Origin of laccase gene structural diversity in edible mushrooms

Christophe Billette, Thierry Gibard, Marie Foulongne-Oriol, Jean-Michel Savoie

ORIGIN OF LACCASE GENE STRUCTURAL DIVERSITY IN EDIBLE MUSHROOMS

CHRISTOPHE BILLETTE, THIERRY GIBARD, MARIE FOULONGNE-ORIOL, JEAN-MICHEL SAVOIE
INRA, UR1264, Mycologie et Sécurité des Aliments,
BP 81, F-33883 Villenave d'Ornon,
France
billette@bordeaux.inra.fr

ABSTRACT

Laccase genes have been found in fungi, plants, insects and bacteria. In Basidiomycetes, the number of laccase genes ranges from 0 to 17. The role of these genes is not well known. It seems to be important in fungal interaction, development, melanine synthesis, human and plant pathogenesis, [ectomycorrhizal association and nutrition of the fungi. Their role as lignin-modifying enzymes is controversial. Laccase phylogeny already published is not congruent with species phylogeny. Phylogeny of gene is generally deduced from analyse of aligned sequences of nucleotides or amino acids. Presence of indels, position of introns can give complementary information. To better understand the origin and evolution of Basidiomycete laccase genes, we studied the structure of diverse multi copper oxidase genes in different phylum of Eukaryotes, in basidomycetes and in *Agaricus bisporus*. A mutation has been conserved in all basidiomycete laccases and constitutes a molecular synapomorphy for basidiomycete laccase genes. At least three main laccase subfamilies (A, B and C) are present in Polyporales and Agaricales. In *A. bisporus* the 4 new laccase genes we have sequenced (lcc3-lcc6) belong to a new sub-family for this species. Cladogram of laccase genes from edible mushrooms built with Dollo parsimony of intron position is congruent with dendrograms built with homologous amino acid sequences, but is not identical. Some consequences on the evolution of laccase genes and their diversity are discussed.

Keywords: Laccase phylogeny, laccase sub-families, intron position, *Agaricus bisporus*, basidiomycetes, protein evolution.

INTRODUCTION

Laccase are polyphenol oxydase belonging to two protein families that are involved in various systems of interactions between fungi and their biotic and abiotic environment [1-4]. The multi copper oxidase family (mco) is the large one and contains also other enzymes as ferroxidase, ascorbate oxidase, ferroxidase-laccase [1]. The second family, recently discovered, contains proteins with the DUF152 domain, present in a hypothetical Bacteroides of the bovine rumen and in *Eschericia coli* [2].

Laccases of scomycetes and basidiomycetes form two clades belonging to the mco family. Laccases have the ability to oxidize a wide range of substrates [6]. Some laccases play a role in sclerotization (insects), in lignification process (LMCO of plantes) but there roles in basidiomycetes are not well elucidated (see Kues and Ruhl for review [1]). During a long time, an important role of basidiomycete laccases has been attributed to delignification, but it is also contested [7]. Most of fungal laccases are secreted and some of them seem to play a role in recycling recalcitrant organic matter and in the detoxification of natural environments [8]. Some are involved in ectomycorrhizal association [4], in plant pathogenesis [6] or in fungal-fungal

interaction [2, 9]. For instance, to improve *A. bisporus* strains for their resistance to *Leucanicillium fungicola*, one possible track is to better understand the mechanism of interaction of the two fungi. Thus, laccase genes appeared to be good candidates. At the beginning of this work, two laccase genes, lcc1 and lcc2, were known [10]. We have discovered and annotated 4 new laccase genes in *A. bisporus*.

Intron positions are not suitable information for phylogenetic analysis at long evolution distances [11]. Nevertheless, Kilaru *et al.* [3] on *Coprinopsis cinerea* laccases, Tavares *et al.* [12] and Boulet *et al.* [13] on *Arabidopsis, Caenorhabditis and Drosophila* and Matheny *et al.* [14] for basidiomycota phylogeny, showed that they could be very useful in phylogeny in complement of analyses based on protein or nucleotide sequences. Moreover, the intron positions could help to elucidate general mechanisms of gene family evolution. Indeed intron presence/absence is a relatively very slowly evolving character [15] and selective forces acting on it differ to those acting on protein sequences.

The study of intron positions and protein structure would help to go further in the origin and evolution of basidiomycete laccases. That also may help us to deduce putative laccase functions. We supposed that mco family is a very old family that has experienced several intron gain and loss waves. We also supposed that no parallel intron insertion has occurred in laccase genes. In the present work, we compared the phylogeny of basidiomycetes species using laccase genes, based on protein sequence or on intron positions. We discovered insertion of 5 amino acids present in all basidiomycetes laccases but absent in all other mco. We also discovered a *Lentinula edodes* gene with presence or absence of three introns depending of the strain. Consequences for the evolution of laccase genes and their diversity are discussed.

MATERIALS AND METHODS

Sequences used. Sequences used for phylogeny analyses (Table 1) were obtained from GenBank (http://www.ncbi.nlm.nih.gov/), Joint Genome Institute (http://genome.jgi-psf.org/programs/fungi/index.jsf) and Broad Institute (http://www.broadinstitute.org/scientific-community/data). For phylogenetic analyses, we had chosen all published genomic DNA Laccase sequences to have intron positions. The translated coding sequences of these genes were used for the amino acid phylogeny.

Two Polyporus laccase genes and three different *C. cinerea* laccase genes (two belonging from two different well defined clades of subfamily A and lcc8) were used as controls in the analyses. Three ferroxidase gene sequences have been used as out group.

Identification of new *A. bisporus* genes. *Agaricus bisporus* strain U1-7 (having the same mating type allele than H93) is a homocaryotic strain coming from a protoplast of the cultivar U1 [16, 17]. DNA extraction was made with the Nucleon Phytopure extraction kit (GE Healthcare) from lyophilised mycelium. The *lcc2* cDNA sequence [10] was used to generate primers and to amplify and sequence the gene. Inverse PCR [18] was used to obtain the 3' end of the gene. One (or more) restriction enzyme was chosen for inverse PCR, cutting the known sequence. Primers were constructed with Primer 3 on the 3' side of the restriction site, present on the already known sequence, to orientate inverse PCR. When AK7/AK8 primers from D'Souza *et al.* [19] were used to amplify *lcc1* and *lcc2*, two bands were obtained in electrophoresis. *Lcc1* and *lcc2* corresponded to the same band. So, the second band was expected to be another laccase gene (renamed hereafter as *lcc3*). This bans had been sequenced. To obtain a larger part of the gene, a new degenerated primer (AK9 : TGRCARTGGARGAACCAKGG) was constructed on part of a simplified l4 signature of laccase sequences (P-W-F-(LF)-H-C-H). The totality of the gene was obtained by inverse PCR in both 3' and 5' directions.

Table 1 : Gene list, intron numbers and GenBank entries of fungal laccases and ferroxidases used in the two phylogenetic analyses.

Enzymes	Code	clades	Organism	gDNA accession number and region	Protein accession number	Gene name	Number of introns
Laccases		Agaricales					
	Abi	Agaricoid clade (VI)	*Agaricus bisporus*	L10664	AAC18877	lcc1	14
				FJ838791	ACZ06558	lcc2	14
				EU580106	ACE73659	lcc3	12
				FJ392313	ACR19861	lcc4	12
				FJ617019	ACU52699	lcc5	12
				FJ878811	ACZ57763	lcc6	12
	Lbi		*Laccaria bicolor*	NW_001889910 <486085..>488315	XP_001886681	lcc1	13
				NW_001889940 <72836..>75100	XP_001889567	lcc2	13
				NW_001889873 complement(109817.. 112096)	XP_001874989	lcc3	13
				NW_001889884 194127..196383	ACN49091	lcc4	13
				DS547137 <116239..>118647	EDR01587	lcc5	14
				NW_001889885 complement(429758.. 431860)	XP_001881925	lcc6	10
				NW_001889885 complement(432396.. 434513)	XP_001881926	lcc7	10
				NW_001889885 complement(435301.. 437493)	XP_001881927	lcc8	10
				NW_001889938 <142966..145306	XP_001889429	lcc9	13
	Cci		*Coprinopsis cinerea*	BK004113	DAA04508	lcc3	13
				BK004118	DAA04513	lcc8	8
				BK004117	DAA04512	lcc7	12
				BK004126	DAA04521	lcc16	18
				BK004127	DAA04522	lcc17	19
	Pos	Pluteoid clade (II)	*Pleurotus ostreatus*	AJ344434	CAC69853	poxa3	21
				Z22591	CAA80305	pox1	19
				Z49075	Q12739	pox2	19
				FM202669	CAR48257	pox3	10
				FM202670	CAR48258	pox4	19
				AJ005017	CAA06291	poxA1b	15
				FM202671	PSEUDO: CAR48259	pox5	18
	Led	Marasmioid clade (IV)	*Lentinula edodes*	AB055157	BAB84354	lcc1	27
				AB055158	BAB84355	lcc2=lac1	13
				FJ473386	ACR24356	=lac1 without 3 introns	10
				AB055159	BAB84356	lcc3	11
				AB543788	BAJ12091	lcc5	20
				AB543787	BAJ12090	lcc6	27
				-	BAH80447	lcc4	nd
	Fve		Flammulina velutipes	AY485826	AAR82931	laccase	9
	Pru	Polyporales	*Panus rudis*	AY839935	AAW28932	lacA	12
	Rmi		*Rigidoporus microporus (Fomes lignosus)*	-	CAE81289	lcc1	nd
Ferroxidases	Lbi		*Laccaria bicolor*	NW_001889910 complement(<222411.. >224748)	XP_001886726	lcc11	9
	Apo		*Auricularia polytricha*	AY616035	AAT73205	lac1	13
	sce		*Saccharomyces cerevisiae*	L25090	AAA64929	Fet3	0

Two clones sequenced from a genomic library available in our lab corresponded to partial laccase genes sequences. By inverse PCR we obtained the totality of these new genes (*lcc4* and *lcc5*). *Lcc6* was obtained from an extra band during one of these inverse PCRs.

Amino acid phylogeny. The phylogenetic analysis of laccase sequences was performed on the Phylogeny website platform (www.phylogeny.fr). Sequences were aligned with MUSCLE (v3.7) with default settings. After alignment, ambiguous regions (i.e. containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b) with default settings. The phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.0 aLRT) with default settings. Reliability for internal branch was assessed using the bootstrapping method (500 bootstrap replicates). Graphical representations of the phylogenetic tree were performed with TreeDyn (v198.3).

Intron position phylogeny. In GenBank, intron positions are indicated on gDNA sequences and protein sequence is deduced by translating CDS. The positions of introns of the encoding gene were determined on aligned protein sequences by an annotation on a BioEdit file [20]. The convention presented in Fig. 1 for schematizing the exact position of the intron on each codon.

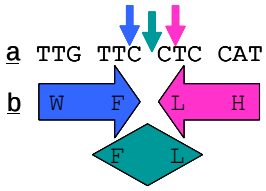


Figure 1: Convention used on BioEdit (version 7.0.5.3) to schematize intron position in a gene on the corresponding protein sequence.
a DNA sequence with three arrows indicating three possible positions of introns in the codons, b corresponding protein sequence with arrows and diamond overlapping two amino acid letters to indicate intron position in the encoding gene.

Dollo phylogeny. A matrix was built with all intron positions in all sequences. Paup was used to perform the Dollo parsimony analysis [21]. One thousand bootstrap replicates were used. We made the approximation that the ancestor gene had no intron at all.

Analysis of insertions in laccase genes. T-Coffee (Phylogeny.fr) was used to align sequences. The alignment were visualised with Jalview tool.

RESULTS AND DISCUSSION

Obtention of *lcc2* gDNA sequence and identification of four new laccase genes in *A. bisporus*. The complete genomic sequences of *lcc2*, *lcc3*, *lcc4* and *lcc5* and the partial sequence of *lcc6* have been obtained (GenBank accession number : FJ838791, EU580106, FJ392313, FJ617019 and FJ878811 respectively). It is not surprising to find at least six laccase genes in this species as seventeen have been found in *C. cinerea* [3] which is in the same Agaricoid clade.

The sequences of the six laccase genes of *A. bisporus* were compared between each other. Intron positions in lcc2 were the same as in lcc1 except for the 3' UTR intron 15 which has not been found in lcc2. The 3' end of lcc4 is still hypothetical, an intron could permit the prolongation of the protein (lcc4bis end: CPIWDSEPNFVKHAATMILDPLINFAFGPIFPVYIL

LRFPTSLICMKPNAFH). Sequencing the mRNA will be necessary to control the end of the protein. The study of intron positions of these genes revealed the presence of a new laccase sub-family in *A. bisporus* (Fig. 2). This has led us to compare the intron positions of numerous laccase genes and especially from *C. cinerea* as two laccase sub-families had been found in this species [3]. To better understand the phylogenetic relationships of these genes we have build and compared a phylogeny based on amino acid homologous sequences and a phylogeny based on the conservation of intron positions.

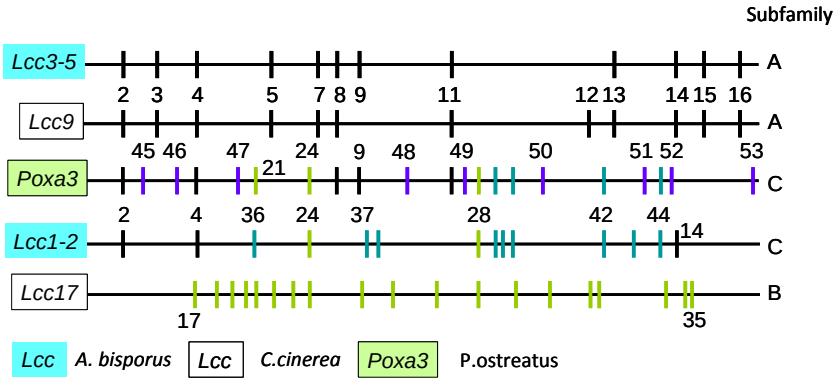


Figure 2: Relative positions of introns in different laccases genes showing the presence of two sub-families in Agaricus bisporus. Nomenclature of introns 2 to 35 [3], 36 to 44 (*A. bisporus lcc1* and *lcc2*), 45 to 53 (*P. ostreatus poxa3*).

Congruence but not identity between protein sequences and intron positions based phylogeny. Fig. 3A shows the maximum likelihood phylogeny of 38 protein sequences of laccases from edible mushrooms. The ferroxidase genes added as outgroup represented a clade. A second clade was composed by laccases of the subfamily B identified by Kilaru [3]. In a third clade (renamed hereafter as "clade C"), there were the large subunit of heterodimeric proteins encoded by *A. bisporus lcc1*, *lcc2* and *P.ostreatus poxa3* cluster. This clade wais confirmed by the analysis of intron position (Fig. 3B). We propose to name the proteins belonging to clade C "laccase subfamily C". In a fourth clade ("clade A"), we found the three control proteins from *C. cinerea* belonging to subfamily A [3]. The proteins contained in clade A were renamed "laccase subfamily A". Clade A corresponds to subfamily 1 in Kues and Ruhl [1], clade B and C to subfamily 2. Here clade A and C clustered, with good bootstrap values. The two most divergent proteins found in this last clade were from Polyporales. Even if the bootstrap values were very low, it was the first time that Polyporales laccases were placed at a basal position of the clade A in a phylogenic tree of laccases belonging to Polyporales and Agaricales. With this positioning, laccase phylogeny is congruent with the phylogeny of species and the origin of laccase subfamily A should be anterior to the divergence of Polyporales and Agaricales.

In this phylogenetic tree *A. bisporus* laccases are divided into two clades, A and C. All the new *A. bisporus* laccases belong to a subfamily unknown before in this species. *Agaricus bisporus* laccases of clade A formed a specific clade. This suggests that these four genes appeared by duplication events after the divergence between the ancestor of *A. bisporus* and the ancestor of the other species of Agaricoid clade VI (ie. *C. cinerea* and *L. bicolor*). These laccases clustered, but with very low bootstrap, with *C. cinerea* lcc7 encoded protein. *Pleurotus ostreatus* laccases were divided into three groups, two in clade A and one in clade C. *Laccaria bicolor* laccases were divided into two groups of clade A as previously shown [4]. *Lentinula edodes* laccases were divided into three groups of clade A.

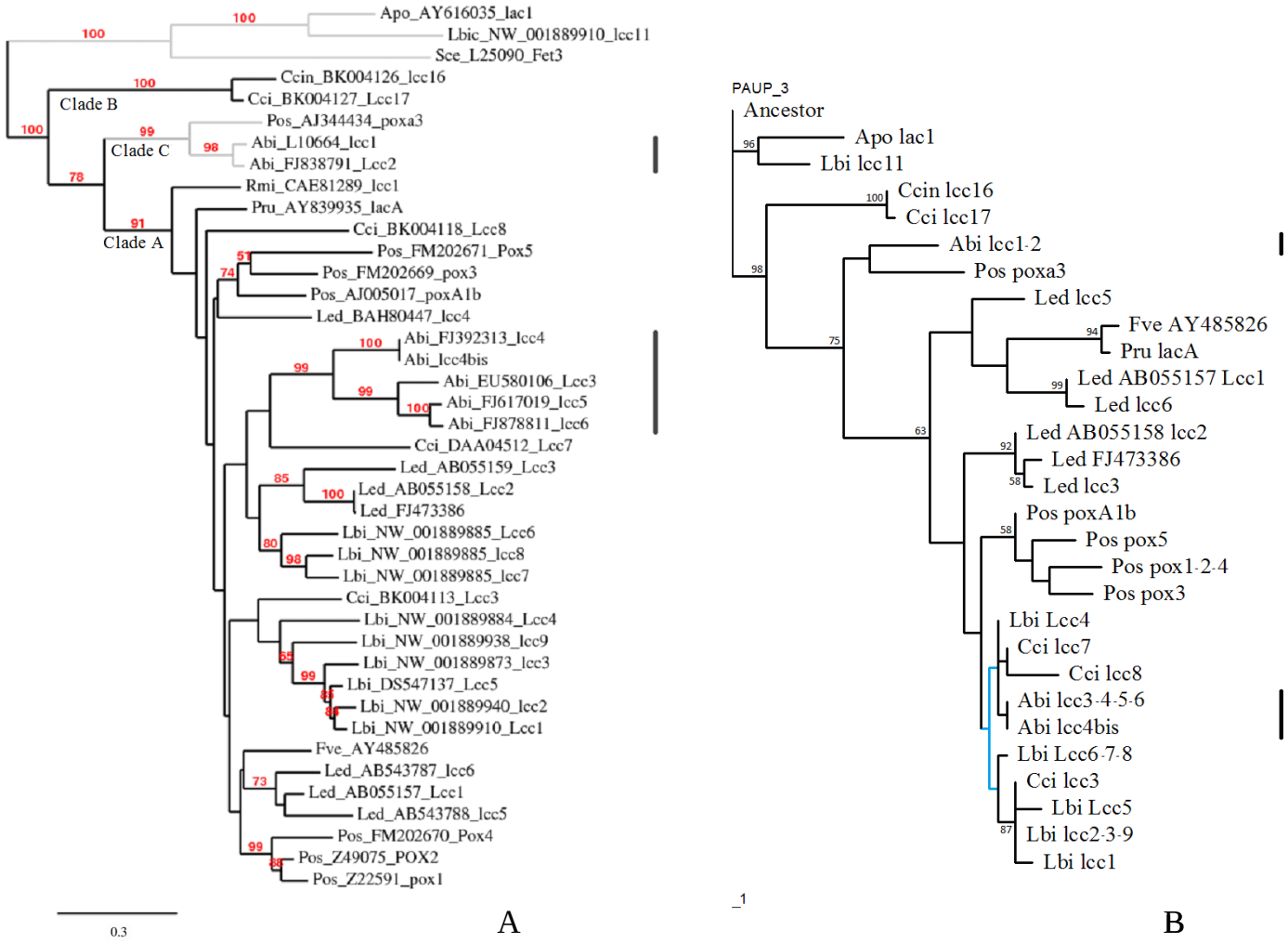


Figure 3: A- Maximum likelihood (ML) tree of edible mushroom laccases. B- One of the eight most parsimonious trees obtained with the Dollo parsimony of intron position of edible mushrooms laccase genes. Blue bare indicate Agaricoid clade VI laccase subfamily A genes.
The scale at the bottom indicates a unique event: insertion or deletion of an intron. Bootstrap values are from 500 replicates in A, 1000 in B. The vertical bars indicate the positions of *A. bisporus* laccases.
Lcc1-2 =lcc1, lcc2; pox1-2-4 = pox1, pox2, pox4. *P. ostreatus* gene nomenclature corresponds to: Pox1(LACC9), Pox2 (LACC10), Pox3 (LACC4), Pox4 (LACC1), Pox5 (LACC11), PoxA1b (LACC6) and Poxa3 (LACC2)

The bootstrap values for the three clades A, B and C wre very high, 91, 99 and 100 respectively. A laccase of clade A had more similarity to any another laccase of this clade that to a laccase of clade B or C. However Polyporales and Agaricales laccases were in clade A and only Agaricales laccases were in clade B and C. As phylogenic studies had demonstrated that Polyporales order was not nested inside Agaricales order [22], we can deduce that the separation between these three clades was anterior to the divergence between Agaricales and Polyporales. In the recent genome sequence of a Russulale, *Heterobasidion annosum* (http://genome.jgi-psf.org/Hetan2/Hetan2.home.html), a putative laccase sequence belonging to subfamily C have been found. This also indicates that the origin of subfamily C might be anterior to the divergence of Russulales and Agaricales. Wahleithner *et al.* [23] demonstrated the presence of large insertions in three of the four *Thanateforus cucumeris* laccases, suggesting the existence of other old sub-families in Cantharellales.

Fig. 3B shows one out the eight most parsimonious trees obtained with the Dollo parsimony of intron positions of the laccase genes used for the phylogenetic tree of protein

presented in Fig. 3A. Two proteins for which gDNA was not available, *R. microsporus* lcc1, *L. edodes* lcc4 and one that has no intron, *S. cereviciae* Fet3, are absent of this second tree. In the 37 genes 112 intron positions were identified. The intron-late theory seems to be more acceptable than the intron-early one supposing that there were 112 different introns in an ancestral laccase gene. Moreover, as we have studied new laccase genes and thus found new intron positions, we may expected to find other new intron positions with the future finding of new laccase genes. Stajich and *al.* [24] stated that an invasion by splicosomal introns took place at the time of fungus-animal ancestor, followed by recurrent intron loss occurring in all fungal clades. In this study we considered that the ancestor of all laccase genes had no intron at all; this constitutes for us the plesiomorphic character. Then introns had invaded laccase genes bit by bit, and then some introns disappeared after gene duplication and after speciation. This phenomenon is well analysed by the Dollo parsimony model in which convergences are excluded but reversions are admitted [25].

The phylogenetic tree of Fig. 3B is globally congruent with the one of Fig. 3A. The ferroxidase clade positioned as outgroup with a bootstrap value of 96%. These two genes share only three intron positions between each other but they share no position with laccase genes. The three principal laccase clades A, B and C, are present, with bootstrap values of 63 and 100% for clade A and B, but a low value for clade C. This was explained by the presence, in others of the eight most parsimonious trees, of clades in which *P. ostreatus poxa3* clustered with the clade A. But two intron positions are common to clade C and absent to clade A: n°42 and 44 (Fig.1). All laccase subfamily A genes from Agaricoid clade VI (*A. bisporus, C. cinerea* and *L. bicolor*) clustered together in Fig. 3A but not in Fig. 3B. Subfamily A laccases of *A. bisporus* had identical intron positions except for *lcc4* where a supplementary intron position was proposed (*lcc4bis*). These *A. bisporus* laccase genes clustered with *C. cinerea* lcc7 as in Fig. 3A but also with Cci_lcc8 which was joined here to Cci_lcc7 and with Lbi_lcc4. All these genes possess intron positions 7 and 11 (see Kilaru *et al.* [3] and Ahlawat and Billette [26] for nomenclature of introns). Here we see an important divergence between the results of the two analyses (Fig. 3) for the position of Cci_lcc8 and Lbi_lcc4, but Lbi_lcc4 clustered with Cci_ lcc7 as previously found in a Neighbour joining analysis done by Courty *et al.* [4]. This similarity reinforces the position of Lbi_lcc4 in the same clade as Cci_lcc7 and Abi_lcc3-4-5-6. Another clade with a bootstrap value of 87% was formed by Cci_lcc3 and five *L. bicolor* genes. This clade was congruent in the two analyses, except for the position of Lbi_lcc4. All the genes of this clade possess intron position 6 and 10. Positions of Lbi_lcc6-7-8 were different in the two analyses. In that case, intron position analysis was not efficient as these genes possess neither intron positions 6,7,10 and 11 which allowed distinguishing the two clades described above. We can name these two clades "the clade of Cci_lcc7" and "the clade of Cci_lcc3". *Volvariella volvacea lcc3*, studied by Ahlawat and Billette [26], also possess intron positions 6 and 10 and belong probably to the Cci_lcc3 clade. So the divergence between these two clades is probably older that Agaricoid clade VI.

Three intron positions are present in two to four *P. ostreatus* laccase subfamily A genes and are absent in all other genes. Consequently, in Fig. 3B, contrary to Fig. 3A, *P. ostreatus* laccase subfamily A genes clustered together (bootstraps value 58%) and outside the cluster formed by Agaricoid clade VI laccase subfamily A genes. This is more congruent with the phylogenetic tree of the species proposed by Matheny [27] or Binder [28] than all already published laccase phylogenetic trees. In Fig. 3B, all Marasmioid (*L. edodes* and *F. velutipes)* genes were outside the clade formed by *P. ostreatus* and Agaricoid laccase subfamily A. This is congruent with Binder's phylogeny of species [28] unlike the corresponding part of the tree in Fig. 3A and laccase trees published earlier.

Moreover *L. edodes* laccase genes were in two clades. In one of them there are *F. velutipes* and *P. rudis* (Polyporales) laccase genes. Four intron positions are shared by *P. rudis*

lac A and *F. velutipes* laccase, three of them are shared with *L. edodes*. As they are present in Agaricales and Polyporales, these positions constitute a plesiomorphic (ancestral) character for these two orders. The position of *P. rudis* lacA inside a clade of Agaricales laccase genes is a problem. This clad emight be constituted by genes that evolved very little. Te presence of few and mostly ancestral introns in Pru_lacA (12 introns) and Fve_laccase (9 introns) might have resulted in a wrong place of Pru_lacA in the phylogenetic tree, unlike in Fig. 3A. This problem is not resolved here.

Identification of a conserved, basidiomycete specifique laccase mutation. Fig. 4 shows a difference of 5 amino acid positions between aligned sequences of basidiomycete laccases and other multicopper oxidases. Very early divergent multi copper oxidases (mco) such as plants LMCO, insect laccases, archae or bacterial mco and two sub-unit mco don't possess most of these 5 amino acids positions. Two exceptions have been found among 240 mco sequences examined, a mco from the Ustilaginomycotina *Malassezia globosa* (XP_001729535) with a much bigger insertion at the same place and a mco from *Rhizopus oryzae* (RO3G_07290.3 at the Broad institute www.broadinstitute.org/annotation/genome/rhizopus_oryzae/MultiHome.html) that has the same 5 amino acid position as basidiomycete laccases, but with no apparent homology. that results probably from an independent mutation. All basidiomycetes *bona fide* laccases actually published possess these 5 amino acids positions with some substitutions and none of ascomycete laccase actually known possess them. Consequently the absence of these five amino acids positions is the plesiomorphic state of this character and the presence of these 5 amino acids positions is a synapomorphy for all *bona fide* basidiomycete laccases. Neither the exact position of the insertion nor the exact number of amino acids inserted (4 or 5) were already known.

The presence of these 5 amino acids results from an insertion of 15 bp in the common ancestor of all Basidiomycetes possessing a laccase gene *sensu stricto, i.e.* all homobasidiomycetes clades. This putative insertion, as it is present in *Thanatephorus cucumeris,* took place before the divergence between chantarelloid fungi, Auriculariales and all other homobasidiomycetes [14]. It is the first time that a structural difference is demonstrated between basidiomycete laccases and all the other laccases, and especially ascomycetes laccases. Some parts of the laccase sequences contain variable indels depending of the gene in the same species or depending of the species: e.g. the substrate binding loops. Nevertheless it's the first time that a difference is detected at the level of a whole fungal clade.

To our knowledge, these amino acid positions have never been described as important for laccase catalytic process, neither for substrate affinity. If this mutation has been conserved in all homobasidiomycetes it can result from the neutral evolution and the genetic drift or it gives a selective advantage. Which type of advantage it may give (consequence on the catalytic site, on the substrate affinity) will have to be looked for. Consequence of this mutation on the structure and function of the protein has to be elucidated. Consequence on the evolution of the protein is also interesting: has the mutation given rise to accelerate evolution in other parts of the protein? The origin and date of the mutation should be elucidated.

The consequence of this mutation, if there is one, on the lignin degradation or on the detoxification of its by-products by basidiomycetes, may give us a new explanation for the different fitness of ascomycetes and basidiomycetes. This discovery could also have an interest for biotechnological applications: e.g. to improve ascomycete laccases. Many hypotheses have now to be tested.

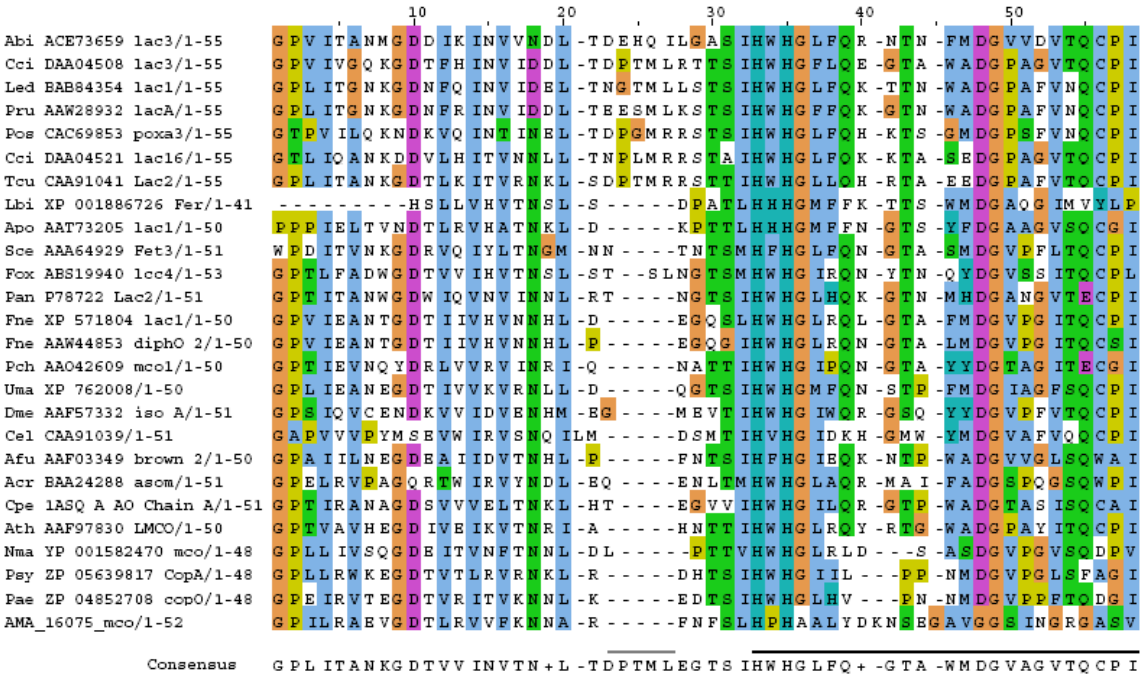
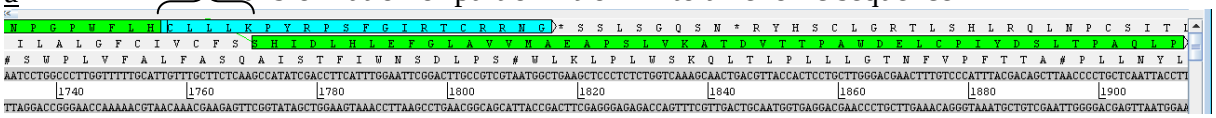


Figure 4 : Alignement of partial sequences of basidiomycete laccases and other multi copper oxidases showing the position of a putative insertion of 5 amino acids in the first ones (grey line) just before the L1 signature of laccases (black line).

Sequence names consist of species code (first letter of genus and first two letters of species name), GenBank protein accession number and protein name (when necessary, laccase is abbreviated to lac and the number used for the gene is added, ferroxidase to Fer, diphenol oxidase to diphO, ascorbate oxidase to AO, copper oxidase to copO, multi copper oxidase to mco) when it exist or gene name. For species name see Table 1 or see below. **Basidiomycetes laccases** sub-family **A** - Abi_lac3, Cci_lac3, Led_lac1, Pru_lacA, sub-family **C** - Pos_poxa3, sub-family **B** - Cci_lac16, **Cantharellales** sub-family -Tcu_lac2 (*Thanatephorus cucumeris*), **ferroxidase** - Lbic_Fer, Apo_lac1, Sce_Fet3, **ascomycete laccases** - Fox_lcc4 (*Fusarium oxysporum*), Pan_Lac2 (*Podospora anserina*) **ferroxidase** - **laccase** - Fne_lac1(*Filobasidiella neoformans*), Fne_diphO, Pch_mco1 (*Phanerochaete chrysosporium*), Uma (*Ustilago maydis*), **insect laccase** - Dme_isoA (*Drosophila melanogaster*), **nematode multi copper oxidase** Cel (*Caenorhabditis elegans*), **fungal pigment** - Afu_brown (*Aspergillus fumigatus*), **ascorbate oxidase** – Acr_asom (*Acremonium sp.*), Cpe_AO (*Cucurbita pepo*), **Plant LMCO** - Ath_LMCO (*Arabidopsis thaliana*), **Archae mco** - Nma_mco (*Nitrosopumilus maritimus*), **Bacteria copper resistance protein** – Psy_CopA (*Pseudomonas syringae*), **Bacteria copper oxidase** – Pae_copO (*Paenibacillus sp.*), **two subunit fungal mco** – AMA_mco (*Allomyces macrogynus* http://www.broadinstitute.org/annotation/genome/multicellularity_project/MultiHome.html)

First time discovery of the presence and absence of three introns in a gene in the same species. We compared some *L. edodes* laccase protein sequences from various origins and available in Genbank database (AAF13037[29], AAF13038[29], BAB83132, BAB84355, BAB83133, AAT99286, ACR24356). These sequences were initially described as different laccases. We observed that all these sequences are probably encoded by different alleles of the same gene *lac1* published first by Zhao *et al.* [29]. However, we observed some differences between the first published sequences [29] and the GenBank sequences. Moreover these GenBank releases were without intron sequences and, probably introns have not been correctly spliced during the annotation. ACR24356 has 99% identity with BAB84355 at the protein level (2 / 533 amino acid differences), but the encoding gene has 3 introns less (10 instead of 13), the three last introns (positions 14, 15 and 16 in Fig. 2). Moreover, BAB83133 don't have the three last introns but had a mutation in intron 14 that leads to disappearance of part of it and to transformation of another part to an exonic sequence, this resulted to a frame shift leading to an early stop codon (Fig 5).

a Transformation of part of intron 14 to an exonic sequence



b

NPGPWFLHCHIDLHLEFGLAVVMAEAPSLVKATDVTTPAWDELCPIYDSLTPAQLP

c

NPGPWFLHCHIDLHLEFGLAVVMAEAPSLVKATDVTTPAWDELCPIYDSLTPAQLP

Figure 5 : Alignment of 3 *Lentinula edodes* lac1 alleles showing two types of mutations affecting introns. In a is presented the BAB83133 protein sequence on the top line (NPGP … RRNG). A mutation in intron 14 leads to disappearance of part of it and to transformation of another part to an exonic sequence. The resulting frame shift stops the protein very early (blue segment). The three first lines correspond to the three reading frames obtained with the gDNA sequence present in the fourth line (Artemis software). In b ACR24356 sequence. In c, BAB84355 sequence with the position of the three introns sharing the gene coding for this sequence**.**

The conservation of the synteny of the flanking genes or the all genome sequencing might confirm that these are alleles of the same gene *lac1*. It is the first time, to our knowledge, that the presence and absence of three introns are found in one gene in the same species. Presence and absence of one intron in the same species have been found very rarely within a living organism [30, 31].

In addition to the interest of this *L. edodes* gene for his splicing variants, it might be very instructive to study it in different strains, as it is an opportunity to study the mechanism of the putative linked insertion or deletion of three introns in a gene. This could help to better understand the mechanism of appearance or disappearance of splicosomal introns which are still not well understood.

CONCLUSION

The identification of a new laccase sub-family (A) in addition to laccase sub-family C in *A. bisporus*, as in *P. ostreatus*, has led us to study intron positions and protein structure of mushroom laccases. Phylogeny based on intron positions compared to phylogeny of species could give complementary information to that based on amino acid sequences. We propose the existence of at least three very old laccase sub-families (A, B and C) in Agaricales and Polyporales in addition of Cantharellales laccase sub-families. The specific laccase roles in each sub family are no yet known but clade specific laccase functions will probably be find in the future. The expression of *lcc1*, *lcc2* and *lcc3* have been studied by Largeteau *et al.* [32] to find if one of them was more expressed on sporophores contaminated by *Leucanicillium fungicola.* None of these genes where over expressed under this condition compared to healthy sporophores. Comparative position of each laccase sub-family on chromosomes and all genome sequences may give confirmations of the phylogenetic hypotheses proposed here, as well as new data. A five amino acid insertion was probably at the origin of all basidiomycetes *bona fide* laccases. We identified a *L. edodes* gene with presence or absence of 3 introns depending of the allele. Studies of intron positions and comparative analysis of protein structure in laccase sub-family in relation to species phylogeny may generate many original data on the evolution of this protein and gene family and also on evolution of introns. New hypotheses have been proposed and are now to be tested.

ACKNOWLEDGMENT:
Serge Moukha taught us inverse PCR technique. Many students helped to obtain the complete new *Agaricus* laccase sequences: Florie Fauché, Emilie Lespinasse, Katia Lehn, Melissa Elie.

REFERENCES

[1]. Kues U., Ruhl M; (2011). Multiple Multi-Copper Oxidase Gene Families in Basidiomycetes-What for? *Current Genomics* 12: 72-94.
[2]. Savoie J.M., Mata G., Mamoun M. (2001). Variability in brown line formation and extracellular laccase production during interaction between white-rot basidiomycetes and *Trichoderma harzianum* biotype Th2. *Mycologia* 93: 243-248.
[3]. Kilaru S., *et al.* (2006). The laccase multi-gene family in *Coprinopsis cinerea* has seventeen different members that divide into two distinct subfamilies. *Current Genetics*. 50: 45-60.
[4]. Courty P.E. *et al.* (2009). Phylogenetic analysis, genomic organization, and expression analysis of multi-copper oxidases in the ectomycorrhizal basidiomycete *Laccaria bicolor*. *New Phytol.*. 182: 736-750.
[5]. Beloqui A. *et al.* (2006). Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen - Biochemical properties, structural analysis, and phylogenetic relationships. *J. Biol. Chem.* 281: 22933-22942.
[6]. Hoegger P.J. *et al.* (2006). Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. *FEBS J.* 273: 2308-2326.
[7]. Lundell T.K. *et al.* (2010). Lignin-modifying enzymes in filamentous basidiomycetes - ecological, functional and phylogenetic review. *J. Basic Microbiol.* 50: 5-20.
[8]. Qasemian L. *et al.* (2011). Halotolerant laccases from *Chaetomium* sp., *Xylogone sphaerospora* and *Coprinopsis* sp. isolated from a Mediterranean coastal area: potential to transform polycyclic aromatic hydrocarbons. submitted.
[9]. Savoie J.M. *et al.* (2004) Oxidative processes in *Agaricus bisporus* dry bubbles. *Mush. Sci.* 16: 527-535.
[10]. Perry C.P. *et al.* (1993). Identification of two laccase genes in the cultivated mushroom *Agaricus bisporus*. *J. Gen. Microbiol.* 139: 1209-1218.
[11]. Rogozin I.B. *et al.* (2003). Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution. *Current Biology* 13: 1512-1517.
[12]. Tavares R. *et al.* (2000). Organization and structural evolution of four multigene families in *Arabidopsis thaliana*: AtLCAD, AtLGT, AtMYST and AtHD-GL2. *Plant Mol. Biol.* 42: 703-717.
[13]. Boudet N. *et al.* (2001). Evolution of intron/exon structure of DEAD helicase family genes in Arabidopsis, Caenorhabditis, and Drosophila. *Genome Res.* 11: 2101-14.
[14]. Matheny, P.B. *et al*., (2007) Contributions of rpb2 and tef1 to the phylogeny of mushrooms and allies (Basidiomycota, Fungi). *Mol. Phylogenetics Evol.* 43(2): 430-451.
[15]. Irimia M., Roy S.W. (2008). Spliceosomal introns as tools for genomic and evolutionary analysis. *Nucleic Acids Research* 36: 1703-1712.
[16]. Callac P. *et al.* (1993). Morphological, genetic, and interfertility analyses reveal a novel, tetrasporic variety of *Agaricus bisporus* from the sonoran desert of California. *Mycologia* 85: 835-851.
[17]. Kerrigan R.W. *et al.* (1994). The heterothallic life-cycle of *Agaricus bisporus* var *burnettii* and the inheritance of its tetrasporic trait. *Exp. Mycol.* 18: 193-210.
[18]. Ochman H. *et al.* (1988). Genetic Applications of an Inverse Polymerase Chain-Reaction. *Genetics* 120: 621-623.

[19]. D'Souza T.M *et al*. (1996) Isolation of laccase gene specific sequences from white rot and brown rot fungi by PCR. *Appl. Environ. Microbiol*. 62: 3739-3744.
20. BioEdit (URL:http://www.mbio.ncsu.edu/BioEdit/).
[21].Swofford D.L. (1998). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Met hods), Version 4. Sinauer Associates, Sunderland,Massachusetts.
[22]. Garcia-Sandoval R. *et al.* (2011). Molecular phylogenetics of the Gloeophyllales and relative ages of clades of Agaricomycotina producing a brown rot. *Mycologia* 103: 510-524.
[23]. Wahleithner J.A. et al. (1996). The identification and characterization of four laccases from the plant pathogenic fungus *Rhizoctonia solani*. *Current Genetics* 29: 395-403.
[24]. Stajich J.E. et al. (2007). Comparative genomic analysis of fungal genomes reveals intron-rich ancestors. *Genome Biology* 8 :R223
[25]. Darlu P., Tassy P. (1993) La Reconstruction phylogénétique. Concepts et Méthodes. – ed. Masson.
[26]. Ahlawat O.P., Billette C (2011). Positioning of introns in different laccase genes, a relevant tool for solving phylogenetic position ambiguity of *Volvariella volvacea* laccase genes. in 7th International conference on mushroom biology and mushroom products (ICMBMP7). 2011. Arcachon, France.
[27]. Matheny P.B. *et al*. (2006). Major clades of Agaricales: a multilocus phylogenetic overview. *Mycologia* 98: 982-995.
[28]. Binder M. et al. (2010). Amylocorticiales ord. nov and Jaapiales ord. nov.: Early diverging clades of Agaricomycetidae dominated by corticioid forms. *Mycologia* 102: 865-880.
[29]. Zhao J., Kwan H.S. (1999). Characterization, molecular cloning, and differential expression analysis of laccase genes from the edible mushroom *Lentinula edodes*. *App. and Environ. Microbiol*. 65: 4908-4913.
[30]. Llopart A. *et al*. (2002). Intron presence-absence polymorphism in Drosophila driven by positive Darwinian selection. *PNAS* 99: 8121-8126.
[31]. Omilian A.R. *et al.* (2008). Intron presence-absence polymorphisms in Daphnia. *Mol. Biol. Evol.* 25: 2129-2139.
[32]. Largeteau M.L. *et al.* (2010). Expression of phenol oxidase and heat-shock genes during the development of *Agaricus bisporus* fruiting bodies, healthy and infected by *Lecanicillium fungicola*. *Appl. Microbiol. and Biotech*. 85: 1499-1507.