Mycorrhiza (2009) 19:549–558
DOI 10.1007/s00572-009-0252-3

ORIGINAL PAPER

Variation in host specificity and gene content in strains from genetically isolated lineages of the ectomycorrhizal fungus *Paxillus involutus s. lat.*

Jenny Hedh · Tomas Johansson · Anders Tunlid

Received: 5 November 2008 / Accepted: 22 April 2009 / Published online: 12 May 2009
© Springer-Verlag 2009

Abstract Ectomycorrhizal fungi are known to vary in host range. Some fungi can enter into symbiosis with multiple plant species, while others have restricted host ranges. The aim of this study was to examine variation in host specificity among strains from the basidiomycete *Paxillus involutus s. lat*. Recent studies have shown that this fungus consists of at least four genetically isolated lineages, phylogenetic species (PS) I (which corresponds to the morphological species *Paxillus obscurosporus*), PS II (*P. involutus s. str.*), PS III (*Paxillus validus*), and PS IV (not yet supported by any reference material). Thirty-five *Paxillus* strains of PS I to IV were examined in microcosms for their capacity to infect birch (*Betula pendula*) and spruce (*Picea abies*). Seventeen strains were compatible and formed mycorrhizae with both tree species. Seven strains were incompatible with both birch and spruce. The gene content in three pairs of incompatible and compatible strains PS I, II, and III were compared using microarray-based comparative genomic hybridizations. Of 4,113 *P. involutus* gene representatives analyzed, 390 varied in copy numbers in at least one of the three pairwise comparisons. Only three reporters showed significant changes in all three pairwise comparisons, and none of these were changed in a similar way in three comparisons. Our data indicate that changes in host range have occurred frequently and independently among strains in *P. obscurosporus*, *P. involutus s. str.*, and *P. validus*. No evidence was obtained demonstrating that these changes have been associated with the gain or loss of similar genes in these three species.

Keywords Comparative genomics hybridization (CGH) · Gene duplications · Host specificity · Ectomycorrhizae

Electronic supplementary material The online version of this article (doi:10.1007/s00572-009-0252-3) contains supplementary material, which is available to authorized users.

J. Hedh · T. Johansson · A. Tunlid (✉)
Department of Microbial Ecology, Lund University,
Ecology Building,
223 62 Lund, Sweden
e-mail: anders.tunlid@mbioekol.lu.se

Introduction

There are an estimated 7,000–10,000 fungal and about 8,000 plant species that can form ectomycorrhizae (EM; Taylor and Alexander 2005). The majority of these fungi are basidiomycetes with some ascomycetes and a few zygomycetes (Smith and Read 1997).

Records of plant–sporocarp associations have demonstrated that there is a large variation in the host range of EM fungi. Some species have a narrowly restricted host range (i.e., high specificity), while others can form mycorrhizae with diverse host plants (low specificity; Molina et al. 1992; Trappe 1962). Furthermore, the variation in host specificities between strains within EM-forming species can be extensive (Cairney 1999; Smith and Read 1997). For example, in vitro studies of isolates from *Pisolithus tinctorius*, *Paxillus involutus*, and *Laccaria bicolor* have shown that they can differ markedly in their ability to form EM and to stimulate the growth of the host plant (Burgess et al. 1994; Gafur et al. 2004; Laiho 1970; Wong et al. 1990). Studies in laboratory settings have also indicated that the host specificity of ectomycorrhizal fungi does not appear to be absolute; strains can display a range of capacities for infecting a host species (Burgess et al. 1994). The ability of an ectomycorrhizal fungus to infect a host depends on a number of environmental conditions. Accordingly, associations that are formed under laboratory

conditions reflect a potential to infect a given host species and could differ from "ecological specificity" observed in nature (Molina and Trappe 1982). Studies on the specificity of EM associations can also be complicated due to the fact that the capacity of fungal isolates to form EM can decrease during prolonged maintenance in axenic cultures (Cairney 1999; Laiho 1970). It should also be remembered that plants vary in their receptivity towards symbionts, and the host genotype may influence the formation of EM (Lamhamedi et al. 1990).

The genomic mechanisms that could account for variations in host specificity of EM fungi are not known. Generally, such phenotypic differences could result from variations in gene content from quantitative differences in gene expression or from structural differences in gene products (Tunlid and Talbot 2002). We have examined some of these mechanisms in strains of *P. involutus* with different host specificity (Le Quéré et al. 2004, 2006). The strains analyzed included NAU, which is incompatible with poplar and birch (*Betula pendula*), and the two compatible strains MAJ and ATCC 200175. By using a cDNA microarray, we showed that the compatible and incompatible strains differed in gene copy numbers and in plant-induced gene expression patterns. Furthermore, analysis of sequence information from two of the differentially expressed genes indicated that they appear to have evolved at an enhanced rate, indicating different selection pressures in compatible and incompatible strains. Thus, several genomic mechanisms could possibly be involved in generating variation in host specificity of *P. involutus*.

Mating-type tests, morphological analyses, and, more recently, multilocus sequencing have shown that *P. involutus s. lat.* represents a group of species (Fries 1985; Hahn and Agerer 1999; Hedh et al. 2008). Based on the concordance of five gene genealogies of European isolates of *P. involutus s. lat.* can be separated into four distinct, genetically isolated lineages, phylogenetic species (PS) I, II, III, and IV (Hedh et al. 2008). Sequencing of reference material showed that PS I corresponds to *Paxillus obscurosporus*, PS II *to P. involutus s. str.*, and PS III to *Paxillus validus.* PS IV has not yet been supported by any reference material (Hedh et al. 2008). The strains used in the previous host specificity experiments, MAJ and NAU, belong to PS III. In this study, we have examined the host specificity of a larger number of strains from the *P. involutus s. lat.* species group. Several newly isolated strains were included in the experiments. The aim was to investigate whether variations in host specificity could be found in several of the *P. involutus s. lat.* lineages, and if so, whether the genomic mechanisms that could account for these phenotypic differences are similar or different. The gene content in several pairs of compatible and incompatible strains were compared using a newly designed cDNA microarray containing a significantly larger number of gene reporters as compared with a previously used array.

Materials and methods

Fungal isolates

In the years 2003 and 2004, 126 strains of *P. involutus s. lat.* were isolated by collection of fruiting bodies in the southern part of Sweden. Of these, 22 were selected for microcosm experiments with birch and spruce (*Picea abies*) together with another 13 isolates already in our stock (Table S1, supplementary material), for a total of 35 isolates. The isolates were maintained on agar containing 0.5× modified Melin–Norkrans (MMN) medium (Brun et al. 1995).

Screening for host compatibility

Screening for host compatibility with birch was performed using a previously described assay (Brun et al. 1995). The strains of *P. involutus s. lat*. were grown on 0.5× MMN agar medium for 10–14 days. Seeds of birch (Skuleskogen, Sweden) were surface sterilized, transferred to water agar (0.7% (*w*/*v*)), and incubated for 7–14 days until germination occurred. The birch seedlings were transferred to the edge of the growing mycelium (six seedlings/plate), and the syntheses were incubated for 3 weeks. Assays for testing compatibility with spruce were done essentially as previously described (Duddridge 1986; Finlay et al. 1988). Spruce seedlings were generated from seeds (Munkfors, Sweden), surface sterilized, and incubated on water agar (0.7% (*w*/*v*)) for 14 days. The seedlings were transferred to Petri dishes containing a mixture of peat, 0.5× MMN, and vermiculite (4 vols. vermiculite, 1 vol. peat, 2 vol. 0.5× MMN, combined and autoclaved). Three seedlings were transferred to each Petri dish. The Petri dishes were inoculated by the transfer of four plugs of actively growing mycelium from the outer edge of *P. involutus s. lat.* colonies growing on 0.5× MMN agar plates. The dishes were sealed, and the system was incubated for 8 weeks. Greenhouse settings for the birch- and the spruce-synthesis experiments were 60% air humidity and 18/6 h and 18–20/16°C day/night cycle. In both experiments, the numbers of infected seedlings were counted, as visualized by the presence of a pseudoparenchymatous mantle.

Extraction of DNA

The strains ATCC 200175, SE03071001, SE03071610, MAJ, NAU, and Pi01SE were grown on cellophane-covered agar plates containing Gamborg B-5 basal liquid medium (Sigma–Aldrich) supplemented with glucose

(2.5 g l^{-1}). The strains were incubated for approximately 3 weeks. The mycelium was then transferred to the surface of Gamborg B-5 basal liquid medium supplemented with glucose (2.5 g l^{-1}) and grown for 7–10 days (Le Quéré et al. 2006). The mycelium was harvested, ground in liquid nitrogen, and stored in a −80°C freezer until use. DNA was extracted from 1 g of starting material using the E.Z.N.A.® Fungal DNA Kit (Omega Bio-Tek) according to the manufacturer's protocol, except that the incubation time for RNase treatment was increased to one hour. The DNA was precipitated by standard means and quantified by spectrophotometry. Agarose gel electrophoresis showed that the sizes of the DNA fragments ranged between 200 and 1,500 bp.

Microarray arrays and genomic hybridization

A custom-made cDNA microarray were constructed by printing cDNA PCR products (Johansson et al. 2004, unpublished) on UltraGAPS coated slides (Corning) using a 48-pin configured MicroGrid II array printer (BioRobotics) controlled by the MicroGrid TAS Application Suite (ver. 2.2.0.6). Following the terminology used in the context of DNA microarrays, the printed PCR products are in the following text referred to as reporters. The array, which contains a total of 23,232 printed reporters, representing replication of a uniset of 4,911 reporters: 4,891 experimental and 20 control reporters. Experimental reporters were obtained from 11 different cDNA libraries representing various mycorrhizal and non-mycorrhizal conditions of *P. involutus s. str.* (strain ATCC2000175) and *B. pendula*. In total, 19,188 expressed sequence tags (ESTs) were generated from these libraries which were assembled into 4,891 contigs that putatively represent unique transcripts (Johansson et al., unpublished). Of the experimental reporters, 3,082 were of fungal origin, 779 of plant origin, and 1,030 of unknown origin (from mixed mycorrhizal libraries). Forty PCR-amplified genomic fragments corresponding to different regions of a cosmid clone containing a 33-kb genomic region from *P. involutus s. str.* (Le Quéré et al. 2002) were also spotted on the arrays. Each reporter was printed at least in quadruplicate together with a number of control reporters. Information on the cDNA microarray ("Lund Univ Mycorrhiza_Betula pendula-Paxillus involutus 23 k, version 2.1") including the accession numbers to sequences is available at the ectomycorrhizae BL-EBI ArrayExpress database (http://www.ebi.ac.uk/arrayexpress), accession number A-MEXP-853.

Microarray analyses were performed using genomic DNA from six strains of *Paxillus* and hybridizations were performed with direct contrasts within each phylogenetic group: SE03071001 and SE03071610 (PS I); ATCC 200175 and Pi01SE (PS II); and MAJ and NAU (PS III). In total, six dual-label microarray hybridizations including dye swaps were performed. A total of 4 μg DNA for each strain and each dye (Cy3 or Cy5) was used as template for labeling. For dual-color labeling, the Bioprime comparative genomics hybridization (CGH) labeling system (Invitrogen) was used, and the labeled samples were purified using the Bioprime purification module (Invitrogen). After purification, the Cy3- and Cy5-labeled samples were combined in an amber Eppendorf tube. The preparation of the samples and the pre- and post-hybridization washes of the printed slides were done using the Pronto!™ Universal Microarray Kit (Corning). The manufacturer's protocol was followed except that the first three post-hybridization washes were replaced by soaking the slides in post-hybridization solution for 10 min at 50°C (cover slips removed). The slides were hybridized using the MAUI® Hybridization System v1.1 at 42°C. The slides were dried by a short centrifugation and then scanned, as described previously (Johansson et al. 2004). The entire raw dataset is deposited and available at the EMBL-EBI ArrayExpress database (http://www.ebi.ac.uk/) under the accession number E-MEXP-1499.

Processing of microarray data

Microarray data images were inspected manually, and low-quality spots were excluded from further analyses. After inspection and reduction of reporters known to be of plant origin, data for 4,113 reporters remained in the dataset. Levels in signal intensities were estimated using two successive mixed models (Wolfinger et al. 2001):

$$\log_2\left(Y_{ijkg}\right) = T_i + D_j + TD_{ij} + S_k + SD_{jk} + \varepsilon_{ijkg} \qquad \text{(M1)}$$

$$R_{ijkg} = T_{ig} + D_{jg} + TD_{ijg} + S_{kg} + SD_{ijg} + \xi_{ijkg} \qquad \text{(M2)}$$

Model M1 was used to normalize all the data, and model M2 was then fit separately to one gene at a time. The term Y_{ijkg} represents the raw intensity measurement from the ith strain (ATCC 200175, SE03071001, SE03071610, MAJ, NAU, Pi01SE), the jth dye channel (Cy5 and Cy3), the kth slide (1, ..., 6) for the gth gene (1, ..., 4,113). The term R_{ijkg} represents the residual computed as $\log_2(Y_{ijkg})$ minus the fitted effects from model M1. The symbols T, D, and S represent the effects of strain, dye, and slide effect, respectively. Double symbols represent the corresponding interaction effects. The terms S and SD in model M1 are considered to be random effects, as are terms S and SD in model M2; others are fixed effects, and ε and ξ are stochastic errors. All the random effect terms, including the errors, are assumed to be normally distributed with means

of zero and effect-specific variance components. Estimates of normalized signal intensities are based on fitted values from model M2 and were fitted using PROC MIXED in SAS (SAS/STAT Software Version 9.1; SAS Institute Inc., Cary, NC, 1999).

Clustering analyses and filtering for amplified and deleted genes

Normalized signal intensities ($\log_2$) for each strain were evaluated by comparison of data to the strain ATCC 200175. Fold changes were used for hierarchical clustering (average linkage and uncentered correlation) using Cluster 3.0 (http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster), and results were viewed using Java TreeView (http://jtreeview.sourceforge.net; Eisen et al. 1998; Fig. 1). Genes being amplified or deleted were identified based on $\log_2$ fold changes in pairwise comparisons. A gene was included based on two criteria: (1) the $\log_2$ fold change for a pairwise comparison was ≥0.5 or ≤−0.5 and (2) the standard deviation (SD) for the $\log_2$ fold change of the comparison for at least one of the pairs was ≤0.3.

Sequence analysis

For the 4,113 putative genes analyzed on the array, a homology search was carried out using the tblastx algorithm (Altschul et al. 1990) with an E-value threshold of 1E-10 against the UniProt sequence database (Apweiler et al. 2004). Gene ontology (GO; Ashburner et al. 2000) annotations were inferred by retrieving information from the UniProt entry corresponding to the highest tblastx score. Using the full GO (www.geneontology.org), all the classified genes were mapped to all their parent terms in the yeast GO Slim (www.geneontology.org/GO.slims). A search for putative fungal homologues of the *P. involutus s. str.* reporters was also performed by querying (in November 2007) the following databases using the blastx algorithm (Altschul et al. 1990) with a threshold value of 1E-5: Fungal genome database at Broad Institute (containing 375,232 protein sequences from 34 fungal genomes) (http://www.broad.mit.edu/annotation/) and Pfamseq database (Finn et al. 2006) based on UniProt 9.7 (containing 4,084,128 protein sequences, Release 22). The probability *P* of observing the number of genes within a given category by chance was estimated using the hypergeometric distribution.

Results and discussion

Host compatibility

Thirty-five strains of the *P. involutus* species group, including isolates collected during the years 2003 and 2004 as well as a few isolates kept in culture for longer periods of time, were examined for their ability to form ectomycorrhizae with birch and spruce seedlings (Table S1, supplementary material). In total, 25 of the examined strains were found to be compatible with birch, but their ability to form ectomycorrhizae varied extensively. Ten of the examined strains did not infect the birch seedlings and were considered to be incompatible with birch. The proportion of birch seedlings infected by the compatible strains varied between 17% and 100%. Data for host compatibility with spruce were obtained for 29 of the 35 *Paxillus* strains. The number of replicates in the experiments with spruce was smaller than in those for birch. Nineteen of them were found to be compatible (9% to 100% of the seedlings were infected), and ten were incompatible. Seven out of the 35 examined strains infected neither birch nor spruce and formed a group from which incompatible (−) strains were selected.

Strains that were compatible to both birch and spruce (+) were identified in all four of the previously identified phylospecies of *P. involutus s. lat.* Incompatible (−) strains were found among PS I, II, and III, among isolates not maintained in culture for a prolonged time (collected in

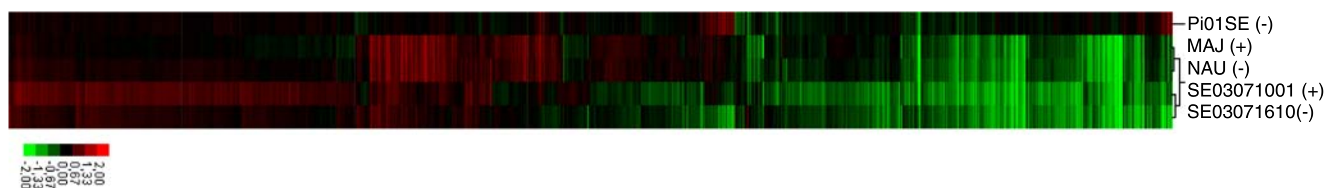

Fig. 1 Fold changes ($\log_2$) of normalized hybridization intensities from comparative genomic hybridization (CGH) analyses of five strains of *Paxillus involutus* (Pi01SE, MAJ, NAU, SE03071001, and SE03071610), as compared to the ATCC 200175 reference strain. The strains were either compatible (+) or incompatible (−) to birch and spruce (Table 1). The microarray used was constructed from PCR-amplified cDNA derived from a collection of EST clones obtained from the strain ATCC 200175. A total of 4,113 gene reporters were included for analysis, and the panel shows a hierarchical clustering of $\log_2$ fold changes, uncentered in both dimensions. The *scale* indicates levels of fold changes and ranges from *greenish* (negative fold change) to *reddish* (positive fold change) with intensities of >0.5 and <−0.5 considered as indicating deleted or duplicated genes, respectively, in relation to the reference strain

2003 and 2004) as well as among those kept in culture for long periods, including the previously studied NAU.

Pairs of compatible (+)/incompatible (−) strains of the *P. involutus* species group were selected from each of PS I, II, and III for comparative genomic analysis (Table 1). The strains from PS I (SE3071610 (+) and SE3071001 (−)) were collected in the same week in 2003 at the same locality (Lund, Sweden), growing less than 30 m apart from each other. SE03071610 were situated in a small grove, and SE03071001 were growing on a grass field. Their ITS regions are identical. The ATCC 200175 strain of PS II has been used in numerous laboratory experiments for more than 20 years. Microcosm studies have shown that this strain has a very wide host range and can infect a large number of tree species including birch, spruce, and pine (Brun et al. 1995; Ek et al. 1994; Finlay et al. 1988). The incompatible strain selected from PS II (Pi01SE) was isolated from a pine forest in Sweden in 2002. The identity of the ITS regions of ATCC 200175 and Pi01SE is 99.3%. The host specificity of the PS III strains MAJ (+) and NAU (−) have previously been examined in microcosms using poplar and birch (Gafur et al. 2004; Le Quéré et al. 2004). These experiments have shown that MAJ is compatible with both poplar and birch, whereas NAU is incompatible with both tree species (Gafur et al. 2004). The identity of their ITS regions is 99.5%.

Analysis of hybridization signals

The differences in gene content between the compatible (+) and incompatible (−) strains of PS I, II, and III were analyzed by array CGH. The cDNA microarray used was constructed from gene reporters originating from the ATCC 200175 strain of *P. involutus s. str.* (Materials and methods). In total, the hybridization signals of 4,113 cDNA probes were analyzed putatively representing 4,113 *P. involutus s. str.* genes, which can be compared with the 1,075 probes present on the arrays used in the previous studies (Le Quéré et al. 2004, 2006). The fraction of genes covered by the 4,113 probes can be estimated to correspond to approximately 53% of the total number of genes in the fungal genome, assuming a gene content of 7,700 (Le Quéré et al. 2002).

When the $\log_2$ ratios of the hybridization signals for the various strains relative to the ATCC 200175 strain were analyzed by clustering, Pi01SE was together with the reference strain found in one group, MAJ and NAU in a second group, and SE3071610 and SE3071001 in a third group (Fig. 1). This partitioning is in agreement with the distinct separation of the strains into three genetically isolated species based on the concordance of the gene genealogies of five nuclear genes (Hedh et al. 2008).

We have previously shown that the analyses of CGH data from such distant groups as PS I, II, and III can be susceptible to significant problems associated with biases introduced during the normalization procedures (Le Quéré et al. 2006). The normalization of the hybridization signals can be skewed because of the presence of a large number of variable genes, yielding weak signals. The problem can be limited by restricting the analysis to comparisons between closely related strains such as those found within the phylogenetic species of *P. involutus s. lat.* (Le Quéré et al. 2006). In the following analyses, the hybridization signals of the (+)/(−) strains

Table 1 *Paxillus* strains used for comparative genomic hybridizations (CGH)

PS[a]	Strain	Origin[b]	Location	Habitat	Possible host species[c]	Compatibility in microcosm[d]	ITS[e]
I	SE03071610	SE	Skåne	Public lawn	*Populus*, *Tilia*, *Acer*, *Fagus*, *Quercus*, *Sorbus*	*Betula* (+), *Picea* (+)	AY585912
I	SE03071001	SE	Skåne	Public lawn	*Populus*, *Cornus*, *Tilia*, *Fagus*, *Quercus*, *Sorbus*, *Acer*	*Betula* (−), *Picea* (−)	AY585910
II	ATCC 200175	GB	Scotland	*Unknown*	*Betula*	*Betula* (+), *Picea* (+), *Pinus* (+), *Fagus* (+)	AY585913
II	Pi01SE	SE	Skåne	Sandy pine forest	*Pinus*	*Betula* (−), *Picea* (−)	AY585918
III	MAJ	FR	*Unknown*	*Unknown*	*Populus*	*Populus* (+), *Picea* (+), *Betula* (+)	AY585917
III	NAU	FR	*Unknown*	*Unknown*	*Quercus*	*Betula* (−), *Populus* (−), *Picea* (−)	AY585915

All strains examined for host compatibility with birch and spruce are listed in Table S1 (supplementary material).

[a] Phylogenetic species (PS) according to Hedh et al. (2008)

[b] *GB* Great Britain, *SE* Sweden, *FR* France

[c] Woody plants within a 10-m radius

[d] A plus sign (+) indicates a compatible strain and a minus sign (−) an incompatible strain

[e] BL/DDJB/GenBank accession numbers for sequence information covering the ITS region

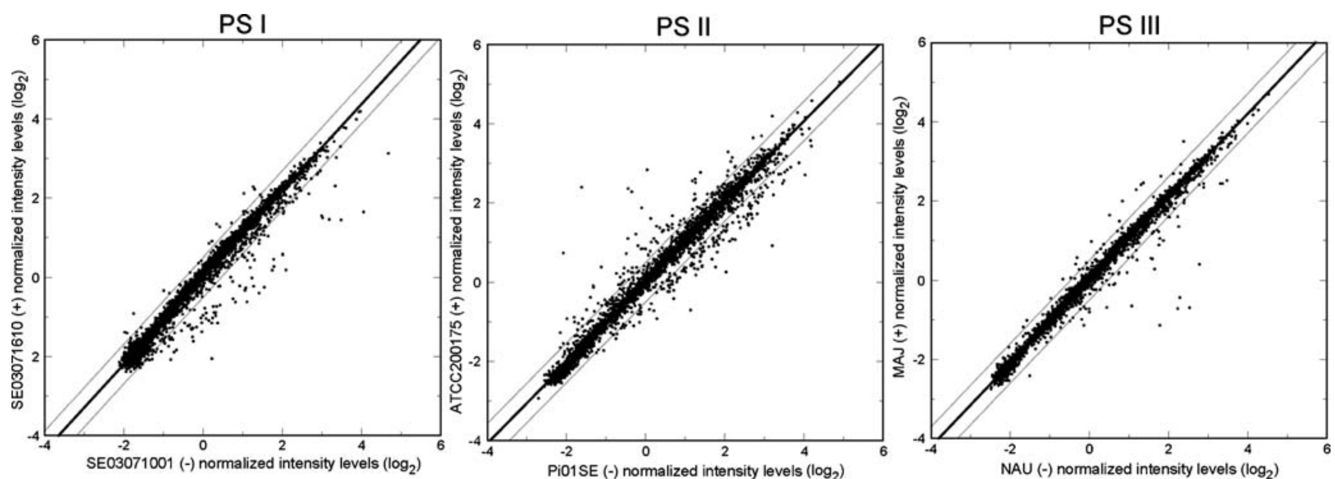

Fig. 2 Divergence of gene content in compatible (+) versus incompatible (−) strains within each of the phylogenetic species (PS) I, II, and III of *Paxillus involutus s. lat*. The scatter plots show the normalized intensity levels (log_2 scale) of 4,113 gene reporters analyzed. The *solid, diagonal line* shows genes with almost identical hybridization signals in the compatible and the incompatible isolates. The *dotted lines* at $y=x+1$ and $y=x-1$ correspond to a log_2 fold change between the two isolates of +0.5 and −0.5, respectively. Functional annotations of the divergent genes are listed in Table S2 (supplementary material). Characteristics of the isolates are given in Table 1

within PS I, II, and III were, therefore, directly compared (Fig. 2).

The differences in hybridization signals in CGH experiments depend on many factors including sequence divergence between the DNA of the compared samples and variation in gene copy numbers. Analyses of 17 nuclear genes from strains within PS II (three strains) and III (two strains) have shown that a majority of them display an intraspecific sequence identity of at least 97.7% (Le Quéré et al. 2006). The minimum intraspecific sequence identity of four nuclear genes (*rabA*, *hydA*, *gpiA*, and *actA*) was 99.1% in PS I (nine strains), 98.5% for PS II (33 strains), and 98.3% for PS III (three strains; Hedh et al. 2008). Previous studies have shown that a difference in sequence identity of 98–100% can generate a change in the log_2 fold values of the hybridization signals of −0.3 to +0.3 (Le Quéré et al. 2006). Thus, fold changes outside this range will primarily be associated with differences in gene copy number and not with sequence divergence (Le Quéré et al. 2006). Accordingly, we have, when comparing strains within a PS in the following analyses, interpreted variation in hybridization signals that are ≥0.5 or ≤−0.5 as related to differences in gene content. However, we cannot exclude the possibility that such signals can be due to genes displaying that are highly divergent in sequences. Further analyses are needed to clearly distinguish between the processes of sequence divergence and variation in gene copy numbers.

Gene content in compatible and incompatible isolates

The CGH analysis showed that the gene content of the closely related compatible (+) and incompatible (−) strains differed significantly (Table 2). The proportion of genes that varied in copy number when comparing the compatible (+) and incompatible (−) isolates within PS I, II, and III was 3.5%, 6.3%, and 1.3%, respectively. In total, 390 genes varied in copy number in at least one of the three pairwise comparisons. Within this nonredundant set, only three gene reporters showed significant changes in all three pairwise comparisons, while 58 (14.9%) showed significant changes in at least two of the comparisons (Fig. 3). None of the three genes showing significant changes for all three pairs were diverging in the same direction.

From similarity searches of all 390 genes varying in copy numbers, 203 genes (52%) displayed sequence similarities against sequences in the Fungal genome database at the Broad Institute; 138 (35%) contained sequences with similarities to protein family motifs in the Pfam database; and 129 (33%) were homologous to sequences present in the UniProt sequence database. The total number of genes with homologues to the Fungal genome database, Pfam database, and UniProt sequence database among the 4,113 reporters on the microarray were 2,400 (59%), 1,227 (70%), and 2,544 (38%), respectively. Accordingly, genes with no homologues were slightly overrepresented in the cohort of duplicated genes as compared to their abundance among all the genes on the array ($P<0.05$).

Based on the sequence homologies, 88 of the 390 duplicated genes could be annotated into functional categories (Table S2, supplementary material). Among those, several displayed sequence similarities to members of large gene families, including the major facilitator superfamily (Pfam domain PF07690; Goffeau et al. 1997) and the cytochrome P450 (PF00067) gene family. The P450

Table 2 The number of genes found to be amplified or deleted in the pairwise comparisons of compatible (+) and incompatible (−) isolates within the three phylogenetic species (PS) I to III of *Paxillus involutus*

PS	Strains	Total	Amplified	Deleted
I	SE03071610 (+)/SE03071001 (−)	142	21	121
II	ATCC 200175 (+)/Pi01SE (−)	260	114	146
III	MAJ (+)/NAU (−)	53	23	30
	Sum	455	159	297
	Unique	390	154	269

The strains are described in Table 1. The procedure for identifying amplified or deleted genes is shown in Fig. 2. Genes being amplified or deleted were identified based on $\log_2$ fold changes in hybridization signals in pairwise comparisons. A gene was included based on two criteria: (1) the $\log_2$ fold change for a pairwise comparison was ≥0.5 or ≤−0.5 and (2) the standard deviation for the $\log_2$ fold change of the comparison for at least one of the pairs was ≤0.3 (Le Quéré et al. 2006)

gene family is among the largest gene families in the basidiomycete *Phanerochaete chrysosporium* (Martinez et al. 2004). Members of this family have many different functions, including oxidative metabolism of endogenous and xenobiotic compounds (Nelson 1999).

Other genes that varied in copy numbers among the analyzed strains of *P. involutus s. lat.* encoded motifs of protein kinases (PF00069) and the Ras family of small GTPases (PF00071; Table S2, supplementary material). Notably, genes encoding protein kinases and Ras GTPases are among the largest and most expanded gene families in the genome of the ectomycorrhizal fungus *Laccaria bicolor* (Martin et al. 2008). Proteins of both families are known to have important roles in signaling pathways controlling many important cellular processes like morphological changes, cell cycle transitions, secretion, and stress responses (Manning et al. 2002; Ridley 2001; Westfall et al. 2004). The nonredundant set of amplified and deleted genes also contained two hydrophobin genes, *hydA* and *hydE*. Hydrophobins are small secreted proteins known to play important roles in a range of different processes related to growth and development in fungi (Wösten 2001). We have recently shown that the hydrophobin gene family in *P. involutus s. lat.* contains at least seven members including *hydA* and *hydE* (Rajashekar et al. 2007). The *P. involutus s. lat. hyd* gene family is dynamic; some members are maintained for long periods of time whereas others appear to be rapidly lost because of deleterious mutations. Among the cohort of amplified or deleted genes were also a number of genes displaying sequence homologies to genes involved in the uptake and metabolism of nitrogen compounds. They included peptidases (PF01546, PF00450), amino acid permeases (PF00324), and amino transferases (PF00155).

Comparison with the previous CGH study

In the previous CGH study of *P. involutus* (Le Quéré et al. 2006), 21 genes were identified that varied in copy numbers between MAJ and NAU, of which 14 were also significant in the present study. In the comparison between ATCC

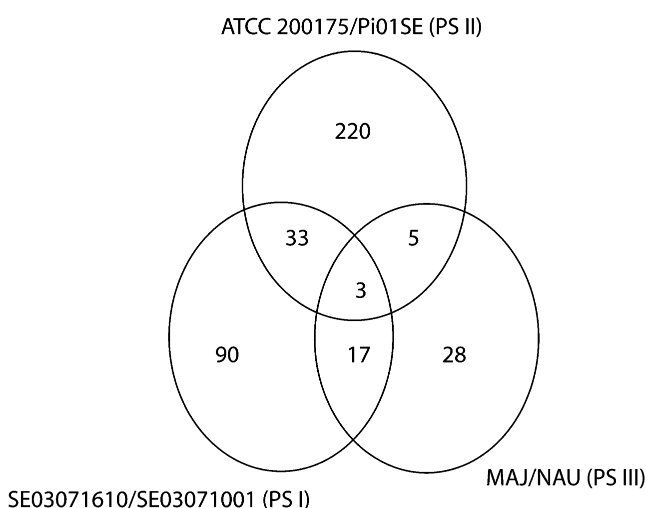

Fig. 3 A Venn diagram showing the distribution of 390 divergent fungal gene reporters in the pairwise comparisons of compatible and incompatible strains within phylogenetic species (PS) I, II, and III

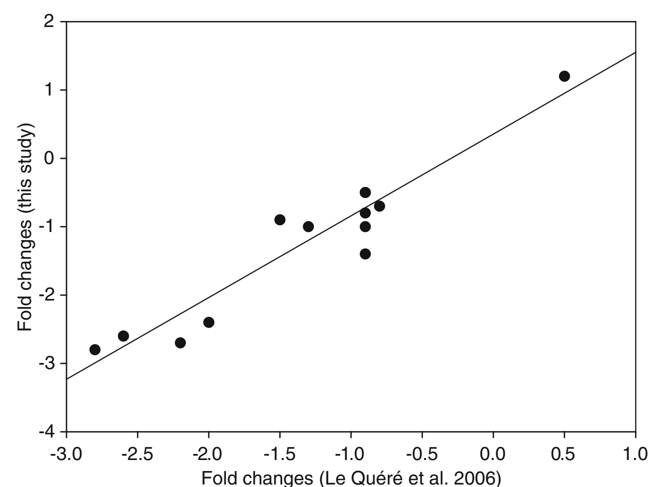

Fig. 4 A comparison of the $\log_2$ fold changes for 14 gene reporters in the strains MAJ and NAU obtained in this and a previous study (Le Quéré et al. 2006)

200175 and Pi01SE, 41 genes were identified to vary in copy numbers in the previous analysis, of which 22 were recognized in the present study. The discrepancies between the two studies could be explained by the fact that the intensity of hybridizations signals was slightly higher in the study by Le Quéré et al. (2006) as compared to the current study. Thus, the reporters that did not have significant fold changes in the present study had very low hybridization intensity levels in at least one of the analyzed samples (log_2 residual values below −2.0). However, the fold values for reporters with hybridization signals above these levels were very similar in the two studies as shown by comparing the log_2 fold changes in the MAJ and NAU hybridization experiments (Fig. 4).

Conclusions

Measures of ectomycorrhizae host specificities based on experiments in microcosms like the ones used in this study reflect the potential of the fungal isolates to infect a given tree species. Hence, it needs to be remembered that the observed ability of the examined *P. involutus s. lat.* strains to form mycorrhizal associations may differ from their performance in nature. Most likely, the incompatible strains have a capacity to form EM associations with various host tree species in the field, and they had not completely lost their ability form EM. Nevertheless, our experiments indicate that strains with narrow host range appear to have evolved repeatedly and independently among strains with wider host ranges in the *P. involutus* species group. The low capacity to establish symbiosis with birch and spruce was not due to prolonged culturing in the laboratory because incompatible strains also were found among the newly isolated specimens. The direction of these changes, i.e., whether strains with a narrow host range have evolved from those with wider host ranges or vice versa is not known since the phylogenetic relationships between the analyzed strains cannot be resolved based on ITS sequences. However, the finding of compatible and incompatible strains growing in the same habitat and with identical ITS sequences suggest that changes in host specificities could evolve rapidly and under similar ecological conditions.

Comparisons of genomes using DNA microarrays provide a rapid and cost-effective method to assess the presence, absence, and divergence of sequences in closely related organisms (Gresham et al. 2008). In fungi, this method has been used to characterize genome variability in *Saccharomyces cerevisiae* (Winzeler et al. 2003; Carreto et al. 2008; Dunn et al. 2005), *Candida albicans* (Moran et al. 2004), and *Cryptococcus neoformans* (Hu et al. 2008). In agreement with the CGH data presented for *P. involutus*, these studies have demonstrated that strains and closely related species of fungi can display an extensive variation in gene copy numbers. Additionally, variation in copy numbers is more likely to be found in genes belonging to certain functional categories. In *S. cerevisiae*, these categories comprise transporters, proteins of the major facilitator superfamily, and genes involved in drug response (Dunn et al. 2005; Carreto et al. 2008). Transporters and members of the major facilitator superfamily were also among genes varying in copy numbers in *P. involutus.* When interpreting the CGH data of the *P. involutus*, it has to be remembered that the examined strains in this study are dikaryotic and each cell of the mycelium contains two haploid nuclei. Due to the fact that the genome of *P. involutus* has not yet been sequenced, the microarray probes cannot be mapped to a specific haploid nucleus. Hence, it is not known to what extent deletions and amplifications in the two haploid genomes, respectively, contribute to the observed variations in gene copy numbers among the analyzed isolates.

We have previously shown that an incompatible and a compatible strain within PS III differ in gene content and in the expression patterns of plant-induced genes. Furthermore, we demonstrated that the sequences of several proteins have diverged at different rates in the two strains, probably because of changes in their selection pressures (Le Quéré et al. 2004, 2006). As a first attempt to examine whether incompatible strains have evolved by similar or different genomic mechanisms in the different lineages of *P. involutus*, we compared the differences in gene content between incompatible and compatible and closely related strains from three genetically isolated lineages (PS I, II, and III). Despite the fact that the contrasted strains differed significantly in gene content, none of the 390 variable gene reporters were changed in a similar way in the three pairs. Accordingly, our analysis provides no evidence that the emergence of incompatible phenotypes with narrow host potential among PS I, II, and III is associated with the gain or loss of similar genes. However, the possibility of parallel changes in gene content cannot be ruled out because the cDNA array used in this work contains only a portion (approximately 53%) of the total number of genes assumed to be present in the genome of *P. involutus*. It should also be noted that parts of the variations in gene content between the strains might not be associated with phenotypic adaptations but rather may be the result of nonadaptive forces such as genetic drift and mutations (Lynch 2007).

Acknowledgements This study was supported by grants from the Swedish Research Council (VR) and the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS). Custom microarrays were produced at the DNA Microarray Resource Center at the BioMedical Center B10 in Lund, and DNA sequencing was performed at the Center of Genomic Ecology at the Ecology Building in Lund. We acknowledge Dr. Björn Canbäck and Balaji Rajashekar for help with bioinformatics. We thank Dr. Susanne Erland for valuable discussion on this work.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS (2004) UniProt: the Universal Protein knowledgebase. Nucl Acids Res 32:D115–D119

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29

Brun A, Chalot M, Finlay RD, Söderström B (1995) Structure and function of the ectomycorrhizal association between *Paxillus involutus* (Batsch) Fr. and *Betula pendula* (Roth.). I. Dynamics of mycorrhiza formation. New Phytol 129:487–493

Burgess T, Dell B, Malajczuk N (1994) Variation in mycorrhizal development and growth stimulation by 20 *Pisolithus* isolates inoculated on to *Eucalyptus grandis* W. Hill ex Maiden. New Phytol 127:731–739

Cairney JWG (1999) Intraspecific physiological variation: implications for understanding functional diversity in ectomycorrhizal fungi. Mycorrhiza 9:125–135

Carreto L, Eiriz MF, Gomes AC, Pereira PM, Schuller D, Santos MA (2008) Comparative genomics of wild type yeast strains unveils important genome diversity. BMC Genomics 9:524

Duddridge JA (1986) The development and ultrastructure of ectomycorrhizas III. Compatible and incompatible interactions between *Suillus grevillei* (Klotzch) Sing., and 11 species of ectomycorrhizal host in vitro in the absence of exogenous carbohydrate. New Phytol 103:457–464

Dunn B, Levine RP, Sherlock G (2005) Microarray karyotyping of commercial wine yeast strains reveals shared, as well as unique, genomic signatures. BMC Genomics 6:53

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863–14868

Ek H, Andersson S, Arnebrant K, Söderström B (1994) Growth and assimilation of NH^{4+} and NO^{3-} by *Paxillus involutus* in association with *Betula pendula* and *Picea abies* as affected by substrate pH. New Phytol 128:629–637

Finlay RD, Ek H, Odham G, Söderström B (1988) Mycelial uptake, translocation and assimilation of nitrogen from ^{15}N-labelled ammonium by *Pinus sylvestris* plants infected with four different ectomycorrhizal fungi. New Phytol 110:59–66

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer EL, Bateman A (2006) Pfam: clans, web tools and services. Nucl Acids Res 34:D247–D251

Fries N (1985) Intersterility groups in *Paxillus involutus*. Mycotaxon 24:403–410

Gafur A, Schützendübel A, Langenfeld-Heyser R, Fritz E, Polle A (2004) Compatible and incompetent *Paxillus involutus* isolates for ectomycorrhiza formation in vitro with poplar (*Populus* x *canescens*) differ in H_2O_2 production. Plant Biol 6:91–99

Goffeau A, Park J, Paulsen IT, Jonniaux JL, Dinh T, Mordant P, Saier MH Jr (1997) Multidrug-resistant transport proteins in yeast: complete inventory and phylogenetic characterization of yeast open reading frames with the major facilitator superfamily. Yeast 13:43–54

Gresham D, Dunham MJ, Botstein D (2008) Comparing whole genomes using DNA microarrays. Nat Rev Genet 9:291–302

Hahn C, Agerer R (1999) Studies on the *Paxillus involutus*-complex. Nova Hedwigia 69:241–310

Hedh J, Samson P, Erland S, Tunlid A (2008) Multiple gene genealogies and species recognition in the ectomycorrhizal fungus *Paxillus involutus*. Mycol Res 112:965–975

Hu G, Liu I, Sham A, Stajich JE, Dietrich FS, Kronstad JW (2008) Comparative hybridization reveals extensive genome variation in the AIDS-associated pathogen *Cryptococcus neoformans*. Genome Biol 9:R41

Johansson T, Le Quéré A, Ahrén D, Söderström B, Erlandsson R, Lundeberg J, Uhlén M, Tunlid A (2004) Transcriptional responses of *Paxillus involutus* and *Betula pendula* during formation of ectomycorrhizal root tissue. Mol Plant Microbe Interact 17:202–215

Laiho O (1970) *Paxillus involutus* as a mycorrhizal symbiont of forest trees. Acta Forest Fenn 106:1–73

Lamhamedi MS, Fortini JA, Kope HH, Kropp BR (1990) Genetic varaition in ectomycorrhiza formation by *Pisolithus arhizus* on *Pinus pinaster* and *Pinus banksiana*. New Phytol 115:689–697

Le Quéré A, Johansson T, Tunlid A (2002) Size and complexity of the nuclear genome of the ectomycorrhizal fungus *Paxillus involutus*. Fung Genet Biol 36:234–241

Le Quéré A, Schützendübel A, Rajashekar B, Canbäck B, Hedh J, Erland S, Johansson T, Tunlid A (2004) Divergence in gene expression related to variation in host specificity of an ectomycorrhizal fungus. Mol Ecol 13:3809–3819

Le Quéré A, Astrup-Eriksen K, Rajashekar B, Schützendübel A, Canbäck B, Johansson T, Tunlid A (2006) Screening for rapidly evolving genes in the ectomycorrhizal fungus *Paxillus involutus* using cDNA microarrays. Mol Ecol 15:535–550

Lynch M (2007) The frailty of adaptive hypotheses for the origins of organismal complexity. Proc Natl Acad Sci U S A 104:8597–8604

Manning G, Plowman GD, Hunter T, Sudarsanam S (2002) Evolution of protein kinase signaling from yeast to man. Trends Biochem Sci 27:514–520

Martin F, Aerts A, Ahren D, Brun A, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buee M, Brokstein P, Canback B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbe J, Lin YC, Legue V, Le TF, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Oudot-Le Secq MP, Peter M, Quesneville H, Rajashekar B, Reich M, Rouhier N, Schmutz J, Yin T, Chalot M, Henrissat B, Kues U, Lucas S, Van de PY, Podila GK, Polle A, Pukkila PJ, Richardson PM, Rouze P, Sanders IR, Stajich JE, Tunlid A, Tuskan G, Grigoriev IV (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature 452:88–92

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F, Coutinho PM, Henrissat B, Berka R, Cullen D, Rokhsar D (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22:695–700

Molina R, Trappe JM (1982) Patterns of ectomycorrhizal host specificity and potential among pacific northwest conifers and fungi. Forest Sci 28:423–458

Molina R, Massicotte H, Trappe JM (1992) Specificity phenomena in mycorrhizal symbiosis: community-ecological consequences and practical implications. In: Allen MF (ed) Mycorrhizal functioning, an interactive plant-fungal process. Chapmann and Hall, New York, pp 357–423

Moran G, Stokes C, Thewes S, Hube B, Coleman DC, Sullivan D (2004) Comparative genomics using *Candida albicans* DNA microarrays reveals absence and divergence of virulence-associated genes in *Candida dubliniensis*. Microbiology 150:3363–3382

Nelson DR (1999) Cytochrome P450 and the individuality of species. Arch Biochem Biophys 369:1–10

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

Ridley AJ (2001) Rho family proteins: coordinating cell responses. Trends Cell Biol 11:471–477

Smith SE, Read DJ (eds) (1997) Mycorrhizal symbiosis, 2nd edn. Academic, San Diego

Taylor AF, Alexander I (2005) The ectomycorrhizal symbiosis: life in the real world. Mycologist 19:102–112

Trappe JM (1962) Fungus associates of ectotrophic mycorrhizae. Bot Rev 28:538–606

Tunlid A, Talbot NJ (2002) Genomics of parasitic and symbiotic fungi. Curr Opin Microbiol 5:513–519

Westfall PJ, Ballon DR, Thorner J (2004) When the stress of your environment makes you go HOG wild. Science 306:1511–1512

Winzeler EA, Castillo-Davis CI, Oshiro G, Liang D, Richards DR, Zhou Y, Hartl DL (2003) Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays. Genetics 163:79–89

Wolfinger RD, Gibson G, Wolfinger ED, Bennett L, Hamadeh H, Bushel P, Afshari C, Paules RS (2001) Assessing gene significance from cDNA microarray expression data via mixed models. J Comp Biol 8:625–637

Wong KKY, Montpetit D, Piche Y, Lei J (1990) Root colonization by four closely related genotypes of the ectomycorrhizal basidiomycete *Laccaria bicolor* (Maire) Orton - comparative studies using electron microscopy. New Phytol 116:669–680

Wösten HA (2001) Hydrophobins: multipurpose proteins. Annu Rev Microbiol 55:625–646