Mutualism and asexual reproduction influence recognition genes in a fungal symbiont

M.A. van der Nest[1], E.T. Steenkamp[2], P.M. Wilken[1], J. Stenlid[3], M.J. Wingfield[1], B.D. Wingfield[1] and B. Slippers[1]*

[1]Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa; [2]Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa; [3]Department of Forest Mycology and Plant Pathology, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Corresponding author:

*E-mail: bernard.slippers@fabi.up.ac.za
Tel: +27-12-420-3906
Fax: +27-12-420-3960
Present address:
Department of Genetics
Forestry and Agricultural Biotechnology Institute (FABI)
74 Lunnon Road, Hillcrest
Pretoria
0002
South Africa

Research highlights:

- We examined the mating system of the *Amylostereum* symbionts of Siricid woodwasps.
- Generally, polymorphism in their *rab1* pheromone receptor genes was limited.
- Surprisingly, *rab1* evolution was not linked to positive or balancing selection.
- Suppressed recombination and purifying selection appear to drive evolution of *rab1*.

Key words:
Symbiosis, *Amylostereum areolatum*, *Sirex noctilio*, pheromone receptor and evolution.

ABSTRACT

Mutualistic symbiotic interactions between microbes and insects are wide-spread in nature. Alignment of the reproductive interests of the organisms involved in a symbiotic interaction, typically involves clonal reproduction of the microbial symbiont that is transmitted vertically by the insect partner. In this study, we used the *Amylostereum* fungus-*Sirex* woodwasp mutualism to

consider whether the prolonged association and predominantly asexual mode of reproduction have affected the mating system of the fungal partner. DNA sequence information for the pheromone receptor gene *rab1* encoded at the *mat-B* locus, as well as the eukaryotic translation elongation factor 1α gene and the ribosomal RNA internal transcribed spacer region were utilized. The identification of *rab1* alleles in *A. chailletii* and *A. areolatum* populations revealed that this gene is more polymorphic compared to the other two regions, although the diversity of all three regions was lower than the corresponding genes in other free-living Agaricomycetes. Our results indicated that suppressed recombination might potentially be implicated in the diversification of *rab1*, while no evidence of balancing selection was detected. In addition, positive selection was detected at only two codons suggesting that purifying selection rather than positive selection represents an important driving force in the evolution of *rab1*. Therefore, the long-term symbiotic relationship with their insect partners has not only affected the diversity at this mating-type locus, but it has also influenced the manner in which selection drives and maintains this diversity in *A. areolatum* and *A. chailletii*.

1 INTRODUCTION

Most eukaryotic organisms reproduce sexually to generate offspring even though it is more costly than asexual reproduction (e.g., Barton and Charlesworth 1998; Otto 2003). This costly reproductive strategy could be maintained because recombination acts to provide advantageous genotypes necessary for adaptation to changing environments or because recombination acts to eliminate deleterious mutations (e.g., Zeyl and Bell 1997; Taylor et al. 1999; Neiman et al. 2010). Both of these hypotheses are consistent with the fact that the absence of sexual reproduction decreases the overall fitness of an organism and could ultimately lead to extinction (Butlin 2006; Paland and Lynch 2006; Howe and Denver 2008). An observation in fungi is that most asexually reproducing populations retain some level of sexual reproduction, despite its cost, thus generally favouring a system of mixed modes of reproduction (Taylor et al. 1999; Hsueh and Heitman 2008).

Notwithstanding the advantages, mixed modes of sexual and asexual reproduction appear not to be feasible in all organisms. For example, the fungal symbionts of insects mainly reproduce asexually and are transmitted from mother to offspring in a vertical fashion (Chapela et al. 1994; Judson and Normark 1996). These modes of reproduction and transmission ensure co-dependence between the symbiotic partners, but could lead to a reduction in genetic diversity (Rispe and Moran 2000; Mira and Moran 2002). The absence of sexual recombination could also result in an accumulation of mildly deleterious mutations, increased genetic drift, more rapid sequence evolution (i.e., excess of amino acid substitutions), a shift in nucleotide base composition due to mutational bias and genome erosion (Rispe and Moran 2000; Kaltenpoth et al. 2010). Nevertheless, several ancient lineages of asexually reproducing organisms still exist in successful symbiotic relationships (Welch and Meselson 2001; Jany and Pawlowska 2010). In

these relationships, it is thought that selection by the host might limit the accumulation of deleterious mutations in the symbiont (Kaltenpoth et al. 2010).

In this study, we considered the ancient and obligate symbiotic relationship between the wood-rotting fungus *Amylostereum areolatum* and its hymenopteran Siricid insect partner, *Sirex noctilio*. In this relationship, the fungus is necessary for the development of the larvae, while the woodwasp spreads the asexual spores and/or mycelium of the fungus (Vasiliauskas et al. 1998, Thomsen and Koch 1999), thereby facilitating vertical transmission of *A. areolatum* (Madden 1981). Like other fungal symbionts of insects (Chapela et al. 1994; Judson and Normark 1996), *A. areolatum* can also reproduce sexually and has a tetrapolar mating system, *i.e.,* the genes governing sexual recognition in the fungus are present on two unlinked mating type loci (*mat-A* and *mat-B*) (Boidin and Lanquetin 1984; van der Nest et al. 2008; 2009). In fungi with tetrapolar mating systems, the *mat-A* locus harbours genes that encode homeodomain proteins (functional transcriptional factors), while the *mat-B* locus harbours genes that encode peptide pheromones and pheromone receptors (e.g., Brown and Casselton 2001; Heitman et al. 2007). The sexual sporocarps of these fungi are, however, rarely found in nature and usually only in the native range of the insect and fungus (Vasiliauskas and Stenlid 1999; Slippers et al. 2003; Nielsen et al. 2009). The population biology of these fungi also suggests that, like other insect symbionts, they rely on the woodwasp for the effective spread of asexual spores (Vasiliauskas et al. 1998; Vasiliauskas and Stenlid 1999; Thomsen and Koch 1999).

The evolutionary forces acting on the *mat* loci are thought to drive divergence between *mat* alleles, thus ensuring compatibility between individuals in a population (e.g., May et al. 1999; Devier et al. 2009). A form of balancing selection, known as negative frequency-dependent selection, probably acts to preserve the characteristically high allelic and nucleotide diversities at

these loci (May et al. 1999). This mechanism involves the selection for rare alleles, because individuals carrying such rare alleles will be sexually compatible with a larger proportion of other individuals in the population (May et al. 1999; Ruggiero et al. 2008). Diversity at the *mat* loci of eukaryotes is also promoted by accelerated evolutionary rates, which is evident in the increased frequency of non-synonymous substitutions located in these regions (Vicoso et al. 2008; Devier et al. 2009). These high rates of substitution could be ascribed to suppressed recombination and/or positive selection, where the former prevents the loss of mutations and the latter acts to maintain beneficial amino acid substitutions (Uyenoyama 2005; Menkis et al. 2008; Vicoso et al. 2008).

The overall aim of this study was to determine whether the symbiotic relationship between *A. areolatum* and *S. noctilio*, together with a predominantly asexual mode of reproduction in the fungus, has influenced the evolution of the genes determining sexual recognition in the fungal partner. Two specific questions were addressed: i) How does the pattern and extent of polymorphism at a *mat* locus compare to those in other regions of the genome? (ii) Which evolutionary forces most likely influence the patterns and rates of polymorphism at the *mat* loci? To answer these questions we utilized DNA sequence information for the pheromone receptor gene (*rab1*) encoded at the *mat-B* locus, as well as the eukaryotic translation elongation factor 1α (Tef-1α) gene and the ribosomal RNA (rRNA) internal transcribed spacer (ITS) region, which includes the spacers ITS1, ITS2 and the 5.8S rRNA gene. The *rab1* gene has been demonstrated previously to be involved in sexual recognition in *A. areolatum* (van der Nest et al. 2008). For comparison, we included data from the closely related species *A. chailletii* that more frequently reproduces sexually (Vasiliauskas et al. 1998; Vasiliauskas and Stenlid 1999; Slippers et al.

2001). Finally, these systems were also compared with those of other free-living Agaricomycetes.

2 MATERIALS AND METHODS

2.1 Fungal strains

Heterokaryotic isolates of *A. areolatum* (CMW16848) and *A. chailletii* (NAc3) were included for the identification and characterization of pheromone receptor genes. Additionally, twenty-five isolates each of *A. areolatum* and *A. chailletii* obtained from various culture collections were included to investigate the allelic variation and diversity of pheromone receptor genes in naturally occurring isolates (Supplementary Tables 1 and 2). These isolates were selected to capture the known diversity of the fungi and were collected from South Africa, Brazil, Argentina, Australia, New Zealand, United States of America, Canada, France, Sweden, United Kingdom, Switzerland, Denmark, Norway, Austria, Italy, Greece and Lithuania. Working cultures of these heterokaryons were maintained on potato dextrose agar (PDA) (24 gL^{-1} of PDA, 1 gL^{-1} glucose, and 1 gL^{-1} yeast extract) (Biolab, Johannesburg, South Africa). All of the isolates used in this study are also maintained at 4 ºC in the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa. Genomic DNA was collected from the isolates using the method described by Zhou et al. (2004).

2.2 PCR, cloning and nucleotide sequencing of the rab1 gene

A large portion of the the *rab1* gene sequence for *A. areolatum* (CMW16848) was available from a previous study (van der Nest et al. 2008), while that for *A. chailletii* was identified using degenerate PCR primers (br1-F and br1-R; Supplementary Table 3) designed by James et al.

(2004b). All PCRs were performed on an Eppendorf thermocycler (Eppendorf AG, Germany) using reaction mixtures containing 1 ng/µl DNA, 0.2 mM of each of the four dNTPs, 1.5 mM $MgCl_2$, 0.5 µM of each primer and 2.5 U FastStart *Taq* (Roche Diagnostics, Mannheim). Thermal cycling conditions consisted of an initial denaturation step at 94 °C for 2 min followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, extension at 72 °C for 30 s, and a final extension at 72 °C for 10 min. The resulting PCR products were purified using polyethylene glycol (PEG) precipitation (Steenkamp et al., 2006) and the purified PCR products were cloned using the pGEM-T Easy vector System I (Promega Corporation, Madison, USA). PCR products were cloned in order to obtain haplotype phases of all sequences derived from the heterokaryotic isolates. The cloned products were amplified from individual colonies with plasmid-specific primers (Steenkamp et al., 2006), after which the PCR products were purified using PEG precipitation. The purified products were then sequenced with the plasmid-specific primers, Big Dye Cycle Sequencing kit version 3.1 (Perkin-Elmer, Warrington, UK) and an ABI3700 DNA analyzer (Applied Biosystems, Foster City, USA).

To obtain the sequences upstream and downstream of these fragments in *A. areolatum* and *A. chailletii*, nested PCR primers designed with Primer 3 (cgi v0.2) (http://www.genome.wi.mit.edu/genome _software/other/primer3.html) and PCR-based genome walking (Siebert et al. 1995) were used. The nested primers used for genome-walking included RAB1-4 (Supplementary Table 3) for *A. areolatum* and RAB5 (Supplementary Table 3) for *A. chailletii*. The remaining portion of the pheromone receptor gene for *A. chailletii* was obtained using a primer (RAB6; Supplementary Table 3) based on the sequence of *A. areolatum*. These PCR products were amplified, cloned and sequenced as described above.

All sequence files were analyzed with Chromas Lite 2.0 (Technelysium) and BioEdit version 7.0.2.5 (Hall 1999). They were also compared to those in the protein database of the National Centre for Biotechnology Information (www.ncbi.nih.nlm.gov) using BlastX. To predict the features in the secondary structure of the Rab1 protein, we used TOPCON (http://topcons.cbr.su.se/), which calculates consensus predictions using a Hidden Markov Model and inputs five commonly used topology prediction methods (Bernsel et al., 2009).

2.3 Allelic variation and diversity of rab1, Tef-1α and ITS

To identify unique *rab1* alleles in *A. areolatum* and *A. chailletii*, we used two approaches to discover polymorphisms among a diverse set of 25 isolates of each species. The one approach entailed analysis a 186-base pair (bp) region of *rab1*, which was previously shown to be polymorphic (van der Nest et al. 2008). This fragment was amplified and sequenced using primer set RABF+RABR (Supplementary Table 3). The second approach involved PCR-RFLP (restriction fragment length polymorphism) analysis of a 682-bp fragment of the *rab1* gene. For this purpose, PCR products were generated with primers RAB1-470F and RAB1-1800R (Supplementary Table 3) for the 25 isolates of *A. areolatum*. The amplicons were then digested with the enzyme *Eco*RV (Roche Diagnostics) and visualised with agarose gel (Roche Diagnostics) electrophoresis (Sambrook et al. 1989; van der Nest et al. 2008).

Based on the polymorphisms observed, a set of 13 isolates for each of *A. areolatum* and *A. chailletii*, were selected. For these isolates, the 682-bp portion of the *rab1* gene was sequenced for *A. areolatum*, as well as for *A. chailletii* using primer set RAB7+RAB8 (Supplementary Table 3). For comparative purposes, portions of the two housekeeping loci, ITS and Tef-1α, were also amplified and sequenced for the 13 isolates of each species. Primer set ITS1+ITS4 (Supplementary Table 3; White et al. 1990) was used to amplify of the ITS region of both

species. For the Tef-1α region, the primer set TEFac1+TEFac2 (Supplementary Table 3) was used for *A. chailletii* and primer set TEFaa1+TEFaa2 (Supplementary Table 3) for *A. areolatum*. Other than those used for the PCR-RFLPs, all the Tef-1α and ITS PCR products, as well as both of the 186- and 682-bp fragments of *rab1*, were purified, cloned and at least 5 clones per individual were sequenced, as described above.

Following sequence analysis with Chromas Lite and BioEdit, sequence alignments for each locus were produced using MAFFT version 5.85 (http://mafft.cbrc.jp/alignment/server/) (Katoh et al. 2002). For each of the datasets, nucleotide diversity (π; Nei and Li 1979) was determined using the software package DnaSP version 5.10 (Librado and Rozas 2009), while allelic frequencies were calculated using GENEPOP software version 1.2. Because a limited number of clones were sequenced per individual, the possibility of underestimating allelic diversity could not be excluded.

2.4 Molecular evolution of rab1

To identify the evolutionary forces acting on the pheromone receptor genes of *A. areolatum* and *A. chailletii*, all the unique *rab1* alleles identified in this study were examined. The *A. areolatum rab1* alleles were amplified and sequenced using primer set RAB9+RAB10 (Supplementary Table 3) for isolates CMW8900 and CMW2822, while the identified *A. chailletii rab1* alleles were amplified and sequenced using primer set RAB11+RAB12 (Supplementary Table 3) for isolates LIIAc116, DAC2, US2 and It1.8.

The CODEML program in the PAML version 3.14 package (Yang and Nielson 2002) was used to determine patterns of selective pressure acting on the pheromone receptor alleles identified in the two fungi. The phylogenetic tree required by CODEML was generated by

subjecting a MAFFT-generated nucleotide alignment of *rab1* to a maximum likelihood (ML) analyses using PhyML version 3.0 software (Guindon and Gascuel 2003). This ML analysis employed gamma correction (G) to account for among site rate variation, a proportion of invariable sites (I) and the HKY (Hasegawa et al. 1987) nucleotide substitution model as indicated by jModeltest version 0.1.1 and the Akaike Information Criterion (Posada 2008).

Positive selection was evaluated by computing ω across all the sites for each of the loci (Yang et al. 2000; Devier et al. 2009), where ω reflects the non-synonymous (*dN*)/synonymous (*d*S) substitution rate ratio (Yang and Nielson, 1998). To test for variation of selective pressures across the codons, goodness of fit was calculated for the different site-specific models proposed by Yang et al. (2000). Statistical significance was calculated with likelihood ratio tests (LRT), which entailed analysis of the χ^2 distribution of $2\Delta\ln$ (*i.e.*, twice the log likelihood difference between the two models) values for the different models (Yang and Nielson 1998), where the degrees of freedom were equal to the differences in number of parameters between the two models (Yang et al. 2000).

To determine whether balancing selection acts on *rab1* to maintain rare alleles over long evolutionary times (Vieira et al. 2008), a phylogenetic tree based on the amino acid sequences of the pheromone receptors present in *A. areolatum*, *A. chailletii* and sequences from other Basidiomycetes and Ascomycetes available in GenBank was constructed (See Figure 3 for accession numbers for the pheromone receptors of other fungi). The amino acid sequences were aligned using MAFFT and an ML phylogeny inferred with PhyML, which utilized the LG (Le and Gascuel, 2008) model of amino acid substitution, I and the observed amino acid frequencies, as indicated by ProtTest 2.4 (Abascal et al. 2005). Branch support was determined using PhyML with the same best-fit model and 1000 bootstrap replicates. Both the tree and dataset have been

submitted to TreeBASE (http://www.treebase.org/treebase/index.html) and the Study Accession URL is: http://purl.org/phylo/treebase/phylows/study/TB2:S12966.

DnaSP was used to study the extent of recombination within the *rab1*, ITS and Tef-1α regions examined. The recombination parameter R was calculated (Hudson 1987), while the minimum number of recombination (R_M) events during the history of the species (Hudson and Kaplan 1985) was estimated using neutral coalescence simulations, based on the number of segregating sites, intermediate levels of recombination and 10000 replications (Librado and Rozas 2009). The extent of recombination within each species was also compared by examining single-locus phylogenies for incompatibility. The latter analyses were based on separate trees inferred from the DNA sequence information for the three loci of the 13 selected isolates of each of *A. areolatum* and *A. chailletii*. MAFFT-generated datasets were subjected to PhyML analyses using best-fit model parameters, as described before. ML analysis of the *rab1* dataset employed the HKY with I and G, the ITS dataset employed the TrNeF model (Tamura and Nei 1993; Posada 2008), and the Tef-1α dataset employed the TrN model (Tamura and Nei 1993; Posada 2008). For each dataset, branch support was determined using the respective best-fit models and 1000 bootstrap replicates. To assess congruencies between the resulting gene trees, a strict consensus tree was computed using Mega software version 4.0.2 (Kumar et al. 2008). The partition homogeneity test using PAUP version 4.0b10 (Swofford 2000) was used to examine the null hypothesis of recombination in *A. areolatum* and *A. chailletii* (Houbraken et al. 2008). Significance was assigned by comparing the summed tree length from the actual data to those from 100 artificial datasets.

3 RESULTS

3.1 PCR, cloning and nucleotide sequencing of the rab1 gene

It was possible to identify and sequence the complete pheromone receptor gene *rab1* (1539 bp) in *A. areolatum*, as well as a large portion of the gene in *A. chailletii* (1265 bp). Typical of pheromone receptors, the inferred amino acid sequences for the *Amylostereum* Rab1 (Fig. 1) harboured the seven transmembrane-spanning helices that are characteristic of the rhodopsin-like superfamily of G protein-linked receptors, as well as extracellular and cytoplasmic loop domains and a long cytoplasmic tail (e.g., James et al. 2004a; Raudaskoski and Kothe 2010). The *rab1* gene also contained five introns (Fig. 1), which are comparable with those reported for the pheromone receptors of *Coprinopsis cinerea* that has four or five, *Coprinellus disseminatus* that has five and *Schizophyllum commune* that has three introns (Vaillancourt et al. 1997; James et al. 2006). The position of introns 2-5 closely corresponded to those in previously identified receptor sequence introns, while intron 1 in the *rab1* N-terminal appears to represent a novel position (Fig. 1). The position and size of the introns present in the *rab1* genes in *A. chailletii* and *A. areolatum* were similar, except that the last intron in *A. chailletii* was larger than in *A. areolatum*.

3.2 Allelic variation and diversity of rab1, Tef-1α and ITS

A total of 8 and 6 polymorphisms were observed within the 186-bp alignment of *rab1* for the 25 representative isolates of each of *A. areolatum* and *A. chailletii*, respectively (Supplementary Tables 1 and 2). For 13 of the 25 representative isolates of *A. areolatum* and of *A. chailletii*, a larger portion of the *rab1* gene, as well as a portion of Tef-1α and ITS were also analysed. Following cloning of each PCR product, sequencing and alignment, the datasets for ITS, Tef-1α

and *rab1* in these representative isolates consisted of 606, 474 and 682 nucleotides, respectively. Based on the sequence alignments, two types of Tef-1α were identified. The GenBank accession numbers for the two types of Tef-1α identified in *A. areolatum* are HQ864714 and HQ864713, respectively. Because the one type contained a premature stop codon, it was considered a pseudogene and was not included in subsequent analyses.

Among the Tef-1α, ITS and *rab1* sequences of the 13 representatives of *A. areolatum* and of *A. chailletii*, a total of 19 (Supplementary Table 4) and 24 (Supplementary Table 5) nucleotide polymorphisms, respectively, were identified. However, in both species, about half of the polymorphisms were located in the *rab1* dataset. As a result, the nucleotide diversity or π-values (0.004 and 0.02, respectively) for the *rab1* regions in *A. areolatum* and *A. chailletii* were generally higher than the π-values for the ITS and Tef-1α regions (Table 1).

Among the 13 representatives for each species, the number of alleles identified in the three genes (*rab1*, ITS and Tef-1α) investigated ranged from 2 to 6 for *A. areolatum* and 3 to 5 for *A. chailletii* (Figure 2A). With regards to *rab1*, the same three alleles that were identified within the set of 13 *A. areolatum* isolates were also detected in the set of 25 *A. areolatum* isolates collected in different regions of the world (Supplementary Table 4). In contrast, an additional *rab1* allele was detected in the global collection of 25 *A. chailletii* isolates. The *A. chailletii* isolates included in this study thus harboured at least four pheromone receptor alleles (Supplementary Tables 2 and 5), although some apparently occur at very low frequencies (Fig. 2A). These findings were also supported by the PCR-RFLP analysis with *Eco*RV.

The GenBank accession numbers for the *A. areolatum* and *A. chailletii* ITS sequences are KC865539-KC865592, the *A. areolatum* and *A. chailletii rab1* sequences are KC865479-

KC865530, while the accession numbers for *A. areolatum* and *A. chailletii* Tef-1α are KC865425-KC865478.

Heterozygous and homozygous genotypes were detected in both *A. chailletii* and *A. areolatum* for all three gene regions investigated (Fig. 2B). Nine, 8 and 5 unique ITS, Tef-1α and *rab1* genotypes were identified for the *A. chailletii* individuals. Among the *A. areolatum* isolates respectively 3, 7 and 3 ITS, Tef-1α and *rab1* genotypes were identified. Compared to *A. areolatum*, the *A. chailletii* individuals thus harboured considerably more unique ITS (9 *vs*. 3) and *rab1* (5 *vs*. 3) genotypes. *Amylostereum chailletii* also included many more genotypes represented by a single individual (*i.e.*, 6 *vs*. 1 ITS genotypes and 2 *vs*. 0 *rab1* genotypes) in comparison with *A. areolatum*. Some of the identified genotypes were over-represented. For example, 76.9 % of the *A. areolatum* individuals shared the same ITS genotype (Fig. 2B). These over-represented genotypes were also homozygotic, with both nuclei of the heterokaryon having the same allelic state.

3.3 Molecular evolution of rab1

A large proportion of the three *A. areolatum rab1* (1579 bp) alleles, as well as the five *A. chailletii rab1* (1448 bp) alleles were sequenced (GenBank accession numbers KC865531-865538) (Fig. 1). In *A. areolatum*, alleles shared 99.0-99.9 % nucleotide sequence identity with 15 polymorphic sites (Supplementary Table 6). The alleles in *A. chailletii* shared more than 98.0 % nucleotide sequence identity with each other (Supplementary Table 6), while they shared only 83 % nucleotide sequence identity with the *A. areolatum rab1* alleles. This suggests that all the alleles belong to the same sub-locus. This follows from the fact tthat the pheromone receptors in *C. cinerea* belonging to the same sub-locus are > 60 % similar, while those in different sub-loci are < 32-35 % similar (Riquelme et al. 2005).

Comparison of the *rab1.1* allele with the *rab1.2* and the *rab1.3* alleles in *A. areolatum* revealed two polymorphic sites that result in non-synonymous amino acid substitutions (present in the long cytoplasm carboxy-terminal intracellular tail) (Fig. 1). The single polymorphic site identified in the *rab1.2* and *rab1.3* alleles represented a silent substitution, which suggests that the *rab1.2* and *rab1.3* alleles are functionally equivalent in sharing the same mating type specificity. Several non-synonymous substitutions were identified among the *rab1* alleles present in *A. chailletii* (Fig. 1), including one in transmembrane region 1, one in the transmembrane region 5, two in the intracellular loop 3 and two in the long cytoplasm carboxy-terminal intracellular tail.

Calculations using model M3 with variable selective pressures acting on the codons and models that assume no selection (M0) provided a better fit for *rab1* (Table 2). The log likelihood difference between M3 and M0 was 4.57 with a significant χ^2 distribution test result ($P < 0.05$), while the log likelihood differences were not significant for M2 and M1, M5 and M1, M6 and M1 or for M8 and M7. Therefore, the positive-selection models (M2, M5, M6 and M8) did not provide a better fit in comparison to models that assume no positive selection (M1 and M7). Even though the models that assume positive-selection did not provided a better fit, all of these models identified two positively selected codons (codons 256 and 373) respectively located in the intracellular loop EL3 region and in the long cytoplasm carboxy-terminal intracellular tail. However, at the 95 % level, only models M5 and M6 identified these codons as being subject to positive selection (Table 2).

The *rab1* alignment (consisting of 337 amino acids) and ML phylogeny included sequences of the *A. areolatum* and *A. chailletii rab1* alleles identified in this study, as well as sequences from the pheromone receptors present in other Basidiomycetes and Ascomycetes (Fig. 3). All of

the *Amylostereum* pheromone receptors grouped in the same major cluster. The putative pheromone receptors of *A. areolatum* and *A. chailletii* appeared to be each other's closest neighbours and they grouped together with proteins from *C. cinerea* (including CcSTE3.3, Rcb3B43, AAQ96346) and *L. bicolor* (including LbSTE3.3, 192000).

No evidence of recombination within *rab1*, ITS or Tef-1α for *A. areolatum* and *A. chailletii* was detected. This is because of the low values for the recombination parameters (R and R_M) (Table 3) that were observed for all three loci in both species, suggesting that meiotic recombination between segregating sites is limited (Garg et al. 2007). However, evidence of recombination among individuals was detected in this study. Two distinct clades were identified for all three the genes, with one clade including only *A. areolatum* isolates and the other only *A. chailletii* isolates (Supplementary Figures 1A, 1B and 1C). Clonal lineages were observed between the three trees for both the species. For example, the alleles of the *A. areolatum* isolates from the Southern Hemisphere grouped together for all three the genes (Supplementary Figures 1A, 1B and 1C). However, consistent with the results of the partition homogeneity test ($P < 0.001$), incongruences were also observed among the *rab1*, Tef-1α and ITS phylogenies (Supplementary Figures 1A, 1B and 1C). For example, the alleles of the *A. areolatum* isolates obtained from Southern Hemisphere countries grouped together with different isolates for each of the gene trees. No structure was evident in the strict consensus tree for either *A. areolatum* and *A. chailletii* (Supplementary Figure 1D), suggesting that recombination is restricted to within-species interactions.

4 DISCUSSION

This study represents the first attempt to investigate the evolution of a mating type locus of a fungus involved in an obligate mutualism with an insect. The *Amylostereum-Sirex* association has impacted significantly on the biology and evolution of the fungal partner, similar to that observed in other symbionts. This is reflected in the predominantly asexual mode of reproduction of *A. areolatum* and *A. chailletii* and their overall low genetic diversity. The data presented here clearly demonstrate higher levels of diversity in the *rab1* pheromone receptor in *A. areolatum* and *A. chailletii* compared to ITS and Tef-1α. However, the diversity of all three genes was lower than the corresponding genes in other free-living Agaricomycetes. Rather than positive or balancing selection, purifying selection represents an important driving force in the evolution of *rab1* in *A. chailletii* and *A. areolatum*. Our study thus suggests that the long-term symbiotic relationship with their insect partners has not only affected the diversity at this locus, but it has also impacted on the manner in which selection drives and maintains this diversity in *A. areolatum* and *A. chailletii*.

Comparison of the inferred amino acid sequences of the *Amylostereum rab1* with those in other Agaricomycetes revealed that they share the same structure and conserved domains (e.g., James et al. 2004a; Raudaskoski and Kothe 2010). The Rab1 protein is thus likely to function in the same way as pheromone receptors of other fungi. In yeast, a conformational change in the receptor causes the release of the G protein bound to the pheromone receptor protein when the pheromone binds to the pheromone receptor protein. This results in the activation of the MAP kinase signalling that activates specific transcriptional factors involved in mating and heterokaryon formation (Marsh et al. 1991). To confirm that this occurs in *Amylostereum*, it will be necessary to do transcript profiling and functional analyses.

Identification of naturally occurring pheromone receptor *rab1* alleles in *A. chailletii* and *A. areolatum* revealed that these genes are multi-allelic and polymorphic. This is consistent with what is known for other fungi with tetrapolar mating systems (e.g., May et al. 1999; Kothe et al. 2003; Riquelme et al. 2005, Devier et al. 2009). Both *A. areolatum* and *A. chailletii* displayed nucleotide diversities for *rab1* that are higher than the values observed for their corresponding ITS and Tef-1α sequences. This is similar to what has previously been found for other eukaryotes (e.g., May et al. 1999; James et al. 2001; Devier et al. 2009) where the genes controlling sexual recognition are more polymorphic and diverse than the rest of the genome. The pheromone receptor encoded by *rab1* in the fungi examined is thus subject to selection for diversification.

Previous work on sex-related genes has shown that their diversification may be driven by balancing selection and suppressed recombination (Meyer and Thomsen 2001; Uyenoyama 2005). Balancing selection is known to maintain fungal mating type alleles distributed throughout populations at roughly equal frequencies (James et al. 2001). However, we found no evidence for balancing selection operating on the *rab1* alleles of *Amylostereum*. All of the *rab1* alleles in *A. areolatum* and *A. chailletii* clustered in a species-specific manner, closely with the pheromone receptor genes CcSTE3.3 in *C. cinerea* and LbSte3.3 in *L. bicolor* (Niculita-Hirzel et al. 2008; Martinez et al. 2009). Neither species thus harboured traces of trans-species polymorphism, which is a hallmark of balancing selection (Vieira et al. 2008; Devier et al. 2009). Nevertheless, no evidence of recombination within the region of the *rab1* analysed was detected, which is in agreement with the notion that allelic diversity of eukaryotic recognition loci is influenced by suppressed recombination (e.g., Uyenoyama 2005; Menkis et al. 2008). Our

results thus suggest that suppressed recombination and not balancing selection could explain the higher levels of diversity in the *rab1* gene of *A. chailletii* and *A. areolatum*.

Positive selection is also known to maintain high levels of diversity of sex-related genes (e.g., Civetta and Singh 1998, Karlsson et al. 2008). As expected, we detected the effects of positive selection on the *Amylostereum rab1*, albeit only at two codons of the gene. One of the codons is situated in the third extracellular (loop EL3) of the inferred Rab1 receptor protein. In fact, a large proportion (27 %) of the non-synonymous substitutions occurring between the *A. areolatum* and *A. chailletii rab1* alleles were also located in this region. These results are consistent with previous reports that the first and third extracellular loops of the pheromone receptor are usually more variable, because these regions interact with the mating pheromones (Reneke et al. 1988, Niculita-Hirzel et al. 2008). The second codon potentially under positive selection is located in the long C-terminal cytoplasmic tail following the last transmembrane domain, a region that also harboured a large proportion of the non-synonymous substitutions between *A. areolatum* and *A. chailletii*.

It is likely that one or more of the non-synonymous substitutions identified in *rab1* may determine mating type specificity. This could be true for either of the codons under positive selection, or any other of the non-synonymous substitutions between *A. areolatum* and *A. chailletii* detected in this study. For example, the non-synonymous substitutions located in the third extracellular loop region EL3 of the pheromone receptor that interacts with the mating pheromones (Reneke et al. 1988, Niculita-Hirzel et al. 2008) could alter the mating type specificity. Non-synonymous substitutions located in the third intracellular region (loop IL3) and the C-termini cytoplasmic tail have also been implicated in mating type specificity as both these regions of the receptor interact with G proteins and affect G protein signalling (Hegner et al.

1999; Gola et al. 2000; Marsh et al. 1991; Karlsson et al. 2008). In *A. areolatum* and *A. chailletii,* the possibility cannot be excluded that another pheromone receptor closely linked to *rab1* may determine mating type specificity, although *rab1.1* and *rab1.2* have previously been shown to co-segregate with mating type specificity in *A. areolatum* (van der Nest et al. 2009). Future genome sequence-based studies will determine whether these substitutions or those located at other receptors confer mating type specificity in *Amylostereum*.

The results presented here suggest that the entire *rab1* gene is under purifying selection and not positive selection (Table 2), which is different to what might have been expected (e.g., Civetta and Singh 1998, Karlsson et al. 2008). The ω-values for *rab1* in these fungi were significantly below 1, indicating that purifying selection was responsible for eliminating most amino acid substitutions that might have arisen. Similar results have been obtained for *Microbotryum* spp. (Devier et al. 2009). These fungi also have a tetrapolar mating type system, but they predominantly represent self-fertilizers that require only two pheromone receptor alleles and apparently do not need positive selection to generate and maintain additional alleles. A similar situation may exist in *A. areolatum*, which also appears to have only two functional *rab1* alleles, based on differences at the amino acid level, that could alter mating type specificity. The possibility that there are additional alleles at this locus that were missed in this study cannot be excluded, even though we sampled from various regions of the world. Nevertheless, our results suggest that purifying selection represents an important driving force in the evolution of *rab1* in *A. chailletii* and *A. areolatum*.

The nucleotide diversities for *rab1*, ITS and Tef-1α in *A. areolatum* and *A. chailletii* were all much lower than those observed for the regions examined in other fungi (e.g., Ciampi et al. 2009, Engh et al. 2010). For example, the *rab1* alleles of *A. areolatum* and *A. chailletii* shared

high levels of sequence identity (between 97% and 99 % nucleotide identity), while pheromone receptors from the same sub-group in *S. commune* share only 90 % sequence identity (Wendland and Kothe 1996; Gola and Kothe 2003). Also, the *A. areolatum* and *A. chailletii rab1* nucleotide diversity values were considerably lower than those reported for *Coprinellus disseminatus* (James et al. 2006) and *S. lacrymans* (Engh et al. 2010). Similarly, the observed π-values for the ITS and Tef-1α regions in *A. areolatum* and *A. chailletii* were significantly lower than the values observed for these genes in *S. lacrymans* (Engh et al. 2010). The low nucleotide diversities observed for the *Amylostereum rab1* gene thus appear to be associated with limited genetic diversity at other unlinked loci. This is consistent with previous observations that the overall genetic diversity of *A. areolatum* and *A. chailletii* is generally very low (e.g., Vasiliauskas et al. 1998; Vasiliauskas and Stenlid 1999). Compared to free-living Agaricomycetes, *A. areolatum* and *A. chailletii* have a low overall genetic diversity as well as a low genetic variation at their *mat-B* locus. This low diversity could be due to the close association between these fungi and their woodwasp vectors.

Collectively, the results of this study suggest that the patterns of polymorphism observed at *rab1* in *A. areolatum* and *A. chailletii* are determined by the combined effects of the selection and demographic processes brought about by the unique lifestyle of these fungi. For example, inbreeding and/or asexuality in *A. areolatum* and *A. chailletii* could explain the low nucleotide diversity of the genes revealed in this study (Glémin et al. 2006; Haag and Roze 2007; Haudry et al. 2008). Similar observations have also been made for other microbial symbionts with reduced levels of recombination and effective population sizes (Rispe and Moran 2000; Mira and Moran 2002). However, recombination does not appear to be completely absent in *A. areolatum* and *A. chailletii* because the results showed that both species are characterized by unique multi-locus

genotypes among individuals (Supplementary Tables 4 and 5) and by incongruencies among the phylogenies inferred from the three loci studied (Supplementary Figure 1). Our results thus highlight the fact that, despite a predominant asexual mode of reproduction, *A. areolatum* and *A. chailletii* also reproduce sexually (Milgroom 1996; Otto 2003; Haag and Roze 2007; Houbraken et al. 2008). This is consistent with the notion that asexual reproduction allows for the maintenance of the symbiotic relationships of these fungi, while occasional sexual reproduction ensures their fitness and abiltiy to adapt to change. During such sexual interactions, polymorphism at the *mat* loci ensure compatibility among partners, which is the consequence of suppressed recombination that drives diversification of these loci. However, the predominantly asexual lifestyle of these fungi has apparently relaxed the need for positive selection to maintain additional *mat* alleles.

5 ACKNOWLEDGEMENTS

We thank the National Research Foundation (NRF), members of the Tree Pathology Co-operative Programme (TPCP) and the THRIP initiative of the Department of Trade and Industry (DTI), South Africa for financial support.

6 REFERENCES

Abascal F, Zardoya R, Posada D, 2005. ProtTest, selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104-2105.

Barton NH, Charlesworth B, 1998. Why sex and recombination. *Science* **281**: 1986-1990.

Bernsel A, Viklund H, Hennerdal A, Elofsson A, 2009. TOPCONS, consensus prediction of membrane protein topology. *Nucleic Acids Research (Webserver issue)* **37**: W465-8.

Boidin J, Lanquentin P, 1984. Le genre *Amylostereum* (Basidiomycetes) intercompatibilités partielles entre espèces allopartriques. *Bulletin de la Société mycologique de France* **100**: 211-236.

Brown AJ, Casselton LA, 2001. Mating in mushrooms: Increasing the chances but prolonging the affair. *Trends in Genetics* **7**: 393-400.

Butlin R, 2006. Comment on "transitions to asexuality result in excess amino acid substitutions". *Science* **313**: 1389b.

Chapela IH, Rehner SA, Schultz TR, Mueller UG, 1994. Evolutionary history of the symbiosis between fungus-growing ants and their fungi. *Science* **266**: 1691-1694.

Ciampi MB, Gale LR, de Macedo Lemos EG, Ceresini PC, 2009. Distinctively variable sequence-based nuclear DNA markers for multilocus phylogeography of the soybean- and rice-infecting fungal pathogen *Rhizoctonia solani* AG-1 IA. *Genetics and Molecular Biology* **32**: 840-846.

Civetta AR, Singh S, 1998. Sex-related genes, directional sexual selection, and speciation. *Molecular Biology and Evolution* **15**: 901-909.

Devier B, Aguileta G, Hood ME, Giraud T, 2009. Ancient trans-specific polymorphism at pheromone receptor genes in Basidiomycetes. *Genetics* **181**: 209-223.

Engh IB, Skrede I, Sætre G-P, Kauserud H, 2010. High variability in a mating type linked region in the dry rot fungus *Serpula lacrymans* caused by frequency-dependent selection? *BMC Genetics* **11**: 64.

Garg S, Alam MT, Das MK, Dev V, Kumar A, Dash AP, Sharma YD, 2007. Sequence diversity and natural selection at domain I of the apical membrane antigen 1 among Indian *Plasmodium falciparum* populations. *Malaria Journal* **6**: 154-163.

Glémin S, Bazin E, Charlesworth D, 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proceedings of the Royal Society of London Series B* **273**: 3011-3019.

Gola S, Hegner J, Kothe E, 2000. Chimeric pheromone receptors in the basidiomycetes *Schizophyllum commune*. *Fungal Genetics Biology* **30**: 191-196.

Gola S, Kothe E, 2003. The little difference, in vivo analysis of pheromone discrimination in *Schizophyllum commune*. *Current Genetics* **42**: 276-283.

Guindon S, Gascuel O, 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**: 696-704.

Haag CR, Roze D, 2007. Genetic load in sexual and asexual diploids, segregation, dominance and genetic drift. *Genetics* **176:** 1663-1678.

Hall T, 1999. BioEdit, A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95-98.

Hasegawa M, Kishino H, Yano T, 1985. Dating of human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**: 160-174.

Haudry A, Cenci A, Guilhaumon C, Paux E, Santoni S, David J, Glémin S, 2008. Mating system and recombination affect molecular evolution in four *Triticeae* species. *Genetic Research CAMB* **90**: 97-109.

Heitman J, Kronstad JW, Taylor JW, Casselton LA, 2007. Sex in Fungi: Molecular determination and evolutionary implications, ASM Press, Washington, DC.

Houbraken J, Varga J, Rico-Munoz E., Johnson S, Samson RA, 2008. Sexual Reproduction as the Cause of Heat Resistance in the Food Spoilage Fungus *Byssochlamys spectabilis* (Anamorph *Paecilomyces variotii*). *Applied and Environmental Microbiology* **74**: 1613-1619.

Howe DK, Denver DR, 2008. Muller's ratchet and contemporary mutation in *Caenorhabditis briggsae* mitochondrial genome evolution. *BMC Evolutionary Biology* **8**: 62.

Hsueh Y-P, Heitman J, 2008. Orchestration of sexual reproduction and virulence by the fungal mating-type locus. *Current Opinion in Microbiology* **11**: 517–52.

Hudson RR, Kaplan NL, 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**: 147-164.

Hudson RR, Kreitman M, Aguade M, 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153-159.

James TY, Kües U, Rehner SA, Vilgalys R, 2004a. Evolution of the gene encoding mitochondrial intermediate peptidase and its cosegregation with the A mating-type locus of mushroom fungi. *Fungal Genetics Biology* **41**: 381-390.

James TY, Liou S-R, Vilgalys R, 2004b. The genetic structure and diversity of the A and B mating type genes from the tropical oyster mushroom, *Pleurotus djamor*. *Fungal Genetics and Biology* **41**: 813-825.

James TY, Moncalvo J-M, Li S, Vilgalys R, 2001. Polymorphism at the ribosomal DNA spacer and its relation to breeding structure of the widespread mushroom *Schizophyllum commune*. *Genetics* **157**: 149-161.

James TY, Srivilai P, Kües U, Vilgalys R, 2006. Evolution of the bipolar mating system of the mushroom *Coprinellus disseminatus* from its tetrapolar ancestors involves loss of mating-type-specific pheromone receptor function. *Genetics* **172**: 1877-1891.

Jany J-L, Pawlowska TE, 2010. Multinucleate spores contribute to evolutionary longevity of asexual Glomeromycota. *The American Naturalist* **175**: 424-435.

Judson OP, Normark BB, 1996. Ancient asexual scandals. *Tree* **11**: 41-46.

Kaltenpoth M, Goettler W, Koehler S, Strohm E, 2010. Life cycle and population dynamics of a protective insect symbiont reveal severe bottlenecks during vertical transmission. *Evolutionary Ecology* **24**: 463-477.

Karlsson M, Nygren K, Johannesson H, 2008. The evolution of the pheromonal signal system and its potential role for reproductive isolation in heterothallic *Neurospora*. *Molecular Biology and Evolution* **25**: 168-178.

Katoh K, Misawa K, Kuma K, Miyata T, 2002. MAFFT, a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**: 3059-3066.

Kothe E, Gola S, Wendland J, 2003. Evolution of multispecific mating-type alleles for pheromone perception in the agaricomycetes fungi. *Current Genetics* **42**: 268-275.

Kumar S, Dudley J, Nei M, Tamura K, 2008. MEGA, A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* **9**: 299-306.

Le SQ, Gascuel O, 2008. An improved general amino acid replacement matrix. *Molecular Biology and Evolution* **25**: 1307-1320.

Librado P, Rozas J, 2009. DnaSP v5, A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451-1452.

Lukens L, Yicunt H, May G, 1996. Correlation of genetic and physical maps at the a mating-type locus of *Coprinus cinereus*. *Genetics* **144**: 1471-1477.

Madden JL, 1981. Egg and larval development in the woodwasp, *Sirex noctilio*. *Forests Australian Journal of Botany* **14**: 25-30.

Marsh L, Neiman AM, Herskowitz I, 1991. Signal transduction during pheromone response in yeast. *Annual Review of Cell and Developmental Biology* **7**: 699-728.

Martinez D, Challacombe J, Morgenstern I, David H, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P, Hammel KE, Vanden Wymelenberg A, Gaskell J, Lindquist E, Sabat G, Sandra Splinter B, Larrondo LF, Canessa P, Vicuna R, Yadav J, Doddapaneni H, Subramanian V, Pisabarro AG, Lavín JL, Oguiza JA, Master E, Henrissat B, Coutinho PM, Harris P, Magnuson JK, Baker SE, Bruno K, Kenealy W, Hoegger PJ, Kües U, Ramaiya P, Lucas S, Salamov A, Shapiro H, Tu H, Chee CL, Misra M, Xie G, Teter S, Yaver D, James T, Mokrejs M, Pospisek M, Grigoriev IV, Brettin T, Rokhsar D, Berka R, Cullen D, 2009. Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 1954–1959.

May GS, Badrane H, Vekemans X, 1999. The signature of balancing selection, Fungal mating compatibility gene evolution. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 9172-9177.

Menkis A, Jacobson DJ, Gustafsson T, Johannesson H, 2008. The mating-type chromosome in the filamentous ascomycete *Neurospora tetrasperma* represents a model for early evolution of sex chromosome. *Plos Genetics* **4**: e1000030.

Metzger KJ, Thomas MA, 2010. Evidence of positive selection at codon sites localized in extracellular domains of mammalian CC motif chemokine receptor proteins. *BMC Evolutionary Biology* **10**: 139.

Meyer D, Thomsen G, 2001. How selection shapes variation of the human major histocompatibility complex, a review. *Annals of Human Genetics* **65**: 1–26.

Milgroom MG, 1996. Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology* **34**: 457-477.

Mira A, Moran NA, 2002. Estimating population size and transmission bottlenecks in maternally transmitted endosymbiotic bacteria. *Microbial Ecology* **44**: 137-143.

Nei M, Li W-H, 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* **76**: 5269-5273.

Neiman M, Hehman G, Miller JT, Logsdon JM, Taylor DR, 2010. Accelerated mutation accumulation in asexual lineages of a freshwater snail. *Molecular Biology and Evolution* **27**: 954-963.

Niculita-Hirzel H, Labbé J, Kohler A, le Tacon F, Martin F, Sanders IR, Kües U, 2008. Gene organization of the mating type regions in the ectomycorrhizal fungus *Laccaria bicolor* reveals distinct evolution between the two mating type loci. *New Phytologist* **180**: 329-342.

Nielsen C, Williams DW, Hajek AE, 2009. Putative source of the invasive *Sirex noctilio* fungal symbiont, *Amylostereum areolatum*, in the eastern United States and its association with native siricid woodwasps. *Mycological Research* **113**: 1242-1253.

Otto SP, 2003. The Advantages of Segregation and the Evolution of Sex. *Genetics* **164**: 1099-1118.

Paland S, Lynch M, 2006. Transitions to asexuality results in excess amino acid substitutions. *Science* **311**: 990-992.

Posada D, 2008. jModelTest, Phylogenetic Model Averaging. *Molecular Biology and Evolution* **25**: 1253-1256.

Raudaskoski M, Kothe E, 2010. Basidiomycete mating type genes and pheromone signalling. *Eukaryotic Cell* **9**: 847-859.

Reneke JE, Blumer KJ, Courchesne WE, Thorner J, 1988. The carboxy-terminal segment of the yeast α-factor receptor is a regulatory domain. *Cell* **55**: 221-234.

Riquelme M, Challen MP, Casselton LA, Brown AJ, 2005. The origin of multiple *b* mating specificities in *Coprinus cinereus*. *Genetics* **170**: 1105-1119.

Rispe C, Moran NA, 2000. Accumulation of deleterious mutations in endosymbionts, Muller's ratchet with two levels of selection. *American Naturalist* **156**: 425-441.

Ruggiero MV, Jacquemin B, Castric V, Vekemans X, 2008. Hitch-hiking to a locus under balancing selection, high sequence diversity and low pollution subdivision at the S-locus genomic region in *Arabidopsis halleri*. *Genetic Research CAMB* **90**: 37-46.

Sambrook J, Fritsch EF, Maniatis T, 1989. Molecular cloning, a laboratory manual. Cold Spring Harbour Laboratory Press, New York.

Siebert PD, Chenchik A, Kellog DE, Lukyanov KA, Lukyanov SA, 1995. An improved PCR method for walking in uncloned genomic. DNA. *Nucleic Acids Research* **23**: 1087–1088.

Slippers B, Coutinho TA, Wingfield BD, Wingfield MJ, 2003. A review of the genus *Amylostereum* and its association with woodwasps. *South African Journal of Science* **99**: 70-74.

Slippers B, Wingfield BD, Coutinho TA, Wingfield MJ, 2002. DNA sequence and RFLP data reflect geographical spread and relationships of *Amylostereum areolatum* and its insect vectors. *Molecular Ecology* **11**: 1845 –1854.

Slippers B, Wingfield MJ, Wingfield BD, Coutinho TA, 2000. Relationships among *Amylostereum* species associated with siricid woodwasps inferred from mitochondrial ribosomal DNA sequences. *Mycologia* **92**: 955-963.

Slippers B, Wingfield MJ, Wingfield BD, Coutinho TA, 2001. Population structure and possible origin of *Amylostereum areolatum* in South Africa. *Plant Pathology* **50**: 206-210.

Steenkamp ET, Wright J, Baldauf SL, 2006. The protistan origins of animals and fungi. *Molecular Biology and Evolution* **23**: 93-106.

Swofford DL, 2000. PAUP*, phylogenetic analysis using parsimony (*and other methods), version 4.0b4a. Sinauer Associates, Sunderland, MA.

Tabata M, Harrington TC, Chen W, Abe Y, 2000. Molecular phylogeny of species in the genera *Amylostereum* and *Echinodontium*. *Mycoscience* **41**: 585-593.

Tamura K, Nei N, 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10**: 512-526.

Taylor JW, Jacobson DJ, Fisher MC, 1999. The evolution of asexual fungi, Reproduction, speciation and classification. *Annual Review of Phytopathology* **37**: 197-246.

Thomsen IM, Koch J, 1999. Somatic compatibility in *Amylostereum areolatum* and *A. chailletii* as a consequence of symbiosis with Siricid. *Mycological Research* **103**: 817-823.

Uyenoyama MK, 2005. Evolution under tight linkage to mating type. *New Phytologist* **165**: 63-70.

Vaillancourt LJ, Raudaskoski M, Specht CA, Raper CA, 1997. Multiple genes encoding pheromones and pheromone receptor define the Bβ1 mating-type specificity in *Schizophyllum commune*. *Genetics* **146**: 541-551.

Van der Nest MA, Slippers B, Steenkamp ET, de Vos L, van Zyl K, Stenlid J, Wingfield MJ, Wingfield BD, 2009. Genetic linkage map for *Amylostereum areolatum* reveals an association between vegetative growth and sexual and self recognition. *Fungal Genetics and Biology* **46**: 632-641.

Van der Nest MA, Slippers B, Stenlid J, Wilken PM, Vasaitis R, Wingfield MJ, Wingfield BD, 2008. Characterization of the systems governing sexual and self-recognition in the white rot agaricomycetes *Amylostereum areolatum*. *Current Genetics* **53**: 323-336.

Vasiliauskas R, Stenlid J, 1999. Vegetative compatibility groups of *Amylostereum areolatum* and *A. chailletii* from Sweden and Lithuania. *Mycological Research* **103**: 824-829.

Vasiliauskas R, Stenlid J, Thomsen IM, Vasiliauskas R, Stenlid J, Thomsen IM, 1998. Clonality and genetic variation in *Amylostereum areolatum* and *A. chailletii* from northern Europe. *New Phytologist* **139**: 751-758.

Vicoso B, Haddrill PR, Charlesworth, B, 2008. A multispecies approach for comparing sequence evolution of X-linked and autosomal sites in *Drosophila*. *Genetic Research CAMB* **90**: 421-431.

Vieira J, Fonseca NA, Santos RAM, Habu T, Tao R, Vieira CP, 2008. The number, age, sharing and relatedness of S-locus. *Genetic Research CAMB* **90**: 17-26.

Welch DBM, Meselson MS, 2001. Rates of nucleotide substitution in sexual and anciently asexual rotifers. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 6720-6724.

Wendland J, Kothe E, 1996. Allelic divergence of two mating type genes encoding the pheromone receptor Bar1 of Bα specificity in the basidiomycete *Schizophyllum commune*. *FEMS Microbiology Letters* **145**: 451-455.

White TJ, Bruns T, Lee S, Taylor J, 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In, Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds), *PCR Protocols. A Guide to Methods and Applications*. Academic Press, San Diego, pp. 315-322.

Yang Z, 2007. PAML 4, phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**: 1586-1591.

Yang Z, Nielson R, 1998. Synonymous and non-synonymous rate variation in nuclear genes of mammals. *Journal of Molecular Evolution* **46**: 409-418.

Yang Z, Nielson R, 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* **19**: 908-917.

Yang Z, Nielson R, Goldman N, Pedersen AM, 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**: 431-449.

Yang Z, Wong WSW, Nielsen R, 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution* **22**: 1107-1118.

Zeyl C, Bell G, 1997. The advantage of sex in evolving yeast populations. *Nature* **388**: 465-468.

Zhou XD, De Beer ZW, Ahumada R, Wingfield BD, Wingfield MJ, 2004. Ophiostomatoid fungi associated with two pine-infesting bark beetles from Chile. *Fungal Diversity* **15**: 253-266.

Figures

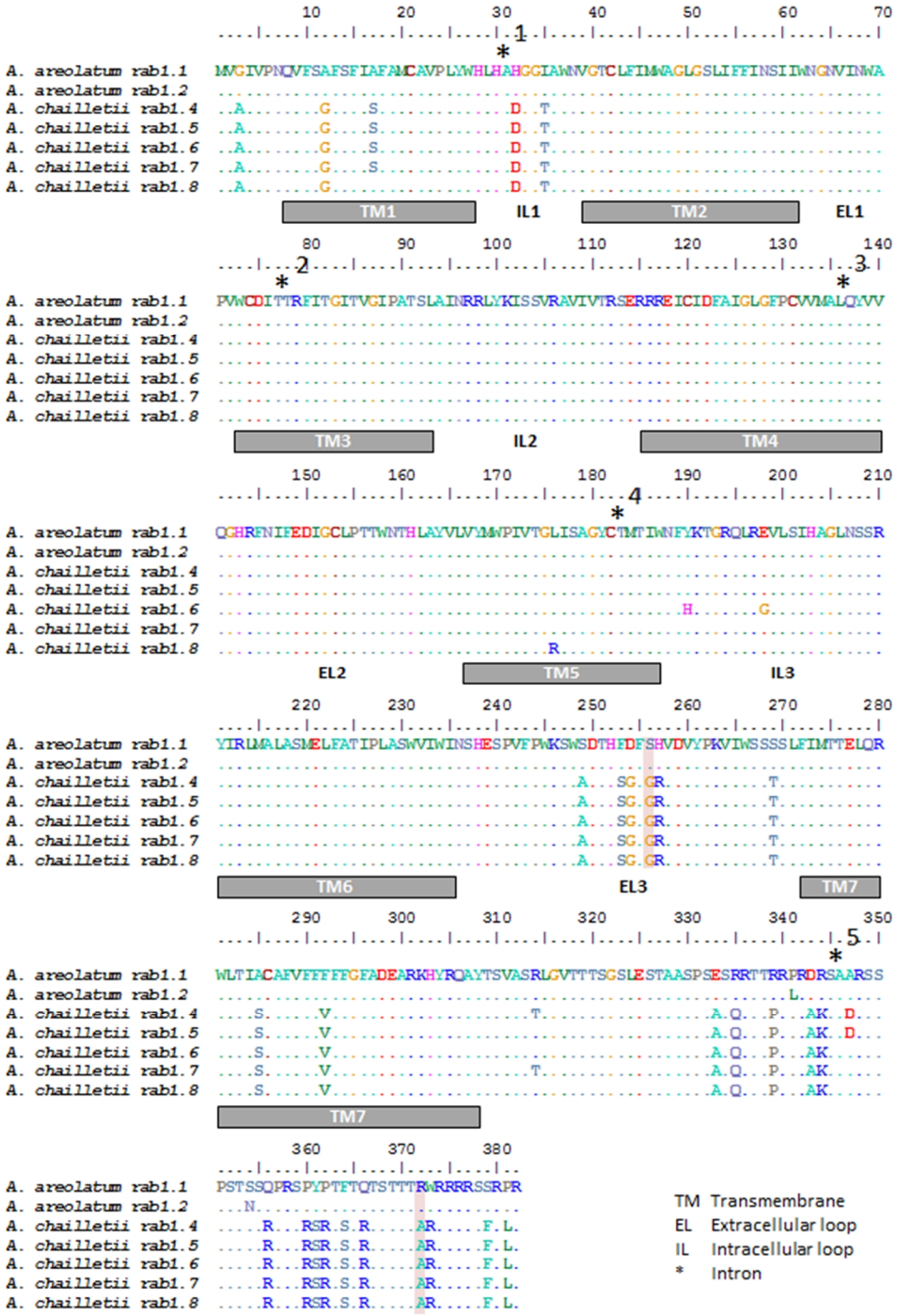


Figure 1. The predicted amino acid sequences for the various alleles of the putative pheromone receptor *rab1* in *A. areolatum* and *A. chailletii*. Protein domains were predicted with TOPCON (http://topcons.cbr.su.se/) and codons under positive selections are indicated in the shaded boxes.

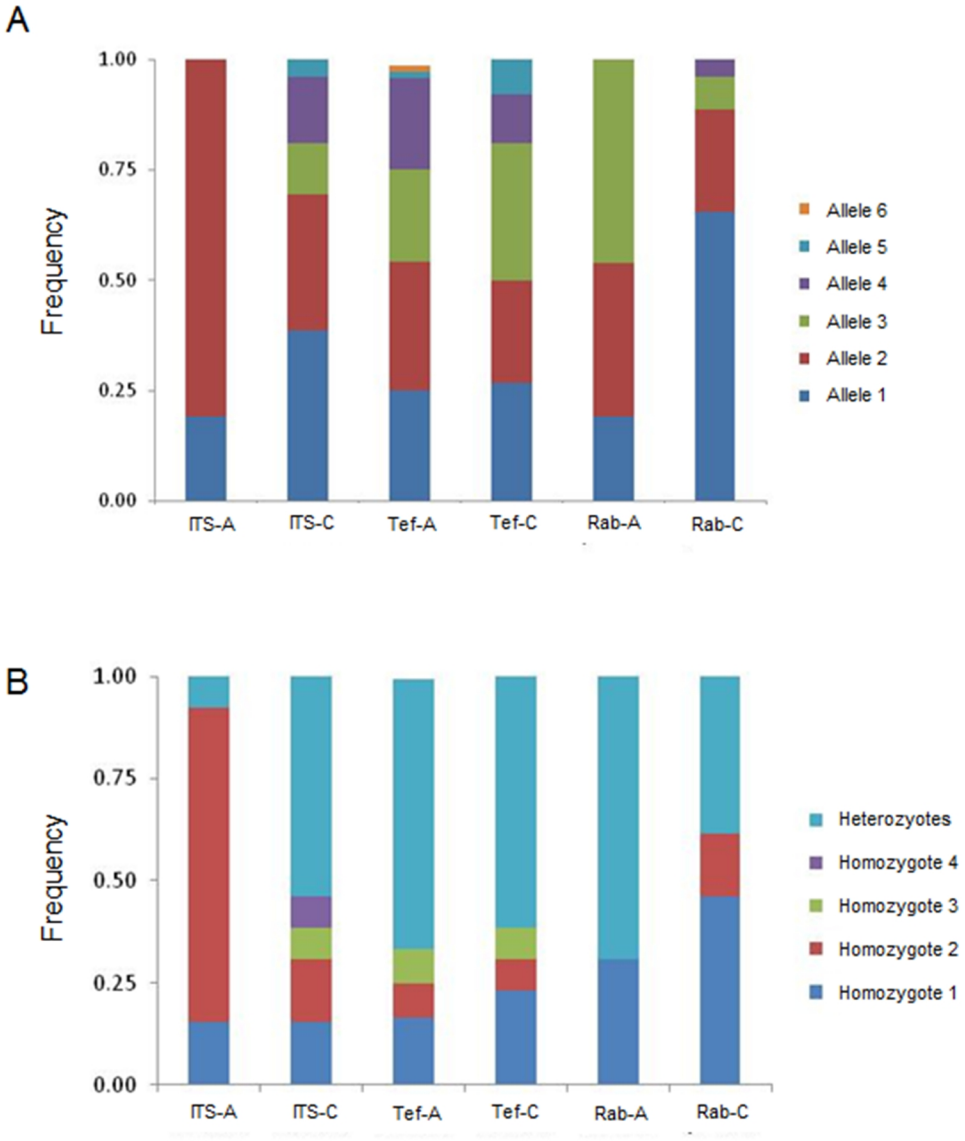


Figure 2. Stack histograms of allele frequencies (A) and genotype frequencies (B) at individual loci. ITS-A, ITS of *A. areolatum*; ITS-C, ITS of *A. chailletii*; Tef-A, Tef-1α of *A. areolatum*; Tef-C, Tef-1α of *A. chailletii*; Rab-A, *rab1* of *A. areolatum*; Rab-C, *rab1* of *A. chailletii*.

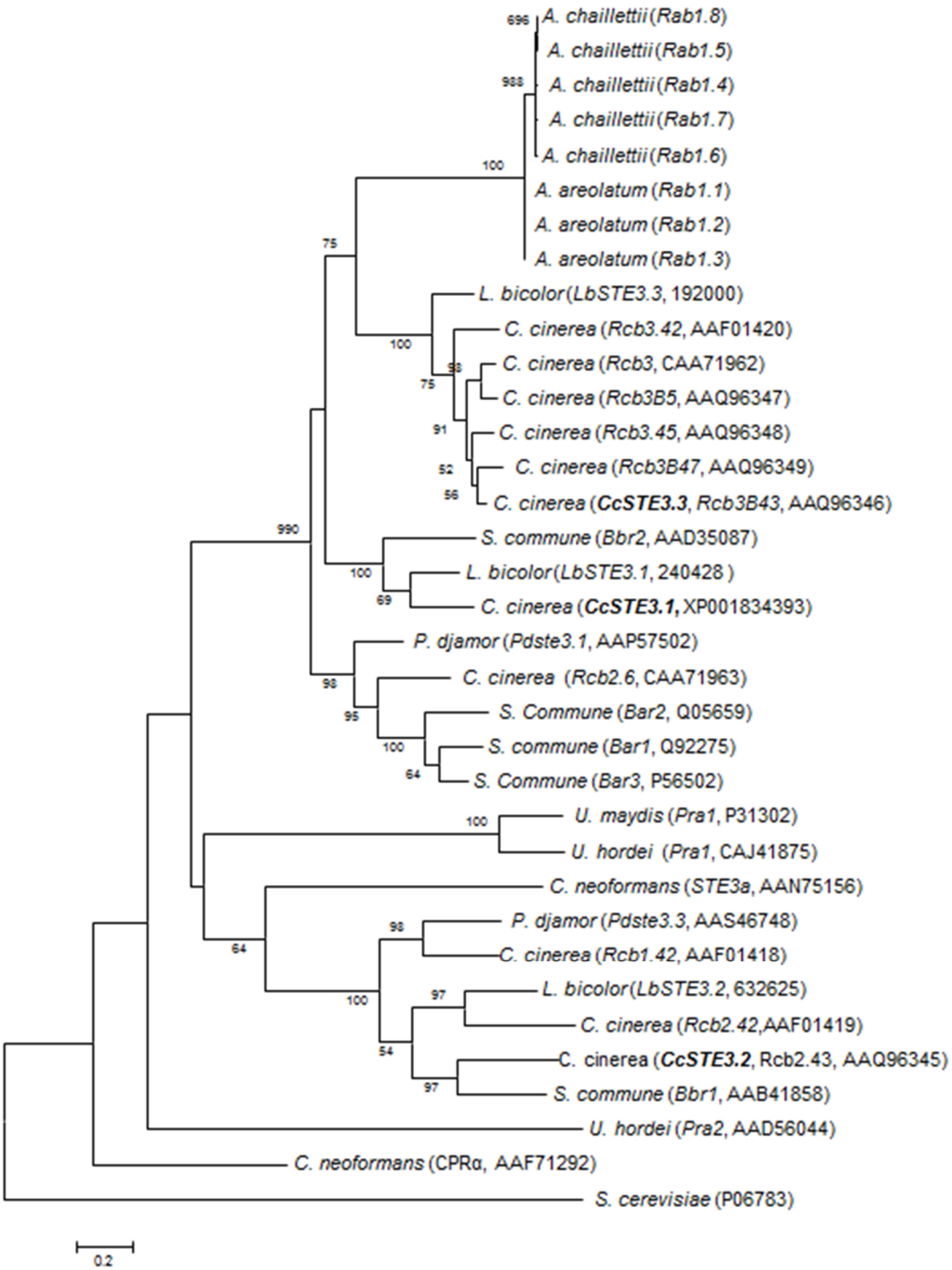


Figure 3. An amino acid-based maximum likelihood phylogeny of the pheromone receptors present in *A. areolatum*, *A. chailletii* and other Agaricomycetes. Percentage bootstrap (100 replicates) values greater than 50% are shown below the tree branches. All of the *rab1* alleles in *A. areolatum* and *A. chailletii* clustered in a species-specific manner, closely with the pheromone receptor genes CcSTE3.3 in *C. cinerea* and LbSte3.3 in *L. bicolor* (Niculita-Hirzel et al. 2008; Martinez et al. 2009).

TABLES

Table 1. Heterozygosity tests[a] of the three genes present in the 13 heterokaryons each of *A. areolatum* and *A. chailletii*.

Locus	Species	A [b]	Π[c]	π_n/π_s [d]
ITS	*A. areolatum*	2	0.0004	-
	A. chailletii	5	0.0020	-
EF	*A. areolatum*	5	0.0050	0.00*
	A. chailletii	5	0.0060	0.00*
rab1	*A. areolatum*	3	0.0042	0.00*
	A. chailletii	4	0.0200	0.00*

[a] All estimates were determined with GENEPOP version 4. Significantly different values are indicated with an asterisk ($P<0.05$).

[b] Average number of alleles per locus.

[c] Mean number of pair-wise differences among sequences (Nei and Li, 1979).

[d] Ratio of non-synonymous nucleotide variation to synonymous nucleotide variation.

Table 2. Likelihood scores and parameter estimates for the site-specific models (Yang et al., 2000) evaluated in this study.

Model [a]	***ln* Likelihood**[b]	Sites with P ($\omega > 1$) >0.95 NEB*	***dN/dS***[c]	**Estimates of parameters**[d]
M0 (One-ratio)	-2100.89	n.a.	0.0682	ω_1=0.0682
M1 (Neutral)	-2096.32	n.a.	0.1321	ω_0= *0.01309* p_0= 0.89171 ω_1=*1* p_1= *0.10829*
M2 (Selection)	-2096.32	-	0.1200	ω_0=0.01308 p_0= 0.89171
				ω_1=*1* p_1= *0.04405*
M3(Discrete)	-2096.14	n.a.	0.1283	ω_0= 0.00000 p_0= *0.81906* ω_1= *0.48426*
				p_1= *0.16723* ω_2= *3.44796* p_2=0.01371
M5 (Gamma)	-2096.31	256 G* 372 A*	0.1194	*a*= 0.06624 *b*= 0.36890
M6 (Double gamma)	-2096.14	256 G*	0.1283	p_0= *0.89259* a_0= 0.04008 b_0= 1.13382
		372 A*		p_1= *0.10741* a_1= *2.73793* b_1= 2.73793
M7 (β distribution)	-2096.30	n.a.	0.1183	*p*= *0.01742* *q*= *0.11677*
M8 (β + positive selection)	-2096.30		0.1260	p_0= *0.96972* *p*= *0.05439* *q*= 0.44414
				ω=1.000

[a] Site-specific models implemented in the CODEML program in PAML version 3.14 package (Yang and Nielson, 2002).

[b] Model likelihoods used for calculating statistical significance with likelihood ratio tests (LRT), which entailed analysis of the χ^2distribution of 2Δln (*i.e.*, twice the log likelihood difference between the two models) for the different models (Yang and Nielson, 1998).

[c] The non-synonymous (*dN*)/synonymous (*dS*) substitution rate ratio (Yang and Nielson, 1998).

[d] Parameters estimated for each model according to those proposed by Yang (2007).

*NEB = Naive empirical Bayes

Table 3. Results of the recombination tests as determined with DnaSP 5.10 (Librado and Rozas, 2009).

Genes	Species	R^a	$R_M{}^b$	Coalescence simulations[c]		
				Confidence interval[d]	P ($R_M \leq$ observed R_M)[e]	Avg $R_M{}^f$
Rab1	*A. areolatum*	0.0010	1.0000	0.0, 0.0	2.8190	0.0002
	A. chailletii	0.0007	0.9940	0.0, 1.0	0.9944	0.1709
ITS	*A. areolatum*	0.0010	0.0001	0.0, 0.0	2.8068	0.0000
	A. chailletii	0.5000	0.9920	0.0, 1.0	4.2558	0.1728
Tef-1α	*A. areolatum*	7.0990	0.6417	0.0, 2.0	1.2937	0.3911
	A. chailletii	0.0200	0.9995	0.0, 0.0	1.5337	0.0005

[a] Estimate of the population recombination parameter R excluding alignment gaps and haploid corrections.

[b] The observed minimum number of recombination events in the data.

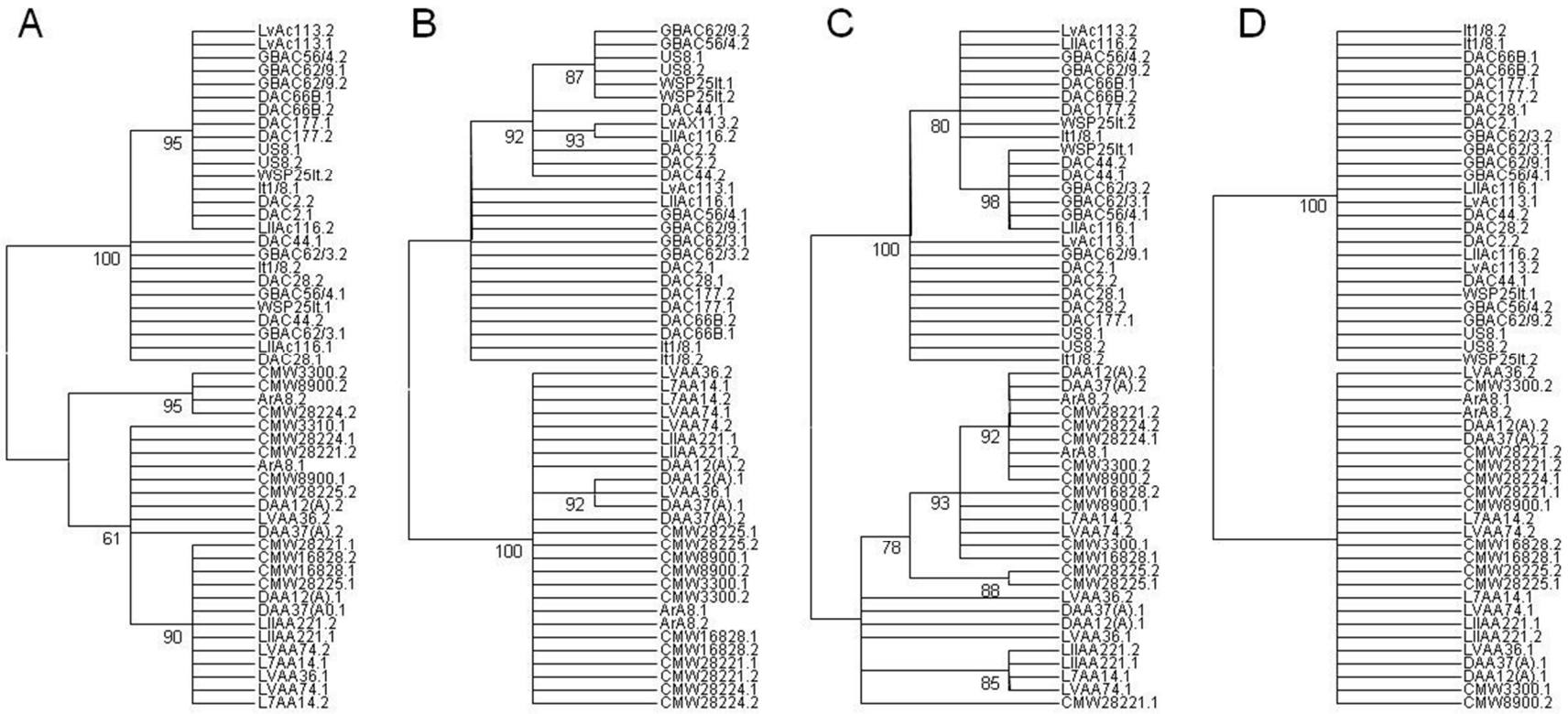
[c] Neutral coalescence simulations given the number of segregating sites, with an intermediate level of recombination and 10000 replications.

[d] The confidence interval (lower limit, upper limit) for R_M.

[e] The probability that R_M is less than or equal to the observed RM.

[f] The average value of R_M.

SUPPLEMENTARY MATERIAL



Supplementary Figure 1. Maximum likelihood gene trees based on the *rab1* (A), ITS (B) and Tef-1α (C) sequences for the isolates of *A. areolatum* and *A. chailletii* included in this study. The strict consensus (D) for the three trees was inferred in MEGA. For heterozygotic isolates, the alleles are indicated by the digits 1 or 2 following an isolate number. Percentage bootstrap (100 replicates) values greater than 50% are shown below the tree branches.

Supplementary Table 1. The nucleotide polymorphisms *rab1* sequences in the 25 *A. areolatum* isolates.

Isolate number	Collector	Geographic origin	Nucleotide polymorphisms[a]							
			62	74	78	89	91	95	150	159
			*	*	*	*	*	*		
CMW16848	R. Vasaitis	Austria	C	T	G	C	G	T	A	A
			C	T	G	T	G	T	A	A
LIIAA5(LK)	R. Vasaitis	Lithuania	T	C	A	T	T	G	G	G
			T	C	A	T	T	G	G	G
LVAA36	R. Vasaitis	Sweden	C	T	G	C	G	T	A	A
			C	T	G	T	G	T	A	A
L7AA14	R. Vasaitis	Denmark	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
LVAA74	R. Vasaitis	Lithuania	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
LIIAA221	R. Vasaitis	Lithuania	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
DAA12(A)	I. Thomsen	Denmark	C	T	G	C	G	T	A	A
			C	T	G	T	G	T	A	A
DAA37(A)	I. Thomsen	Denmark	C	T	G	C	G	T	A	A
			C	T	G	T	G	T	A	A
CMW28225	I. Thomsen	Denmark	C	T	G	C	G	T	A	A
			C	T	G	T	G	T	A	A
DAA547(B)	I. Thomsen	Denmark	C	T	G	C	G	T	A	A
			C	T	G	C	G	T	A	A
CMW4644	B. Slippers	Australia	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A

CMW8898	B. Slippers	Brazil	**C**	**T**	**G**	**C**	**G**	**T**	**A**	**A**
			C	T	G	T	G	T	A	A
CMW8900	B. Slippers	South Africa	**T**	**C**	**A**	**T**	**T**	**G**	**G**	**G**
			C	**T**	**G**	**C**	**G**	**T**	**A**	**A**
CMW3300	B. Slippers	New Zealand	**T**	**C**	**A**	**T**	**T**	**G**	**G**	**G**
			C	**T**	**G**	**C**	**G**	**T**	**A**	**A**
CMW3310	G.B. Rawlings	France	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
ArA8	B. Slippers	Argentina	**T**	**C**	**A**	**T**	**T**	**G**	**G**	**G**
			C	**T**	**G**	**C**	**G**	**T**	**A**	**A**
CMW16828	B. Slippers	Austria	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
At-II-28	B. Slippers	Austria	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
CMW28221	H. Solheim	Norway	C	T	G	T	G	T	A	A
			C	**T**	**G**	**C**	**G**	**T**	**A**	**A**
DAA547B	H. Solheim	Norway	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
CMW28223	O. Holgenrieder	Switzerland	**T**	**C**	**A**	**T**	**T**	**G**	**G**	**G**
			C	T	G	T	G	T	A	A
CMW28224	O. Holgenrieder	Switzerland	**T**	**C**	**A**	**T**	**T**	**G**	**G**	**G**
			C	**T**	**G**	**C**	**G**	**T**	**A**	**A**
AtII-18	B. Slippers	Austria	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
At-23	B. Slippers	Austria	C	T	G	T	G	T	A	A
			C	T	G	T	G	T	A	A
AtIII-4	B. Slippers	Austria	**C**	**T**	**G**	**C**	**G**	**T**	**A**	**A**
			C	T	G	T	G	T	A	A

[a] Polymorphisms located in introns are indicated with asterisks.

Supplementary Table 2. The nucleotide polymorphisms *rab1* sequences in the 25 *A. chailletii* isolates.

Isolate number	Collector	Geographic origin	Nucleotide polymorphisms					
			87	92	103	130	139	160
			*	*	*			
A.ch.536	R. Vasaitis	Sweden	**C**	**T**	**C**	**C**	**T**	**C**
			C	**T**	**C**	**C**	**T**	**C**
LvAc22C	R. Vasaitis	Lithuania	C	T	C	C	C	C
			C	T	C	C	C	C
LvAc113	R. Vasaitis	Lithuania	C	T	C	C	C	C
			C	T	C	C	C	C
LIIAc116	R. Vasaitis	Lithuania	C	T	C	C	C	C
			C	T	C	C	C	C
GBAC56.4	D. Redfern	UK	C	T	C	C	C	C
			C	T	C	C	C	C
GBAC62.11	D. Redfern	UK	**C**	**T**	**C**	**C**	**T**	**C**
			C	T	C	C	C	C
GBAC62.9	D. Redfern	UK	C	T	C	C	C	C
			C	T	C	C	C	C
GB AC62.3	D. Redfern	UK	**C**	**T**	**C**	**C**	**T**	**C**
			C	T	C	C	C	C
GBAC62	D. Redfern	UK	C	T	C	C	C	C
			C	T	C	C	C	C
DAC2	I. Thomsen	Denmark	C	T	C	C	C	C
			C	T	C	C	C	C
DAC7	I. Thomsen	Denmark	C	T	C	C	C	C
			C	T	C	C	C	C
DAC28	I. Thomsen	Denmark	C	T	C	C	C	C
			C	T	C	C	C	C

DAC35	I. Thomsen	Denmark	**C**	**T**	**C**	**C**	**T**	**C**
			C	T	C	C	C	C
DAC44	I. Thomsen	Denmark	C	T	C	C	C	C
			C	T	C	C	C	C
DAC66B	I. Thomsen	Denmark	C	T	C	C	C	C
			C	T	C	C	C	C
DAC66D	I. Thomsen	Denmark	**C**	**T**	**C**	**C**	**T**	**C**
			C	T	C	C	C	C
DAC78	I. Thomsen	Denmark	C	T	C	C	C	C
			C	T	C	C	C	C
DAC177	I. Thomsen	Denmark	C	T	C	C	C	C
			C	T	C	C	C	C
DAC1941	I. Thomsen	Denmark	**C**	**T**	**C**	**C**	**T**	**C**
			C	T	C	C	C	C
US1	H.H. Burdsall	US	**A**	**C**	**G**	**T**	**C**	**T**
			C	T	C	C	C	C
US8	H.H. Burdsall	US	**A**	**C**	**G**	**T**	**C**	**T**
			C	T	C	C	C	C
CMW3299	R.F. Cain	Canada	C	T	C	C	C	C
			C	T	C	C	C	C
GRWSP3	B. Slippers	Greece	C	T	C	C	C	C
			C	T	C	C	C	C
WSP25It	B. Slippers	Italy	C	T	C	C	C	C
			C	T	C	C	C	C
It1.8	B. Slippers	Italy	C	T	C	C	C	C
			C	T	C	C	C	C

[a] Polymorphisms located in introns are indicated with asterisks

Supplementary Table 3. List of the primers used for PCR.

Primer	Sequence 5'to 3'	Reference
rab1	TTATGAAGCGGGTCGCTACAAG	This study
RAB2	TACAGCCATGACCATCTGGAACTT	″
RAB3	CGTGTACCCCAAGGTCATCT	″
RAB4	CCGAGTCTCGACGTA CAACC	″
RAB5	CCAGACGTCGACCTCACAT	″
RABF	CTGGCCTACGTCCTCGTCTA	″
RABR	GTATGTAGCGGCTGGAGTTG	″
RAB7	TGGCCTACGTCCTCGTCTAT	″
RAB8	AGGGGTTGTACGTTGAGACG	″
RAB9	GTCTTCTCCGCCTTCTCCTTC	″
RAB 10	CCAGACGTCGACCTCACAT	″
RAB 11	TTGTTTTCCATGTCCCCTTC	″
RAB 12	ATTCCGTAGGCGAAAACTCC	″
RAB1-470F	TCTTGGGCTGACTTTTCC	van der Nest et al., 2008
RAB 1-1800R	GGCAGGTAGATCGAGGTTG	van der Nest et al., 2008
BR1-F	TGGCATMTNCARGCNTGGAAYTC	James et al., 2004b
BR1-1R	GCGAGNRNCATNAGNCGNAKGTA	James et al., 2004b
ITS1	TCCGTAGGTGAACCTGCGG	White et al., 1990
ITS4	TCCTCCGCTTATTGATATGC	White et al., 1990
TEFac1	TCCTGGAGAGGAAGACGAAG	This study
TEFac2	GTTCGAGGCTGGTATCTCCC	″
TEFaa1	AGACGTCCTGGAGAGGAAGG	″
TEFaa2	GGTATCTCCAAGGACGGTCA	″

Supplementary Table 4. The nucleotide polymorphisms in the Tef-1α, ITS and Rab1 sequences in the 13 *A. areolatum* isolates.

Isolate number	Geographic origin	Nucleotide polymorphisms																		
		ITS	Tef-1alpha type 1						*rab1*											
		540	81	147	267	312	330	341	62	74	78	89	91	95	150	159	288	375	569	605
				*		*	*	*	*	*	*	*	*	*					*	*
LVAA36	Sweden	T	T	C	C	A	A	T	C	T	G	C	G	T	A	A	C	C	G	T
		T	T	C	C	A	A	T	C	T	G	T	G	T	A	A	C	C	G	T
L7AA14	Lithuania	C	T	C	C	G	A	T	C	T	G	T	G	T	A	A	C	C	G	T
		C	C	T	T	A	G	T	C	T	G	T	G	T	A	A	C	C	G	T
LVAA74	Lithuania	C	T	C	C	G	A	T	C	T	G	T	G	T	A	A	C	C	G	T
		C	C	T	T	A	G	T	C	T	G	T	G	T	A	A	C	C	G	T
LIIAA221	Lithuania	C	T	C	C	G	A	T	C	T	G	T	G	T	A	A	C	C	G	T
		C	T	C	C	G	A	T	C	T	G	T	G	T	A	A	C	C	G	T
DAA12(12)	Denmark	T	T	C	C	A	A	T	C	T	G	C	G	T	A	A	C	C	G	T
		C	C	T	T	A	G	C	C	T	G	T	G	T	A	A	C	C	G	T
DAA37(A)	Denmark	T	C	T	T	A	G	C	C	T	G	C	G	T	A	A	C	C	G	T
		T	T	C	C	A	A	T	C	T	G	T	G	T	A	A	C	C	G	T
CMW28225	Denmark	C	T	C	C	A	A	T	C	T	G	C	G	T	A	A	C	C	G	T
		C	T	C	C	A	A	T	C	T	G	T	G	T	A	A	C	C	G	T

CMW8900	South Africa	C	C	T	T	A	G	T	T	C	A	T	T	G	G	G	T	T	T	C
		C	C	T	T	A	G	C	C	T	G	C	G	T	A	A	C	C	G	T
CWM3300	New Zealand	C	C	T	T	A	G	T	T	C	A	T	T	G	G	G	T	T	T	C
		C	C	T	T	A	G	C	C	T	G	C	G	T	A	A	C	C	G	T
ArA8	Argentinia	C	C	T	T	A	G	C	T	C	A	T	T	G	G	G	T	T	T	C
		C	C	T	T	A	G	C	C	T	G	C	G	T	A	A	C	C	G	T
CMW16828	Austria	C	C	T	T	A	G	T	C	T	G	T	G	T	A	A	C	C	G	T
		C	C	T	C	A	G	T	C	T	G	T	G	T	A	A	C	C	G	T
CMW28221	Norway	C	C	T	T	A	G	C	C	T	G	T	G	T	A	A	C	C	G	T
		C	T	C	C	A	A	T	C	T	G	C	G	T	A	A	C	C	G	T
CMW28224	Switzerland	C	N	N	N	N	N	N	T	C	A	T	T	G	G	G	T	T	T	C
		C	N	N	N	N	N	N	C	T	G	C	G	T	A	A	C	C	G	T

[a] Polymorphisms located in introns are indicated with asterisks

Supplementary Table 5. The nucleotide polymorphisms in the Tef-1α, ITS and Rab1 sequences in the 13 *A. chailletii* isolates.

Isolate number	Geographic origin	Nucleotide polymorphisms[a]																							
		ITS						Tef-1α								*rab1*									
		97	144	193	220	504	512	129	135	162	325	343	358	486	498	87	92	103	130	160	260	280	391	553	615[b]
								*	*		*	*				*	*	*						*	*
LvAc113	Lithuania	A	A	-	G	G	C	C	T	A	C	C	C	C	T	C	T	C	C	C	A	G	C	insert	-
		A	G	-	C	G	C	C	T	A	C	C	C	C	C	C	T	C	C	C	A	G	C	Insert	-
LIIAc116	Lithuania	A	G	-	C	G	C	C	T	A	C	C	C	T	C	C	T	C	C	C	G	G	C	Insert	insert
		A	G	-	C	G	C	T	-	G	T	T	T	C	T	C	T	C	C	C	A	G	C	Insert	-
GBAC56.4	UK	A	G	-	G	A	C	C	T	A	C	C	C	T	C	C	T	C	C	C	A	G	C	Insert	-
		A	A	A	G	G	C	T	-	G	T	T	T	T	C	C	T	C	C	C	G	G	C	Insert	insert
GBAC62.9	UK	A	G	-	G	A	C	C	T	A	C	C	C	C	T	C	T	C	C	C	A	G	C	Insert	-
		A	A	A	G	G	C	C	T	A	C	C	C	C	C	C	T	C	C	C	A	G	C	Insert	-
DAC2	Denmark	A	G	-	C	G	C	C	T	A	C	C	C	C	T	C	T	C	C	C	A	G	C	Insert	-
		A	A	A	G	G	C	C	T	A	C	C	C	C	T	C	T	C	C	C	A	G	C	Insert	-
DAC28	Denmark	A	G	-	C	G	C	C	T	A	C	C	C	C	T	C	T	C	C	C	G	G	C	Insert	insert
		A	A	-	G	G	C	C	T	A	C	C	C	C	T	C	T	C	C	C	G	G	C	Insert	insert
DAC44	Denmark	G	G	-	G	A	T	T	-	G	T	T	T	T	C	C	T	C	C	C	G	G	C	Insert	insert
		A	G	-	G	A	C	T	-	G	T	T	T	T	C	C	T	C	C	C	G	G	C	Insert	insert

DAC66B	Denmark	A	A	-	G	G	C	C	T	A	C	C	C	T	C	C	T	C	C	C	A	G	C	Insert	-
		A	A	-	G	G	C	C	T	A	C	C	C	T	C	C	T	C	C	C	A	G	C	Insert	-
DAC177	Denmark	A	A	-	G	G	C	C	T	A	C	C	C	T	C	C	T	C	C	C	A	G	C	Insert	-
		A	A	A	G	G	C	C	T	A	C	C	C	C	T	C	T	C	C	C	A	G	C	Insert	-
US1	US	G	G	-	G	A	T	n	N	n	n	n	n	n	n	C	T	C	C	C	A	G	C	Insert	-
		A	G	-	G	A	C	n	N	n	n	n	n	n	n	A	C	G	T	T	G	A	G	-	insert
US2	US	A	G	-	G	A	C	C	T	A	C	C	C	C	T	C	T	C	C	C	A	G	C	Insert	-
		A	G	-	G	A	C	C	T	A	C	C	C	C	T	C	T	C	C	C	A	G	C	Insert	-
WSP25It	Italy	A	G	-	G	A	C	C	T	A	C	C	C	T	C	C	T	C	C	C	G	G	C	insert	insert
		A	G	-	G	A	C	T	-	G	T	T	T	T	C	C	T	C	C	C	A	G	C	Insert	-
It1.8	Italy	A	A	A	G	G	C	C	T	A	C	C	C	T	C	C	T	C	C	C	A	G	C	Insert	-
		A	A	A	G	G	C	C	T	A	C	C	C	C	T	C	T	C	C	C	G	A	G	-	insert

[a] Polymorphisms located in introns are indicated with asterisks. Deletions are indicated with hyphens and unknown bases with "n".

[b] insert sequence "**TT**"

[c] insert sequence "**TCGGCTCCTACCCGCACATGT**"

Supplementary Table 6. Comparison of percentage identity of the *rab1* alleles present in *A. areolatum* and *A. chailletii*.

Locus	**Species**	*Rab1.1*	*Rab1.2*	*Rab1.3*	*Rab1.4*	*Rab1.5*	*Rab1.6*	*Rab1.7*
Rab1.1	*A. areolatum*	-						
Rab1.2	*A. areolatum*	98.9*	-					
Rab1.3	*A. areolatum*	98.9	99.8	-				
Rab1.4	*A. chailletii*	83.5	83.6	83.5	-			
Rab1.5	*A. chailletii*	81.9	82.1	81.9	98.4	-		
Rab1.6	*A. chailletii*	83.0	83.1	83.0	98.6	97.0	-	
Rab1.7	*A. chailletii*	83.2	83.3	83.2	99.4	97.9	98.8	-

* The sequence identity is the number of exactly matching residues (expressed as a percentage).