Divergence in gene expression related to variation in host specificity of an ectomycorrhizal fungus

Le Quéré, Antoine; Schutzendubel, Andres; Rajashekar, Balaji; Canbäck, Björn; Hedh, Jenny; Erland, Susanne; Johansson, Tomas; Tunlid, Anders

Published in:
Molecular Ecology

DOI:
10.1111/j.1365-294X.2004.02369.x

Published: 2004-01-01

Citation for published version (APA):
Divergence in gene expression related to variation in host specificity of an ectomycorrhizal fungus

ANTOINE LE QUÉRÉ,† ANDRES SCHÜTZENDÜBEL,‡ BALAJI RAJASHEKAR, BJÖRN CANBÄCK, JENNY HEDH, SUSANNE ERLAND, TOMAS JOHANSSON and ANDERS TUNLID

Department of Microbial Ecology, Lund University, Ecology Building, SE-223 62 Lund, Sweden

Abstract

Ectomycorrhizae are formed by mutualistic interactions between fungi and the roots of woody plants. During symbiosis the two organisms exchange carbon and nutrients in a specific tissue that is formed at the contact between a compatible fungus and plant. There is considerable variation in the degree of host specificity among species and strains of ectomycorrhizal fungi. In this study, we have for the first time shown that this variation is associated with quantitative differences in gene expression, and with divergence in nucleotide sequences of symbiosis-regulated genes. Gene expression and sequence evolution were compared in different strains of the ectomycorrhizal fungus Paxillus involutus; the strains included Nau, which is not compatible with birch and poplar, and the two compatible strains Maj and ATCC200175. On a genomic level, Nau and Maj were very similar. The sequence identity was 98.9% in the 16 loci analysed, and only three out of 1075 genes analysed by microarray-based hybridizations had signals indicating differences in gene copy numbers. In contrast, 66 out of the 1075 genes were differentially expressed in Maj compared to Nau after contact with birch roots. Thirty-seven of these symbiosis-regulated genes were also differentially expressed in the ATCC strain. Comparative analysis of DNA sequences of the symbiosis-regulated genes in different strains showed that two of them have evolved at an enhanced rate in Nau. The sequence divergence can be explained by a decreased selection pressure, which in turn is determined by lower functional constraints on these proteins in Nau as compared to the compatible strains.

Keywords: comparative genomics, ectomycorrhiza, gene expression, host specificity, microarray

Received 22 June 2004; revision received 9 September 2004; accepted 9 September 2004

Introduction

It has been estimated that at least 6000 species of fungi, primarily basidiomycetes with some ascomycetes and zygomycetes, form mutualistic relationships with woody plants (Malloch et al. 1980). These ecologically very important associations are called ectomycorrhizae and are the dominant mycorrhizal type associated with trees in temperate and boreal ecosystems (Smith & Read 1997). In ectomycorrhizae, the fungal partner obtains photosynthetic sugars from the host plant while, in return, the plant receives mineral nutrients from the fungus. The exchange of nutrients occurs in a specific symbiotic tissue that is formed between the fungal hyphae and the host roots. This tissue consists of a mantle, which develops from the fungal hyphae surrounding the root, and a Hartig net, which is formed by the hyphae penetrating between the outer cells of the root (Smith & Read 1997).

Analyses of DNA-based phylogenies have shown that the ancestors of the ectomycorrhizal homobasidiomycetes were free-living and that mycorrhizal symbionts have evolved repeatedly from saprophytic precursors (Hibbett et al. 2001). Many of these basidiomycetes form symbioses with several host species. Furthermore, it is known that fungi can differ markedly in their ability to form ectomycorrhizae and to promote the growth of the host plant (Smith & Read 1997). Screenings, mainly based on records...
of sporocarp-plant associations, have shown that ectomycorrhizal fungi include both generalist and specialist species (Trappe 1962). Studies have also demonstrated that the magnitude of intraspecific differences in host preferences can be as large as the between-species differences (Smith & Read 1997; Cairney 1999). This variation is associated with differences in the ability to form the mantle and Hartig net. The molecular basis for the differential infectivity is not clear, but it has been suggested to reside in the capacity to produce hormones, extracellular enzymes and not yet identified recognition factors (Cairney 1999). Furthermore, it has been shown that the patterns of polypeptides produced by strains of the ectomycorrhizal fungus Pisolithus tinctorius are correlated with the infectivity of the strain (Burgess et al. 1995).

The genomic mechanisms that could account for the variation in host preferences between species or strains of ectomycorrhizal fungi are, however, not known. Generally such phenotypic differences could be the result of variations in gene content, of quantitative differences in gene expression, or of structural differences in gene products (King & Wilson 1975; Wray et al. 2003). DNA microarray technology has opened up new possibilities of comparing transcript abundance between closely related species and/or strains, and of identifying genes that are associated with morphological and physiological divergence (Ferea et al. 1999; Jin et al. 2001; Enard et al. 2002). In addition, DNA microarray-based comparative genomic hybridization can be used to assess genome rearrangements like amplification or deletion at single gene resolution, which might play an important role in adaptive evolution (Hughes et al. 2000; Dunham et al. 2002).

We have used a cDNA microarray to examine the differences in gene expression between incompatible and compatible strains of the ectomycorrhizal fungus Paxillus involutus (Basidiomycetes; Boletales). This fungus is widespread in the northern hemisphere and forms ectomycorrhizae with many species of coniferous and deciduous trees (Wallander & Söderström 1999). Strains of P. involutus are known to differ in their ability to form ectomycorrhizal associations with various hosts (Laith 1970; Gafur et al. 2004). In this study we included a strain Nau that was originally isolated from oak in France (Marschner et al. 1999). It was recently shown that Nau, in contrast to several other strains of P. involutus, does not infect poplar in vitro (Gafur et al. 2004). For the comparative analysis, we identified the closest evolutionary relative of Nau to be the poplar-compatible strain Maj. In addition, we included a third strain ATCC 200175 (henceforward called ATCC), in which gene regulation during the formation of ectomycorrhizae on birch (Betula pendula) has previously been studied (Johansson et al. 2004). The cDNA array contained cDNA probes putatively representing 1075 P. involutus genes. The probes were obtained from a nonredundant set of expressed sequence tags (ESTs) produced either from cDNA libraries of the ectomycorrhizal tissue formed between the ATCC strain and birch mycelium or from saprophytically grown mycelia of the ATCC strain (Johansson et al. 2004). The fraction of P. involutus genes that is represented on the array can be estimated to correspond to approximately 14% of the total number of genes in the fungal genome assuming a gene content of 7700 (Le Quéré et al. 2002).

Materials and methods

Biological material, media and culture conditions

The strains of Paxillus (Table 1) were maintained on Modified Melin-Norkrans (MMN) agar medium (Brun et al. 1995). An axenic system for the formation of ectomycorrhizae between Betula pendula (Skuleskogen, SkogForsk, Sweden) and Paxillus involutus was used with reduced amounts of carbon (0.25 g glucose/L) (Brun et al. 1995). The association was grown under controlled conditions of 22 °C day/15 °C night temperatures, a day/night period of 15 h/9 h, and 60% humidity. The material was harvested after 14 days. After fixation the fungal-root tissues were examined under a light microscope (Axioplan, Carl Zeiss) (Gafur et al. 2004).

cDNA microarray, labelling and hybridization analysis

Microarrays had previously been constructed (Johansson et al. 2004; Le Quéré & Wright et al. manuscript in preparation) using a nonredundant set of 2159 ESTs, either of fungal (1075 ESTs) or plant (1074 ESTs) origin. The plant ESTs, and 10 ESTs with uncertain origin, were excluded from this investigation. The mycorrhizal tissue (Maj and ATCC) or root tips with attached hyphae (Nau) was harvested and total RNA was isolated as described (Johansson et al. 2004). Antisense RNA (aRNA) was synthesized by T7 polymerase in vitro transcription using one round of amplification (MessageAmp, Ambion (Europe) Ltd). The aRNA was analysed by gel electrophoresis and the titre was determined spectrophotometrically. The aRNA targets were labelled (Cy3 or Cy5) and the microarray slides were hybridized and scanned as previously described (Le Quéré, Wright et al. manuscript in preparation). The experiments were designed as two-sample comparisons (ATCC and Maj; ATCC and Nau; Maj and Nau) using three independent biological replicates, including technical and dye-swapped control experiments.

Statistical analysis

Data images were manually inspected and low-quality spots were excluded from further analyses. Reporters
### Table 1  
Strains of *Paxillus* used for ITS phylogeny

<table>
<thead>
<tr>
<th>Order</th>
<th>Species and strain name</th>
<th>Location</th>
<th>Habitat</th>
<th>Woody plants within a 10-m radius</th>
<th>Compatibility in laboratory</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Paxillus involutus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Pi11</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Picea</td>
<td></td>
<td>AF167698†</td>
</tr>
<tr>
<td>2</td>
<td>Pi12</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Fagus, Quercus</td>
<td></td>
<td>AF167699†</td>
</tr>
<tr>
<td>3</td>
<td>Pi3</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Salix</td>
<td></td>
<td>AF167695†</td>
</tr>
<tr>
<td>4</td>
<td>PiM3</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Fagus</td>
<td></td>
<td>AF167700†</td>
</tr>
<tr>
<td>5</td>
<td>Pi9H</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Picea</td>
<td></td>
<td>AF167696†</td>
</tr>
<tr>
<td>6</td>
<td>SE03-09–2808</td>
<td>Sweden, Småland</td>
<td>Pinus heath forest</td>
<td>Picea, Picea, Betula</td>
<td>Betula</td>
<td>AY589920*</td>
</tr>
<tr>
<td>7</td>
<td>Pi10H</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Betula</td>
<td></td>
<td>AF167697†</td>
</tr>
<tr>
<td>8</td>
<td>Pi01SE</td>
<td>Sweden, Viken</td>
<td>Pine sand forest</td>
<td>Pinus</td>
<td></td>
<td>AY589918*</td>
</tr>
<tr>
<td>9</td>
<td>Pi08BE</td>
<td>Belgium, Maatheide</td>
<td></td>
<td></td>
<td></td>
<td>AY589919*</td>
</tr>
<tr>
<td>10</td>
<td>ATCC200175</td>
<td>Scotland</td>
<td></td>
<td>Betula</td>
<td></td>
<td>AY589913*</td>
</tr>
<tr>
<td>11</td>
<td>SE03-10–0903</td>
<td>Sweden, Torna Hällestad</td>
<td>Fagus forest</td>
<td>Sorbus, Fagus, Betula, Quercus</td>
<td></td>
<td>AY589916*</td>
</tr>
<tr>
<td>12</td>
<td>PA03-09–1401</td>
<td>Sweden, Hässleholm</td>
<td>Mixed forest</td>
<td>Pinus</td>
<td></td>
<td>AY589914*</td>
</tr>
<tr>
<td>13</td>
<td>SE03-10–0501</td>
<td>Sweden, Kjuekull</td>
<td>Mixed tree stand on pasture</td>
<td>Fagus (Fraxinus), Betula, Quercus, Corylus</td>
<td></td>
<td>AY589921*</td>
</tr>
<tr>
<td>14</td>
<td>HW03-09–2501</td>
<td>Sweden, Torna Hällestad</td>
<td>Road side outside spruce hedge</td>
<td>Betula, Picea</td>
<td></td>
<td>AY589922*</td>
</tr>
<tr>
<td>15</td>
<td>SE03-07–1610</td>
<td>Sweden, Lund</td>
<td>Small grove on public lawn</td>
<td>Populus (Cornus), Tilia, Acer, Fagus, Quercus, Sorbus, Corylus</td>
<td></td>
<td>AY589912*</td>
</tr>
<tr>
<td>16</td>
<td>Pi12</td>
<td>Germany, Bavaria</td>
<td></td>
<td></td>
<td></td>
<td>AY589911*</td>
</tr>
<tr>
<td>17</td>
<td>PiM4</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Betula</td>
<td></td>
<td>AY589913†</td>
</tr>
<tr>
<td>18</td>
<td>PiM1</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Quercus</td>
<td></td>
<td>AY589920†</td>
</tr>
<tr>
<td>19</td>
<td>PiM2</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Populus</td>
<td></td>
<td>AY58994†</td>
</tr>
<tr>
<td>20</td>
<td>SE03-07–1622</td>
<td>Sweden, Lund</td>
<td>Small grove on public lawn</td>
<td>Populus (Cornus), Tilia, Acer, Fagus, Quercus, Sorbus</td>
<td></td>
<td>AY589911*</td>
</tr>
<tr>
<td>21</td>
<td>Pi1</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Tilia</td>
<td></td>
<td>AY58990†</td>
</tr>
<tr>
<td>22</td>
<td>SE03-07–1001</td>
<td>Sweden, Lund</td>
<td></td>
<td>Populus (Cornus), Tilia, Acer, Fagus, Quercus, Sorbus</td>
<td></td>
<td>AY589910*</td>
</tr>
<tr>
<td>23</td>
<td>Nau</td>
<td>France</td>
<td></td>
<td></td>
<td></td>
<td>AY589915*</td>
</tr>
<tr>
<td>24</td>
<td>Maj</td>
<td>France</td>
<td></td>
<td></td>
<td></td>
<td>AY589917*</td>
</tr>
<tr>
<td>25</td>
<td>Pv2</td>
<td>Canada</td>
<td></td>
<td></td>
<td></td>
<td>AY589918*</td>
</tr>
<tr>
<td></td>
<td><strong>Paxillus vernalis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>304</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Alnus</td>
<td></td>
<td>AY589919*</td>
</tr>
<tr>
<td>27</td>
<td>133</td>
<td>Germany, Bavaria</td>
<td></td>
<td>Alnus</td>
<td></td>
<td>AY589920†</td>
</tr>
<tr>
<td></td>
<td><strong>Melanogaster broomeianus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Mbr1</td>
<td></td>
<td></td>
<td>Tilia</td>
<td></td>
<td>AY589921*</td>
</tr>
</tbody>
</table>

Data taken from: *this study, †(Jarosch & Bresinsky 1999), and ‡NCBI.
related to \( B. \) pendula (Le Quéré, Wright et al. manuscript in preparation) were excluded and only data for reporters representing fungal and control genes were extracted. For those spots remaining, the raw fluorescence intensities for each channel on each slide were collected. For each channel, the mean background fluorescence was calculated. After local background correction for each spot, the reporters yielding intensities less than twice the background intensity were excluded and the fluorescence for the remaining reporters was multiplied to give a common channel mean of 5000 fluorescence units for each slide. As a result, data for 1022 reporters remained in the data set. The statistical approach, the mixed model analyses of variance (ANOVA) (Wolfinger et al. 2001), served two purposes: first, normalization of the data to remove systemic biases that may have affected all genes simultaneously, such as differences in the amount of RNA that was labelled for a particular replicate of a treatment; and second, assessment of the contribution of biological and experimental sources of error to the variation in the expression of each individual gene. This procedure used differences in normalized expression levels, rather than ratios, as the unit of analysis of expression differences. We subjected the corrected \( \log_{2} \) transformed measures \( y_{ijg} \) for the gene \( g \) \((g = 1, \ldots, 1022)\), which included scores for 50 573 fungal spot measures, to a normalization model of the form: \( y_{ijg} = \mu + A_i + D_j + (A \times D)_i \beta + \mathbf{e}_{ijg} \), where \( \mu \) is the sample mean, \( A_i \) is the effect of the \( i \)th array \((i = 1–11)\), \( D_j \) is the effect of the \( j \)th dye \((Cy3 \text{ or } Cy5)\), \( (A \times D)_i \beta \) is the array–dye interaction (channel effect) and \( \mathbf{e}_{ijg} \) is the stochastic residual. We then subjected the residuals from this model, which can regarded as a crude indicator of the relative expression level (and are referred to in the text as ‘normalized expression levels’), to 1015 gene-specific models of the form: \( y_{ijg} = \mu + A_i + S_j + D_j + \mathbf{e}_{ijg} \), where \( S_j \) is the \( j \)th strain \([ATCC, Maj, \text{ or } Nau; \text{ two degrees of freedom} (d.f.)]\). In the gene models, which were fitted using PROC MIXED in SAS (SAS/STAT Software Version 8, SAS Institute Inc.), the \( A_i \) variable controls for spot effects and is random (10 d.f., leaving 8 d.f. for the residual error). The microarray data are available at the EBI-EMBL ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) (accession number E-MEXP-179).

The overall distance between the expression profiles of two strains was calculated using the hybridization values of the transcriptome profiles as described by Enard et al. (2002). Briefly, the distance between the samples was calculated by summing the \( \log_{2} \) values of the hybridization-signal ratios for genes that were expressed in all comparisons. The resulting distance matrix was used to build neighbour-joining trees (Saitou & Nei 1987) using the PHYLP program (Felsenstein 1993). In a similar way, the genomic distances between the strains were calculated using data from comparative genomic hybridizations. In this experiment, DNA from the different strains was cohybridized (ATCC and Maj; ATCC and Nau; Maj and Nau) to cDNA microarrays that were identical to those used for the expression studies (Le Quéré, Rajashekar et al. manuscript in preparation).

**DNA sequencing**

For complete DNA sequencing of the cDNAs representing candidate genes, plasmid clones were selected from a collection of EST clones constructed in the pTriplEx vector (BD Clontech) and maintained as bacterial lysates (Johansson et al. 2004). The bacterial lysate was used for retransformation into Escherichia coli DH5\( \alpha \) and transformants were verified by standard procedures. Plasmids were prepared and used as starting material for DNA sequencing using the dideoxy chain-termination method, employing the BigDye Terminator Kit (Applied Biosystems) and either using the pTriplEx2-specific universal forward primer P104 (\( 5'-GGGAACGGCCATTTGTTG-3' \)), the reverse primer T23V (\( 5'-T23V-3', V = A, G \text{ or } C \)), or template-specific primers. The polymerase chain reaction (PCR) products were purified by isopropanol precipitation and finally loaded onto an ABI3100 DNA sequencer (Applied Biosystems). In addition, genomic regions of 17 different loci were PCR amplified and sequenced from different Paxillus strains using template specific primers (Table 2). Sequence information was validated and assembled using the SEQUENCER 3.0 software (Gene Codes Corporation).

**Phylogenetic analysis**

To examine the phylogenetic relationship between the Paxillus strains (Table 1), the internal transcript spacer region (ITS) of the rDNA was amplified by PCR and sequenced in both directions using the primer pair ITS1–ITS4 (White et al. 1990). To construct phylogenetic trees of the \( \text{cphC}, \text{cchA} \) and \( \text{rabA} \) genes, we retrieved homologous genes from GenBank (Benson et al. 2004), the COGEME EST database (Soanes et al. 2002) and from the genome database of Phanerochaete chrysosporium (http://www.jgi.doe.gov) using BLAST (Altschul et al. 1990) searches with the ATCC strain as query sequence. The sequences were then aligned using CLUSTAL-W (Thompson et al. 1994). Regions with ambiguous sites were removed manually from the amino acid or nucleotide sequence alignments by SEAVIEW (Galtier et al. 1996). Phylogenetic trees of nucleotide sequence alignments for the ITS region, \( \text{cphC} \) and \( \text{cchA} \) genes were calculated by the maximum likelihood method using PAUP (Swofford 1998). A bootstrap of 100 replicates was generated in all three cases. The software MODELTEST (Posada & Crandall 1998) was used to evaluate the appropriate models and parameter values used in the likelihood analysis. An unrooted phylogenetic tree of the \( \text{rabA} \) gene was constructed using PHYLO_WIN software (Galtier et al. 1996). Protein
alignments were made using CLUSTAL-W (Thompson et al. 1994) and a neighbour-joining method (Saitou & Nei 1987) with 500 bootstrap replications was used to develop the phylogeny.

Calculation of substitution rates

Nucleotide sequence information for 13 Paxillus coding regions (Table 2) was translated and the protein alignments for each gene were made by CLUSTAL-W (Thompson et al. 1994). Then the protein alignment was used as a template to align the corresponding coding nucleotide sequences by TRANALIGN of emboss (Rice et al. 2000). The rates of nonsynonymous (\(d_N\)) and synonymous (\(d_S\)) nucleotide substitutions per site were estimated using the CRANN software (Creevey & McInerney 2003). The presumed species tree used in the input to CRANN was based on the ITS phylogeny.

### Table 2 Genes of Paxillus used for estimating per cent identity* and \(d_N/d_S\) analysis†

<table>
<thead>
<tr>
<th>Gene</th>
<th>GenBank homologue</th>
<th>Strains analysed</th>
<th>Alignment length</th>
<th>Regulation in ECM</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>actA</td>
<td>Actin 1 (gi:14194451) [Schizophyllum commune]</td>
<td>6‡</td>
<td>482</td>
<td>+ ‡</td>
<td>AY585949,6027–3188</td>
</tr>
<tr>
<td>β-tubA</td>
<td>beta tubulin 2 (gi:23477228) [Suillus bovinus]</td>
<td>6‡</td>
<td>316</td>
<td>+ ‡</td>
<td>AY585948,6022–2656</td>
</tr>
<tr>
<td>calA</td>
<td>Calmodulin (gi:18150814) [Paxillus involutus]</td>
<td>5§</td>
<td>432</td>
<td>+ §</td>
<td>AY586017–2188, AAD17455***</td>
</tr>
<tr>
<td>cipC1</td>
<td>CipC protein (gi:20429042) [Emericella nidulans]</td>
<td>5§</td>
<td>256</td>
<td>+ §††</td>
<td>AY586008–12†††, CAC87272***</td>
</tr>
<tr>
<td>cipC2</td>
<td>CipC protein (gi:20429042) [Emericella nidulans]</td>
<td>6‡</td>
<td>234</td>
<td>+ ‡</td>
<td>AY585947,6003–7†††</td>
</tr>
<tr>
<td>cchA</td>
<td>copper chaperone (gi:25956321) [Trametes versicolor]</td>
<td>4††</td>
<td>156</td>
<td>+ ††</td>
<td>AY586013–16†††, AAN75572***</td>
</tr>
<tr>
<td>ppiA</td>
<td>putative cyclophilin (gi:16943775) [Pleurotus ostreatus]</td>
<td>6‡</td>
<td>444</td>
<td>+ ‡</td>
<td>AY585941,60–64§§</td>
</tr>
<tr>
<td>gpiA</td>
<td>Glc-6-P isomerase (gi:40713647) [Agaricus bisporus]</td>
<td>6††</td>
<td>1483</td>
<td></td>
<td>AY585946,5998–6002§§</td>
</tr>
<tr>
<td>locA</td>
<td>lectin (gi:18478668) [Xerocomus chrysenteron]</td>
<td>5§</td>
<td>354</td>
<td>+ †††</td>
<td>AY585973–77†††, AAL73235***</td>
</tr>
<tr>
<td>gstA</td>
<td>glutathione S-transferase (gi:1353751) [Naegleria fowleri]</td>
<td>5§</td>
<td>509</td>
<td>+ †††</td>
<td>AY585993–97†††, XP_391216***</td>
</tr>
<tr>
<td>hetC1</td>
<td>het-c2 protein (gi:2133321) [Podospora anserina]</td>
<td>6††</td>
<td>508</td>
<td>+ ††</td>
<td>AY585945,88–92§§</td>
</tr>
<tr>
<td>hxtA</td>
<td>hexose transporter (gi:6715095) [Aspergillus parasiticus]</td>
<td>5§§</td>
<td>1532</td>
<td>+ §§</td>
<td>AY585978–82§§, XP_391358***</td>
</tr>
<tr>
<td>hspA</td>
<td>small heat shock protein (gi:21950718) [Laccaria bicolor]</td>
<td>6‡</td>
<td>453</td>
<td>+ §§</td>
<td>AY585944,83–87§§</td>
</tr>
<tr>
<td>micA</td>
<td>mitochondrial carrier protein (gi:19075302)</td>
<td>4**</td>
<td>770</td>
<td>+ ‡</td>
<td>AY585943,70–72§§</td>
</tr>
<tr>
<td>ndkA</td>
<td>nucleoside diphosphate kinase (gi:32405858) [Schizosaccharomyces pombe]</td>
<td>6‡</td>
<td>427</td>
<td>+ ‡</td>
<td>AY585942,65–69§§</td>
</tr>
<tr>
<td>ptrA</td>
<td>inorganic phosphate transporter (gi:586531) [Saccharomyces cerevisiae]</td>
<td>6‡</td>
<td>473</td>
<td>+ ‡</td>
<td>AY585940,55–59§§</td>
</tr>
<tr>
<td>rhaA</td>
<td>small GTPase (gi:710852) [Mus musculus]</td>
<td>5§§</td>
<td>485</td>
<td>+ †††</td>
<td>AY585950–4†††, EAK82432***</td>
</tr>
</tbody>
</table>

ECM, ectomycorrhiza.

*All genes (except micA) from ATCC, Maj and Nau strains were used for per cent identity estimation.
†β-tubA, gpiA, hspA and micA were not used for \(d_N/d_S\) analysis.
‡Isolates ATCC200175, Pi01SE, Pi08BE, Maj, Nau and Pi01DE.
§Isolates ATCC200175, Pi01SE, Pi08BE, Maj and Nau.
¶Isolates ATCC200175, Pi08BE, Maj and Nau.
**Isolates ATCC200175, Pi01SE, Pi08BE and Pi01DE.
††Taken from Johansson et al. 2004.
‡‡Taken from Johansson et al. (manuscript in preparation).
§§Taken from Johansson et al. (manuscript in preparation).
¶¶This study.
***Used as outgroup in \(d_N/d_S\) analysis.

Results

Phylogeny

To identify the closest evolutionary relative to Nau, an ITS phylogenetic tree using a number of strains and isolates of Paxillus was constructed (Fig. 1). The strain Maj that has been shown to be compatible with poplar (Gafur et al. 2004) was recognized as the closest relative to Nau. The Maj and Nau strains formed a clade with high bootstrap support (value 89) within the so-called ‘park’ group of Paxillus involutus strains (Fries 1985; Jarosch & Bresinsky 1999). The ATCC strain was found in another clade with high bootstrap support (value 83). This clade contained strains collected in various coniferous or mixed forests. This comprises the ‘forest’ group of P. involutus (Fries 1985; Jarosch & Bresinsky 1999).

The close genetic relationship between Maj and Nau was also evident from the DNA sequences of 16 loci (Table 2),
which showed a 98.9% identity between Maj and Nau. The identities between ATCC and Maj, and ATCC and Nau in these loci were 95.6 and 95.2%, respectively.

Phenotypic differences

After 2 weeks of interaction, the ATCC and Maj strains of *P. involutus* colonized the roots of birch seedlings and developed a typical ectomycorrhizal tissue, as visualized by the hampered elongation and swelling of the fine roots, and the presence of a pseudoparenchymatous mantle. In addition, the hyphae of these strains had penetrated between the outer epidermal cells of the root and formed the Hartig net where nutrients and carbohydrates are exchanged between the two symbiotic partners (Fig. 2a,b,d,e) (Brun et al. 1995). In contrast, the Nau strain did not develop such a mycorrhizal interface (Fig. 2c,f). Thus the differences in compatibility between Maj and Nau on birch were similar to those observed on poplar (Gafur et al. 2004). It is not known whether the examined isolate of Nau has completely lost its ability to form ectomycorrhiza with all potential hosts including oak.

Differences in gene expression

The patterns of gene expression for the ATCC, Maj and Nau strains after 14 days of contact with the birch were compared using cDNA microarrays. In total, 926 (86%), 747 (70%), and 747 (70%) of the 1075 fungal genes spotted on the cDNA arrays produced signal intensities above the twofold average background threshold in the hybridizations comparing ATCC and Maj, ATCC and Nau and Maj and Nau, respectively. The relative rates of evolutionary change in the transcriptomes of the three strains were estimated by summing the log2 values of the ratios of the hybridization signals for genes that were commonly expressed in all comparisons (Fig. 3). These analyses showed that the transcriptome distance between ATCC and Maj was slightly shorter than the distance between ATCC and Nau.

Following statistical analyses, we identified 66 genes that differed significantly in expression levels between the closely related compatible Maj and the incompatible Nau strain (Table 3). A subset of these genes [in total 37 genes, 36 being up-regulated and one down-regulated (EST clone CD273016)] was also differentially expressed in the compatible ATCC strain as compared to Nau. We considered that these commonly regulated genes in Maj and ATCC were associated with the colonization and development of the symbiotic ectomycorrhizal tissue. A visual presentation of the expression patterns of the 37 symbiosis-related genes is presented in Fig. 4 (left panel). Notably, 20 of them
have previously been reported to be regulated in the ecto-
mycorrhizal root tissue formed between the ATCC strain
and birch (Johansson et al. 2004) (Le Quéré, Wright et
al. manuscript in preparation).

Rapidly evolving genes
To assess the possibility that the sequences of regulated
genes have evolved at different rates in the compatible
and incompatible strains, we calculated the \(d_n\) and \(d_s\) values
for six regulated genes (cipC1, cipC2, cchA, lecA, gstA and
rabA). In addition, we included seven genes in the analysis
that had been found to be regulated in the interaction
between the ATCC strain and birch (Fig. 5). Out of a total
of 126 pairwise comparisons, 58 showed a \(d_n/d_s\) ratio below
0.6, which is indicative of purifying selection associated
with functional constraints. In 61 comparisons the f requencies
of \(d_n\) \(d_s\) or both had a zero value. In the remaining
seven comparisons the \(d_n/d_s\) ratio exceeded one. These were
only found in two genes (cchA and cipC1). Interestingly,
all the observations were between Nau and different
compatible strains. However, the differences between the

Table 3  Significant differences in gene expression between three
strains of Paxillus involutus during interaction with birch seedlings
(Betula pendula) as estimated by the mixed model ANOVA*

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Up</th>
<th>Down</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATCC and Maj</td>
<td>24</td>
<td>77</td>
<td>101</td>
</tr>
<tr>
<td>Maj and Nau</td>
<td>63</td>
<td>3</td>
<td>66</td>
</tr>
<tr>
<td>ATCC and Nau</td>
<td>66</td>
<td>38</td>
<td>104</td>
</tr>
</tbody>
</table>

*The number of genes with \(P < 0.05\) (Bonferroni corrected).

Analyses were made on 1075 fungal genes and three different
strains (ATCC, Maj and Nau).

Fig. 4  Visual representation of strain variation at the transcrip-
tome and genome levels for genes found to be related to
mycorrhizal efficiency in Paxillus involutus. The left panel
(Transcrip tome) shows log2-transformed relative expression levels
(residuals) from the mixed model ANOVA for 37 genes (out of 1075)
that were identified as commonly regulated in both the compatible
strains, ATCC and Maj, but not in the incompatible Nau strain.
Genes were hierarchically clustered using a matrix of correlation
coefficients (clust re 3.0). Higher and lower
expression levels than the calculated mean are indicated by red
and green, respectively. The next panel (Genome) shows data from
comparative genomic hybridization (array-CGH) of the ATCC,
Maj and Nau strains. The raw hybridization intensities were
normalized, log2-transformed and centred according to Le Quéré,
Rajashekar et al. (manuscript in preparation). The genes
were placed in order in accordance with the hierarchical clustering
of the expression data. Genes are identified by EST clone names and
the GenBank description for putative homologues. The red circles
and the blue diamonds (Le Quéré, Wright et al. manuscript in
preparation) mark genes that have previously been found to be
regulated in the interaction between the ATCC strain and birch.
The identities of each of the clones in boldface have been confirmed
by resequencing. Genes examined for sequence divergence (cipC1,
cipC2, cchA, lecA, gstA and rabA) are marked (cf. Table 2 and
Fig. 5).
and $d_s$ values were not statistically significant using the Z-test.

In *P. involutus*, we identified two *cipC* genes, one encoding the CipC1 protein (108 amino acids) and one the CipC2 protein (107 amino acids) with a sequence identity of 47%. A phylogenetic analysis of fungal *cipC* genes shows that the *cipC1* and *cipC2* genes of different strains of *P. involutus* were separated in two well-resolved clades with bootstrap support values of 95 and 97, respectively (Fig. 6). The duplication of the ancestral *cipC* gene has occurred rather recently, after the divergence of *Paxillus* from its ancestor within the bolete clade, as inferred from the phylogeny of ribosomal DNA sequences (Hibbett et al. 2001). Most importantly, the analysis shows that the *cipC1* sequence from Nau, diverged before the corresponding sequences from the compatible strains. The most probable explanation for this topology is the observed rate increase in the Nau strain. In contrast, the *cipC2* portion of the tree shows the same topology as the ITS tree (Fig. 1) where Nau and Maj are the most closely related strains. Based on the alignment of the CipC1 protein sequences (83 amino acids), we identified 15 substitutions between Nau and the closely related compatible strain Maj. The $d_n/d_s$ ratio was estimated at 1.32 for the corresponding nucleotide sequences. When comparing Maj with a more distant strain like the compatible ATCC, there were no amino acid substitutions whatsoever. Furthermore, alignment of intron sequences of the *cipC1* gene revealed nine substitutions between Nau and Maj, and none between Maj and ATCC.

The other gene that showed a $d_n/d_s$ ratio above one (1.76 when comparing Nau and Maj), was the *cchA* gene. Also in this case, the Nau gene has evolved rapidly resulting in 11 amino acid substitutions between Nau and Maj, but only three between the more distantly related Maj and ATCC.

The phylogenetic tree showed a well-supported *P. involutus* clade (bootstrap support value of 95) with Nau and Maj diverging before the other strains (Fig. 7).

Fig. 5 Comparison of rates of nonsynonymous ($d_n$) and synonymous ($d_s$) nucleotide substitutions per site. Each point represents a pairwise comparison made between genes (*cipC1*, *cchA*, or rest (actA, β-tubA, calA, ppiA, gpiA, lcaA, getA, hetC1, hxtA, hspA, micA, ndkA, ptrA, rabA)) from five different strains of *Paxillus involutus* (ATCC, Pi01SE, Pi08BE, Maj and Nau) and one strain of *P. flamentosus* (Pi01De). The highlighted gene names in the scatter plot refer to the following comparisons: (0) ATCC vs. Pi01SE; (1) ATCC vs. Pi08BE; (2) ATCC vs. Maj; (3) ATCC vs. Nau; (4) Pi01SE vs. Pi08BE; (5) Pi01SE vs. Maj; (6) Pi01SE vs. Nau; (7) Pi08BE vs. Maj; (8) Pi08BE vs. Nau; (9) Maj vs. Nau. The diagonal line shows the neutral expectation when $d_n$ is equal to $d_s$. See Table 2 for additional details.

Fig. 6 Phylogeny of *cipC* proteins. The maximum likelihood tree for *cipC* proteins based on nucleotide sequence information is shown. Values at nodes represent the bootstrap support value in per cent for 100 replicates. Accession numbers of the protein/nucleotide sequences from top to bottom are: AY586003, AY586005, AY586004, AY586006, AY586007, AY585947, AY586008, AY586011, AY586015, AY586009, AY586012, BU963955, Scaffold_4 (490068–490202, 490254–490334, http://www.jgi.doe.gov), BU963954, BU963951, CB010355, CB011958, CB010733, Scaffold_50 (61264–61419, 61464–61526, http://www.jgi.doe.gov), EAA71391, FsCon[0257] (COGEME database), UmCon[1708] (COGEME database), CAC87227, XP_323537, EAA52181, UmCon[0073] (COGEME database), CpCon[0039] (COGEME database), Cd64941, BcCon[008] (COGEME database), mg[0591] (COGEME database), BM863150 and CK326079.
The variation in expression profiles between the compatible and incompatible strains was not reflected in the genomic distances as analysed by comparative genomic hybridizations (Fig. 3). Analyses of the genomic array data showed that Maj and Nau were very similar and were equally distant from the ATCC strain. A statistical analysis using the mixed model ANOVA showed that only 0.3% (three out of 1075) of the genes had hybridization signals indicating differences in gene copy numbers between Maj and Nau. These three genes were not found to be among the symbiosis-regulated genes. A similar analysis showed that approximately 200 out of 1075 genes had significantly different signal intensities when comparing ATCC with either Maj or Nau. Ten of these variable genes were found among the symbiosis-regulated genes (Fig. 4).

Among the symbiosis-regulated genes was a homologue (lecA) to a family of saline-soluble, low molecular weight fungal lectins (carbohydrate-binding proteins) (Balogh et al. 2003). The function of these proteins is unclear but it has been proposed that they are involved in development, storage of nutrients and defense reactions (Balogh et al. 2003). Two glutathione S-transferase (GST) homologues were also up-regulated. One of them (gstA) displayed a high degree of homology to a GST from the human pathogen *Naegleria fowleri* (identity 49%). GSTs are recognized as detoxification enzymes and are involved in the removal of reactive oxygen species such as H$_2$O$_2$ (Sheehan et al. 2001). H$_2$O$_2$ is an important component of the stress response in plants following their interaction with pathogens, and its production stimulates the expression of a number of cell-protecting enzymes including GSTs (Lamb & Dixon 1997). Notably, H$_2$O$_2$ has been shown to accumulate in the outer mantle during the compatible interaction between Maj and poplar, but not in the incompatible association between Nau and poplar (Gafur et al. 2004). Small GTPases, including Rho, Rab and cdc42, are important regulators for cell growth, cell cycle progression and organization of the actin cytoskeleton and they may be involved in the morphological changes accompanying the formation of ectomycorrhizae (Johansson et al. 2004). A more detailed phylogenetic analysis revealed that the rabA is homologous to members of a Rab family including Ypt31 and Ypt32 from *Saccharomyces cerevisiae* (Pereira-Leal & Seabra 2001) (Fig. 8). Ypt31 and Ypt32 are probably involved in intra-Golgi transport or in the formation of transport vesicles at the most distal Golgi compartment (Jedd et al. 1997).

We also obtained data indicating that the lack of compatibility with birch and poplar in Nau has been associated with an enhanced rate of sequence evolution in two of the differentially expressed genes, cipC1 and cchA. Although the $d_s/d_\alpha$ ratio of these genes exceeded one, which would indicate positive selection (Nei & Kumar 2000), $d_\alpha$ was not significantly different from $d_s$. More probably, the sequence divergence of cipC1 and cchA reflects a decreased selection pressure associated with the lower functional constraints on these proteins in Nau.

The cipC1 gene product of *P. involutus* displayed high sequence identity to a cononamycin-induced protein (cipC) first identified in *Aspergillus nidulans*. The function of this protein is not known but its abundance is significantly increased in the presence of the antibiotic cononamycin,
which stimulates hyphal hyper-branching (Melin et al. 2002). CipC homologues have also been reported in several EST studies of ectomycorrhizal fungi (Peter et al. 2003). The second gene with an enhanced rate of evolution is \textit{ATX1}, which translates into a protein of 70 amino acids and displays significant sequence similarity to a number of fungal metallochaperones including the antioxidant 1 (atxl1p) protein of \textit{S. cerevisiae} (identity 44%). The \textit{ATX} gene was originally identified as a gene conferring protection against reactive oxygen species (Lin & Culotta 1995). Subsequently, it was shown that atxl1p is a cytosolic copper chaperone that transfers copper from the plasma membrane copper transporter \textit{Ctr1p} to the intracellular copper-transporting \textit{P}-type \textit{ATPase \textit{Cc}2\textit{p}}, located in the Golgi compartment of the secretory pathway (Pufahl et al. 1997). \textit{Cc}2\textit{p} translocates copper ions to secreted copper-containing enzymes such as the iron oxidase \textit{Fet3p}. Other enzymes that can be provided with copper by this pathway are the copper-containing laccases (Uldschmid et al. 2003), which are thought to be involved in a number of processes related to the function of the ectomycorrhizal symbiosis (Burke & Cairney 2002).

The underlying differences for the changes in the expression of the symbiosis-regulated genes are not known. However, duplications and/or deletions of genes are probably not involved, because all of the differentially expressed genes were found in the same copy number in \textit{Maj} and \textit{Nau}. More probably, the observed differences in expression are the result of changes in promoter elements and levels of transcription factors.

### Acknowledgements

This study was supported by grants from the Swedish Research Council, the Foundation for Strategic Research (Genome Research Programme), and the Swedish Council for Environment, Agricultural Sciences and Spatial Planning. Andreas Schützendübel received financial support through a Marie-Curie-Fellowship. Custom microarrays were produced at the SWEGENE DNA Microarray Resource Centre at the Biomedical Centre B10 in Lund, and DNA sequencing was performed at the SWEGENE Center of Genomic Ecology at the Ecology Building in Lund, supported by the Knut and Alice Wallenberg Foundation through the SWEGENE consortium. We thank Dr R. Langenfeld-Heyser and Dr Eva Friman for their help with the microscopic investigations and DNA sequencing, respectively.

### References


The paper is one in a series of ongoing studies of the functional and evolutionary genomics of the ectomycorrhizal fungus *Paxillus involutus*. The work described in this study was part of the PhD programs of Antoine LeQuéré, Balaji Rajashekar and Jenny Hedh. Andres Schützendübel with a PhD in Forest Botany (Georg-August-University) and Bjorn Canbäck with a PhD in Molecular Biology (Uppsala University) participated in this project as postdocs. Susanne Erland is associate professor at Lund University. Her main research interests are on the community structure and genetic diversity of ectomycorrhizal fungi. Tomas Johansson is assistant professor at Lund University. His research focused on gene expression during the development of ectomycorrhizal association. Anders Tunlid is professor of microbial ecology, Lund University.