Gene organization of the mating type regions in the ectomycorrhizal fungus *Laccaria bicolor* reveals distinct evolution between the two mating type loci

Hélène Niculita-Hirzel¹, Jessy Labbé², Annegret Kohler³, François le Tacon², Francis Martin², Ian R. Sanders¹ and Ursula Kües³

¹Department of Ecology and Evolution, University of Lausanne, Biophore Building, CH–1015 Lausanne, Switzerland; ²UMR 1136, Interactions Arbres/Microorganismes, INRA-Nancy, F–54280 Champenoux, France; ³Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute, Georg-August-University Göttingen, D–37077 Göttingen, Germany

**Summary**

- In natural conditions, basidiomycete ectomycorrhizal fungi such as *Laccaria bicolor* are typically in the dikaryotic state when forming symbioses with trees, meaning that two genetically different individuals have to fuse or ‘mate’. Nevertheless, nothing is known about the molecular mechanisms of mating in these ecologically important fungi.
- Here, advantage was taken of the first sequenced genome of the ectomycorrhizal fungus, *Laccaria bicolor*, to determine the genes that govern the establishment of cell-type identity and orchestrate mating.
- The *L. bicolor* mating type loci were identified through genomic screening. The evolutionary history of the genomic regions that contained them was determined by genome-wide comparison of *L. bicolor* sequences with those of known tetrapolar and bipolar basidiomycete species, and by phylogenetic reconstruction of gene family history.
- It is shown that the genes of the two mating type loci, *A* and *B*, are conserved across the Agaricales, but they are contained in regions of the genome with different evolutionary histories. The *A* locus is in a region where the gene order is under strong selection across the Agaricales. By contrast, the *B* locus is in a region where the gene order is likely under a low selection pressure but where gene duplication, translocation and transposon insertion are frequent.

**Key words:** gene order conservation, mating type loci, mushroom-forming fungi, recombination, sex chromosome evolution.


---

**Introduction**

The growth of boreal, temperate and montane forest trees relies on root colonization by ectomycorrhizal fungi, Basidiomycetes and to a lesser extent Ascomycetes, that provide the host plant with nutrients and water in exchange for photosynthetically derived carbohydrates. Basidiomycetes display a life cycle that is predominantly dikaryotic with two separate haploid nuclei in the hyphal cells, the mating process between two monokaryons taking place, in natural conditions, before the establishment of the symbiotic association (Harley & Smith, 1983) or before the infection of the plant by pathogenic Basidiomycetes such as *Ustilago maydis* (for review see Banuett, 2007). In the laboratory, some basidiomycete monokaryotic mycelia have the potential to form ectomycorrhizas (Kropp & Fortin, 1988; Gay et al., 1994) but their symbiotic potential was shown to be lower than that of dikaryons (Kropp & Fortin, 1988; Gay et al., 1994). Consequently, understanding
how the specialized regions of the genome that control mating in fungi (the mating type loci) regulate the formation of the dikaryons, and if they regulate the establishment of the symbiotic association, is important. Although the mating type loci have been studied in a large number of basidiomycete species (for review see Casselton & Olesnicky, 1998; Fraser & Heitman, 2004; Fraser et al., 2007), very little is known about the genetics of mating in ectomycorrhizal fungi. Recently the first draft genome of a basidiomycete ectomycorrhizal fungus, *Laccaria bicolor*, has been released (Martin et al., 2008) and can be compared with other closely related basidiomycete genomes in which the mechanisms that give rise to mating compatibility are understood.

In all Ascomycetes, and some Basidiomycetes, mating compatibility is determined by different alleles of a single mating type (MAT) locus (mating type system called bipolar). In a large number of basidiomycete species (55–65% of the Agaricomycotina) the mating compatibility is controlled by different alleles of two unlinked regions of the genome (Whitehouse, 1949; Raper & Flexer, 1971), commonly known as the A and B loci (called tetrapolar mating system). The number of potentially compatible mating type alleles per mating type locus gives rise to thousands of potentially compatible mating types in some tetrapolar fungi (e.g. c. 12 000 in *Coprinopsis cinerea* and c. 20 000 in *Schizophyllum commune*, Raper, 1966). The population genetic consequences of the tetrapolar versus the bipolar system is a reduction in the amount of interbreeding among the haploid progeny of one individual (*A1B1* can mate with *A2B2*, but not with *A1B1, A1B2* or *A2B1*).

The genes encoded by the A and B loci have been characterized in a number of Basidiomycetes. In all studied higher Basidiomycetes, the A mating type locus contains genes for two types of homeodomain transcription factors (HD1 and HD2) and the genes at the B mating type locus encode lipopeptide pheromones and G-protein-coupled pheromone receptors (STE3-like pheromone receptors with seven transmembrane domains) (for review see Casselton & Olesnicky, 1998; Fraser et al., 2007). More than one copy of these genes is frequently observed in agaricomycete species. The model species *C. cinerea* contains three gene copies for each type of homeodomain transcription factor (gene pair *a, b* and *d*), at least six copies of mating type pheromones and three copies of STE3-like pheromone receptors (subfamily 1, 2 and 3). By contrast, the model species *S. commune* possesses two gene copies for STE3-like pheromone receptors and six pheromone genes and at least two gene copies for each type of transcription factor (for review see Casselton & Olesnicky, 1998; Casselton & Kües, 2007). Each unique combination of alleles at the different subloci specifies a unique mating type, making the subloci redundant in function. Receptors and pheromones diverged rapidly in sequence within and between species. Only the seven transmembrane domains share enough similarity in sequence between species to help the assignment of new STE3-like receptors. By contrast, the mating type pheromones share only a CaaX motif at the C-terminus and this makes them difficult to detect. Similarly, the homeodomain transcription factors diverged in sequence within and between species while certain domains for essential functions remain conserved.

The Agaricomycetes use an elegant strategy to ensure that mating-dependent developmental pathways are activated only after fusion between compatible mates. Two functional domains of their mating type-specific transcription factors are separated into two proteins encoded at the *A* locus by divergently transcribed pairs of the HD1 and HD2 genes, respectively. The products of the HD2 gene family carry the crucial DNA binding domain, the homeodomain motif HD2. By contrast, the products of the HD1 gene family carry an atypical homeodomain motif HD1, with low or no affinity for DNA. In *C. cinerea*, HD1 proteins were shown to also carry the nuclear localization signal (NLS) and the activation domain (AD), thought to be required for transcriptional activation of target genes (Asante-Owusu et al., 1996). Moreover, dimerization motifs were characterized in the N-terminal domains of HD1 and HD2 gene products, respectively. Indeed, heterodimerization of HD1 and HD2 gene products from the different mates plays an important role in bringing together the various domains required for the formation of a functional transcription factor complex (Banham et al., 1995; Spitt et al., 1998). If the full function of HD1/HD2 heterodimers is conserved across Agaricales, the functional motifs described in *C. cinerea* are also expected to be detected in *L. bicolor* homologues.

The *L. bicolor* mating system is considered to be tetrapolar (Fries & Mueller, 1984; Kropp & Fortin, 1988; Doudrick & Anderson, 1989) with at least 45 *A* and 24 *B* mating type alleles (Raffle et al., 1995). In order to understand the mechanisms involved in compatibility and incompatibility during mating in *L. bicolor*, it is essential to know which genes are located at each of the mating type loci. Furthermore, understanding how allelic diversity can be generated at these loci requires the analysis of the evolution of these two loci and, in particular, whether recombination plays a major role in the diversity and arrangement of genes at these loci. Indeed, an increase in recombination rate in the vicinity of mating type loci could reduce the association and potential selective conflicts between mating type genes and linked genes (Hsueh et al., 2006). Alternatively, recombination activity within the gene could mix existing variations in a combinatorial fashion to contribute to the evolution of novel alleles (Badrane & May, 1999). Such a mechanism has been suggested to explain the evolution of shared clusters of genes among major histocompatibility complex alleles (Hughes et al., 1993; Yuhki & O’Brien, 1994; Zangenberg et al., 1995) and alleles of the gametophytic self-incompatibility locus (Awadalla & Charlesworth, 1999; Vieira et al., 2003). Recombination within a gene may also have an advantage if it increases the effectiveness with which deleterious mutations are eliminated (Kondrashov, 1984; Rice, 2002). Such an effect may be most relevant among...
interacting genes (Rice, 1998) or among sex-determining alleles (Beye et al., 1999).

The aim of the present study was to identify the A and B loci in whole-genome sequences of L. bicolor and compare the genes and their arrangement to loci known in other Basidiomycetes such as C. cinerea, Phanerochaete chrysosporium or Pleurotus djamor. We then analysed whether the two loci have evolved with recombination in or around the loci and within genes at these two loci.

Materials and Methods

Strains and culture conditions

A fruiting body of L. bicolor (Maire) P.D. Orton (common name: bicoloured deceiver) was collected in 1976 under Tsuga mertensiana in the Crater Lake National Park (OR, USA) by J. Trappe and R. Molina and deposited at the Forest Service (Corvallis, OR, USA). A subculture of this strain, so-called S238-O, was transferred to INRA-Nancy in 1980 and renamed S238N (Di Battista et al., 1996). For this study, spores of L. bicolor were obtained from sporophores collected under Pseudotsuga menziesii, that had been inoculated with L. bicolor strain S238N in a glasshouse or in a nursery (Di Battista et al., 1996) and germinated according to Fries (1983). Forty-two different sib-monokaryotic mycelia were used in this study, including the H82 line whose genome was sequenced (Martin et al., 2008). All material was stored and subcultured at INRA-Nancy (Champenoux, France). All monokaryotic strains were grown in Petri dishes containing a modified Pachlewski agar medium (Di Battista et al., 1996) and germinated according to Fries (1983).

MAT loci annotation

The L. bicolor genome sequencing, assembly and annotation were described in Martin et al. (2008). All L. bicolor sequences from the monokaryon H82 are available at the Joint Genome Institute website: http://genome.jgi-psf.org/Lacbi1/Lacbi1.home.html. Complete C. cinerea (= Coprinus cinereus) DNA sequences and protein sequences were obtained from the C. cinerea strain Okayama 7 genome database (Coprinus cinereus Sequencing Project, Broad Institute of MIT and Harvard, http://www.broad.mit.edu; Stajich et al., 2006), while sequence data for P. chrysosporium (Martinez et al., 2004) were obtained from the Joint Genome Institute website (http://genome.jgi-psf.org/Phchr1/Phchr1.home.html) and the sequence data for Pleurotus djamor from National Center for Biotechnology Information (NCBI) GenBank (accession numbers AY462111 and AY462110 for A and B locus, respectively). The evolutionary relationships between species used in the present study are shown in Supporting Information Fig. S1. We carried out BLAST translated DNA and amino acid sequence similarity searches of the predicted proteins from C. cinerea and P. chrysosporium mating type regions against the L. bicolor translated genome and vice versa using BLOSUM62 substitution matrices and the DUST and SEG filters for low-complexity regions (Altschul et al., 1997). All reciprocal hits were tested for functional homology by psi-blast searches against all NCBI database (http://www.ncbi.nlm.nih.gov/BLAST/) and SMART (http://smart.embl-heidelberg.de/index2.cgi) searches for putative conserved domains. Genes were annotated in L. bicolor based on homology to the existing annotation in other fungi. Correction on some L. bicolor, C. cinerea and P. chrysosporium gene annotation as the annotation of new genes for P. djamor and new pheromone-like genes in L. bicolor and C. cinerea are detailed in the Supporting Information, Text S1.

Syntenic relationships were illustrated with CHROMOMAPPER (http://www2.unil.ch/biomapper/chromomapper/, Niculita-Hirzel & Hirzel, 2008), COILS program (http://www.ch.embnet.org/software/COILS_form.html; Lupas et al., 1991) was used to locate dimerization motifs at a window size of 14 sites. HMMTOP 2.0 (Tusnády & Simon, 2001), TMHMM (Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark) and PHOBiUS (Käll et al., 2007) programs were used to predict the transmembrane helices in STE3-like pheromone receptor proteins.

Phylogenetic reconstruction

Protein-coding DNA sequences were aligned using the ClustalW algorithm and edited by hand in BIOEDIT version 7.0.5.3 (Hall, 1999). To reconstruct the evolutionary relationships of L. bicolor STE3-like pheromone receptor sequences with that of other fungal species the full-length coding sequences of the STE3-like pheromone receptor genes deduced from the L. bicolor genome (http://genome.jgi-psf.org/Lacbi1/Lacbi1.home.html; Martin et al., 2008), the C. cinerea strain Okayama 7 genome (http://www.broad.mit.edu/annotation/genome/coprinus_cinereus/Home.html), the P. chrysosporium genome (http://genome.jgi-psf.org/Phchr1/Phchr1.home.html), as well as from the Coprinellus disseminatus and P. djamor B locus sequences available at the NCBI GenBank were used. The Cryptococcus neoformans STE3 sequence was added as an outgroup. The accession numbers of all these sequences are given in the Supporting Information section. To estimate phylogenetic relationships, the alignment was analysed using sequential and parallel maximum likelihood-based inference as implemented in the RAxML software, version 4.0.0 (Stamatakis, 2006), with a general time-reversible model of nucleotide substitution and additionally assuming a percentage of invariant sites and gamma-distributed substitution rates at the remaining sites (GTR + I + G). Branch support was inferred from 1000 replicates of nonparametric bootstrapping (Felsenstein, 1985). Additionally, a Bayesian Markov chain Monte Carlo (MCMC) analysis was performed using MRBAYES 3.1 (Ronquist & Huelsenbeck, 2003). We ran two independent
MCMC analyses, each involving six incrementally heated chains over one million generations, using the GTR + I + G model of nucleotide substitution and starting from random trees. Model parameters were not fixed but sampled during MCMC. Trees were sampled every 100 generations, resulting in an overall sampling of 10 000 trees per run, from which the first 5000 trees of each run were discarded. The remaining 5000 trees sampled from each run were pooled and used to compute a majority rule consensus tree to get estimates for the posterior probabilities. Stationarity of the process was assessed using the TRACER software (Rambaut & Drummond, 2003).

Gene expression

Expression of predicted mating type genes was detected in free-living mycelium of *L. bicolor* S238N, ectomycorrhizal root tips (Douglas fir ECM and Poplar ECM) and fruiting bodies of *L. bicolor* S238N using the *L. bicolor* whole-genome expression array data described in Martin *et al.* (2008).

DNA extraction

Mycelium for DNA extraction was prepared by growing isolates at 28°C in 20 ml of Pachlewski liquid medium according to Henrion *et al.* (1992), until stationary phase. The mycelium was removed from the growth medium, rinsed in H₂O, and lyophilized. Approximately 80 mg were used for DNA extraction, using the DNeasy Plant Mini Kit (Qiagen, Courtaboeuf, France) according to the manufacturer's instructions, and DNA was recovered in 50 µl of distilled water. Quality of DNA extracts was tested by polymerase chain reaction (PCR) amplification of the ribosomal DNA internal transcribed spacer (ITS) using the following primers: ITS1 (5′-TCCTCCGCTTATTGATATGC-3′) and ITS4 (5′-TCGGTAAGGTGAACCTGCGG-3′) (Hention *et al.*, 1992).

Recombination analysis

In order to determine the recombination rate between and within *L. bicolor* mating type loci, the inheritance of the allele 1 or allele 2 of each mating type gene present in the dikaryon S238N was determined in 42 sib-monokaryons, (F₁ progeny of the dikaryon S238N) by a allele specific PCR test. The allele 1 corresponds to the monokaryon H82 sequence whose genome was sequenced (Martin *et al.*., 2008). A partial sequence of the second allele of *Lba1* (**Lba1.2**) and *LbSTE3.1* (**LbSTE3.1.2**) were obtained from the dikaryon *L. bicolor* S238N cDNA library constructed and sequenced as described in the Supporting Information section. Contigs with alleles **Lba1.2** and with **LbSTE3.1.2** were generated from the accession numbers EYMTT9U01EFVE6, EYMTT9U01DQOK6 and EYMTT9U01BO6Z1, EYMTT9U01CEA01, EYMTT9U01D5QSG and EYMTT9U01DNINT, respectively.

Two independent PCRs were conducted for each gene in each monokaryon. The DNA from the monokaryon H82 and the dikaryon S238N was used as a positive control. A
PCR was carried out on 15 ng of genomic DNA in 25 µl, with final concentrations of 250 µM of each dNTP, 0.2 µM of each of the primers (Micsynth, Balgach, Switzerland), Qia-gen PCR buffer containing 25 mM MgCl₂, 1x, 1.5 units of Taq DNA Polymerase (Qiagen). Reactions were performed in a T1 Thermocycler (Biometra, Châtel-St-Denis, Switzerland) under the following conditions: initial denaturation at 94°C for 30 s, annealing at specific temperature (Supporting Information, Table S1) for 30 s and extension at 72°C for 90 s. After 40 cycles, a terminal extension of 7 min at 72°C was carried out. Finally the PCR products were run in 1% agarose gel. In order to confirm if the sequence is belonging to the allele 1 or allele 2 of a particular gene, amplicons from five monokaryons were analysed on an ABI3700 to confirm if the sequence is belonging to the allele 1 or allele 2 of a particular gene, amplicons from five monokaryons were analysed on an ABI3700.

Results

Characterization of the L. bicolor A locus

We screened the L. bicolor genome of the H82 monokaryon for HD1 and HD2 sequences and found one gene encoding the HD1 domain (Lba1, ID: 301103) and one gene encoding the HD2 domain (Lba2, ID: 379291). The two genes are linked and divergently transcribed (Fig. 2) as seen in A mating type loci of other basidiomycete species (Casselton & Olesnicky, 1998; James et al., 2004, 2006). As expected, HD1 and HD2 domains were both highly conserved between the Basidiomycetes, including L. bicolor. However, few other motifs were found to be conserved among products of the HD1 gene family (see the Supporting Information, Fig. S2). The absence of motifs conserved among products of the HD2 gene family (see the Supporting Information, Fig. S3) is in accordance with the finding in C. cinerea that almost all sequence C-terminal to the HD2 homeodomain can be deleted without a loss of function (Kües et al., 1994).

The conserved motifs between the HD1 proteins correspond to nuclear localization signals, transactivation domains and dimerization domains already characterized in C. cinerea. Indeed, the two nuclear localization signals (NLS1 and NLS2) present in the C-terminal region of the C. cinerea HD1 proteins (Tymon et al., 1992, Asante-Owusu et al., 1996, Spit et al., 1998) were both detected in the L. bicolor Lba1 protein (Fig. 2, Fig. S2). These observations suggest a functional conservation of both nuclear localization motifs in Agaricales.

Four transactivation domains, AD1–AD4, were predicted in C. cinerea HD1 proteins (Badrane & May, 1999), but only three of them (AD1, AD2 and AD4) appear to be of functional importance (Kües et al., 1994, Asante-Owusu et al., 1996). The similarity in the serine, threonine and proline composition (Mermod et al., 1989; Tymon et al., 1992) and the sequence conservation of five to seven amino acids between AD1 or AD2 transactivation domains of L. bicolor, C. cinerea, C. disseminatus (a1 protein only) and Coprinopsis scobicola HD1 proteins, suggest a functional conservation of these two domains through all these species. By contrast, the low similarity in amino acids composition and the absence of sequence conservation observed between AD4 of L. bicolor and the other fungal species suggest that this transactivation domain might not be functional in L. bicolor (Fig. S2).

Because a heterodimeric interaction between HD1 and HD2 proteins is required for sexual compatibility (Banham et al., 1995; Spit et al., 1998), we looked for coiled/coil motifs using predictive algorithms (Lupas et al., 1991). The robustness of these predictions has already been tested by Badrane & May (1999) on C. cinerea proteins. Two coiled/coil motifs, Di-1 and Di-2, were predicted in the Lba2 protein with a high probability (P = 0.90 and P = 0.98, respectively; Fig. S3), one on each side of the HD2 domain (Fig. 2). Interestingly, only the Di-1 motif is at a similar position in all Agaricales proteins and was shown to be functional in C. cinerea proteins suggesting that only this motif has an important role in Lba1/Lba2 dimerization (Banham et al., 1995).

For the Lba1 protein, three coiled/coil motifs were predicted with distinct probabilities: one on N-terminal and two on C-terminal of HD1 domain (Fig. 2). The N-terminal motif, Di-1, was predicted with a low probability (P = 0.30) as in C. cinerea alleles (P = 0.088–0.35). However, Di-1 was proved to have a functional importance for heterodimerization and protein discrimination in C. cinerea (Banham et al., 1995). Consequently, it seems that the Di-1 motif could also participate in Lba1/Lba2 protein dimerization. The C-terminal coiled/coil motifs, Di-3a and Di-3b, of Lba1 were predicted with a strong probability of 0.90 and 0.98, respectively (Fig. S2) and their positions are conserved between the Lba1 and two proteins of C. disseminatus or Lba1 and most C. cinerea HD1 proteins. Nevertheless, in vitro protein dimerization studies with C. cinerea proteins suggest that Di-3b does not participate in heterodimerization...
Research

Characterization of the *L. bicolor* B locus

In a tetrapolar species, such as *L. bicolor*, the B locus is expected to contain at least one mating type pheromone gene and one STE3-like pheromone receptor gene. The *L. bicolor* genome was screened with all known Agaricales STE3-like sequences and thirteen STE3-like pheromone receptor genes were identified. Their distribution in the *L. bicolor* genome is given in the Supporting Information (Table S2). Only eight of these genes encode all the seven transmembrane helices characteristic of STE3-like receptors (as predicted by hmmtop, tmhmm and phobious programs). The other five genes correspond to partial duplication of pheromone receptor genes and they are not functional (Table S2). Mating type specific pheromone genes (*PhB1*, *PhB2* and *PhB3*; Fig. 3), encoding the characteristic CaaX motifs at the C-terminus of their proteins, were detected only in the close proximity of *LbSTE3.1, LbSTE3.2* and *LbSTE3.3*. These three STE3-like pheromone receptors/mating type pheromone gene pairs are clustered in one locus of scaffold 56 (Fig. 3).

Interestingly, the molecular phylogeny of STE3-like pheromone receptor sequences from *L. bicolor, C. cinerea, C. disseminatus, P. djamor* and *P. chrysosporium* shows that *LbSTE3.1, LbSTE3.2* and *LbSTE3.3* belong to the three independent subfamilies of STE3-like receptors known to be involved in mating in *C. cinerea* (Riquelme et al., 2005). In *C. cinerea*, the genes for these three STE3-like receptors (*CcSte3.1, CcSte3.2b* and *CcSte3.3*) are grouped into a single locus encompassing approx. 9.5 kb, with two or more mating-type-specific pheromone genes being associated with each of them and an extra gene coding for another *CcSte3.2* receptor (*CcSte3.2a*) in between that appeared to be obtained through a gene duplication event (Figs 3, 4). Surprisingly, the order of the genes coding for the three independent subfamilies of STE3-like receptor is conserved between *C. cinerea* strain Okayama 7 and *L. bicolor* H82 monokaryon despite the lack of homology in the surrounding regions between the two species.
Four of the five other full-length *L. bicolor* STE3-like receptors genes (*LbSTE3.4, LbSTE3.6 and LbSTE3.9, LbSTE3.13*) appear to be homologues to three other subfamilies of STE3 receptors that are not known to play a role in mating (Fig. 4). Moreover, no pheromone genes were found in the proximity of these receptors suggesting that they do not assume the *B* locus function. Nevertheless, 45 more pheromone-like genes encoding peptides with a CaaX motifs at their C-terminus, were detected spread throughout the *L. bicolor* genome, of which 36 look functional (Table S3). The unprocessed precursors of all these pheromone-like peptides have a shorter N-terminal domain than the mating type-specific pheromone precursors and, surprisingly, they have multiple homologues in the *C. cinerea* genome (Fig. S4). The sequence conservation of these pheromone-like peptides, as well as of nonmating type STE3-like pheromone receptors, suggests that the function of these shared components of signalling pathways is conserved between *C. cinerea* and *L. bicolor*.

To examine whether all the full-length STE3-like receptors of *L. bicolor* are functional, transcript profiling was carried out using the custom NimbleGen oligoarray. The eight full-length STE3-like receptor genes were all expressed in free-living mycelium, ectomycorrhizal root tips and fruiting bodies (Table S4). But only *LbSTE3.1, LbSTE3.2* and *LbSTE3.6* genes were strongly repressed, as were the three mating type pheromone genes and the transcription factor genes *Lba1* and *Lba2* in the fruiting body of the S238N strain (Table S4). This finding is interesting because the fruiting body is the tissue in which changes from a dikaryotic stage to a diploid stage occurs.

Recombination within and around A and B mating type loci

Analysis by allele-specific PCR test of *A* and *B* allele segregation in S238N progeny showed that the two mating...
type loci are unlinked. Moreover, when the two loci are mapped on linkage group J in the *L. bicolor* genetic map of Doudrick et al. (1995), their position corresponds to that expected for the mating type loci (J. Labbé et al., unpublished). These observations confirmed that the newly characterized *A* and *B* loci correspond to the functional *A* and *B* loci. Interestingly, the *A* and *B* mating type loci are located on chromosomes of very different size – the *A* locus on the largest chromosome and the *B* locus on one of the smallest chromosomes (Martin et al., 2008; J. Labbé et al., unpublished) – as is the case in *C. cinerea* (O’Shea et al., 1998). In *C. cinerea*, the linkage map is similar in length between the two chromosomes containing the mating type loci despite the huge difference between their physical maps (Muraguchi et al., 2003). In consequence, the organization around the *A* locus is expected to be maintained by a low recombination frequency, whereas that around the *B* locus is predicted to be much less conserved because of a high recombination frequency (Casselton & Kies, 2007).

To examine whether the *L. bicolor* *A*- and *B*-containing regions are in genomic regions with different recombination rates, we explored different genomic characteristics that are known to illustrate the recombination frequency such as the length of intergene distance or the presence of G + C rich sequences. Indeed, although recombination events would be counterselected within clusters of essential genes with short intergene distances (Pal & Hurst, 2003; Poyatos & Hurst, 2007), they will be frequent in clusters of genes with large intergene distances (Cromie et al., 2007). Moreover, in the basidiomycete *C. neoformans* frequent recombination events, detected in the border of the mating type locus, are associated with regions with a high G + C content (Hsueh et al., 2006).

To address this issue, we first determined the median intergenic distance. It was found that the intergene spacers had different sizes between the two regions of the genome. The distance between genes was small in the 450 kb surrounding the *A* locus (median = 229 bp; Supporting Information Fig. S5a). By contrast, it was longer in the 350 kb surrounding the *B* locus (median = 1.4 kb; Supporting Information Fig. S5c). Notably, the intergenic distances within the *B* locus (*LbSTE3.1, PbB1, LbSTE3.2, PbB2, LbSTE3.3, PbB3*) were relatively short, those within *B* subloci being shorter than those between *B* subloci (464 bp and 819 bp, respectively; Fig. S5c). This difference in the intergene length between the two regions containing the *MAT* loci suggests differences in recombination rates between these two regions, the *A* locus being in region subject to lower rates of recombination.

Second, we monitored the presence of potential recombination hotspots by determining the G + C content in the two genomic regions with a 4-kb sliding window. No G + C-rich sequences were detected in the proximity of the *A* locus, the closest one being situated 250 kb downstream of the *A* locus (Fig. S5b). However, no G + C peak was detected surrounding the *B* locus either (Fig. S5d). These features would suggest an absence of recombination and, in particular, of hotspot recombination in the proximity of both loci.

Recombination is known to be suppressed within the mating type loci of different basidiomycetes (Fraser et al., 2007) or within the regions encoding the protein domains mediating self/nonself recognition (Badrane & May, 1999). When recombination occurs a local enrichment in G and C can often be observed (Duret et al., 2006). Therefore, we calculated the G + C content in the two *MAT* loci with a small sliding window of 100 bp. Short G + C rich sequences (over 60%) were detected within both loci: in the *B* locus upstream of the *LbSTE3.1, LbSTE3.2 and LbSTE3.3* genes, and in the *A* locus in the 3′ end of gene *Lba2* (Supporting Information, Fig. S6). The authenticity of these recombination sites was tested by an allele-specific PCR test for mating type genes. This test monitored the recombination events occurring at the *MAT* loci of the 42 S238N sib-monoaryktotic progeny between *LbSTE3.1, LbSTE3.2 and LbSTE3.3* genes as well as between and within genes *Lba1* and *Lba2* (Fig. 1, Table 1). Surprisingly, in most of the recombining individuals only one of the two genotypes, expected after a reciprocal exchange, is observed, suggesting that these products are an outcome of a nonreciprocal exchange (gene conversion). The excess of one parental genotype to the other infers the same hypothesis. A high number of recombination events were detected in the 3′ end of the *Lba2* gene and upstream of the *LbSTE3.2 and LbSTE3.3* genes, as predicted by the high G + C content (Table 1). Recombination events occurring upstream of the *LbSTE3.1* gene could not be detected by the PCR test. No crossing-over events were detected in the *A* locus, but one was detected in the *B* locus between the first and second *B* subloci.

**Gene order conservation**

Because gene pairs with short intergene spacers are less likely to have been rearranged during evolution (Poyatos & Hurst, 2007) we expected a higher conservation of the gene order in the *A* mating type containing region than in the *B* mating type containing region, with a conservation of the gene order restricted to the *B* locus itself. This prediction was confirmed by genomic comparison between *L. bicolor* and *C. cinerea*, the closest species to *L. bicolor* with a fully sequenced genome. The syntenic region containing the *A* locus spans more than 350 kb (331 kb in *L. bicolor* and 363 kb in *C. cinerea*) with 119 genes being homologues between the two species (Fig. 5; 36–80% of sequence identity depending on the genes with the exceptions of *HD1* and *HD2* genes that present only 20% and 15% of sequence identity, respectively). Just two insertions of one and three genes and one deletion of one gene (*FMNOR*) are observed in this part of the *L. bicolor* genome compared with the *C. cinerea* genome. The shared gene order is also relatively conserved at a longer phylogenetic distance, such as with *P. djemor* and *P. chrysosporium* (Fig. 5). The inversion events that disturbed the synteny in this region
during evolution are clearly visible. It is interesting to note that all the genes in this conserved region are single copies in the fungal genomes and that several of them, including \textit{pad1} and \textit{ade8}, are critical for fungal fitness in nature (Fig. 5; Kües et al., 2001; James et al., 2004).

In contrast to the \textit{A} locus-containing region, little gene order conservation was observed around the \textit{B} locus between \textit{L. bicolor} and the other Agaricomycotina (Supporting Information, Fig. S7).

Two STE3-like pheromone receptor genes, \textit{LbSTE3.4} and \textit{LbSTE3.6}, are located at the proximity of \textit{LkSTE3.1}, \textit{LkSTE3.2} and \textit{LkSTE3.3} in the \textit{L. bicolor} genome. Similarly, their homologues in \textit{C. cinerea}, \textit{CcSTE3-2151} and \textit{CcSTE3-2153} are located 38 kb downstream of the \textit{C. cinerea} \textit{B} locus. One gene, \textit{UPB2}, coding for a conserved protein of unknown function, is found in \textit{C. cinerea} and \textit{P. djamor} in close vicinity to the \textit{B} mating type genes and is still linked to the \textit{L. bicolor} \textit{B} locus but it is highly duplicated upstream to the pheromone receptor genes of the \textit{L. bicolor} \textit{B} locus (Fig. S7). Two genes linked to the \textit{B} locus in \textit{P. djamor} and \textit{C. cinerea} (\textit{RPS19} and \textit{CLA4}) were not found in \textit{L. bicolor} in the region surrounding the \textit{B} locus, but in another scaffold (scaffold 5) which also contains a gene coding for a nonmating-specific STE3-like pheromone receptor (\textit{LhST3.9}) at a distance of 424 kb. By contrast, two genes located 5' (2083 and 2084) and two genes located 3' (2089 and 2114) to genes \textit{RPS19} and \textit{CLA4} in \textit{C. cinerea} were also found 240 kb upstream of the \textit{L. bicolor} \textit{B} locus (Fig. S7). Interestingly, in addition to gene duplication and translocation events, the surrounding of the \textit{L. bicolor} \textit{B} locus showed an accumulation of transposable elements (Fig. S7). These observations suggest high rearrangement rates in the genomic regions surrounding the \textit{B} locus of \textit{L. bicolor}.

### Discussion

We have investigated the genetics underlying of the tetrapolar mating system of the ectomycorrhizal fungus \textit{L. bicolor}. Using genome comparison, molecular phylogeny, allele-specific PCR tests and gene expression, we demonstrated that the mating specificity in this species is encoded by the two mating type loci known in Agaricomycotina, the \textit{A} mating type locus (homeodomain transcription factor genes) and the \textit{B} mating type locus (pheromones/pheromone receptor genes). The two loci are located in two unlinked genomic regions and have a very different evolutionary history.

### Conservation of mating type gene functions

Comparisons of gene structures between distinct lineages have been instrumental in understanding the conservation of structural motifs necessary for protein function and, further, in developing predictive algorithms that relate secondary structure and protein function (Lupas et al., 1991). We used the sequence variation at the \textit{A} and \textit{B} mating type loci to test the correlation between known function and sequence conservation. We showed that, despite more than 100 million years of divergence (Berkhe & Taylor, 2001), the \textit{L. bicolor} and \textit{C. cinerea} mating type proteins share sequence conservation for all \textit{C. cinerea} functional domains. The \textit{Lba2} protein carries the classical homeodomain that binds target DNA sequences. The \textit{Lba1} protein carries the atypical homeodomain, two nuclear localization domains, as well as at least two of the functional transactivation (AD1 and AD2) domains described in \textit{C. cinerea} HD1 proteins AD1 and AD2 (Kües et al., 1994; Asante-Owusu et al., 1996; Badrane & May, 1999). The

<table>
<thead>
<tr>
<th>Locus</th>
<th>Genotype</th>
<th>Number observed</th>
<th>Observed frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A2C.1 A2N.1 A1N.1 A1C.1</td>
<td>8</td>
<td>20.5%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.2 A2N.2 A1N.2 A1C.2</td>
<td>12</td>
<td>30.8%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.1 A2N.1 A1N.1 A1C.2</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.2 A2N.2 A1N.2 A1C.1</td>
<td>2</td>
<td>5.1%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.1 A2N.2 A1N.2 A1C.2</td>
<td>13</td>
<td>33.3%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.2 A2N.1 A1N.1 A1C.1</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.1 A2N.2 A1N.2 A1C.1</td>
<td>1</td>
<td>2.6%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.2 A2N.1 A1N.1 A1C.2</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.1 A2N.2 A1N.1 A1C.1</td>
<td>3</td>
<td>7.7%</td>
</tr>
<tr>
<td>A</td>
<td>A2C.2 A2N.1 A1N.2 A1C.2</td>
<td>0</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

The parental alleles are noted 1 (H82 monokaryon) and 2 (in bold type) for each locus. N refers to the 5' region of the gene and 'C' to the 3' region of the gene.
position of dimerization domains is also conserved between \textit{L. bicolor} and \textit{C. cinerea} in HD1 and HD2 proteins. Conservation is, however, surprising for the AD3 transactivation domain as this motif is less conserved between \textit{C. cinerea} and its closer relatives, \textit{C. scobicola} and \textit{C. disseminatus}, than between \textit{C. cinerea} and \textit{L. bicolor} (Fig. S2). Moreover, this domain was not found to be functionally essential in \textit{C. cinerea} HD1 proteins (Kües \textit{et al.}, 1994; Asante-Owusu \textit{et al.}, 1996). Site-directed mutagenesis combined with biochemical assays will be necessary to test if this domain provides a specific function to \textit{Lba1} alleles.

In the \textit{B} locus, STE3-like pheromone receptor genes encode for a seven transmembrane protein which couple to heterotrimeric guanine nucleotide-binding proteins (\(\alpha\), \(\beta\), and \(\gamma\)GTPase proteins) to effect intracellular signalling (Casselton \& Olesnicki, 1998). Despite the dissimilarity in primary sequence, the \textit{L. bicolor} STE3-like pheromone receptor genes have a common tertiary structure with those of the other Agaricales, namely, a short \(N\)-terminal extracellular domain, seven hydrophobic \(\alpha\)-helices (the transmembrane domains), three extracellular and three intracellular loops, a \(C\)-terminal intracellular tail (Supporting Information, Fig. S8). The particularity of \textit{L. bicolor} receptors lies in the relatively short \(C\)-terminal tail compared with those of \textit{C. cinerea} or of those of the ascomycete \textit{Saccharomyces cerevisiae}. Such a short terminal tail means that an interaction between the terminal tail and the GTPase \(\alpha\) subunit is unlikely. In the \textit{S. cerevisiae} STE3 receptor, the second and third intracellular loops of the protein were predicted to interact with the GTPase \(\alpha\) subunit (reviewed by Bourne, 1997) and the first and third extracellular loops with the mating type specific pheromone (Sen \& Marsh, 1994). If these functions are conserved in Agaricales, a higher conservation of the domain interacting with the GTPase protein (the second and third intracellular loop of STE3-like receptor) compared with those interacting with pheromones (first and third extracellular loops of STE3-like receptor) is expected, each pheromone and its respective receptor being engaged in long-term coevolution. Surprisingly, the first and third extracellular loops are more conserved than the second intracellular loop between all mating type specific and
nominating type specific STE3-like paralogues. No conservation was observed for the third intracellular loop. These features make it difficult to predict which domains interact with the mating type pheromones. However, they suggest a functional innovation for the STE3-like receptor in Agaricales: the capacity to interact by their extracellular domains with proteins other than the mating type pheromones, for example, the conserved nonmating pheromones whose genes were detected in this study in the L. bicolor genome. The sequence variability at the internal domain could reflect receptor-specific interactions with different classes of GTPase α proteins. Indeed, eight classes of GTPase α genes were found in the L. bicolor genome, four being specific to Agaricales (Martin et al., 2008; S. Duplesis et al., unpublished).

Evolution of mating type loci

The molecular characterization of the L. bicolor A locus highlights that this locus encodes only one pair of homeo-domain transcription factors as in the bipolar P. chrysosporium and the tetrapolar species P. djamor. This distinguishes L. bicolor from C. cinerea, which encodes three functional pairs of homeodomain transcription factors. Interestingly, the number of A alleles observed in C. cinerea populations (124–164 mating-types; Raper, 1966; May & Matzke, 1995) is three times more than in L. bicolor (c. 45 mating-types; Raffle et al., 1995) or P. djamor (c. 58 mating-types; James et al., 2004) populations. This suggests similar rates of evolution of each A sublocus within species under the assumption of a low recombination rate between the A subloci in C. cinerea. The low recombination rate and the small number of rearrangement events observed in the region containing the A locus would explain the homogeneity in the evolutionary rate of the A locus in Agaricales. Situated in a low recombining region with a very low transposable element density, the divergence in the self-recognition domains between the products of the A locus alleles is favoured, the multiple alleles of this region being maintained in the population by balancing selection (May et al., 1999). By contrast, the DNA-binding domains of the HD2 mating type proteins are essential and should not be able to tolerate the accumulation of deleterious mutations. Biased gene conversion has to purge such mutations, as shown in C. cinerea (Badrane & May, 1999) and as suggested by the present study.

The organization of the B locus is more complex in L. bicolor than the organization at the A locus. Three tightly linked subloci, born from tandem duplication of pheromone/receptor units, were identified. These two duplication events have been traced by phylogeny from the common ancestor to C. cinerea and L. bicolor. The B mating type genes of L. bicolor encode proteins that contain structural domains (seven transmembrane domains for the STE3-receptor) and self-recognition domains. Balancing selection is expected to maintain linked allelic diversity between a given class of receptors and a given class of pheromones. Indeed, short intergene spacers counterelect recombination events within subloci. Unexpectedly, the three B subloci must have remained linked for a long time since the general gene order is conserved between L. bicolor and C. cinerea. This suggests that these three genes are functionally important or that recombination was suppressed between B subloci. Nevertheless, in C. cinerea each one of these subloci is sufficient for mating (Riquelme et al., 2005). The conservation of two other STE3-like receptor genes, LbSTE3.4 and LbSTE3.6, in the proximity of the B locus and the similarity of the LbSTE3.6 expression profile to that of the mating type genes in L. bicolor open the discussion on the existence of mating type STE3-like receptors other than those already characterized in C. cinerea.

Unlike the A locus, the B locus is surrounded by genes with relatively long intergene distance. This feature allows sequence translocation, gene duplication and transposon insertion (Fig. S7) that has caused divergences among related species. The rapid evolution of the region surrounding the B locus would favour a species-specific rate of B locus diversification (c. 24 mating-types in L. bicolor, Raffle et al., 1995; c. 79 mating-types in C. cinerea, Riquelme et al., 2005; c. 231 mating-types in P. djamor, James et al., 2004).

It is interesting to note the striking difference between evolution at the A and B loci. Despite typical features characteristic of sex-determining genes (fast-evolving genes, suppression of recombination, punctual gene conversion to protect the functional domains against erosion), the genomic regions containing the two mating type loci seem to have different evolutionary histories. The A locus is in a region where the gene order is under strong selection across the Agaricales. Consequently, the genes linked to the A locus evolve at a constant rate in each species. By contrast, the B locus is in a region where the gene order is likely to be under low selection pressure but where gene duplication, translocation and transposon insertion are frequent. Consequently, the genes surrounding the B locus evolve at a species-specific rate, depending on the recombination rate variation.

Understanding the differences in mating type locus evolution is important for understanding the theoretical basis for sex chromosome formation. Indeed, a widely accepted scheme for the evolution of Y chromosomes in the XX/XY system consists of three major events: acquisition of the sex-determining loci, suppression of recombination, and genetic degeneration driven by evolutionary processes such as Muller’s ratchet (Charlesworth, 1978; Charlesworth & Charlesworth, 2000). However, the evolution of sex chromosomes of a haploid system was predicted to be different from that in a diploid system (Bull, 1983). In a haploid organism, degeneration should not occur, because it would impair essential genes. In L. bicolor, further investigations consisting of a full comparative genomic analysis of the mating type containing regions from compatible monokaryons are needed to determine whether gene loss and degeneration occur in the mating type containing regions of
mating compatible partners. Nevertheless, the frequent gene duplication and inversion events as well as the accumulation of transposable elements combined with gene degenerations observed in the B mating type containing region in the present study show that the B-containing region evolves faster than the A-containing region. These events resemble those that shaped the Cryptococcus mating type region (Frase et al., 2004) and those that shaped the sex chromosomes in animals and plants (Frase & Heitman, 2004).

Acknowledgements

We would like to thank Timothy James and the three anonymous referees for critical comments on an earlier version of this paper. We acknowledge the Joint Genome Institute and the Laccaria Genome Consortium for access to the L. bicolor genome sequence before publication and the Broad Institute for releasing the C. cinerea genomic data to the public. This work was supported by the Swiss National Science Foundation (SNSF Grant No. 3100A0-105790/1), which is gratefully acknowledged. The transcript profiling was funded by INRA and the Network of Excellence EVOLTREE. UK was funded by the Deutsche Bundesstiftung Umwelt (DBU).

References


Carpophilus disseminatus from its tetrapolar ancestors involves loss of mating-type-specific pheromone


Supporting Information

Additional supporting information may be found in the online version of this article.

Text S1 (Complements the Materials and Methods section): Study species; Complement to MAT loci annotation section; Accession numbers of the sequences used in the phylogenetic reconstruction; Construction and sequencing of the dikaryon S238N cDNA library.

Fig. S1 Phylogeny of basidiomycete species compared with Lactaria bicolor in the present study.

Fig. S2 Amino acid alignment of HD1 mating type proteins.
Fig. S3  Amino acid alignment of HD2 mating type proteins.

Fig. S4  Amino acid alignment of pheromone-like and mating type specific pheromone sequences encoding CaaX motifs at their C-terminus in Laccaria bicolor and Coprinopsis cinerea.

Fig. S5  The intergene distances have different size between the genomic regions containing the A and B loci of Laccaria bicolor.

Fig. S6  The G + C content variation in the A and B mating type loci of Laccaria bicolor.

Fig. S7  Schematic comparison of the region surrounding the B locus from Laccaria bicolor, Coprinopsis cinerea, Pleurotus djamor and Phanerochaete chrysosporium.

Fig. S8  Amino acid alignment of STE3-like pheromone receptor sequences.

Table S1  Polymerase chain reaction (PCR) primers used to survey recombination events in F₁ progeny of S238N

Table S2  Distribution of STE3-like pheromone receptor genes in Laccaria bicolor genome

Table S3  Distribution of pheromone-like genes in L. bicolor genome

Table S4  Changes in the expression of transcripts coding for predicted mating-type genes in free-living mycelium and Poplar and Douglas-fir ectomycorrhizal root compared to fruiting bodies

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the New Phytologist Central Office.

About New Phytologist

- New Phytologist is owned by a non-profit-making charitable trust dedicated to the promotion of plant science, facilitating projects from symposia to open access for our Tansley reviews. Complete information is available at www.newphytologist.org.

- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication ‘as-ready’ via Early View – our average submission to decision time is just 29 days. Online-only colour is free, and essential print colour costs will be met if necessary. We also provide 25 offprints as well as a PDF for each article.

- For online summaries and ToC alerts, go to the website and click on ‘Journal online’. You can take out a personal subscription to the journal for a fraction of the institutional price. Rates start at £135 in Europe/$251 in the USA & Canada for the online edition (click on ‘Subscribe’ at the website).

- If you have any questions, do get in touch with Central Office (newphytol@lancaster.ac.uk; tel +44 1524 594691) or, for a local contact in North America, the US Office (newphytol@ornl.gov; tel +1 865 576 5261).