overview/protocols-sample-preparation-information/)). All strains were sequenced using Illumina technology. The two *Serpula lacrymans* strains were sequenced on a Illumina GAII at the SNP&SEQ Technology Platform in Uppsala, Sweden, while *S. himantioides* was sequenced on a Illumina Hiseq 2000 at the JGI ([http://genome.jgi.doe.gov/Serla_varsha1/Serla_varsha1.info.html](http://genome.jgi.doe.gov/Serla_varsha1/Serla_varsha1.info.html)).

The Velvet *de novo* assembler (Zerbino and Birney, 2008) was used to assemble reads into contigs for var. *lacrymans* and var. *shastensis*. JGI assembled *S. himantioides* with the AllPathsLG assembler (Gnerre *et al.* 2011). The CEGMA pipeline was used (Parra *et al.*,...
2007) to estimate completeness of all assemblies (Table 1). Protein coding genes in the three *Serpula* strains were annotated using MAKER2 version 2.27 (Holt and Yandell, 2011).

**Functional annotation.** Genes were given a preliminary description by BLAST alignment towards Uniprot. InterProScan was used for functional annotation and classifications of protein families (Jones *et al.*, 2014). Protein sequences of var. *lacrymans*, var. *shastensis* and *S. himantioides* were obtained from the MAKER2 predictions.

**OrthoMCL clustering.** Homologous proteins of the three *Serpula* strains were clustered using the software OrthoMCL (Li *et al.*, 2003). This tool clusters homologous proteins across the given species using Markov cluster algorithm to group orthologs and paralogs. In total 34,273 protein sequences from three different *Serpula* genomes were compared.

**CAFÉ analysis.** CAFÉ estimates a global birth and death rate of gene families and changes in gene family size across a phylogeny (De Bie *et al.*, 2006). All orthoMCL clusters were used as gene families. CAFÉ was run using a global birth/death parameter ($\lambda$). Rapidly evolving gene families were estimated using the best fit $\lambda$ (0.002) at a p-value threshold of 0.01. The ultrametric tree used for CAFÉ analysis was based on a multi-locus maximum likelihood phylogeny of ten loci from (Balasundaram *et al.*, 2015) that was made ultrametric in the R package APE (Paradis *et al.*, 2004).
Selection pressure. Clusters of single copy orthologs were chosen to screen for branch specific changes in selection pressure. The clusters were aligned with the multiple sequence alignment program PRANK (Loytynoja, 2014) with the ‘codon’ alignment mode, using the species phylogeny (Skrede et al., 2011) as guide tree. PRANK has been shown to provide the most accurate alignments, with the lowest false-positive rates (Fletcher and Yang, 2010). The Codeml from the PAML package (Yang, 2007) was used to identify changes in selection regime. For each group of orthologs, a single dN/dS ratio (\(\omega\)) was estimated for all branches on the tree \(H_0\) and for three instances where each one of the species was allowed to evolve at a separate rate \(H_1\), The best fit model was determined using a likelihood ratio test and p-values were adjusted to control the false discovery rate (FDR) for multiple hypothesis testing using a \(\alpha < 0.05\) (Benjamini et al., 2006). All alignments with a significant shift in selection pressure between species were manually examined to remove questionable alignment regions if present and were then rerun in the above outlined analysis.

Functional enrichment analyses. Functional enrichment analysis was used to characterize the genes present in all the genomes compared to gene families that were inferred to be expanded or contracted by CAFÉ. A Python script was used to perform functional enrichment analysis of PFAM domains using Fisher’s exact test (http://cgrlucb.wikispaces.com/Functional+Enrichment+Analysis).

Annotation of genes of specific functions of interest. To predict the secretome of each species, a bioinformatics pipeline consisting of SignalP 4.1 (Petersen et al., 2011), TargetP 2.0 (Emanuelsson et al., 2007), TMHMM 2.0 (Krogh et al., 2001), PS_scan (Hulo, 2006) and WolfPSort v. 0.2 (Horton et al., 2007), was used, as implemented in Kohler et al. (Kohler et
al., 2015). Besides the annotations generated for the entire proteomes (e.g. CAZymes, PFAM domains), the proteolytic enzymes present in each secretome were also annotated through BLAST searches against the MEROPS database (Rawlings et al., 2014). Carbohydrate-active enzymes were predicted by searching predicted proteomes with the dbCAN tool (Yin et al., 2012; Lombard et al., 2014).

As cytochrome P450 (cytP450) is an important class of enzymes involved in specialized metabolism, the clusters annotated with cytP450 PFAM domains in Interproscan were manually curated. Only those of over 300 residues with both the EXXR and CXG motif were accepted as functional, according to the method of Syed and Mashele (2014). According to cytP450 nomenclature, a similarity of 40% was considered sufficient to classify a predicted protein into a particular family. A similarity of 55% would allow allocation to a sub-family. Those with <40% similarity to named cytP450s were – with those that had no significant matches in the NCBI or Uniprot databases – considered to probably belong to novel cytP450 families.

Data availability. All raw sequence reads, and assembled genomes are available on NCBI at Bioproject PRJNA412961. In addition, the S. himantioides MUCL38935 genome is available at the JGI genome browser (http://genome.jgi.doe.gov/Serla_varsha1/Serla_varsha1). The MAKER2 gene predictions, the OrthoMCL clusters and the alignments used as input to the Codeml analyses have been deposited in the Dryad Digital Repository: http://dx.doi.org/xxxx.
Combative ability. Var. lacrymans, which is found predominately, if not exclusively, inside houses in Europe, was hypothesized to show decreased ability to combat for limited resources since it faces few competitors in this environment. An antagonistic experiment was used to test this hypothesis, where the three Serpula strains of interest were confronted with each other and other brown rot decomposer fungi, pairwise, by growing two well-colonized blocks side by side (see supplementary text for detailed experimental setup). The three Serpula strains, and the three species Antrodia xantha, Coniophora puteana and Fomitopsis pinicola were used. All combinations were repeated 10 times. After the experiment, three small wood pieces from within the wood block were transferred to three new culture plates. The strains that were re-isolated from the wood piece were identified and reported. A Pearson’s \( \chi^2 \) Goodness of Fit test was used to test whether one species had significantly outcompeted another.

Wood decay. The specialized house-living var. lacrymans was expected to decompose spruce especially fast as it is mostly found on spruce in houses, where it is known to grow quickly (Harmsen, 1960). To compare the decomposition ability of the three Serpula strains and A. xantha, F. pinicola and C. puteana, mass loss of wood was determined after 60 days colonization at 20 °C on the three tree species Pinus sylvestris, Picea abies and Abies lasiocarpa (See supplementary text for experimental setup; the three non-Serpula species were not tested on Abies lasiocarpa). The significance of the differences in mass loss among strains and among wood species was tested with ANOVA analyses using R (R Core Team, 2008).
Results

*Genome summary.* The gene prediction pipeline identified a total of 11,352 gene models in var. *lacrymans*, 10,910 gene models in var. *shastensis* and 12,011 gene models in *S. himantioides* (Table 1). Annotated genes were clustered into gene families, of which 6,695 were shared among all three strains, corresponding to approximately 55% to 61% of annotated genes in each genome. Given the close relationship among the three species, the number of singleton clusters inferred for each species was surprisingly high. Of the predicted genes 18% in var. *shastensis*, 23% in var. *lacrymans* and 24% in *S. himantioides* were unique to each of the three lineages. Further analysis of singleton genes showed that singletons predominately represented cases where orthologs were absent in the other two species, either due to gene loss or absence of the corresponding coding region from the respective assemblies (results not shown).

*Analyses of selection.* The genome-wide estimates of selection yielded a mean estimate of $\omega = 0.137$ for *S. himantioides*, $\omega = 0.179$ for var. *lacrymans* and $\omega = 0.234$ for var. *shastensis* (gene clusters with $\omega > 2$ were omitted from these estimates).

Shifts in selective pressure on individual genes between species may pinpoint genes whose functions have contributed to adaptation by each species to their respective realised niches. For the analyses of shifts in selective pressure on a gene-by-gene basis, three series of tests were run, each one with a different species as the foreground branch. After correction for multiple testing, 100, 129, and 265 genes with significantly different $\omega$ between foreground and background branches in var. *lacrymans*, var. *shastensis* and *S. himantioides* were detected, respectively (Figure 2). Among the sets of genes, 43% were annotated with...
PFAM domains while the rest were unannotated. Our functional analyses were only focused on the genes that had PFAM annotations. A full list of significant genes is provided in the supplementary material (Supplementary Table 1).

One of the most pronounced functional signatures detected was the selective shift in many proteins involved in intracellular transport (Table 2) with an elevated $\omega$ in $S.\ lacrymans$ compared to $S.\ himantioides$ (higher $\omega$ in one or both of the $S.\ lacrymans$ varieties). Several of these proteins identified were involved in the transport of vesicles to the Golgi stack for secretion, (see supplementary text for details). In contrast, a protein involved in early endosomal membranes evolved faster in $S.\ himantioides$ than in $S.\ lacrymans$. This, suggested a faster evolution of an endocytic pathway in $S.\ himantioides$ versus an exocytic pathway in $S.\ lacrymans$.

In addition to the genes related to membrane transport, two regulators of actin polymerization (the guanine nucleotide exchange factors, Rho GEF and Ras GEF) and a gene with a role in actin depolymerization (cofilin) evolved significantly faster in var. $lacrymans$ than in $S.\ himantioides$ (Table 2).

Expansion and contraction of gene families. All clusters in the dataset and a rooted tree were used to infer 244 and 262 gene families that were expanded on the var. $lacrymans$ branch and on the var. $shastensis$ branch, respectively, compared to the rest of the tree (Table 2). Only 5 were expanded on the common branch leading to var. $lacrymans$ and var. $shastensis$. Compared to the genomic background, CAFÉ inferred 112 and 135 gene families that expanded significantly faster than expected (based on all clusters) in var. $lacrymans$ and var. $shastensis$, respectively (P-value 0.01). In turn, 596 and 473 gene families were contracted on the var. $lacrymans$ branch and on the var. $shastensis$ branch, respectively, and
332 were contracted on the common branch. Six (var. *lacrymans*) and four (var. *shastensis*) gene families showed significantly higher rates of contraction than the genomic background rate.

Functional enrichment of the expanded and contracted gene families demonstrated a change in copy number for gene families related to specialised metabolism amongst all three strains (Supplementary Table 2). In particular, expansions and contractions in a variety of polyketide synthase (PKS) and nonribosomal peptide synthase (NRPS) related PFAM domains were identified (Supplementary Table 2). One NRPS gene family (cluster 0012) was expanded in var. *lacrymans*, var. *shastensis* and their common branch. This gene family had nine gene copies in var. *shastensis* and var. *lacrymans*, but only one in *S. himantioides*. The opposite pattern was found for a putative PKS-NRPS hybrid protein gene family of unknown function (cluster 0005), where *S. himantioides* had ten copies, var. *lacrymans* eight and var. *shastensis* six copies. Copy number changes in ATP-binding cassette (ABC) transporters were also detected. These were reduced in var. *lacrymans* compared to var. *shastensis* and *S. himantioides*.

Cytochrome P450s showed expansion in *S. lacrymans* compared to *S. himantioides* (Supplementary Table 3). Eighty-nine, 91 and 109 predicted functional cytochrome P450s were identified in *S. himantioides*, var. *shastensis* and var. *lacrymans* respectively. Thus, both var. *shastensis* and var. *lacrymans* have experienced expansion of capacity compared to *S. himantioides*, with an extra five families represented in each. Var. *lacrymans* and var. *shastensis* had the same families except that var. *shastensis* uniquely had one member of CYP5145, and var. *lacrymans* had one member of CYP6001, a family that was not present in either of the other strains. Thus, in both var. *shastensis* and var. *lacrymans* the higher numbers of cytochrome P450 copies were predominantly the result of an increased number of genes from existing families.
Several gene families related to wood decay mechanisms were expanded or contracted (Figure 2; Supplementary Table 2, supplementary text for details). Specifically, the set of CAZymes encoded within the three genomes was very similar, but with a somewhat greater gene complement in *S. himantioides* (see supplementary text for details; Supplementary Table 4). In contrast, an iron reductase with only a CBM1 and a CytB domain was found in *S. lacrymans*, but not in *S. himantioides*. (Supplementary Figure 2).

**Evaluating substrate preference.** Both *S. lacrymans* varieties decomposed more of the spruce wood block than *S. himantioides*, under the conditions tested (50% and 45% vs. 30% mass loss, respectively; Figure 3). There was no significant difference in the amount of decomposition between var. *lacrymans* and var. *shastensis* on spruce or fir ($\chi^2$, p>0.05). Var. *shastensis* failed to grow on pine, but it is unknown whether this is due to its inability to decompose pine, or due to other experimental factors, e.g. the experimental set-up on moist perlite may not have provided enough minerals. Spruce was more readily degraded by all strains, and this was particularly pronounced for var. *lacrymans*, which caused a mass loss of 50% of spruce but only 5% of pine wood blocks. See supplementary material for the mass loss of the additional species (Supplementary Figure 1).

**Evaluating antagonistic behavior.** *Serpula himantioides* was significantly more combative than var. *lacrymans* and var. *shastensis*, as well as the three other brown rot species under the conditions tested (Table 3). *Serpula himantioides* was present in 79% of the re-isolations from the confrontations against the other species (i.e. as 50% would be a deadlock, *S. himantioides* took the substrate of the other species in 29% of the cases). The two *S. lacrymans* varieties were less able to exclude the other species compared to *S.*
*himantioides* in this experiment, (var. *shastensis* was found in 40% and var. *lacrymans* in 41% of the cultures following confrontations, i.e. both lost their substrate in about 10% of the cases). When var. *lacrymans* and var. *shastensis* were confronted with *C. puteana* and *A. xantha*, the outcomes were close to 50% (i.e. a deadlock), but both *S. lacrymans* varieties were excluded by *S. himantioides* and *F. pinicola* (Table 3).
In this study we aimed to identify which features have made the dry rot fungus *Serpula lacrymans* var. *lacrymans* the most successful invasive wood-decay fungus in the built environment by comparing its characteristics to its less invasive relatives. Since the successful establishment of an invasive species typically depends on a range of factors, we investigated the contribution of physiological factors (decomposition and combative ability), as well as underpinning genomic features. We detected numerous genomic signatures that may be linked to var. *lacrymans* invasiveness, including changes in selection pressure and evolution in gene families involved in hyphal growth, transportation, defence and decomposition of wood. Our experimental data suggest that *S. lacrymans* has poor antagonistic abilities towards other brown rot fungi, but that it has high wood-decomposition ability compared to its largely non-invasive relative *S. himantioides*. This suggests that *S. lacrymans* is an ecological specialist while *S. himantioides* is more of an ecological generalist.

One of our main findings is the differences in genome-wide selection pressure, evaluated by changes in rates of non-synonymous to synonymous substitutions (\( \omega \)). The \( \omega \) values suggested that on average the genes of *S. himantioides* experienced stronger purifying selection than those of var. *lacrymans*, and especially those of var. *shastensis*. However, even if \( \omega \) values can detect genes under selection, systematically increased \( \omega \) at the genome-wide level, can also be the result of demographic history (Tajima, 1989). In organisms with small effective population sizes, selection is less effective in removing deleterious mutations which can lead to elevated genome-wide \( \omega \) values. Correspondingly, we suggest that var. *lacrymans* and var. *shastensis*, which have higher average \( \omega \) overall in the genome, may have lower effective population sizes compared to *S. himantioides*. The differences in effective population size is expected as *S. himantioides* is distributed worldwide (Carolse
al., 2011), while var. *shastensis* has limited current distribution and var. *lacrymans* has gone through a domestication process (Kauserud et al., 2007).

In more detail, the genomic analyses revealed a selective shift in genes with functions involved in intracellular transport, growth and reorganization of the cell. Our data suggest that evolutionary changes to these processes may underlie the increase capacity of transportation and growth in var. *lacrymans* which in turn is likely to be a key factor for its success in the built environment. Buildings are a dry habitat, where the water resources are the most limiting factor. Var. *lacrymans* can produce the thickest mycelial cords described in the fungal kingdom, up to 2 cm in diameter (Jennings and Bravery, 1991). In comparison, var. *shastensis* and *S. himantioides* produce smaller cords, and *S. himantioides* has a slower growth rate (Harmsen, 1960). The corded network permits the translocation of intracellular resources, e.g. amino acids and water through vacuolar and vesicle trafficking to ensure complete exploitation of large woody substrates (Watkinson et al., 2006).

Proteins associated with endomembrane system functioning and hyphal growth had different selection pressure between *S. lacrymans* and *S. himantioides*, indicating that changes in resource translocation are important in the adaptation to the different niches. Hyphal growth is dependent on both transport and fusion of secretory vesicles to the plasma membrane and on actin cytoskeleton organization and polarization. Indeed, actin is important for polarized growth and also represents the mechanism for the transport of secretory vesicles that contain materials for the synthesis of new cell wall and membranes in the growing tip (Berepiki et al., 2011). We hypothesise that these genes play a role in the development and maintenance of the mycelial cords, possibly through mediating the re-grouping and re-allocation of resources.
To become a successful colonizer of wood, a fungus has to compete for resources with other decay species. However, the confrontation experiments, where the fungi were growing in fir blocks on moist perlite, revealed that var. *lacrymans* and var. *shastensis* have poor combative abilities compared to other wood decay fungi, at least in this nutrient poor set-up (Table 3). Species inhabiting more extreme environmental habitats may reduce their antagonistic abilities, following the universal adaptive strategy theory (Grime and Pierce, 2012). Thus *S. lacrymans* inhabiting the dry treeline and built environments may have lost the capacity for broad antagonistic responses. This may also explain why var. *lacrymans* usually does not spread from colonized buildings into the natural environment, though a few exceptions have been noted in the Czech Republic (Kotlaba, 1992). In less stressful climates in the boreal and temperate zones, where *S. himantioides* is typically found, interspecific antagonistic interactions may be more important. Hence, under these conditions, it may have been more advantageous to evolve strong combative ability. This is supported by the increased numbers of PFAM domains possibly related to defence in *S. himantioides* compared to *S. lacrymans*, e.g. PKS and ABC transporters. PKS are large synthases particularly involved in the biosynthesis of specialized metabolites with many diverse functions. The gene families are known to expand and contract rapidly in response to adaptation to nutritional and environmental factors, pathogens or interactions with other organisms (Bushley and Turgeon, 2010). ABC transporters are often involved in the efflux of small metabolites (Klein *et al.*, 2011; Karlsson *et al.*, 2015). Furthermore, similar expansions of PKS and ABC transporters have been observed in the mycoparasites *Clonostachys rosea* and *Trichodema virens*, and were suggested to be the reason for their extreme combative ability, by producing and transporting toxic compounds from the cells (Karlsson *et al.*, 2015). *Serpula himantoides* is known to produce antifungal substances, himanimides, that could increase its antagonistic ability (Aqueveque *et al.*, 2002). It is unknown if var. *lacrymans* can
produce these substances. More genomic analyses and experiments using different conditions
are needed to pinpoint the exact function of the larger number of PKS and ABC transporters
in *S. himantioides*, and whether any of these expansions are related to the previously detected
himanimides.

Our growth experiments on wood substrates confirm earlier findings that var. *lacrymans* is a highly effective decomposer of coniferous wood (Harmsen, 1960). In natural
environments, *S. lacrymans* typically occupies large logs of *Abies* or *Tsuga* (Figure 1) and
has developed a unique capacity for rapid decay during a short season of favourable growth
conditions. Resource availability and utilization of nutrients involve a diverse chemistry for
saprotrophic fungi. The varying levels of extractives, such as terpenoids and other phenolic
compounds, and the recalcitrant nature of the carbohydrates of wood imply that specialization
and adaptation to these conditions are essential to utilize this niche. Our findings suggest that
*S. lacrymans* is a more successful decomposer of spruce and fir than pine, and is more
specialized for these specific substrates than *S. himantioides*. A more narrow substrate range
was also suggested in a recent study of var. *lacrymans* and *Gloephyllum trabeum*, where they
found gene expression of a wider CAZyme complement in *G. trabeum* than in *S. lacrymans*
(Presley and Schilling, 2017). Furthermore, the speed and efficiency with which *S. lacrymans*
decomposes spruce, compared to *S. himantioides*, could be related to a more efficient CMF
chemistry. The iron reductase (with a CBM1 domain and a cytochrome B domain) found in
var. *lacrymans* and var. *shastensis*, but not *S. himantioides* has previously been suggested to
have an electron transfer function (Yoshida *et al.*, 2005). Thus, it can target reduced iron
directly to the cellulose substrate for efficient CMF. In previous analyses of *S. lacrymans*,
this iron reductase was specifically pinpointed as important in the early oxidative degradation
steps of the CMF chemistry (Eastwood *et al.*, 2011). This could contribute to more efficient
utilization of carbohydrates from its habitat.
The content of inhibitory extractives is greater in pine wood than in spruce wood (Sjöström, 1993), which makes pine a less favourable food source for fungi. Differential gene expression analyses of a white rot fungus (*Phlebiopsis gigantea*) grown on wood where extractives were removed showed several genes potentially related to the processing of extractives (Hori *et al.*, 2014). These differentially expressed genes encoded glutathione-S transferase, ABC transporters, lipases, cytochrome P450s and aldehyde dehydrogenase. We found accelerated evolution in *S. lacrymans* for aldehyde dehydrogenase, an ABC transporter, and cytochrome P450s, and loss of copies of glutathione-S transferase and ABC transporters. The ability to process a diversity of extractives found in wood and secrete their breakdown products may, therefore, also play an important role in substrate specialization and hence adaptation of *S. lacrymans* to a different habitat. Furthermore, the loss of laccases and the increase of cytochrome P450s in the branch leading to *S. lacrymans* could be related to both community interactions and the processing of toxic phenolic derivatives produced during the decomposition of lignocellulose. Brown rot fungi do not utilize lignin, however, they depolymerize lignin to gain access to the cellulose and hemicellulose. Thus, as part of adapting to a specialized niche *S. lacrymans* may have lost genes important for exploitation of some woody substrates in nature, but rather specialized for a more streamlined decomposition of specific substrates. Cytochrome P450s have been suggested to easily duplicate, and to be important in the colonization of new environments and in the breakdown of novel compounds (Syed *et al.*, 2014). Moreover, it has been suggested that the large gene repertoire of cytochrome P450s evolved in *Phanerochaete chrysosporium* increased its resource availability (Syed and Yadav, 2012), thus the expansion of cytochrome P450s could be related to an expansion of biochemical capacity in var. *lacrymans* as it invades timber wood. Timber wood is similar to the wood encountered naturally by primary decay species,
containing more plant-derived compounds than partially degraded wood that is often available in the forest.

The chemistry of defence and foraging is a recurring issue in our dataset. However, without in-depth functional analysis, it is unclear whether the product moved by a particular ABC transporter or metabolised by a cytochrome P450 gene is of importance to the species’ competitive ability and the decomposition of different substrates. Thus, further analyses of the increased set of cytochrome P450s in *S. lacrymans*, and the increased set of PKS and ABC transporters in *S. himantioides*, can pinpoint in which functions these gene expansions are involved.

Our results indicate that the devastating dry rot fungus is an ecological specialist that has developed highly effective brown rot decay and effective systems for transportation and growth. Common traits identified between genetically related var. *lacrymans* and var. *shastensis* when compared with the sister taxon *S. himantioides* suggest that var. *lacrymans* was pre-adapted to the built environment and that the requirements of the mountainous, dry, treeline habitat and the patchy nutrient environment of a house, including a blend of wood and mineral materials, share similar features important for *S. lacrymans*. This enabled var. *lacrymans* to opportunistically exploit the built environment when given the opportunity by human activity. Particularly, the evolution of the thick cords and rapid growth may be linked to its natural substrates, to maximize resource translocation and effectively decay the enormous logs. The lower combative ability, suggested from both physiological and genomic data and the narrower enzymatic assortment of our selected strains might explain why var. *lacrymans* rarely has been able to move from its new building niche back into temperate and boreal woodlands. As var. *shastensis* is very similar to var. *lacrymans* both in genetic and physiologic features, we conclude it has the potential to invade buildings, but has not done so
because its native range has not been widely exploited by humans and so has not been transferred to the built environment.

Acknowledgements

We thank Ella Thoen for technical help, Anikó Várnai for discussions, and Skui Christmas Tree Plantation for the *Abies lasiocarpa* wood. BVS, JH, HK, LT and IS acknowledge the University of Oslo and Norwegian Research Council (project 221840) for funding. NS thanks the Swedish University of Agricultural Sciences and FORMAS (project 2010-1354) for funding. DE, SM and LB thank the UK Natural Environment Research Council, (award NE/K011588/1) for support. Sequencing of SL200 and SHA17-1 was performed by the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory. The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

and analysed the *S. himantioides* genome at JGI. S.V.B., M.B.D., J.H., C.P. and S.C.M. analysed genomic data. S.V.B, J.H., D.C.E, H.K. and I.S. wrote the paper and all other authors discussed and modified the paper.
References


Figure 1. The dry rot fungus *Serpula lacrymans* and its habitat. *Serpula lacrymans* is one of the most devastating decomposer of houses in temperate and boreal regions worldwide. The species is known to form thick cords and a rapid decay of coniferous wood. In nature the species decompose large logs in dry mountain forests.

Figure 2. The comparative genomic differences among the *Serpula lacrymans* var. *lacrymans*, *Serpula lacrymans* var. *shastensis* and *Serpula himantioides*. A) The number of significantly expanded and contracted gene families, based on analyses using a birth-death model of gene family evolution on all gene clusters. The analyses use a rooted ultrametric tree from a 10 loci maximum likelihood analysis, where *S. himantioides* was the out-group. Thus, only changes in var. *shastensis*, var. *lacrymans* and the branch leading to these two, but not the *S. himantioides* branch were evaluated. B) Phylogenetic sketch trees demonstrating the selection analysis. Each tree highlights a branch and the number of genes with significantly increased or decreased $\omega$ values on that branch compared to the expected based on 5,866 single gene clusters. The null hypothesis is equal rates on all branches.

Figure 3. Decomposition rate of *Serpula lacrymans* var. *lacrymans*, *S. lacrymans* var. *shastensis* and *S. himantioides* on different wood species. Percent mass loss of wood blocks from the three plant species fir (*Abies lasiocarpa*), pine (*Pinus syvestris*) and spruce (*Picea abies*) inoculated by var. *lacrymans*, var. *shastensis* and *S. himantioides* for 60 days. No successful growth was obtained for var. *shastensis* on pine.
Table 1. Summary statistics of the genome assembly, annotation and CEGMA analyses of the three sequenced genomes of *Serpula lacrymans* var. *lacrymans*, *S. lacrymans* var. *shastensis* and *Serpula himantioides*

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th># of Contigs</th>
<th># of Scaffolds</th>
<th>N50</th>
<th>Genome size (Mbp)</th>
<th>Assembler</th>
<th>CEGMA</th>
<th># of predicted genes</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. himantioides</em></td>
<td>MUCL38935</td>
<td>5964</td>
<td>4893</td>
<td>20000</td>
<td>46</td>
<td>AllPathsL</td>
<td>89.5 %</td>
<td>12011 $^\S$</td>
</tr>
<tr>
<td>var. <em>lacrymans</em></td>
<td>SL200</td>
<td>4534</td>
<td>1529</td>
<td>59716</td>
<td>37</td>
<td>Velvet</td>
<td>97.6 %</td>
<td>11352</td>
</tr>
<tr>
<td>var. <em>shastensis</em></td>
<td>SHA17-1</td>
<td>3839</td>
<td>1170</td>
<td>92207</td>
<td>38</td>
<td>Velvet</td>
<td>97.2 %</td>
<td>10910</td>
</tr>
</tbody>
</table>

$^*_{Sequenced by JGI, § Number of genes predicted by Maker annotation tool, however the JGI annotation pipeline predicted 13805 gene models.}$
Table 2. The gene families that are evolving at a significant different rate (p value < 0.05 after FDR) among the different *Serpula* strains and includes a PFAM domain related to intracellular transport.

<table>
<thead>
<tr>
<th>Cluster.No</th>
<th>Description</th>
<th>PFAM ID</th>
<th>Test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2435</td>
<td>Domain_of_unknown_function_(DUF202), SPX_domain</td>
<td>PF02656, PF03105</td>
<td>H(l), L(h)</td>
<td>0.00899, 0.00044</td>
</tr>
<tr>
<td>1272</td>
<td>Cofilin/tropomyosin-type_actin-binding_protein, Variant_SH3_domain</td>
<td>PF00241, PF14604</td>
<td>H(l), L(h)</td>
<td>0.03537, 0.00589</td>
</tr>
<tr>
<td>6654</td>
<td>RasGEF_N-terminal_motif, RasGEF_domain</td>
<td>PF00618, PF00617</td>
<td>H(l), L(h)</td>
<td>0.02021, 0.00370</td>
</tr>
<tr>
<td>6080</td>
<td>SNARE_domain</td>
<td>PF05739</td>
<td>H(l), L(h)</td>
<td>0.00346, 0.00003</td>
</tr>
<tr>
<td>6147</td>
<td>RhoGEF_domain</td>
<td>PF00621</td>
<td>H(l), L(h)</td>
<td>0.00899, 0.00573</td>
</tr>
<tr>
<td>3843</td>
<td>PX_domain</td>
<td>PF00787</td>
<td>L(l), H(h)</td>
<td>0.01602, 0.00220</td>
</tr>
<tr>
<td>1940</td>
<td>Oxysterol-binding_protein</td>
<td>PF01237</td>
<td>H(s), L(h)</td>
<td>0.01365, 0.00607</td>
</tr>
<tr>
<td>3226</td>
<td>PH_domain, FHA_domain, Kinesin_motor_domain</td>
<td>PF00169, PF00498, PF00225</td>
<td>H(s)</td>
<td>0.00279</td>
</tr>
<tr>
<td>6485</td>
<td>WD_domain, _G-beta_repeat</td>
<td>PF00400</td>
<td>L(h)</td>
<td>0.00683</td>
</tr>
<tr>
<td>2827</td>
<td>Sec1_family</td>
<td>PF00995</td>
<td>L(h)</td>
<td>0.02796</td>
</tr>
<tr>
<td>1406</td>
<td>FYVE_zinc_finger, TCP-1/cpn60_chaperonin_family</td>
<td>PF01363, PF00118</td>
<td>L(h)</td>
<td>0.01075</td>
</tr>
</tbody>
</table>

H indicates higher omega, L indicates lower omega. l symbolizes *Serpula lacrymans* var. *lacrymans*, s symbolizes *S. l.* var. *shastensis* and h indicates *S. himantioides*, thus H\(l\) indicates significant higher omega for var. *lacrymans*
Table 3. Results from combat experiments with *Serpula lacrymans* var. *lacrymans*, *S. lacrymans* var. *shastensis*, *S. himantioides* and three other fungal species. The proportion of plates with mycelia from the species named in the column after the confrontation test with the species in the row, i.e. read horizontally, higher than 0.5 wins the confrontation with the vertical strain.

<table>
<thead>
<tr>
<th></th>
<th>var. <em>lacrymans</em></th>
<th>var. <em>shastensis</em></th>
<th><em>S. himantioides</em></th>
<th><em>C. puteana</em></th>
<th><em>A. xantha</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>var. shastensis</em></td>
<td>0.450 (20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. himantioides</em></td>
<td>0.689 (45)*</td>
<td>0.685 (27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. puteana</em></td>
<td>0.430 (43)</td>
<td>0.500 (34)</td>
<td>0.155 (45)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. xantha</em></td>
<td>0.400 (45)</td>
<td>0.355 (38)</td>
<td>0.136 (44)**</td>
<td>0.154 (39)**</td>
<td></td>
</tr>
<tr>
<td><em>F. pinicola</em></td>
<td>0.978 (46)**</td>
<td>0.889 (45)**</td>
<td>0.156 (48)**</td>
<td>0.931 (29)**</td>
<td>0.292 (48)**</td>
</tr>
</tbody>
</table>

Number of plates (n) used in parenthesis. * indicates significant different (*p<0.05, **p<0.005) from expected (E=n/2) by a Person $\chi^2$ Goodness of fit test, df=1.