# The evolutionary study of genomic changes associated with morphological evolution of septal pore cap in Agaricomycotina

**Iizuka, Tomoyo**

Doctor of Philosophy

Department of Genetics

School of Life Science

The Graduate University for Advanced Studies,

SOKENDAI

# Contents

# Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Kazuho Ikeo for his appropriate advice, kindly support and other general things. I would also like to especially thank my committee members, Dr. Naruya Saitou, Dr. Hiroshi Akashi, Dr. Jun Kitano, Dr. Shinya Miyagishima and Dr. Masafumi Nozawa for fruitful discussion. I also appreciate the all members of the Laboratory for DNA Data Analysis, Dr. Sonoko Kinjo, Dr. Kaoru Matsumoto, Dr. Norikazu Kitamura, Dr. Junichi Imoto, Dr. Ikuko Yuyama, Dr. Masa-aki Yoshida, Dr. Masakazu Ishikawa, Mr. Toshitugu Okayama, Ms. Yoko Sakuma, Ms. Yuko Watanabe, Ms. Chie Iwamoto and Ms. Hisako Tashiro for daily discussion and making fulfilling NIG life. JSPS research fellowship (DC2) and KAKENHI grant number 18J13859 greatly supported my works. Some computation in this study was performed using the NIG supercomputer at ROIS National Institute of Genetics. I am thankful to my friends and the members of the association of the young mycologist for warm companionship. Finally, I express my appreciation to my family for the kind understanding and support for this work.

# Abstract

Morphological characters change in various directions during species divergence. During this evolutionary process, similar phenotypes appear independently, called convergent evolution. Since these differences have often played ecological roles, e.g., performing a physiological function, the molecular basis of these evolutionary events is important to understand phenotype-genotype relationships, adaptive evolution and measuring evolutionary repeatability.

In the major groups of filamentous fungi Agaricomycotina, one such example for typical morphological evolution occurs in the septal pore cap (SPC), which is involved in a plugging process of cell-like compartments. SPC is located around the hole of cell-like compartments (pore), and was derived from the endoplasmic reticulum. SPC is classified into three morphological types: perforate, imperforate, and vesiculate types. Perforate SPC has many small holes (perforations) on their SPC. Imperforate SPC has a slightly flattened closed membranous structure. The vesiculate SPC consists of vesicles or tubules arranged in a hemisphere and surrounding the pore. Since perforations of perforate SPC allows passing mitochondria and actin filament, this fact suggests that the difference of SPC types contributes to the difference of the functional performance of SPC. Current integrated results of the species phylogeny in fungi and morphological characterization of SPC showed that vesiculate SPC is the most ancestral trait, and perforate SPC is the most neomorphic trait. Interestingly, perforate SPC emerged from imperforate SPC multiple times independently. This fact indicates that morphological convergent evolution had occurred on the perforate type lineages.

Despite the interesting evolutionary history of the variable morphological characters of SPC as described above, the genetic background

of these evolutionary events remains unknown. In this doctoral thesis, I therefore aim to clarify genomic changes associated with the morphological evolution of SPC using comparative genomics. I conducted three analyses; candidate gene extraction, gene search for understanding sequence evolution correlated with the morphological difference between vesiculate SPC and imperforate SPC, and sequence comparison for understanding morphological convergence from imperforate SPC to perforate SPC. More details about the background and aim of this study are described in Chapter 1.

In chapter 2, to detect candidate genes that correlate with the morphological convergent evolution of Perforate SPC, I conducted genome-wide surveys. Using the 12 fungal genomes, I created an orthologous gene dataset and extracted the candidate gene from the dataset based on the results of the phylogenetic analysis of each ortholog. I found the SPC-related gene *spc33* showed different branching pattern than the well-supported species tree. In particular, amino acid sequences of *spc33* showed convergent substitutions at the same site between diverged species with perforate SPC. These results suggest the possibility of the involvement of *spc33* to morphological convergence of perforate SPC. In chapter 3, in order to understand sequence evolution related to morphological evolution from imperforate SPC to perforate SPC, I verified how convergent amino acid substitutions occurred in the *spc33* of *Rhizoctonia solani* (Rhiso). The possible hypotheses are: 1, Rhiso and other perforate type species acquired the convergent substitutions independently, 2, the same substitutions had gained by transferring homologous region from other perforate type species. To clarify which hypothesis is more possible, I conducted synteny analysis, determination of exon/intron structure and phylogenetic analysis of *spc33*. The results provided no supporting evidence of hypothesis 2 from the genome sequences of Rhiso. Thus, I concluded that the hypothesis 1 is more possible for the cause of convergent substitutions in amino acid sequences of *spc33*. In

Chapter 4, I presume the origin of *spc33* in vesiculate type species since the result of sequence similarity searches showed that vesiculate type species do not have *spc33* in their genomes. The *spc33* was also not detected from the other eukaryotes and all prokaryotes. Thus, I searched a weak sequence homology to assess the ancestral gene of *spc33* by using sequence motifs. To see the track of *spc33* from the genome of outgroup species, synteny analysis was conducted. Based on these results, I discussed the origin of *spc33* and its evolution.

In summary, I identified each genetic difference correlated with the morphological difference between vesiculate SPC and imperforate SPC, and morphological convergence from imperforate SPC to perforate SPC. Also, I clarified the process of these genetic changes. These results provide important advances for understanding the genetic basis of morphological evolution of SPC. These achievements contribute to uncovering the molecular mechanisms of convergent evolution and to finding a missing link between morphological evolution and sequence evolution in SPC. My results present the hints for deep understanding about genetic basis of morphological evolution and convergent evolution. In addition, this study provides an impetus for developments in fungal evolutionary biology and genomics.

# Chapter 1

# General introduction

## 1.1 Morphological diversity and its evolution

Morphological characters change in various directions during species divergence. These differences have often played ecological roles e.g. performing a physiological function (Bertossa 2011). Therefore, clarifying the evolution of morphological characters brings fundamental insights to understand the adaptation and evolutionary history of species (Almén et al., 2016). The recent development of NGS technology greatly helps to expand the targets of studies that provide huge clues to reveal the genetic background of various morphological differentiation, changes in genomic architecture that are associated with speciation, and the genetic basis of ecologically important traits. Non-model organisms that are very useful materials to study such a morphological evolution have also become researchable by the enhancement of genomic information of various organisms (Ekblom and Galindo, 2011).

## 1.2 Parallel evolution and Convergent evolution

The same phenotype sometimes appears independently, calls parallel or convergent evolution. Understanding a genetic basis of these phenomena is considered important to discuss phenotype-genotype relationships, adaptive evolution, measuring evolutionary repeatability, and so on (Stern and Orgogozo, 2009; Castoe et al., 2010; Payne and Wagner, 2019). So far, these independent evolutionary events have been categorized by their evolutionary process. At the phenotypic level, Wray 2002 described

that parallel evolution can be detected when a phenotype changes in the same way independently in two lineages (Figure 1.1a). On the other hand, convergent evolution requires an initial period of divergence, followed by a period when the feature becomes more and more similar (Figure 1.1b). Since phenotypic determination and characterization are greatly influenced by genome sequences, the sequence evolution in genomes (Figure 1.2) also has been focused as well as phenotypic evolution (Yang 2014). Detecting these evolutionary events and linking causal relationships between these two layers of evolution have been regarded as important issues to clarify the evolutionary process of organisms (Stern 2013). However, since we cannot track all changes of evolutionary processes, covering the detailed evolutionary process is difficult to define from our available species and its genome sequences. This detection limit makes hard to distinguish between parallel evolution and convergent evolution described above. In this thesis, I therefore use the term “convergence” for independent evolution on both the phenotypic level and the sequence level.

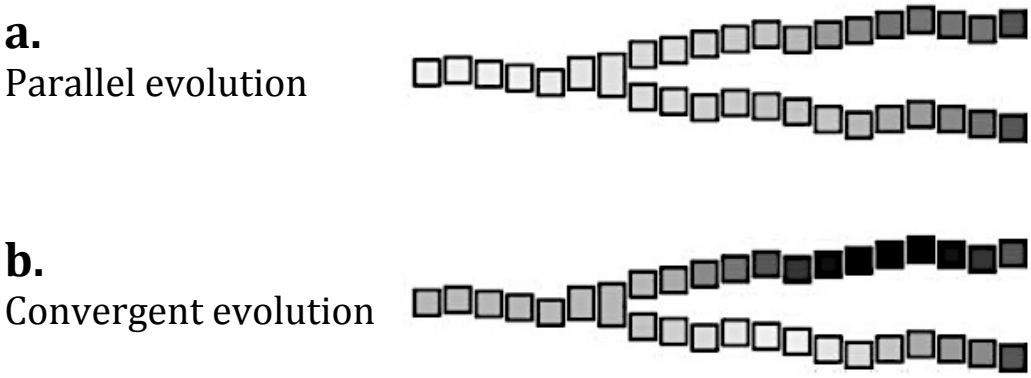



**Figure 1.1.** Parallel evolution (a) and convergent evolution (b) at phenotypic level. Similarities and differences in living organisms are observations, but divergence, conservation, parallelism, and convergence are all inferences about evolutionary history. (Wray 2002, modified)

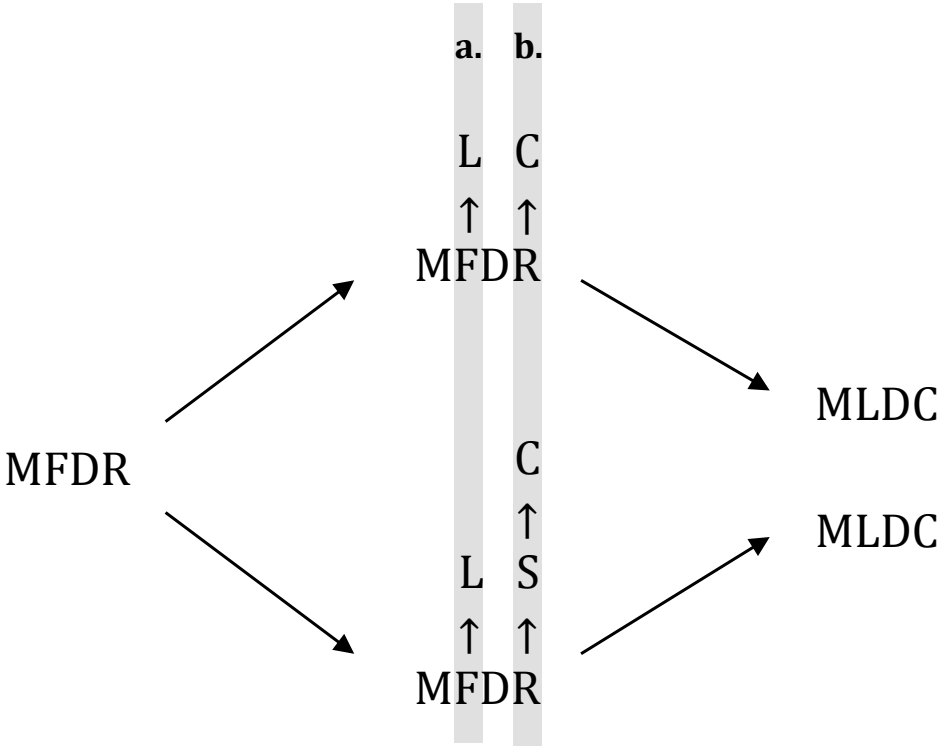



**Figure 1.2.** Parallel evolution (a) and convergent evolution (b) at amino acid sequence level. Both evolutionary mechanisms can produce the same amino acid at a given position in the sequences of homologous proteins.

## 1.3 Ultrastructure, Morphological characters and

## Functions of Septal pore cap in Fungi

To tackle the issues described at 1.1 and 1.2, fungal-specific organelle Septal pore cap (SPC) is one of the most suitable structure. It associates with a plugging process of cell-like compartments in the major groups of filamentous fungi Agaricomycotina (i.e. Tremellomycetes, Dacrymycetes, and Agaricomycetes), including jelly fungi and mushroom-forming fungi (Peer et al., 2010). SPC is located around the hole of cell-like compartments (pore), and its diameter of SPC is 450–600 nm (Figure 1.3). Since some ER targeting markers stain the SPC (van Driel et al., 2008; Müller et al., 1999) and SPCs are connected at their base to the endoplasmic reticulum (ER) (Moore, 1975; Müller et al., 1998a, 2000), it is suggested that SPC was derived from the ER.

Several SPC-types related to the higher-level taxonomic relationships of Agaricomycotina (e.g. McLaughlin et al., 1995; Fell et al., 2001; Hibbett and Thorn, 2001; Wells and Bandoni, 2001; Lutzoni et al., 2004). Figure 1.4 shows the relationship between SPC types and species phylogeny. Perforate SPCs are found in the Agaricomycetes (Hibbett and Thorn, 2001; Wells and Bandoni, 2001) and have many small holes (perforations) on their SPC. The imperforate SPC is found in the Dacrymycetes and Agaricomycetes (Wells and Bandoni, 2001) and consists of a slightly flattened closed membranous structure. The vesiculate SPC can be found in members of the Tremellomycetes, and this SPC type consists of a group of vesicles or tubules arranged in a hemisphere surrounding the pore. Since perforations of perforate SPC allows passing mitochondria and actin filament (Bracker and Butler, 1964; Moore and Marchant 1972), suggests that the difference of SPC types contributes to the difference of the functional performance of SPC.

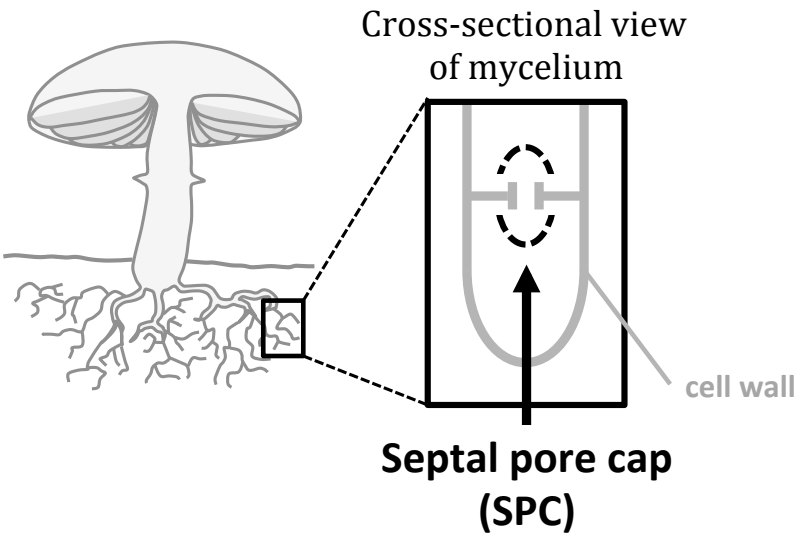



**Figure 1.3.** The schematic diagram of the Septal pore cap (SPC).

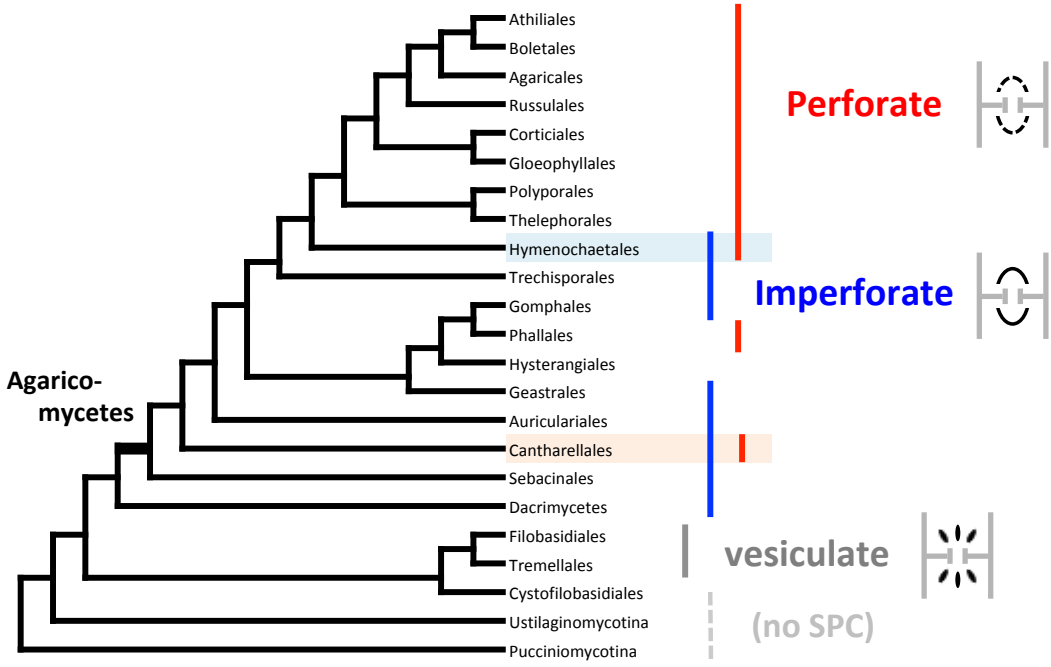



**Figure 1.4.** The phylogeny of Basidiomycota and morphological characters of SPC. The phylogenetic relationship was sited from Hibbett et al. (2014).
The morphological characters of SPC were sited by Driel et al. (2009).

## 1.4 Evolutionary history of morphological characters of Septal pore cap

The earliest research on the morphological evolution of SPC is an attempt to study the process of morphological changes. This study was based on comparing and analyzing the characters of SPCs of various species (Patton and Marchant 1978). Current integrated results of the phylogenetic relationships (Lutzoni et al.,2004; Hibbett et al., 2014) in fungi and morphological observations of SPC show that vesiculate SPC and imperforate SPC may be originated from localized ER around pores, and perforate SPC is the most neomorphic traits (Driel et al., 2009). Interestingly, perforate SPC emerged from imperforate SPC multiple times independently (Figure 1.4). This fact indicates that convergent evolution had occurred in perforate SPC. Despite the accumulation of knowledge about the evolution of SPC at the phenotypic level, the genetic backgrounds of these evolutionary events remain unknown. Identifying specific mutations correlated with the morphological evolution of SPC is needed for further understanding the evolution of SPC.

## 1.5 Aim and Outline of this thesis

As mentioned above, the genetic background of the morphological evolution of SPC is still not understood. In this doctoral thesis, I therefore aim to clarify genomic changes associated with the morphological evolution of SPC using comparative genomics. I conducted three analyses; candidate gene extraction, analysis for understanding sequence evolution correlated with the morphological difference between vesiculate SPC and imperforate SPC, and morphological convergence from imperforate SPC to perforate SPC.

In chapter 2, to detect candidate genes that correlate with the morphological convergent evolution of Perforate SPC, I conducted genome-wide surveys. Using the 12 fungal genomes, I created an orthologous gene dataset and extracted the candidate gene from the dataset based on the results of the phylogenetic analysis of each ortholog. I found the SPC-related gene *spc33* showed different branching pattern than the well-supported species tree. In particular, amino acid sequences of *spc33* showed convergent substitutions at the same site between diverged species with perforate SPC. These results suggest the possibility of the involvement of *spc33* to morphological convergence of perforate SPC.

In chapter 3, in order to understand sequence evolution related to morphological evolution from imperforate SPC to perforate SPC, I verified how convergent amino acid substitutions occurred in the *spc33* of *Rhizoctonia solani* (Rhiso). The possible hypotheses are: 1, Rhiso and other perforate type species acquired the convergent substitutions independently, 2, the same substitutions had gained by transferring homologous region from other perforate type species. To clarify which hypothesis is more possible, I conducted synteny analysis, determination of exon/intron structure and phylogenetic analysis of *spc33*. The results provided no supporting evidence of hypothesis 2 from the genome sequences of Rhiso. Thus, I concluded that the hypothesis 1 is more possible for the cause of convergent substitutions in amino acid sequences of *spc33*.

In Chapter 4, I presume the origin of *spc33* since the result of sequence similarity searches showed that vesiculate type species do not have *spc33* in their genomes. The *spc33* was also not detected from the other eukaryotes and all prokaryotes. Thus, I searched a weak sequence homology to assess the ancestral gene of *spc33* by using sequence motifs. To see the track of *spc33* from the genome of outgroup species, synteny analysis was conducted. Based on these results, I discussed the origin of *spc33* and its

evolution.

In Chapter 5, I summarized the results of Chapters 2, 3 and 4 and made several conclusions.

# Chapter 2

# Candidate genes correlated with morphological evolution of SPC: Sequence evolution associated with phenotypic evolution

## 2.1 Introduction

An important problem in evolutionary biology is what changes occurred at the genome sequence level when deriving morphological diversity. The study of convergent evolution, i.e., the independent evolution of similar phenotypic changes in different lineages, has been said to provide insight into hotspots and key mutations for phenotypic differentiation (Stern and Orgogozo, 2009). In addition, focusing on morphological convergence provides clues for understanding the relationship between phenotypic evolution affected by natural selection and sequence evolution driven by the protein adaptive landscape (Castoe et al., 2010; Stern 2013; Almén et al., 2016). However, genome-wide surveys of convergent evolution are relatively rare and more empirical genome-scale studies are needed to elucidate the genetic basis of morphological convergent evolution.

One such example of traits representing typical morphological convergence is the traits of septal pore caps (SPC) in fungi, which are involved in the construction of mycelia's complex multicellularity (Jedd, 2011; Nguyen et al., 2017). It has been suggested that SPC was newly derived from the endoplasmic reticulum (ER) because SPC was stained by some ER-targeting markers (van Driel et al., 2008; Müller et al., 1999) and SPCs are connected at their base to the ER (Moore, 1975; Müller et al., 1998a, 2000). I focused on a couple of morphological types of SPC: vesiculate, imperforate, and perforate types. Vesiculate SPC is the most ancestral type

of SPC, imperforate SPC had emerged from vesiculate SPC only once, and perforate SPC is remarkable because it independently emerged from imperforate SPC multiple times (van Driel et al., 2009). The difference between SPC types leads to the derivation of different functions for cytoplasmic transport inside the mycelium (Bracker and Butler, 1964; Moore and Marchant, 1972), which implies that SPC types had evolved through adaptation for transportation of various substances. These SPC types are highly conserved at the order level (Hibbett, 2006; van Driel et al., 2009; Oberwinkler et al., 2013; Hibbett et al., 2014) and the understanding genetic background of SPCs' morphological differentiation can substantially contribute to understanding the evolution of higher taxa in Basidiomycota.

Previous SPC studies reported on its functions, fine features, and applicability to fungal classification (Bracker and Butler, 1964; Moore and Marchant, 1972; Lisker et al., 1975; Hibbett, 2006; van Driel et al., 2009); however, despite the long history of the morphological studies of SPC, the genetic basis driving the morphological differentiation of SPC remains unclear. In animals, studies have reported mutations associated with morphological convergent evolution by whole-genome analysis. For example, Hu et al. (2017) found a candidate gene for the evolution of a pseudo-thumb in pandas using positive selection as an indicator. Using positional cloning, Colosimo et al. (2005) found that the Eda pathway is involved in the phenotypic evolution of armor plate in sticklebacks. The genomic background associated with SPCs' morphological convergence should also be elucidated using a similar genomic approach.

In this study, I aimed to clarify the genomic changes associated with the morphological evolution of SPC in terms of comparative genomics. I used publicly available whole-genome sequences from the 1000 Fungal Genomes Project (http://1000.fungalgenomes.org/) to identify evolutionary events at the sequence level. Data for morphological types of SPC were also

accumulated because of their importance for taxonomic classification (Wells, 1994; Müller et al., 1998b, 2000; Lutzoni et al., 2004; Hibbett et al., 2014). By combining these genomic and morphological data, I obtained evidence from 12 fungal genomes for morphological convergence at the amino acid sequence level. Here I report that amino acid substitutions on *spc33* are correlated with the morphological difference between imperforate and perforate SPC.

## 2.2 Materials and Methods

*Genome sequences*

The genome sequence data for 12 species of Basidiomycota, 11 representatives of Agaricomycetes (6 perforate-type species and 5 imperforate-type species), and one representative of Dacrymycetes (imperforate type), were used for this study (Table 2.1). These species are representative of major Basidiomycota species. All sequence data were obtained from the Joint Genome Institute portal (JGI, http://www.genome.jgi.doe.gov) of the United States Department of Energy.

**Table 2.1.** Taxonomic classification and data source for species included in this study.

| Class | Order | Species | SPC type* | Genome assembly size (Mbp) | No. of Genes | Reference |
|---|---|---|---|---|---|---|
| Agaricomycetes | Agaricales | *Galerina marginata* | Perforate | 59.42 | 21451 | Riley et al., 2014 |
| | Corticiales | *Punctularia strigosozonata* | Perforate | 34.17 | 11538 | Floudas et al., 2012 |
| | Gloeophyllales | *Gloeophyllum trabeum* | Perforate | 37.18 | 11846 | Floudas et al., 2012 |
| | Polyporales | *Phanerochaete carnosa* | Perforate | 46.29 | 13937 | Suzuki et al., 2012 |
| | Hymenochaetales | *Rickenella mellea* | Perforate | 46.03 | 18935 | Nagy et al., 2016 |
| | | *Trichaptum abietinum* | Imperforate | 40.61 | 14963 | (Project ID: 1006923) |
| | Trechisporales | *Sistotremastrum niveocremeum* | Imperforate | 35.36 | 13069 | Nagy et al., 2016 |
| | Gomphales | *Ramaria rubella* | Imperforate | 105.46 | 19264 | (Project ID: 1016735) |
| | Cantharellales | *Rhizoctonia solani* AG-1 IB | Perforate | 47.66 | 16497 | Wibberg et al., 2013 |
| | | *Tulasnella calospora* | Imperforate | 62.39 | 19535 | Kohler et al., 2015 |
| | | *Botryobasidium botryosum* | Imperforate | 46.67 | 15034 | Riley et al., 2014 |
| Dacrymycetes | Dacrymycetales | *Calocera viscosa* | Imperforate | 29.10 | 12369 | Nagy et al., 2016 |

* Based on Driel et al., 2009

*Identification of orthologs*

To identify orthologs, the reciprocal best hits (RBH) analysis (Tatusov et al., 1997; Bork et al., 1998; Overbeek et al., 1999) was conducted. I performed all-vs-all blast with BLAST software (blastp, NCBI-blast-2.2.30+) using the representative coding sequences from those remaining after discarding sequences with internal stop codons. The RBH were extracted using original Perl scripts.

*Candidate gene extraction using topological analysis of gene trees*

Candidate genes that correlated with the morphological differences between imperforate and perforate SPCs were extracted from orthologous genes by conducting phylogenetic analyses and assessing the topology of the results. This approach is based on a working hypothesis: i.e., the genetic changes contribute to the independent emergence of perforate SPC. Thus, I extracted genes that show correlated changes with morphological traits by constructing a gene tree for each ortholog. The multiple alignment was created using MUSCLE v3.8.31 (Edger, 2004). Phylogenetic gene trees for each ortholog were generated using two different methods, i.e., the neighbor-joining (NJ) (Saitou and Nei, 1987) and maximum-likelihood (ML) methods (Felsenstein 1981) in MEGA6-CC (Tamura et al., 2013). I then checked the topology of the trees. In the topological analysis, I searched genes that do not support the species phylogeny by their gene tree. When the topology of both NJ and ML trees showed different topology with the species

phylogeny (Hibbett et al., 2014), the gene was obtained as a candidate gene (Figure 2.1). The obtained genes were annotated according to the NCBI BLAST (Johnson et al., 2008). Four databases: UniProt (259, 2017-04) (The UniProt Consortium, 2017), GeneOntology (2017-04-11) (The Gene Ontology Consortium, 2000), *Saccharomyces* Genome Database (R64.2.1, 2014-11-18) (Cherry et al., 2012), and PomBase (30_62, 2017-01-30) (Wood et al., 2012) were also referred to assessing the results. A concatenated tree using an orthologous gene dataset was also constructed using the NJ method to confirm the phylogenetic relationship of the 12 fungal species.

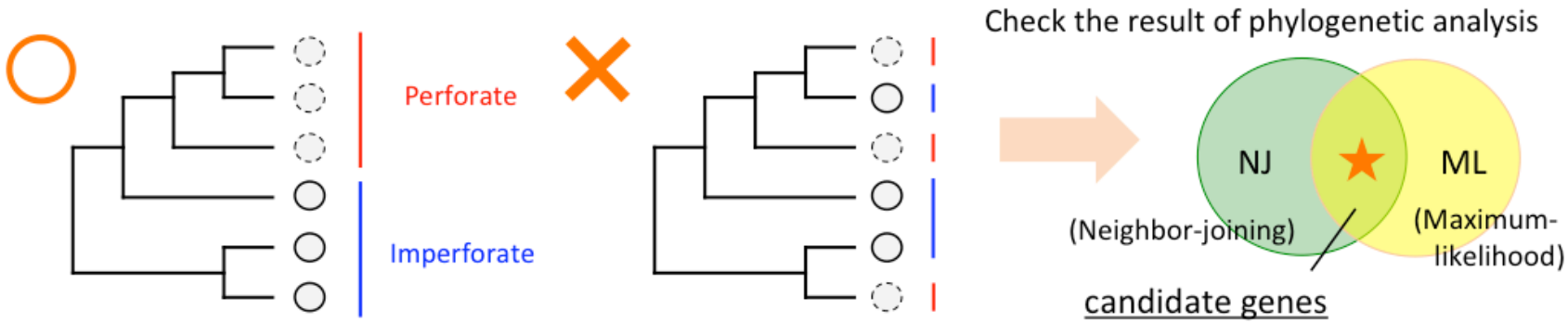



**Figure 2.1.** Schematic diagram of candidate gene extraction algorithm.

*Detection of ancestral sequences and convergent amino acid substitution in candidate genes*

I examined multiple alignments of candidate genes to extract evidence for convergent evolution at the amino acid site level. I compared the amino acid residues of the 12 fungal species to identify SPC type-specific sites that do not share amino acid residues between six imperforate-type species and six perforate-type species.

The ancestral sequences of all orthologs were also estimated to identify convergent amino acid changes associated with the morphological convergence using the maximum likelihood (ML) approach implemented in MEGA6-CC. Convergent substitutions were detected if the data shows a substitution to the exact same amino acid residue in each lineage with the morphological convergence, whereas other species do not have such substitution pattern. Count of convergent substitutions was calculated in each ortholog and compared its ratio among orthologs.

## 2.3 Results

*Candidate genes correlated with convergent evolution*

I focused on the morphological difference between SPC types to examine convergent evolution in fungi. To detect the candidate genes, I first extracted the orthologs of the 12 fungal genome sequences and performed topological analyses using the gene tree for each ortholog. The RBH analysis yielded 2,560 orthologs. In the candidate gene extraction, I found eight genes

of perforate type species *Rhizoctonia solani* (Rhiso) were clustered with other perforate-type species from another lineage (Figure 2.2, Table 2.2). Seven genes were annotated as known genes and another gene as unknown using the referred databases. One of these annotated seven genes was identified as a SPC-related gene, *spc33*.

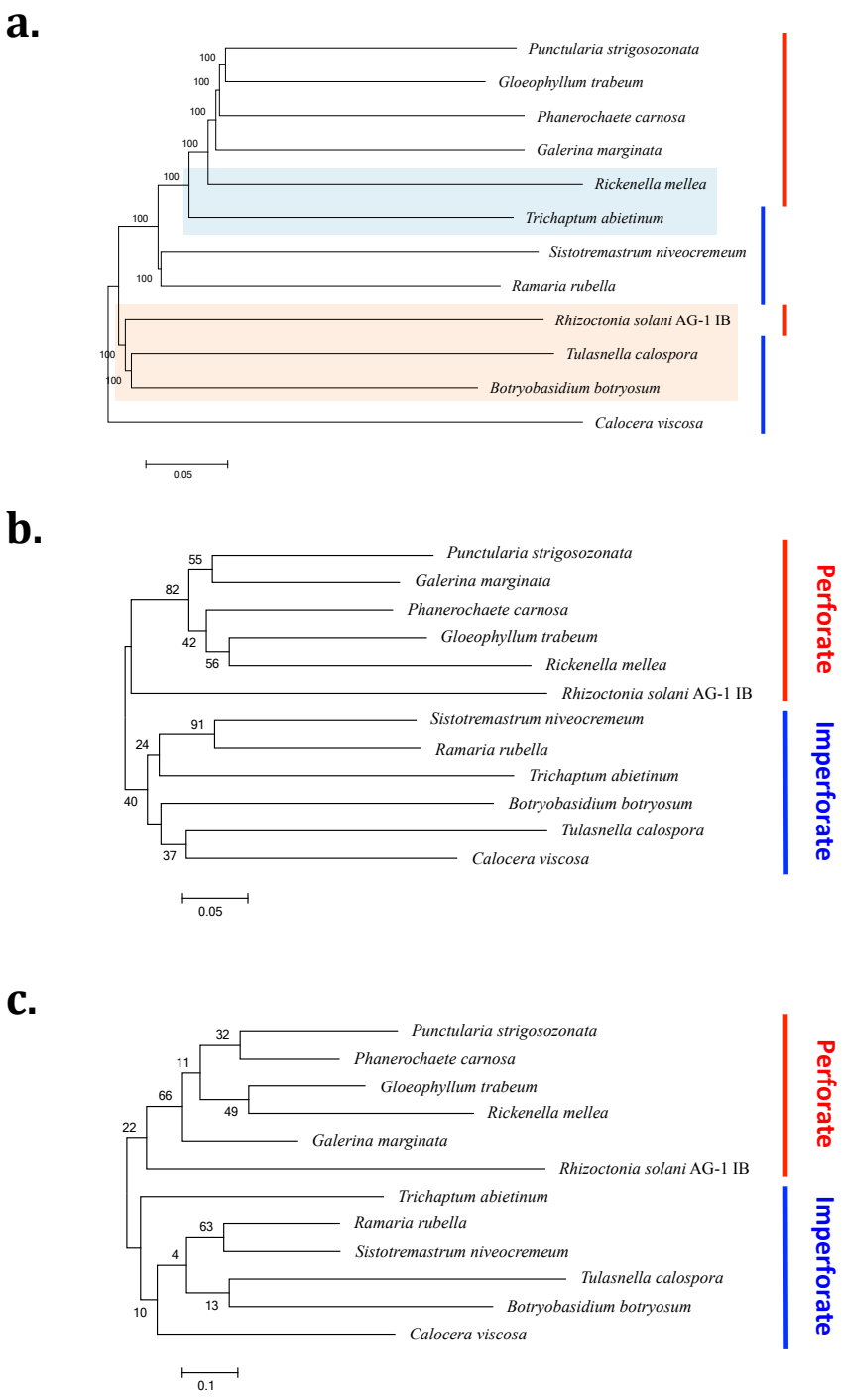



**Figure 2.2.** Phylogenetic relationship of 12 fungal genomes and correlated genes. (a) Phylogenetic tree based on the neighbor-joining (NJ) method of concatenated sequences of 2560 orthologs. Numbers along interior branches represent the bootstrap values with 1,000 replicates. Multiple species were used for *Hymenochaetales* (red) and *Cantharellales* (blue). Gene trees of correlated gene #1 were produced using the (b) NJ and (c) maximum-likelihood (ML) methods. Phylogenetic gene trees of correlated genes #2-#8 were shown in Supplemental Figure 1.

**Table 2.2.** The list of eight candidate genes.

| | BS NJ [1)] | BS ML [1)] | model [2)] | aa [3)] | Gene symbol | Full gene name |
|---|---|---|---|---|---|---|
| #1 | 40 | 22 | LG+G | 596 | *spc33* | 33kDa protein purified from SPCs |
| #2 | 11 | 10 | LG+G | 468 | DSD1 | D-serine dehydratase |
| #3 | 40 | 46 | rtREV+G | 103 | SDH8 | Succinate dehydrogenase |
| #4 | 15 | 12 | LG+G | 318 | n/a | Uncharacterized protein |
| #5 | 9 | 10 | LG+G | 98 | SUMO | Small ubiquitin-like Modifier protein |
| #6 | 37 | 38 | JTT+G+I | 454 | ELP4 | Elongator Acetyltransferase Complex Subunit 4 |
| #7 | 9 | 12 | LG+G+I | 271 | DRE2 | Fe-S cluster assembly protein |
| #8 | 21 | 12 | LG+G | 221 | DUS1 | tRNA-dihydrouridine synthase 1 |



1) Bootstrap value of the neighbor-joining (NJ)/maximum-likelihood (ML) tree supporting a single perforate SPC cluster. 2) Substitution model of ML tree. 3) Amino acid length (*Rhizoctonia solani*).

*SPC type-specific amino acid sites of spc33*

To investigate the relationship between sequence evolution and morphological convergent evolution, I focused on *spc33* to check the differences between amino acid residues. Figure 2.3a shows a noteworthy amino acid substitution pattern of SPC33 shared in perforate-type species. The amino acid substitutions from imperforate-type species to perforate-type species were observed at V/I/L342M/A/F, K357R and M/V359I, which were localized within the short same region of the 17 amino acids, although *spc33* has more than 300 amino acid length.

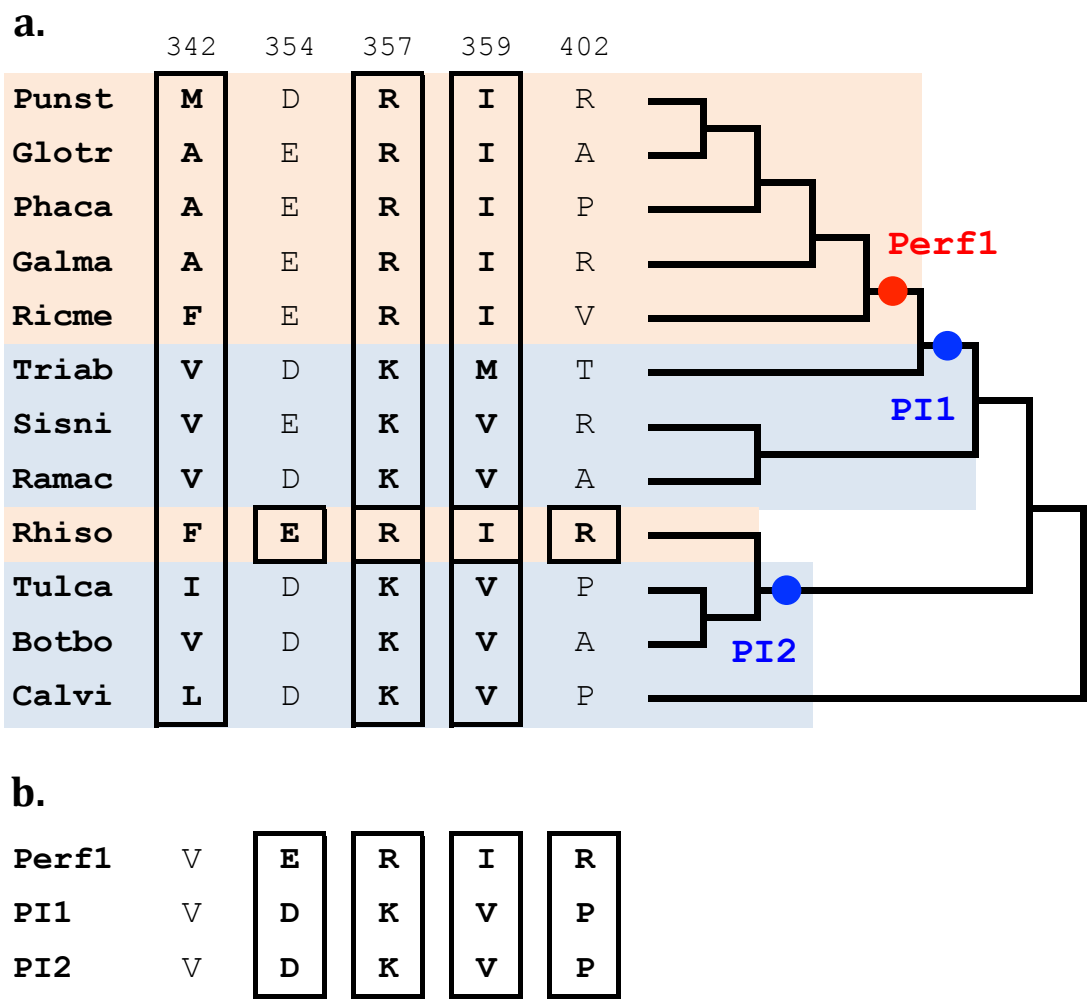



**Figure 2.3.** SPC type-specific mutations in *spc33*. (a) SPC type-specific substitutions in extant species. The abbreviations of the species names are as follow: Punst, *Punctularia strigosozonata*; Glotr, *Gloeophyllum trabeum*; Phaca, *Phanerochaete carnosa*; Galma, *Galerina marginata*; Ricme, *Rickenella mellea*; Triab, *Trichaptum abietinum*; Sisni, *Sistotremastrum niveocremeum*; Ramac, *Ramaria rubella*; Rhiso, *Rhizoctonia solani*; Tulca, *Tulasnella calospora*; Botbo, *Botryobasidium botryosum*; Calvi, *Calocera viscosa*. (b) Ancestral sequences of focused sites. The abbreviations of the branch names are as follow: Perf1, the lineage including five perforate-type species; PI1, perforate–imperforate branching point 1; PI2, perforate–imperforate branching point 2.

*Convergent evolution of spc33 in amino acid sequence levels*

To investigate what changes had occurred in amino acid sequences of *spc33* correlated to the difference in SPC type, I reconstructed the ancestral sequences of *spc33* and checked the substitution process of amino acid changes. I expected if convergent evolution had also occurred at the amino acid sequence level, it can be suggested that phenotypic characterization of perforate SPC is regulated under the limited combination of molecular pathways. The results show that the same pattern of convergent amino acid substitutions was detected at D354E, K357R, V359I, and P402R during the evolution from the lineage of perforate–imperforate branching point 1 (PI1) to the lineage including five perforate-type species (Perf1) and from perforate–imperforate branching point 2 (PI2) to Rhiso (Figure 2.3b). K357R and M/V359I still showed differences in the extant species; therefore, convergent evolution had also occurred at the sequence level. The substitutions of D354E and P402R were also detected; however, these substitution patterns were not conserved in the extant species.

I found that the number of convergent substitutions on *spc33* was much higher than in other orthologs (Figure 2.4). To test whether this observation had not caused by random substitutions, I checked the number of convergent substitutions in all genes of Rhiso and other perforate type lineages. The result showed that the number of convergent-change sites of *spc33* was 0.03 per mutations per 100 amino acid sites (Figure 2.4a). This value was in the top 0.19 % of all orthologs. Furthermore, more than half

orthologs did not have convergent substitution in their amino acid sites (Figure 2.4b), indicating that *spc33* has high convergent substitution late rather than other orthologs. Therefore, convergent substitutions in *spc33* were considered to have occurred according to the morphological convergent evolution of SPC.

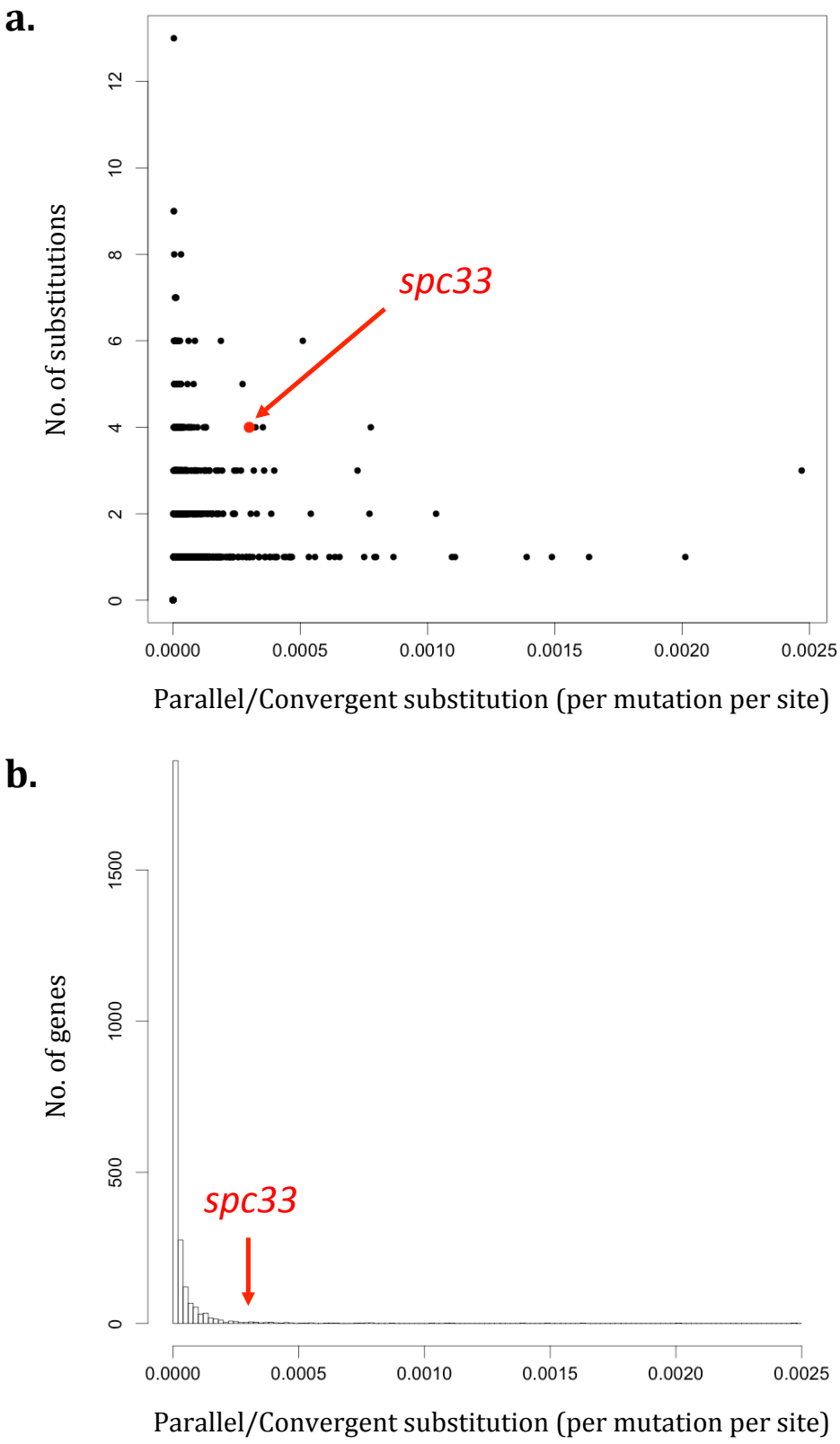



**Figure 2.4.** Convergent substitution frequency described by a) scatter plot with amount of convergent substitutions and b) distribution of proportions of convergent substitution possessed by each ortholog.

## 2.4 Discussion

Van Driel et al. (2009) provided evidence for morphological convergence of SPC and its phylogenetic relationship. Here, I provide evidence for sequence convergence accompanying the morphological convergence of perforate-type SPC in the amino acid sequence level. My results show that *spc33* was clustered by SPC type rather than by species phylogeny (Figure 2.2b, 2.2c). *spc33* in perforate type species Ricme is the most homologous sequence of *spc33* in Rhiso (Supplemental Table 1). It suggests that sequence differences of *spc33* correspond to the morphological convergence of perforate SPC. Five SPC-type specific sites were detected in amino acid sequences of *spc33* based on comparative genome analysis. The amino acid substitutions of V/I/L342M/A/F, K357R and M/V359I were identified in the short same region of *spc33* in perforate-type species from different perforate type lineages (Figure 2.3). Other two convergent substitutions D354E and P402R were also detected from *spc33*. Especially, all sequences showed the same pattern K357R and M/V359I between both two perforate type lineages. Below, I discuss the evidence for convergent evolution in fungal genome sequences and suggest the possible correspondence of *spc33* in the morphological convergent evolution of SPC.

Eight candidate genes were detected as correlated genes with morphological convergence of perforate SPC (Table 2.2). One of these genes was the SPC-related gene, *spc33*. van Peer et al. (2010) demonstrated that *spc33* is a transmembrane protein of SPC by examining *Schizophillum commune* (perforate type). When they constructed a *spc33* knockout mutant of *S.commune*, the SPCs disappear. By combining their report with my findings, that I successfully detected *spc33* as a correlated gene for the morphological convergent evolution of SPC. Since *spc33* is responsible for constructing SPC itself and evolved its amino acid sequence during the

evolution of perforate SPC, these results illustrate a model for the regulation of morphological changes of SPC by the SPC-forming protein SPC33. In the seven other identified genes, however, there was no clear evidence of involvement in SPC. It is known that SPC interacts with many intracellular substances, such as ER, cytoplasm, and cytoskeleton (Bracker and Butler, 1964; Moore and Marchant 1972). Therefore, the relationship between SPC and these seven genes will be revealed when knowledge about the relationship between these intracellular substances and SPCs accumulates.

The five amino acid substitutions related to morphological convergent evolution from imperforate to perforate SPC were identified in my study of *spc33* (Figure 2.3). When the number of convergent substitutions in *spc33* is higher than the number of unrelated convergent substitutions, I may consider the possibility that convergent substitutions in *spc33* are involved in the morphological convergence of perforate SPC. Therefore, I verified the frequency and randomness of convergent substitutions in the orthologous gene dataset. The frequency of convergent substitutions showed the significance of the high convergent substitution ratio of *spc33* (Figure 2.4). These data show that the number of convergent substitutions in *spc33* was significantly higher than for the other orthologs. Therefore, I conclude that convergent substitutions in *spc33* were not random occurrences. Thus, convergent evolution at the sequence level was correlated with the morphological convergent evolution of perforate SPC.

Of course, when convergent evolution occurred at the phenotypic level, there are multiple patterns of mutations that lead to the same phenotype (Almén et al. 2016). However, in the sequences of *spc33*, several convergent amino acid substitutions were detected (Figure 2.3). In site 357, both ancestral branching point PI1 and PI2 showed the same amino acid residue Lysine (Figure 2.3b). On the other hand, the estimated ancestral residue of the site 357 was Arginine at the ancestral branching point Perf1

(Figure 2.3b) and the residue of Rhiso was also changed to Arginine at the emergence of this species (Figure 2.3a). Therefore, the same amino acid substitution from Lysine to Arginine has occurred independently when perforate type species emerged. These substitutions may have been caused by the parallel evolution at the amino acid sequence level. I also checked this site at the nucleotide sequence level of both perforate type lineages. Rhiso acquired the codon CGT (Arg) at the site 357 (Figure 2.5), the two closest imperforate type species (Tulca and Botbo) and the species at the most basal lineage (Calvi) use AAG (Lys). In other perforate type species also use CG at the 1st and the 2nd codon. This result indicates that the substitution pattern of all perforate type species has commonly changed to CGN (Arg) from AAR (Lys). Thus, even at the nucleotide sequence level, I observed the same substitution pattern from AAG to CGT (Figure 2.5). The difference of nucleic acid at the 3rd position of Punst, Glotr, Phaca should be occurred after their diversification. This phenomenon gives us a further question; whether this same substitution pattern has a functional meaning for the *spc33* gene.

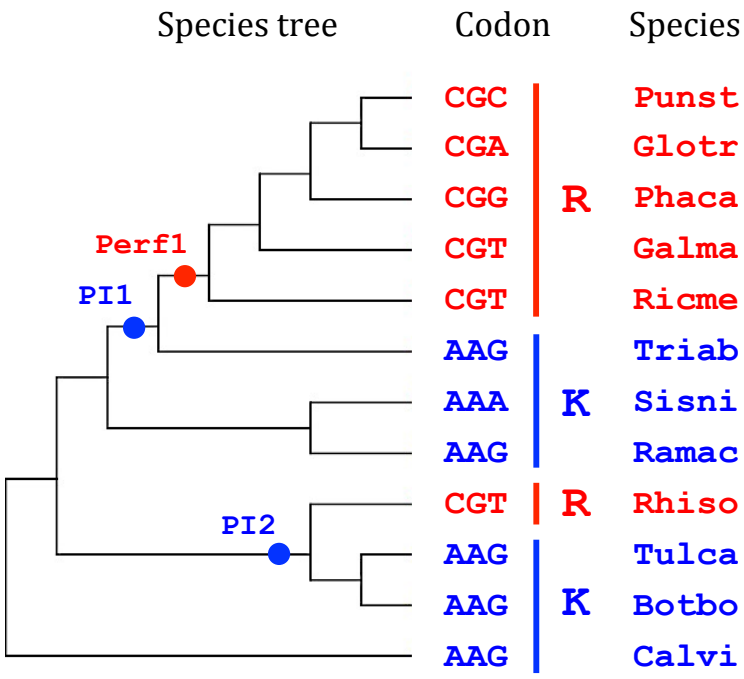



**Figure 2.5.** Codon pattern of site 357 of *spc33*.

It is known that few convergent amino acid substitutions cause phenotypic convergence e.g. the "five-sites" rule for the evolution of red and green color vision in mammals (Nei et al., 1997; Shi and Yokoyama 2003), A67V and G83S in high-altitude-adapted β-globin in birds (Natarajan et al. 2016), I1561V for TTX resistance in common garter snake (Hague et al., 2017). My analysis showed that the substitution pattern K357R from imperforate type species to perforate type species was completely conserved in all fungal species as described above (Figure 2.3). In site 359, the amino acid residue of Rhiso (Ile) was conserved in all other perforate type species (Figure 2.3a). All imperforate type species also have Valine at this site except Triab that has Methionine. Thus, these facts suggest that the convergent substitutions at the site 357 and 359 cause the phenotypic convergence of perforate SPC. In the other three SPC-type specific sites, two substitutions of D354E and P402R showed the same amino acid substitution pattern between Rhiso and other perforate type species (Figure 2.3). However, these substitution patterns were also observed in imperforate type species. Therefore, these substitutions are not considered to be critical patterns for morphological convergence. In site 342, Phenylalanine in Rhiso was observed in the other perforate type species. Even though the residue of this site is not conserved well among species, the substitutions of V/I/L342M/A/F were also identified within the short region of *spc33*. Therefore, these observations may suggest the site 342 is also the candidate of the convergent evolution of perforate SPC.

To understand the genetic background of the morphological evolution of vesiculate SPC, I tried to understand more about the evolution of ancestral states of *spc33*. I searched for *spc33* homologs against sequences of all organisms available in the NCBI database and 41 fungal genome sequences in JGI, including the 12 fungal genomes tested in this study. Surprisingly, the BLASTP results with an E-value cutoff of 1E-5 showed that

vesiculate-type species do not have *spc33* whereas all imperforate- and perforate-type species have *spc33* (Table 2.3). *spc33* homologs were not found in organisms other than perforate- and imperforate-type species (including other fungi, other eukaryotes, and prokaryotes). These present/absent patterns of *spc33* suggest that *spc33* emerged at least before the divergence of imperforate SPC from vesiculate SPC. Furthermore, no known domain was found in *spc33* based on a Pfam 32.0 search. Hence, it is currently difficult to explain the origin of *spc33* that emerged what kind of gene. Considering that the blast search was not detected before the appearance of imperforate-type species, I might be able to assume that *spc33* had emerged just before the emergence of imperforate SPC. This result supports Padamsee et al. (2012), i.e., different genes but not *spc33* should be involved in the development of SPC in vesiculate-type species such as *Wallemia sebi*.

**Table 2.3.** Gene present/absent pattern of *spc33*.

| | Vesiculate (2) | Imperforate (11) | Perforate (28) |
|---|---|---|---|
| present % | 0 | 100 | 100 |

Numbers in parentheses indicate the number of fungal species tested.

In conclusion, I successfully found the evidence for the correlation of convergence between amino acid sequence evolution and the morphological evolution of perforate SPC. Therefore, I found that an amino acid change in the same sequence of the same gene corresponds to morphological convergent evolution. The methods also make the convergent substitution pattern detectable more easily. My findings contribute to both clarifying the genetic basis of morphological convergent evolution and pioneering genome-wide surveys for understanding fungal evolutionary morphology.

# Chapter 3

# Sequence evolution correlated with morphological convergent evolution from Imperforate SPC to Perforate SPC: Independent parallel Amino Acid substitution in *spc33*

## 3.1 Introduction

Understanding the mechanism of genomic changes associated with phenotypic evolution leads to clarifying ability for producing phenotypic variation (evolvability) and evolutionary history of focusing taxa (Stern and Orgogozo, 2009; Castoe et al., 2010; Hague et al., 2017). The previous chapter showed that parallel changes were found from amino acid sequences of *spc33* in association with morphological convergence of perforate SPC. However, the molecular mechanism of such parallel changes remains unknown.

In this chapter, I investigated the cause of the parallel changes including *Rhizoctonia solani* (Rhiso), acquired perforate SPC independently. I focus on two types of hypotheses, “Independent origin” and “Common origin”. “Independent origin” refers to the mutation that occurred independently in each lineage. Independent emergence of particular key mutations at such a particular gene can lead the gain of the ability to produce similar phenotypes between different lineages (Gompel and Prud'homme, 2009). Therefore, when the evidence of “Independent origin” was detected from the sequence of *spc33*, it can be thought that the evolvability of perforate SPC is higher than imperforate SPC. In contrast, “Common origin” of mutation refers to the same substitution had been shared by the “copy-and-paste” event such as horizontal gene transfer (HGT)

between different lineages. When parallel changes at the sequence of *spc33* had occurred by "Common origin", it can be considered that acquisition of perforate type *spc33* was caused by the movement of perforate type *spc33* between different perforate type lineages. It has been said that HGT takes an important role in rapid evolution and adaptation (Jain et al., 2003). Therefore, in this case, we can suggest that morphological convergent evolution of perforate SPC might be accelerated rather than "Independent origin" evolution.

Here I aimed to clarify whether the molecular mechanism of parallel substitutions in *spc33* is "Independent origin" or "Common origin". To solve this problem, I performed three analyses of *spc33*; synteny analysis, determination of exon/intron structure and phylogenetic analysis of *spc33*. By using 12 fungal genomes used in Chapter 2, I compared synteny to compare the genomic position of *spc33* among species. In this analysis, if the same orthologous groups were detected from *spc33*-containing scaffolds, I judged synteny around *spc33* as conserved. When "Independent origin" is supported, the synteny of perforate type species in different lineages Rhiso should also be conserved. If the conservation pattern of exon/intron structure differs with the same pattern accompanied by species phylogeny, it can be considered that substitution events of *spc33* had occurred independently. Finally, to understand the sequence similarity of the amino acid sequences other than key sites and the influence of key sites on the topology of gene trees, I compared the topology of two *spc33* phylogenetic trees. Here I report that "Independent origin" was well supported based on my analysis.

## 3.2 Materials and methods

*Fungal genomes*

Genomes of 12 fungal species that were used in Chapter 2 were subsequently used in this study.

*Synteny analysis*

I compared the gene arrangement of scaffolds where *spc33* is located. To detect *spc33*-containing scaffold, BLASTP was conducted against whole fungal scaffolds. Whole genes located in the *spc33*-containing scaffold were also extracted based on sequence similarity with CDS by BLASTP search. Orthologous gene clusters of the detected genes were constructed by using OrthoMCL (Li et al., 2003) with default parameters. Then, I compared the order and combination of orthologs around *spc33* among species to check the preservability of synteny. The data was visualized by R.

*Determination of exon/intron structure*

In *spc33* orthologs, each exon/intron structure was determined by comparing DNA sequences, mRNA sequences and CDS sequences of *spc33* using clustalW and JGI genome browser. After detecting exons and introns, each gene structure was compared and visualized by R.

*Examine sequence similarity of spc33 other than key sites*

Two types of phylogenetic trees were compared. One is the tree that was removed three key sites detected in Chapter 2 from multiple alignments of *spc33*. Another tree was not removed any sites before phylogenetic analysis. Gene trees were constructed by the maximum-likelihood method using the estimated model as the best amino acid substitution model LG+G calculated in MEGA7 (Kumar et al., 2016).

## 3.3 Results

*Conservation pattern of synteny*

To compare the synteny around *spc33* among 12 fungal genomes, orthologous gene clusters were obtained from *spc33*-containing scaffolds of each fungus. Figure 3.1 shows the synteny of the orthologs detected from eight to twelve species located between 60-kb upstream and 60-kb downstream of *spc33*. The result shows that the orthologous groups were detected within this region. Even Rhiso, which acquired perforate SPC independently has the same members of orthologs, indicates that the same syntenic region exists among these scaffolds.

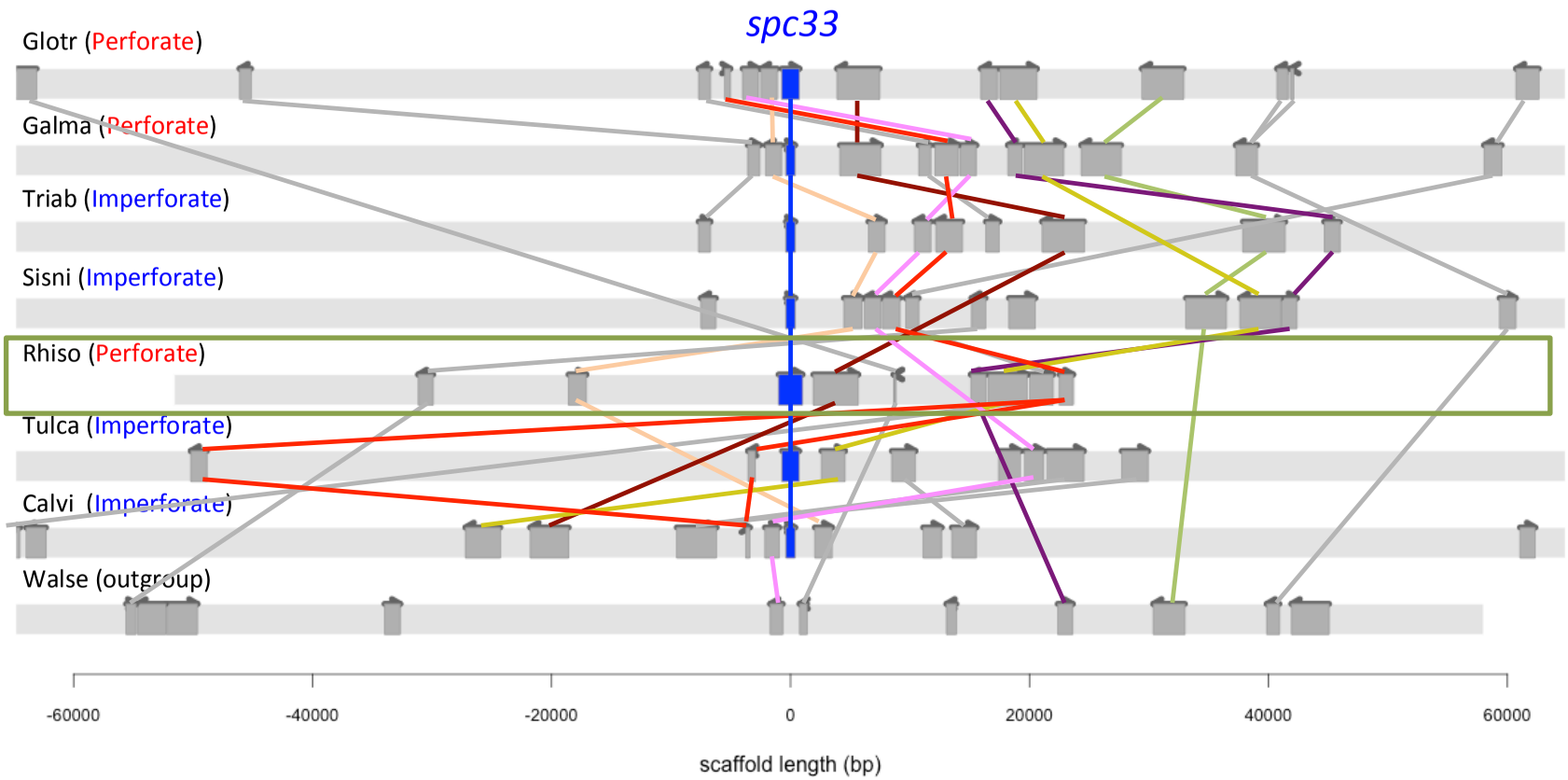



**Figure 3.1.** Synteny aound *spc33*.

*Exon/intron organization and its difference among 12 fungal species*

I predicted the exon/intron structure of *spc33* and compared them among species. I found that the exon/intron pattern of *spc33* was changed accompanied by speciation (Figure 3.2). Half of the species have two exons, whereas other species have three or four exons. The First exon of Rhiso and Tulca codes only start codon. The first exon (second exon in the case of Rhiso and Tulca) shows the most conservation of gene length. However, Rhiso acquired an extremely long second exon. Since *spc33* of Rhiso has short 5' UTR and the amino acid sequences show homology from the middle region of the sequence with the other species (Figure 3.3), Rhiso can be considered to have newly acquired the transcription initiation region on the 5' region of *spc33*.

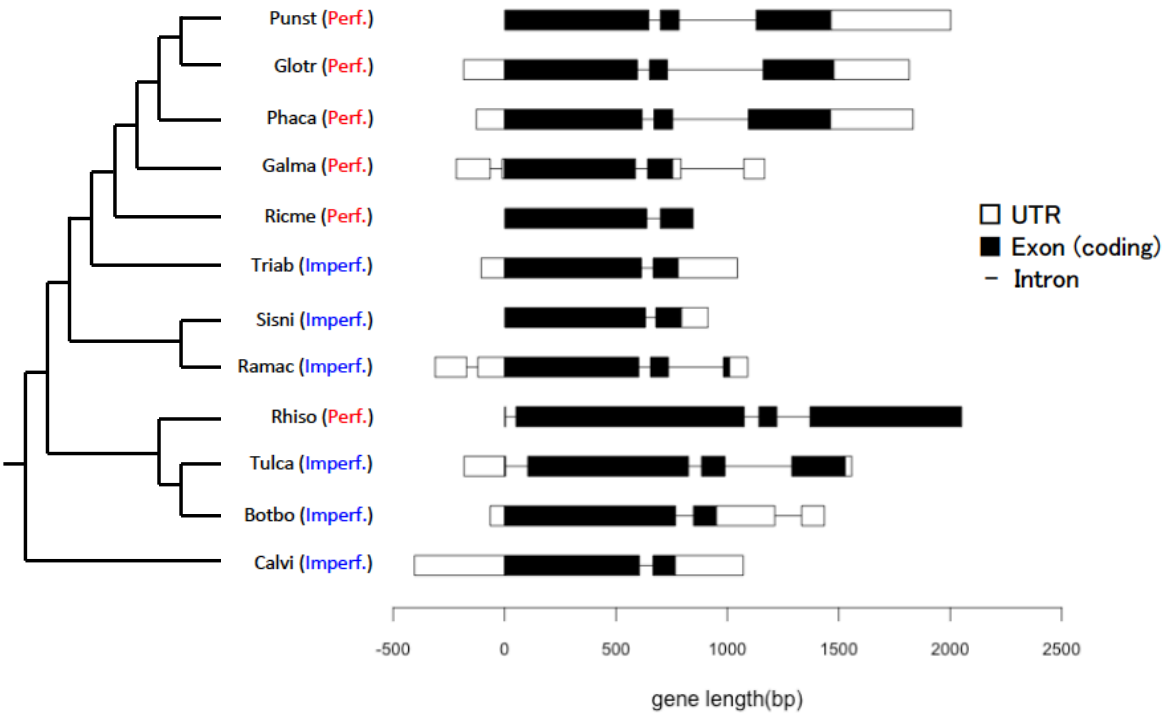



**Figure 3.2.** Comparison of intron and exon lengths in *spc33*.

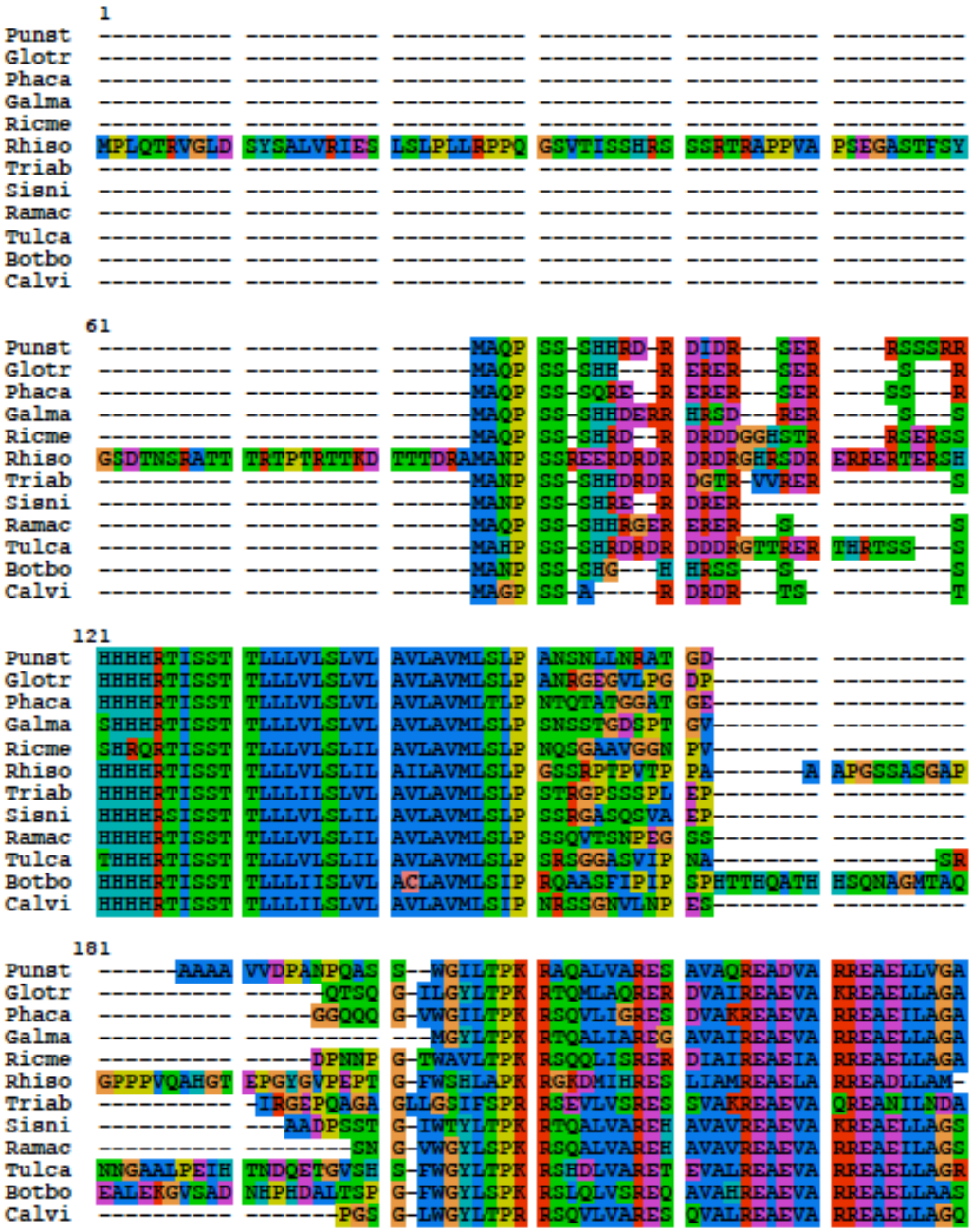



**Figure 3.3.** Multiple alignment of 5' region of *spc33*.

*Difference of topology of spc33 gene tree between with and without key sites of multiple alignments for constructing the tree*

To examine the sequence similarity of the amino acid sequences other than key sites, I constructed two gene trees, one with key sites and the other without key sites. Interestingly, two gene trees showed different tree topologies. The topology of the tree using all amino acids was clustered by morphological traits, as shown in Chapter 2 (Figure 3.4a). On the other hand, the gene tree constructed without three key sites was clustered by species phylogeny (Figure 3.4b). Tree topology was changed just removing three sites. These results suggest that these key sites correspond to the change of topology. Thus, the key sites lead the characterization correlated in morphological differences of SPC in *spc33*.

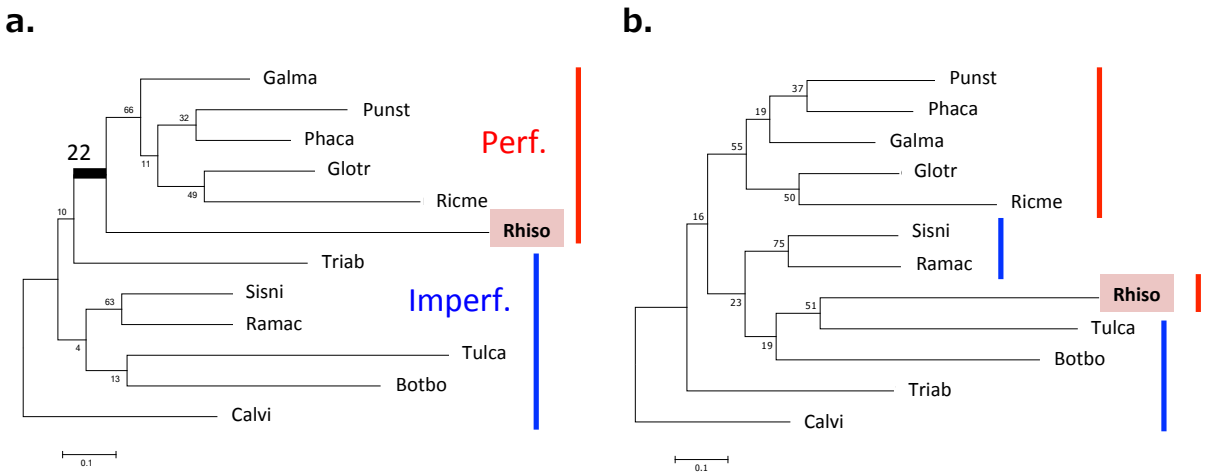



**Figure 3.4.** Phylogenetic gene tree of *spc33*. (a.) gene tree with 3 key sites. (b.) gene tree without 3 key sites.

## 3.4 Discussion

In this chapter, in order to clarify whether the molecular mechanism of parallel substitutions in *spc33* is “Independent origin” or “Common origin”, synteny analysis, comparison of exon/intron structure and phylogenetic analysis were performed. My analysis showed that all tests supported the “Independent origin” origin, meaning parallel substitutions in *spc33* were newly acquired in each perforate type lineages. Below, I will discuss the possible evolutionary event occurred for the sequence evolution of *spc33* and morphological convergence of perforate SPC.

The result of synteny analysis showed that gene repertory around *spc33* is conserved among species (Figure 3.1). This finding suggests that the gene *spc33* had gained at the common ancestor of perforate type species and imperforate type species, and the presence of *spc33* has been maintained in the Rhiso genome. Thus, the possibility of HGT can be denied. This possibility was also denied by the existence of introns in Rhino's *spc33* (Figure 3.2). However, the possibility of gene conversion after HGT is still remained. Here the result of the phylogenetic analysis showed that the *spc33* gene tree constructed without three key sites was clustered by species phylogeny (Figure 3.4b). Therefore, based on the synteny analysis and the phylogenetic analysis, the possibility of gene conversion after HGT was denied. Therefore, this data supports “Independent origin” as a cause of parallel substitutions in *spc33*. On the contrary, if parallel substitutions of *spc33* were caused by “Common origin”, synteny should not be conserved. When the perforate type *spc33* was acquired by interspecific gene transfer e.g. HGT, it is unlikely that *spc33* was inserted into the actual same homologous position of the genome of the different lineage. Therefore, I can reject the possibility of “Common origin” as a genetic mechanism of convergent substitutions of *spc33*. It is interesting whether the acquisition of

these substitutions is adaptive or not. The problem of whether restrictions of amino acid change of *spc33* was caused by the phenotypic level or the molecular level is also an attractive issue to be solved.

I found the evidence of gene shuffling within the species, but not found the evidence of the gene transfer between species. The combinations of orthologs were highly conserved, while the order of the orthologs at scaffolds was quite shuffled among species. This fact suggests that homologous recombination within the species frequently occurred. It is known that simple sequence repeat (SSR) promotes recombination within species (Levinson and Gutman 1987). Actually, abundant SSRs were found from *spc33*-containing scaffolds by using RepeatMasker 4.0.8 (http://www.repeatmasker.org) (Supplemental Table 2). On the other hand, when I tried to detect the evidence of interspecific recombination events, no signals such as transposable elements were found from species comparison. These results suggest that gene shuffling （Figure 3.1） might be caused by SSR and its recombination, not by HGT.

The amount and length of exons and introns showed similar pattern among closely related species (Figure 3.2). It indicates that this pattern had been changed through the species divergence rather than the morphological evolution of SPC. Since key sites are located on the 3'-terminal side of exon 1, the main function of *spc33* can be supposed to be in exon 1. The length of Rhiso's *spc33* became longer by sifting translational start sites to the upstream region (Figure 3.3). Further investigation is required for understanding whether or not this region causes functional changes in Rhiso's *spc33*.

I found that the topology of *spc33* gene tree changes just by deleting three amino acid key sites (Figure 3.3). Amino acid sites other than the key sites of Rhiso's *spc33* were shown to share similar substitution patterns with speciation. Phylogenetic analysis using the partial-deletion option also

supported the similar substitution pattern accompany with species phylogeny (Supplemental Figure 2). Since the length of *spc33* is 183 aa (complete deletion), much greater than three, these data suggest that key sites play an important role in the morphological convergence of perforate SPC. Further analysis is required for investigating whether or not key sites have a great influence on the properties of protein SPC33.

In conclusion, it became clear that the hypothesis "Independent origin" was supported as the cause of the parallel changes in amino acid sequences of *spc33*. Thus, I propose that the perforate SPC was independently acquired in each convergent lineage from imperforate SPC without a lateral genetic transfer event. This finding provides insights into the molecular mechanism and repeatability of the morphological convergence of perforate SPC.

# Chapter 4

# Sequence evolution correlated with morphological evolution from vesiculate SPC to imperforate SPC: Estimation of the origin of *spc33*

## 4.1 Introduction

In the morphological evolution, examining mutations in gene level provides huge clues for understanding the causes of related changes in phenotypic level. A protein-coding gene *spc33* is essential for SPC formation in perforate type species (Peer et al., 2010), indicating that this gene has a large clue in clarifying the genetic basis of morphological evolution of SPC. My previous study showed *spc33* is found in genomes of imperforate type species and perforate type species, but not in genomes of vesiculate type species (Chapter2). Since *spc33* expresses in imperforate type species, it can be considered that *spc33* had appeared and started to function with the acquisition of imperforate SPC. However, the origin of the *spc33* sequence remains unknown.

In this chapter, I therefore aimed to elucidate the ancestral sequence/motif of *spc33* for clarifying the origin of *spc33*. To define an ancestral gene, homologous sequences with *spc33* were detected by multiple bioinformatics analysis tools that have been suggested as effective approaches (Tautz and Domazet-Lošo, 2011). For inferring an ancestral gene by detecting the homologous region including novel gene in the focusing lineage and ancestral gene in the outgroup, synteny among vesiculate, imperforate and perforate SPC type species were examined. Since a gene must have an ancestral sequence and/or functional motif, genes located in the same region among related species are likely to be ancestral sequences.

Therefore, I predicted that if the outgroup also has the same orthologs with upper lineages, there should be the original sequences of the novel gene *spc33*. Another possibility to find ancestral sequences is based on sequence similarity with the novel gene. If the sequence that has similarity with *spc33* was detected from the outgroup genome, the sequence is the candidate ancestral gene. Thus, I conducted sequence similarity search of *spc33* by using BLAST (Altschul et al., 1990) and its derivative PSI-BLAST (Altschul et al., 1997).

Here I showed that MRCK Ser/Thr protein kinase is the ancestral gene of *spc33*. I also discussed the origin of function and its evolution of *spc33*, especially how the ancestral gene had been evolved to *spc33* at the molecular level by comparing amino acid sequences with *spc33*.

## 4.2 Material and method

*Fungal genomes*

Genomes of three fungal species *Wallemia sebi* (Walse, vesiculate), *Calocera viscosa* (Calvi, Imperforate), and *Rhizoctonia solani* AG-1 I B (Rhiso, Perforate) were used in this study. The genome of vesiculate type species Walse was collected from JGI, as same as other genomes described in Chapter 2.

*BLAST search and Pfam analysis*

Homologous sequences of *spc33* were searched by using BLAST to detect the gene family of *spc33*. Calvi, the most basal species among perforate and imperforate type species was used as a query. The result of BLASTP search of *spc33* against NCBI database with the E-value cutoff of $10^{-5}$ was retrieved. Pfam search was conducted to find a domain within *spc33*. *spc33* of Calvi was searched against PFAM 31.0 database with the

e-value cutoff of 1.0.

*PSI-BLAST search*

To find a sequence with the position-specific similarity of *spc33* from vesiculate type species, PSI-BLAST against NCBI database was carried out. *spc33* of Calvi was used as the original query. Other *spc33* were not added into a query for further blast even if these homologs hit at the 1st or later blast to prevent new hits are occupied by *spc33* homologs. The result of PSI-BLAST with the e-value cutoff of $10^{-5}$ was retrieved.

*Synteny analysis*

To detect the region where *spc33* had been gained while the evolution from vesiculate type species to imperforate type species, synteny analysis was performed. First, I compared synteny between Walse and Calvi to detect homologous scaffold using VistaDot and the Vista synteny viewer (Frazer et al., 2004) in JGI. Orthologs located in homologous scaffolds of three species were detected by the method as described in Chapter 3. The homologous region including *spc33* from Rhiso and Calvi was identified in Walse. Then, the ancestral gene located in the homologous region of *spc33* was extracted from Walse genome.

*Identifying orthologs and inferring phylogenetic relationship of ancestral gene*

To count the number of homologs of the ancestral gene in fungal genomes, I conducted BLAST search against 12 fungal genomes as described in Chapter 2. The result of BLASTP search of the ancestral gene against fungal genomes with an E-value cutoff of $10^{-5}$ was retrieved. After that, phylogenetic analysis was carried out to examine the phylogenetic relationship between the ancestral gene and *spc33*. Multiple alignment was

constructed by MUSCLE (Edger, 2004). The gene tree was constructed by the maximum-likelihood method using the best model LG+G calculated in the previous chapter.

## 4.3 Results

*Homology and motif search of spc33*

To identify genes belonging to the same gene family as *spc33*, BLAST search and Pfam search were conducted. Surprisingly, any homolog and functional motif were detected. The N-terminal region of *spc33* was predicted to be a transmembrane region as described in Peer et al. (2010). No *spc33* gene family was identified in vesiculate type species, non-SPC fungi, other eukaryotes, and prokaryotes.

*Ancestral gene of spc33*

The ancestral gene of *spc33* was predicted by using PSI-BLAST and synteny analysis. Since the ancestral sequence of *spc33* was not detected by the direct search of sequence similarity, PSI-BLAST was carried out. In the result, PSI-BLAST detected protein kinase (SYZ62991.1) and SMC family ATPase (WP_071017341.1) from non-fungal species (Table 4.1). The resulted e-value by PSI-BLAST was 3E-18 for both genes. Interestingly, protein kinase was detected the same as PSI-BLAST (Figure 4.1). Any other gene was not detected as the candidate from the ancestral region in Walse.

**Table 4.1.** List of probable ancestral genes of *spc33* calculated by PSI-BLAST.

| Name | Accetion No. | Score | E-value |
|---|---|---|---|
| protein_kinase | SYZ62991.1 | 94.1 | 3E-18 |
| SMC family ATPase | WP_071017341.1 | 94.1 | 3E-18 |

The protein kinase detected by synteny analysis was KELK-motif containing myotonic dystrophy kinase-related Cdc42-binding kinase (MRCK) Ser/Thr protein kinase. On the other hand, the protein kinase detected by PSI-BLAST also contained the functional motif of Ser/Thr protein kinase (Table 4.2). These results indicate the ancestral gene of *spc33* is a Ser/Thr protein kinase. Interestingly, I found this kinase does not have the transmembrane region located in *spc33*. This fact suggests that *spc33* had newly acquired a transmembrane domain through the evolution from Ser/Thr protein kinase to *spc33*.

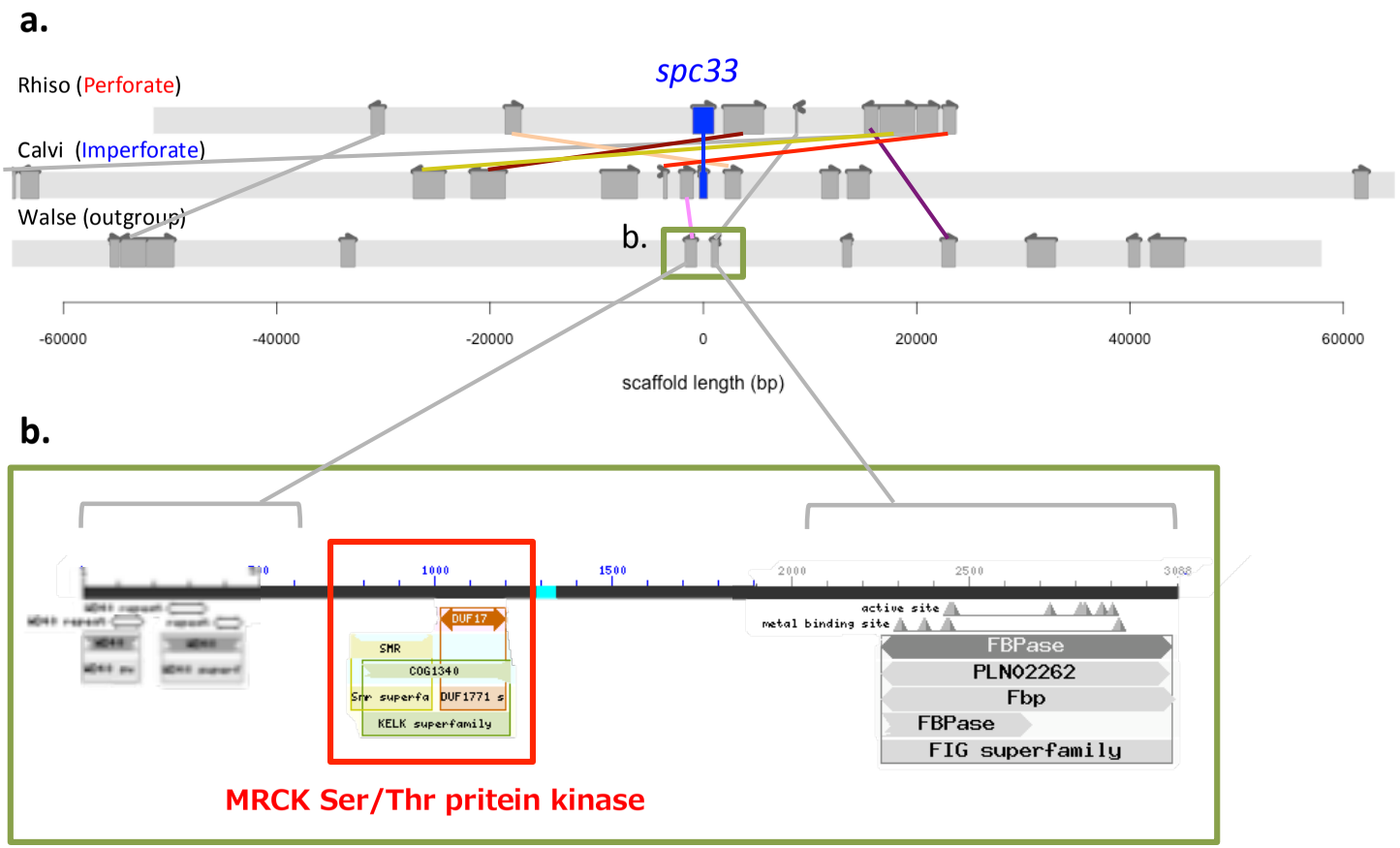



**Figure 4.1.** Result of synteny analysis. a.) synteny between Rhiso, Calvi and Walse. spc33 was shoun in blue. Other orthologous groups were shown in gray. The green box indicates the ancestral region where spc33 have been gained in Walse. b.) probable ancestral gene detected by synteny analysis.

**Table 4.2.** Functional motifs detected from protein kinase SYZ62991.1.

| Name | Description | Interval |
|---|---|---|
| PHA03247 | large tegument protein UL36 | 70-443 |
| STKc_PAK | Catalytic domain of the Serine/Threonine Kinase, p21-activated kinase | 1162-1266 |
| PHA03207 | serine/threonine kinase US3 | 1303-1498 |
| PTKc_PDGFR_beta | Catalytic domain of the Protein Tyrosine Kinase, Platelet Derived Growth Factor Receptor beta | 1366-1501 |
| PKc_MAPKK_plant_like | Catalytic domain of Plant dual-specificity Mitogen-Activated Protein Kinase Kinases and similar protein | 1438-1746 |
| Pkinase | Protein kinase domain | 1443-1744 |
| STKc_obscurin_rpt1 | Catalytic kinase domain, first repeat, of the Giant Serine/Threonine Kinase Obscurin | 1450-1583 |
| S_TKc | Serine/Threonine protein kinases, catalytic domain | 1451-1744 |
| SPS1 | Serine/threonine protein kinase | 1451-1791 |
| PTZ00024 | cyclin-dependent protein kinase | 1461-1503 |

*Protein family of Ser/Thr protein kinase and its phylogenetic relationship*

Homologs of MRCK Ser/Thr protein kinase were detected from 12 fungal genomes. The result of BLAST search showed each fungus has 2~11 homologs within the genome (Figure 4.2). Walse also has three homologs of MRCK Ser/Thr protein kinase. None of the kinase homologs were located in the same scaffold as *spc33*. The gene tree containing these homologs and *spc33* clearly showed separate clusters of *spc33* and kinases (Figure 4.3). The *spc33* orthologs in imperforate and perforate type species ware firmly clustered and were distinguished from all other kinases with a long branch length. The phylogenetic relationship of kinases was not clustered by species.

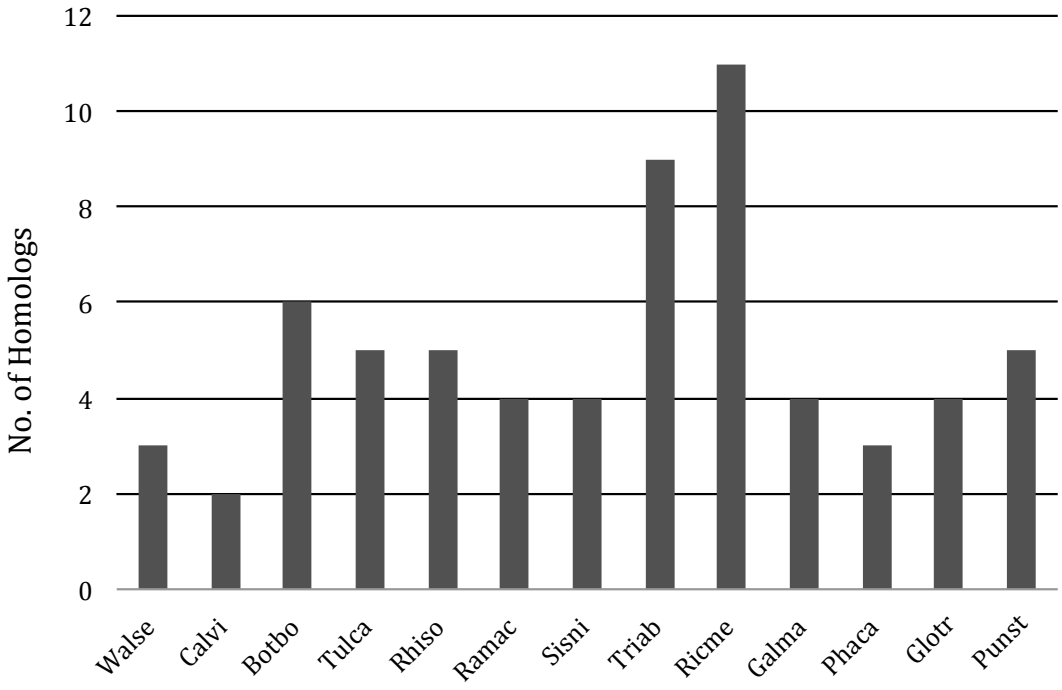



**Figure 4.2.** Amount of MRCK Ser/Thr protein kinase homologs in 13 fungal species.

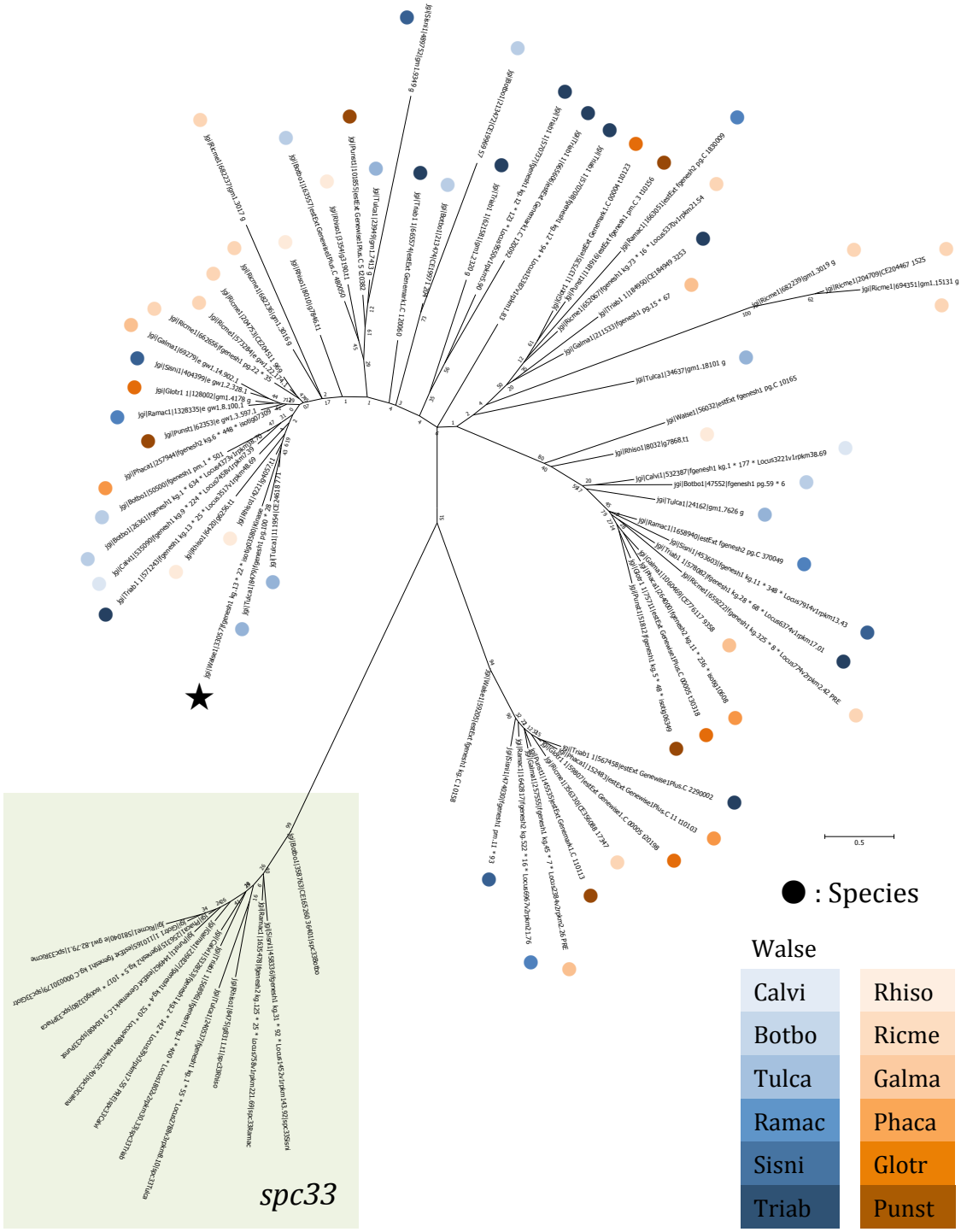



**Figure 4.3.** Phylogenetic tree of Ser/Thr protein kinase homologs and *spc33*. Each color circles in kinases indicate species of each sequences. MRCK Ser/Thr protein kinase detected by synteny analysis is marked with black star.

## 4.4 Discussion

Ser/Thr protein kinase was detected as an ancestral gene of *spc33* based on two completely different approaches synteny analysis (Figure 4.1) and PSI-BLAST (Table 4.1). Kinases are enzymes that chemically add phosphate groups to protein molecules. The family of eukaryotic protein kinases accounts for about 2% of the protein-coding genes of the entire human genome (Mannig et al., 2002). One of them MRCK Ser/Thr protein kinase that was detected from the vesiculate type species is known to be involved in actin-myosin regulation and activities (Unbekandt and Olson 2014). Besides, the protein SPC33 is localized on the surface of SPCs and is in contact with actin filaments through cytoplasmic communication (Moore and Marchant 1972). These findings suggest that the actin regulatory factor that possessed by the ancestral gene of *spc33*, MRCK Ser/Thr protein kinase may also affect the function of *spc33*. Interestingly, I found that three key amino acid sites of *spc33* shown in Chapters 2 and 3 were aligned with the functional motif of the MRCK Ser/Thr protein kinase (Supplemental Figure 3). This finding suggests that this short region of *spc33* may also have kinase activity or a derived function close to the kinase. In particular, although the pores of Perforate SPC pass actin filaments, the imperforate SPC does not have such holes, and does not pass actin filaments (Moore and Marchant, 1972). It is very interesting to investigate if this difference was due to an actin regulator derived from MRCK Ser/Thr protein kinase.

The appearance of *spc33* from the kinase means, in other words, one gene as the kinase lost the kinase function. If the gene loss occurs in a single copy gene in the genome, the influence of phenotypic change will be significant. On the other hand, if the gene has many homologs and a mechanism for compensating the function is provided, the effect on the phenotype is likely to be small even if gene loss occurs once. In this study, I

showed that many homologs of MRCK Ser/Thr protein kinase were found in the genomes of all tested fungi, including perforate, imperforate and vesiculate type species (Figure 4.2). This suggests that even if one kinase was changed to *spc33*, the degree of phenotypic change by the functional change of the kinase is not large.

Phylogenetic analysis showed that *spc33* is highly differentiated from the sequence of the original kinase (Figure 4.3). This implies that the kinase has gotten many mutations through the changing to *spc33*. Interestingly, transmembrane region was not found from amino acid sequences of MRCK Ser/Thr protein kinase even though *spc33* contains the transmembrane region (Peer et al., 2010). It also suggests that a major evolutionary event has occurred that alters the protein structure. Liu et al. (2004) showed that domain recombination is not common for the transmembrane regions of membrane proteins, a majority of integral membrane proteins containing only a single transmembrane domain. This fact suggests that the transmembrane domain in *spc33* had been gained by sequence insertion or substitution rather than recombination.

In conclusion, I successfully identified the ancestral gene of *spc33* that appeared with the acquisition of Imperforate SPC and obtained clues about how *spc33* evolved from the ancestral gene due to changes in sequence and function.

# Chapter 5

## Conclusions

In my doctoral thesis, I have aimed to clarify genomic changes associated with the morphological evolution of SPC in terms of comparative genomics. As a result of investigating the amino acid substitution pattern of orthologous genes among 12 fungal species, SPC-forming gene *spc33* was identified as the gene that most reflects the evolutionary history of SPC (Chapter 2). Gene present/absent pattern, amino acid changes and the molecular background of its genetic changes of *spc33* connected to both two patterns of evolution; the morphological difference between vesiculate SPC and imperforate SPC (Chapter 4), and the morphological convergence from imperforate SPC to perforate SPC in multiple lineages (Chapter 3). Below, I will summarize detected genomic changes and possible evolutionary history in each morphological evolution.

In the morphological difference between vesiculate SPC and imperforate SPC, I clarified that the gene gain had occurred with the morphological differentiation. I found vesiculate type species do not have *spc33* that is essential to form SPC in perforate type species. Furthermore, by analyzing synteny around *spc33* and site-specific similarity of *spc33*, MRCK Ser/Thr protein kinase in vesiculate type species was successfully detected as an ancestral gene of *spc33*. These findings indicate that vesiculate SPC has a different formation process. Since this kinase has a function related to actin-myosin regulation, and one of the remarkable differences between imperforate SPC and perforate SPC is the existence of perforation that allows actin filaments to pass through the SPC, it is possible that *spc33* interacts with actin filament located in the vicinity of SPC.

In the evolution from imperforate SPC to multiple lineages of perforate SPC, I discussed what kind of mutation occurred in the amino acid sequences of *spc33* and how *spc33* functions to cause the morphological convergence. From the sequences of *spc33*, five substitutions that correlated to the morphological parallel evolution of perforate SPC were detected. Among these sites, three key sites showed convergent changes that are related to the morphological convergence. In these substitutions, the possibility of "Common origin" e.g. horizontal gene transfer was denied by conducting synteny analysis, exon/intron pattern inference and constructing gene trees with and without key substitutions. Therefore, I conclude that perforate SPC was independently acquired in each lineage from common traits imperforate SPC without a lateral genetic transfer event. The amount of these changes was significantly higher than other orthologs. Therefore, it is possible that parallel substitutions in *spc33* occurred not randomly but accordingly occurred with morphological evolution. Moreover, three key sites were localized within 17 amino acid residues that are the homologous region with the actin-myosin regulatory region of MRCK Ser/Thr protein kinase. This finding suggests the new hypothesis that the key sites-containing region was derived from the actin-myosin regulatory region of MRCK Ser/Thr protein kinase and be involved in the important function of *spc33*.

From my genome studies of SPC, future prospects have extended to clarify the genetic basis of the morphological evolution of SPC. To verify the involvement of *spc33* to morphological differentiation of SPC, conducting in vivo functional assay is the future important challenge to be solved. In particular, *spc33* gene replacement between perforate type species and imperforate type species is one of the best approaches to tackle the problem. However, I found no promoter was suitable to complete these experiments. Knockout experiments are also challenging because it is suggested that Δ *spc33* strain shows lethal phenotype (Peer 2009, Doctoral thesis). Therefore,

now I am starting the overexpression of various species of *spc33* in the model fungi *Coprinus cinereus* (perforate type). By examining the phenotypic changes of *spc33* transformants, a deeper insight into the genetic basis and mechanisms of the morphological evolution of SPC will be obtained. Molecular mechanisms of the functional differentiation from MRCK Ser/Thr protein to *spc33* are also interesting issues. The function of important motifs in cells should be investigated in more detail to understand whether the actin-myosin regulator of the kinase was taken over and function in *spc33*.

In summary, I identified direct genetic changes correlated with both morphological differences between vesiculate SPC and imperforate SPC, and morphological convergence of from imperforate SPC to perforate SPC. Based on my genome-wide survey, SPC-related gene *spc33* was detected as the candidate gene of the morphological evolution of SPC. The findings of the origin and SPC type-specific amino acid substitutions of *spc33* provide huge clues for understanding the genetic basis of morphological evolution of SPC. These achievements contribute to uncover the molecular mechanisms of convergent evolution and to find a missing link between morphological evolution and sequence evolution. In addition, this study is a leading-edge study in the early days of fungal evolutionary genomics, and will contribute greatly to future fungal evolutionary biology.

## References


Almén, M.S., Lamichhaney, S., Berglund, J., Grant, B.R., Grant, P.R., Webster, M.T., Andersson, L., 2016. Adaptive radiation of Darwin's finches revisited using whole genome sequencing. Bioessays 38, 14–20. https://doi.org/10.1002/bies.201500079

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J Mol Biol. 215, 403-10. https://doi.org/10.1016/S0022-2836(05)80360-2

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389

Bertossa, R.C., 2011. Morphology and behaviour: functional links in development and evolution. Philos Trans R Soc Lond B Biol Sci. 366(1574), 2056-68. https://doi.org/10.1098/rstb.2011.0035

Bork, P., Dandekar, T., Diaz-Lazcoz, Y., Eisenhaber, F., Huynen, M., Yuan, Y., 1998. Predicting function: from genes to genomes and back. J. Mol. Biol. 283, 707–725. https://doi.org/10.1006/jmbi.1998.2144

Bracker, C.E., Butler, E.E., 1964. Function of the septal pore apparatus in *Rhizoctonia solani* during protoplasmic streaming. J. Cell Biol. 21,152–157. https://doi.org/10.1083/jcb.21.1.152

Castoe, A.T., Jason de Koning, A.P., Pollock, D.D., 2010. Adaptive molecular convergence: Molecular evolution versus molecular phylogenetics. Commun. Integr. Biol. 3, 67–69. https://doi.org/10.4161/cib.3.1.10174

Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G., Hirschman, J.E., Hitz, B.C., Karra, K., Krieger, C.J., Miyasato, S.R., Nash, R.S., Park, J., Skrzypek, M.S., Simison, M., Weng,

S., Wong, E.D., 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. 40, D700–D705. https://doi.org/10.1093/nar/gkr1029

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G. Jr., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D., Kingsley, D.M., 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science 307, 1928–1933. https://doi.org/10.1126/science.1107239

van Driel, K.G.A., van Peer, A.F., Grijpstra, J., Wösten, H.A.B., Verkleij, A.J., Müller, W.H., Boekhout, T., 2008. The Septal Pore Cap Protein SPC18 Isolated from the Basidiomycetous Fungus *Rhizoctonia solani* also Resides in Pore-plugs. Eukaryot. Cell 7, 1865–1873. https://doi.org/10.1128/EC.00125-08

van Driel, K.G.A., Humbel, B.M., Verkleij, A.J., Stalpers, J., Müller, W.H., Boekhout, T., 2009. Septal pore complex morphology in the *Agaricomycotina* (*Basidiomycota*) with emphasis on the *Cantharellales* and *Hymenochaetales*. Mycol. Res. 113, 559–576. https://doi.org/10.1016/j.mycres.2008.12.007

Edger, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

Ekblom, R., Galindo, J., 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107, 1–15. https://doi.org/10.1038/hdy.2010.152

Fell, J.W., Boekhout, T., Fonseca, A., Sampaio, J.P., 2001. Basidiomycetous yeasts. *In* The Mycota VII, Systematics and evolution, Part B. McLaughlin, D.J., McLaughlin, E.G., Lemke, P.A. (eds.), Springer-Verlag, Berlin, Germany. https://doi.org/10.1007/978-3-662-46011-5

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376. https://doi.org/10.1007/BF01734359

Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R.A., Henrissat, B., Martínez, A.T., Otillar, R., Spatafora, J.W., Yadav, J.S., Aerts, A., Benoit, I., Boyd, A., Carlson, A., Copeland, A., Coutinho, P.M., de Vries, R.P., Ferreira, P., Findley, K., Foster, B., Gaskell, J., Glotzer, D., Górecki, P., Heitman, J., Hesse, C., Hori, C., Igarashi, K., Jurgens, J.A., Kallen, N., Kersten, P., Kohler, A., Kües, U., Kumar, T.K., Kuo, A., LaButti, K., Larrondo, L.F., Lindquist, E., Ling, A., Lombard, V., Lucas, S., Lundell, T., Martin, R., McLaughlin, D.J., Morgenstern, I., Morin, E., Murat, C., Nagy, L.G., Nolan, M., Ohm, R.A., Patyshakuliyeva, A., Rokas, A., Ruiz-Dueñas, F.J., Sabat, G., Salamov, A., Samejima, M., Schmutz, J., Slot, J.C., St John, F., Stenlid, J., Sun, H., Sun, S., Syed, K., Tsang, A., Wiebenga, A., Young, D., Pisabarro, A., Eastwood, D.C., Martin, F., Cullen, D., Grigoriev, I.V., Hibbett, D.S., 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336, 1715–1719. https://doi.org/10.1126/science.1221748

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., Dubchak, I., 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 32, W273-9. https://doi.org/10.1093/nar/gkh458

Gompel, N., Prud'homme, B., 2009. The causes of repeated genetic evolution. Dev Biol. 332, 36-47. https://doi.org/10.1016/j.ydbio.2009.04.040

Hague, M.T.J., Feldmon, C.R., Brodie, Jr. E.D., Brodie, III E.D., 2017. Convergent adaptation to dangerous prey proceeds through the same first-step mutation in the garter snake *Thamnophis sirtalis*. Evolution 71, 1504–1518. https://doi.org/10.1111/evo.13244

Hibbett, D.S., 2006. A phylogenetic overview of the Agaricomycotina.

Mycologia 98, 917–925. https://doi.org/10.1080/15572536.2006.11832621

Hibbett, D.S., Bauer, R., Binder, M., Giachini, A.J., Hosaka, K., Justo, A., Larsson, E., Larsson, K.H., Lawrey, J.D., Miettinen, O., Nagy, L.G., Nilsson, R.H., Weiss, M., Thorn, R.G., 2014. Agaricomycetes. The Mycota VII Part A Systematics and Evolution 2nd Edition 14. https://doi.org/10.1007/978-3-642-55318-9

Hibbett, D.S., Thorn, R.G., 2001. Basidiomycota: homobasidiomycetes. *In* The Mycota VII, Systematics and evolution, Part A. McLaughlin, D.J., McLaughlin, E.G., Lemke, P.A. (eds.), Springer-Verlag, Berlin, Germany. https://doi.org/10.1007/978-3-642-55318-9

Hu, Y., Wu, Q., Ma, S., Ma, T., Shan, L., Wang, X., Nie, Y., Ning, Z., Yan, L., Xiu, Y., Wei, F., 2017. Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. PNAS 114, 1081–1086. https://doi.org/10.1073/pnas.1613870114

Jain, R., Rivera, M.C., Moore, J.E., Lake, J.A., 2003. Horizontal gene transfer accelerates genome innovation and evolution. Mol. Biol. Evol. 20, 1598-1602. https://doi.org/10.1093/molbev/msg154

Jedd, G., 2011. Fungal evo-devo: organelles and cellular complexity. Trends Cell Biol. 21 ,12–19. https://doi.org/10.1016/j.tcb.2010.09.001

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., Madden, T.L., 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. 36, W5–W9. https://doi.org/10.1093/nar/gkn201

Kohler, A., Kuo, A., Nagy, L.G., Morin, E., Barry, K.W., Buscot, F., Canback, B., Choi, C., Cichocki, N., Clum, A., Colpaert, J., Copeland, A., Costa, M.D., Dore, J., Floudas, D., Gay, G., Girlanda, M., Henrissat, B., Herrmann, S., Hess, J., Hogberg, N., Johansson, T., Khouja, H.R., LaButti, K., Lahrmann, U., Levasseur, A., Lindquist, E.A., Lipzen, A., Marmeisse, R., Martino, E., Murat, C., Ngan, C.Y., Nehls, U., Plett, J.M., Pringle, A., Ohm, R.A., Perotto, S., Peter, M., Riley, R., Rineau, F.,

Ruytinx, J., Salamov, A., Shah, F., Sun, H., Tarkka, M., Tritt, A., Veneault-Fourrey, C., Zuccaro, A., Tunlid, A., Grigoriev, I.V., Hibbett, D.S., Martin, F., 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nature Genet. 47, 410–415. https://doi.org/10.1038/ng.3223

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 33, 1870-4. https://doi.org/10.1093/molbev/msw054

Levinson, G., Gutman, G.A., 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. 4, 203–221. https://doi.org/10.1093/oxfordjournals.molbev.a040442

Li, L., Stoeckert, C.J. Jr., Roos, D.S., 2003. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Res. 13, 2178-2189. https://doi.org/10.1101/gr.1224503

Lisker, N., Katan, J., Henis, Y., 1975. Scanning electron microscopy of the septal pore apparatus of *Rhizoctonia solani*. Can. J. Bot. 53, 1801–1804. https://doi.org/10.1139/b75-209

Liu, Y., Gerstein, M., Engelman, D.M., 2004. Transmembrane protein domains rarely use covalent domain recombination as an evolutionary mechanism. PNAS 101, 3495–3497. https://doi.org/10.1073/pnas.0307330101

Lutzoni, F., Kauff, F., Cox, C.J., McLaughlin, D., Celto, G., Dentinger, B., Padamsee, M., Hibbett, D., James, T.Y., Baloch, E., Grube, M., Reeb, V.R., Hofstetter, V.R., Schoch, C., Arnold, A.E., Miadlikowska, J., Spatafora, J., Johnson, D., Hambleton, S., Crockett, M., Shoemarker, R., Sung, G., Lücking, R., Lumbsch, T., O'Donnell, K., Binder, M., Diederich, P., Ertz, D., Gueidan, C., Hansen, K., Harris, R.C., Hosaka, K., Lim, Y., Matheny, B., Nishida, H., Pfister, D., Rogers, J., Rossman, A., Schmitt, I., Sipman, H., Stone, J., Sugiyama, J., Yahr, R., Vilgalys, R., 2004.

Assembling the fungal tree of life: Progress, classification, and evolution of subcellular traits. Am. J. Bot. 91, 1446–1480. https://doi.org/10.3732/ajb.91.10.1446

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., Sudarsanam, S., 2002. The protein kinase complement of the human genome. Science 298, 1912-34. https://doi.org/10.1126/science.1075762

McLaughlin, D.J., Frieders, E.M., Lü, H., 1995. A microscopist's view of heterobasidiomycete phylogeny. Stud. Mycol. 38, 91-109.

Moore, R.T., Marchant, R., 1972. Ultrastructural characterization of the basidiomycete septum of *Polyporus biennis*. Can. J. Bot. 50, 2463–2469. https://doi.org/10.1139/b72-317

Moore, R.T., 1975. Early ontogenetic stages in dolipore/parenthesome formation in *Polyporus biennis*. J. Gen. Microbiol. 87, 251–259. https://dx.doi.org/10.1099/00221287-87-2-251

Müller, W.H., Montijn, R.C., Humbel, B.M., van Aelst, A.C., Boon, E.J.M., van der Krift, T.P., Boekhout, T., 1998a. Structural differences between two types of basidiomycete septal pore caps. Microbiology 144, 1721–1730. https://doi.org/10.1099/00221287-144-7-1721

Müller, W.H., Stalpers, J.A., van Aelst, A.C., van der Krift, T.P., Boekhout, T., 1998b. Field emission gun-scanning electron microscopy of septal pore caps of selected species in the Rhizoctonia s.l. complex. Mycologia 90, 170–179. https://doi.org/10.1080/00275514.1998.12026896

Müller, W.H., Humbel, B.M., van Aelst, A.C., van der Krift, T.P., Boekhout, T., 1999. The perforate septal pore cap of basidiomycetes. Pp. 120–127. In Plasmodesmata. Structure, function, role in cell communication. van Bel AJE, van Kesteren WJP (eds.), Springer-Verlag, Berlin, Germany. https://doi.org/10.1007/978-3-642-60035-7

Müller, W.H., Stalpers, J.A., van Aelst, A.C., de Jong, M.D.M., van der Krift, T.P., Boekhout, T., 2000. The taxonomic position of *Asterodon*,

*Asterostroma* and *Coltricia* inferred from the septal pore cap ultrastructure. Mycol. Res. 104, 1485–1491. https://doi.org/10.1017/S0953756200002677

Nagy, L.G., Riley, R., Tritt, A., Adam, C., Daum, C., Floudas, D., Sun, H., Yadav, J.S., Pangilinan, J., Larsson, K.H., Matsuura, K., Barry, K., Labutti, K., Kuo, R., Ohm, R.A., Bhattacharya, S.S., Shirouzu, T., Yoshinaga, Y., Martin, F.M., Grigoriev, I.V., Hibbett, D.S., 2016. Comparative Genomics of Early-Diverging Mushroom-Forming Fungi Provides Insights into the Origins of Lignocellulose Decay Capabilities. Mol. Biol. Evol. 33, 959–970. https://doi.org/10.1093/molbev/msv337

Natarajan, C., Hoffmann, F.G., Weber, R.E., Fago, A., Witt, C.C., Storz, J.F., 2016. Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. Science 354, 336–339. https://doi.org/10.1126/science.aaf9070

Nei, M., Zhang, J., Yokoyama, S., 1997. Color vision of ancestral organisms of higher primates. Mol. Biol. Evol. 14, 611–618. https://doi.org/10.1093/oxfordjournals.molbev.a025800

Nguyen, T.A., Cissé, O.H., Wong, J.Y., Zheng, P., Hewitt, D., Nowrousian, M., Stajich, J.E., Jedd, G., 2017. Innovation and constraint leading to complex multicellularity in the Ascomycota. Nat. Commun. 8, 14444. https://doi.org/10.1038/ncomms14444

Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G.D., Maltsev, N., 1999. The use of gene clusters to infer functional coupling. Proc. Natl. Acad. Sci. USA 96, 2896–2901. https://doi.org/10.1073/pnas.96.6.2896

Oberwinkler, F., Riess, K., Bauer, R., Kirschner, R., Garnica, S., 2013. Taxonomic re-evaluation of the *Ceratobasidium-Rhizoctonia* complex and *Rhizoctonia butinii*, a new species attacking spruce. Mycol. Progress 12, 763–776. http://dx.doi.org/10.1007/s11557-013-0936-0

Padamsee, M., Kumar, T.K., Riley, R., Binder, M., Boyd, A., Calvo, A.M.,

Furukawa, K., Hesse, C., Hohmann, S., James, T.Y., LaButti, K., Lapidus, A., Lindquist, E., Lucas, S., Miller, K., Shantappa, S., Grigoriev, I.V., Hibbett, D.S., McLaughlin, D.J., Spatafora, J.W., Aime, M.C., 2012. The genome of the xerotolerant mold *Wallemia sebi* reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. Fungal Genet. Biol. 49, 217–226. https://doi.org/10.1016/j.fgb.2012.01.007

Patton, A.M., Marchant R., 1978. A Mathematical Analysis of Dolipore/Parenthesome Structure in Basidiomycetes. J. Gen. Microbiol. 109, 335-349. https://doi.org/10.1099/00221287-109-2-335

Payne, J.L., Wagner, A., 2019. The causes of evolvability and their evolution. Nature Reviews Genetics 20, 24–38. https://doi.org/10.1038/s41576-018-0069-z

van Peer, A., Wang, F., van Driel, K., Jong, J., Donselaar, E., Müller, W., Boekhout, T., Lugones, L., Wösten, H., 2010. The septal pore cap is an organelle that functions in vegetative growth and mushroom formation of the wood-rot fungus *Schizophyllum commune*. Environ. Microbiol. 12, 833–844. https://doi.org/10.1111/j.1462-2920.2009.02122.x

Riley, R., Salamov, A.A., Brown, D.W., Nagy, L.G., Floudas, D., Held, B.W., Levasseur, A., Lombard, V., Morin, E., Otillar, R., Lindquist, E.A., Sun, H., LaButti, K.M., Schmutz, J., Jabbour, D., Luo, H., Baker, S.E., Pisabarro, A.G., Walton, J.D., Blanchette, R.A., Henrissat, B., Martin, F., Cullen, D., Hibbett, D.S., Grigoriev, I.V., 2014. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. Proc. Natl. Acad. Sci. USA 111, 9923–9928. https://doi.org/10.1073/pnas.1400592111

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Shi, Y., Yokoyama, S., 2003. Molecular analysis of the evolutionary significance of ultraviolet vision in vertebrates. Proc. Natl. Acad. Sci. USA 100, 8308–8313. https://doi.org/10.1073/pnas.1532535100

Stern, D.L., 2013. The genetic causes of convergent evolution. Nat. Rev. Genet. 14, 751–764. https://doi.org/10.1038/nrg3483

Stern, D.J., Orgogozo, V., 2009. Is Genetic Evolution Predictable? Science 323, 746–751. https://doi.org/10.1126/science.1158997

Suzuki, H., MacDonald, J., Syed, K., Salamov, A., Hori, C., Aerts, A., Henrissat, B., Wiebenga, A., VanKuyk, P.A., Barry, K., Lindquist, E., LaButti, K., Lapidus, A., Lucas, S., Coutinho, P., Gong, Y., Samejima, M., Mahadevan, R., Abou-Zaid, M., de Vries, R.P., Igarashi, K., Yadav, J.S., Grigoriev, I.V., Master, E.R., 2012. Comparative genomics of the white-rot fungi, *Phanerochaete carnosa* and *P. chrysosporium*, to elucidate the genetic basis of the distinct wood types they colonize. BMC Genomics 13, 444. https://doi.org/10.1186/1471-2164-13-444

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30, 2725–2729. https://doi.org/10.1093/molbev/mst197

Tatusov, R.L., Koonin, E.V., Lipman, D.J., 1997. A genomic perspective on protein families. Science 278, 631–637. https://doi.org/10.1126/science.278.5338.631

Tautz, D., Domazet-Lošo, T., 2011. The evolutionary origin of orphan genes. Naturereviews 12, 692-702. https://doi.org/10.1038/nrg3053

The Gene Ontology Consortium, 2000. Gene Ontology: tool for the unification of biology. Nature Genet. 25, 25–29. https://doi.org/10.1038/75556

The UniProt Consortium, 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 45, D158–D169. https://doi.org/10.1093/nar/gkw1099

Unbekandt, M., Olson, M.F., 2014. The actin-myosin regulatory MRCK

kinases: regulation, biological functions and associations with human cancer. J. Mol. Med. 92, 217-25. https://doi.org/10.1007/s00109-014-1133-6

Wells, K., Bandoni, R.J., 2001. Heterobasidiomycetes. *In* The Mycota VII, Systematics and evolution, Part B. McLaughlin, D.J., McLaughlin, E.G., Lemke, P.A. (eds.), Springer-Verlag, Berlin, Germany. https://doi.org/10.1007/978-3-662-46011-5

Wells, K., 1994. Jelly fungi, then and now! Mycologia 86, 18–48. https://doi.org/10.1080/00275514.1994.12026372

Wibberg, D., Jelonek, L., Rupp, O., Hennig, M., Eikmeyer, F., Goesmann, A., Hartmann, A., Borriss, R., Grosch, R., Puhler, A., Schluter, A., 2013. Establishment and interpretation of the genome sequence of the phytopathogenic fungus *Rhizoctonia solani* AG1-IB isolate 7/3/14. J. Biotechnol. 167, 142–155. https://doi.org/10.1016/j.jbiotec.2012.12.010

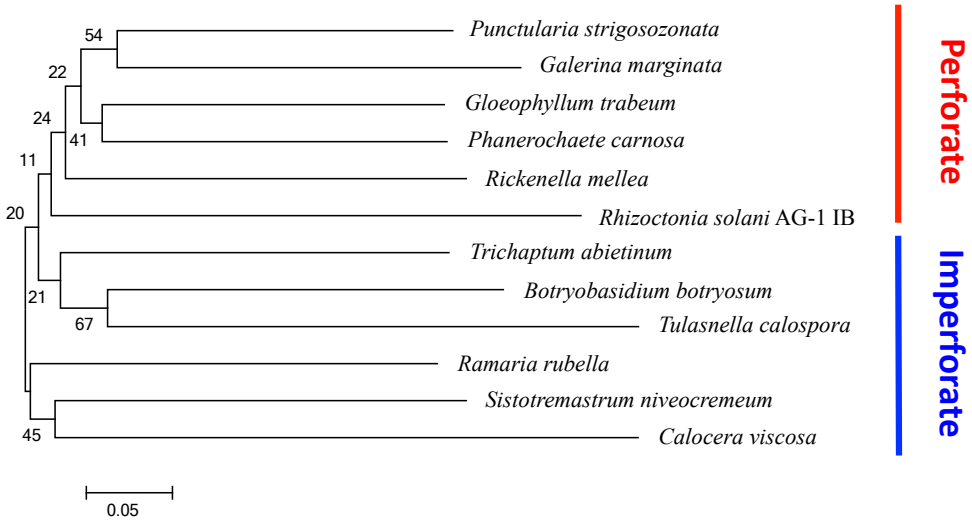
Wood, V., Harris, M.A., McDowall, M.D., Rutherford, K., Vaughan, B.W., Staines, D.M., Aslett, M., Lock, A., Bähler, B., Kersey, P.J., Oliver, S.G., 2012. PomBase: a comprehensive online resource for fission yeast. Nucleic Acids Res. 40, D695–D699. https://doi.org/10.1093/nar/gkr853

Wray, G.A., 2002. Do Convergent Developmental Mechanisms Underlie Convergent Phenotypes? Brain Behav Evol 59, 327–336. https://doi.org/10.1159/000063566

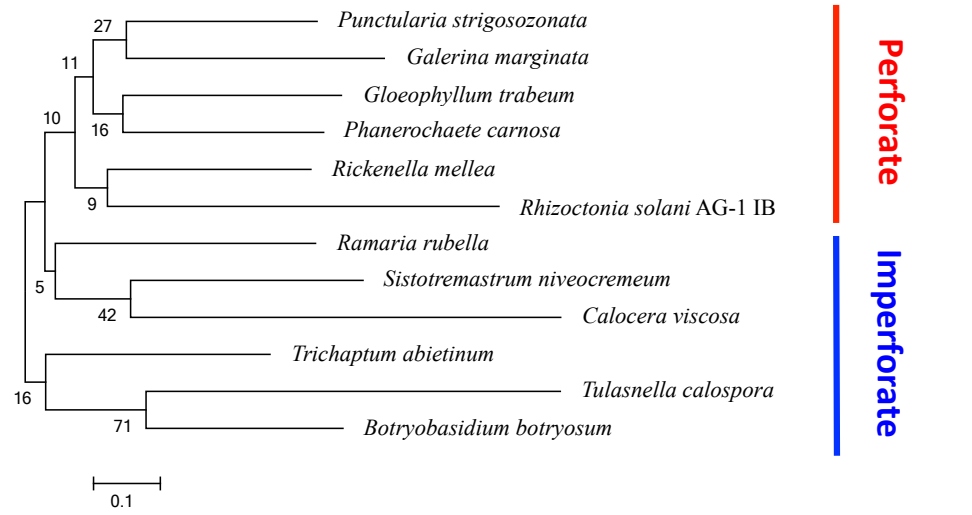
Yang, Z., 2014. Molecular Evolution: A Statistical Approach. Oxford Univ Pr on Demand. https://doi.org/10.1093/acprof:oso/9780199602605.001.0001

# Supplemental figures and tables

#2 NJ

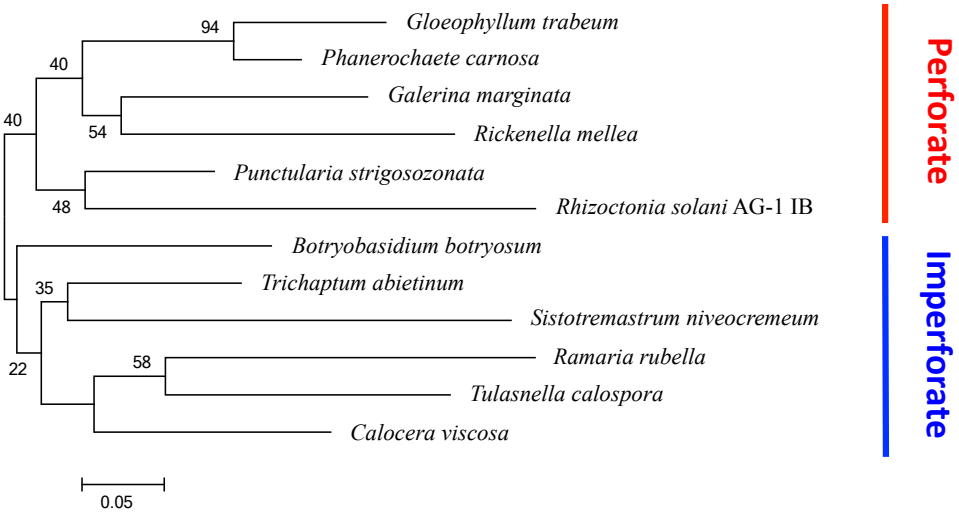



#2 ML

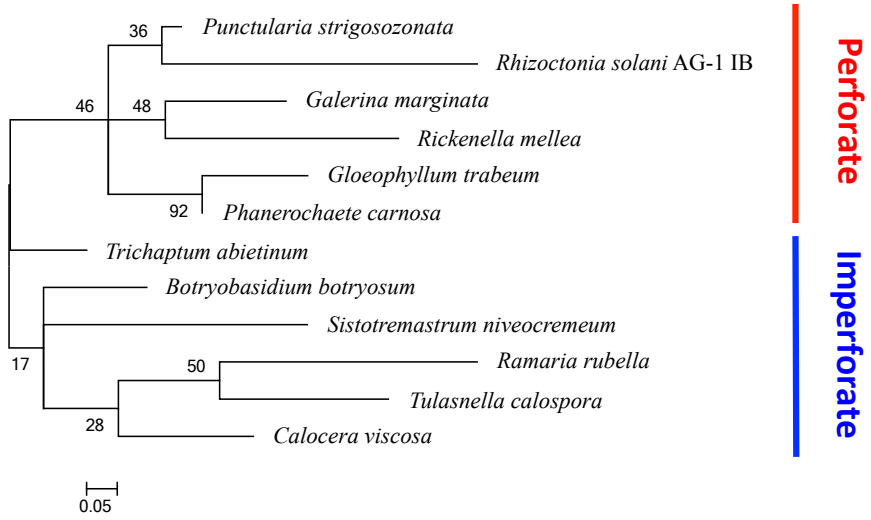



**Supplemental Figure 1.** Phylogenetic relationship of correlated genes #2-#8. The gene trees were produced using the neighbor-joining (NJ) and maximum-likelihood (ML) methods.

#3 NJ

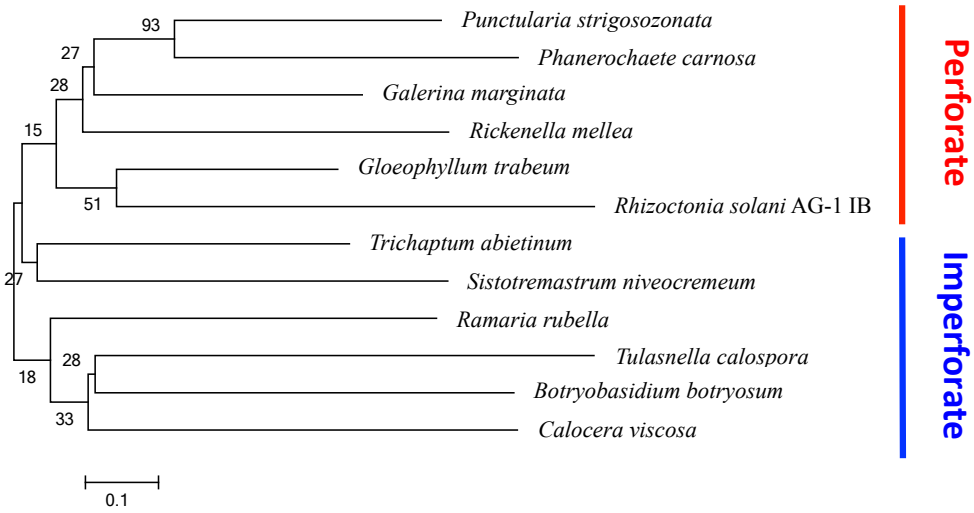



#3 ML

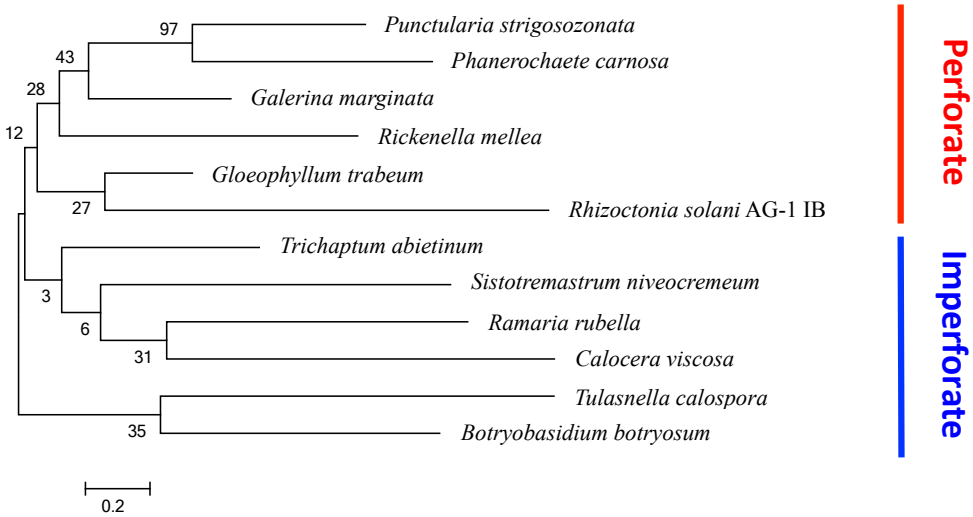



**Supplemental Figure 1.** (continued)

#4 NJ

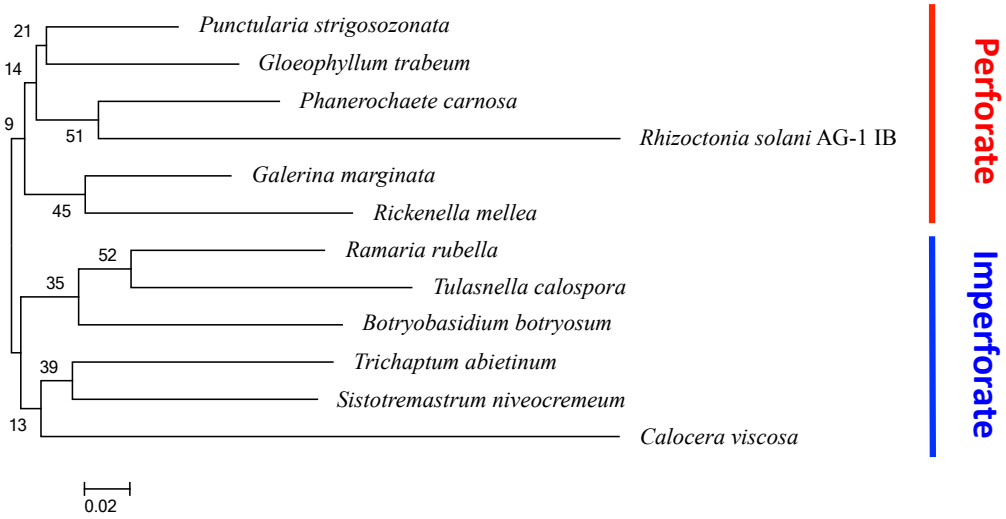



#4 ML

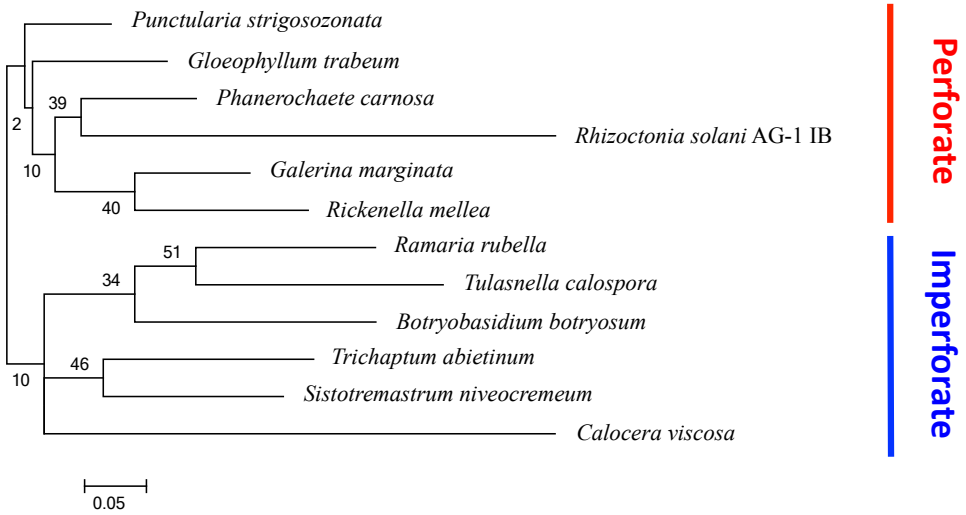



**Supplemental Figure 1.** (continued)

#5 NJ

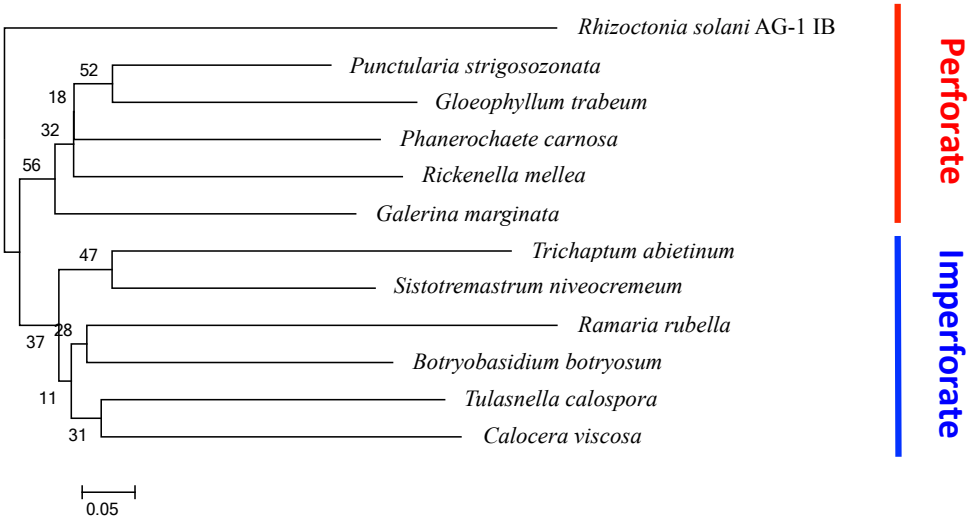



#5 ML

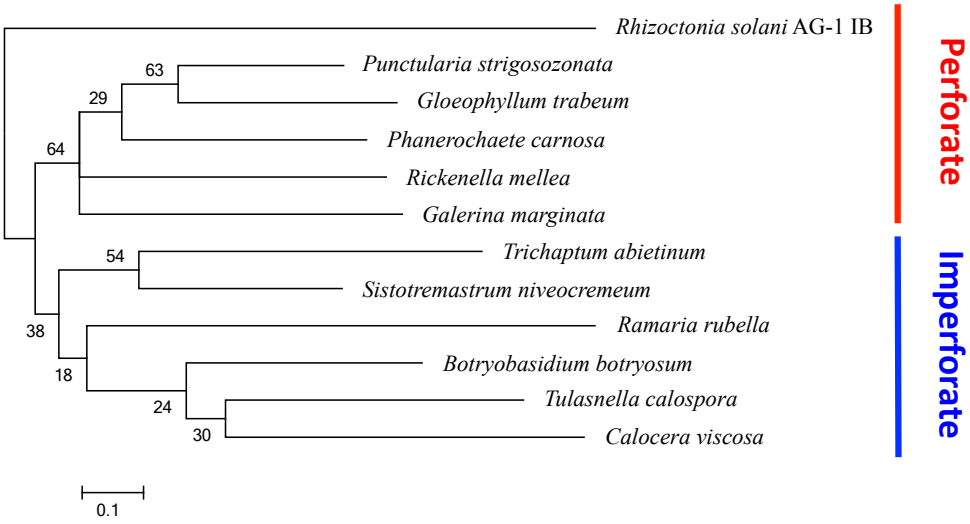



**Supplemental Figure 1.** (continued)

#6 NJ

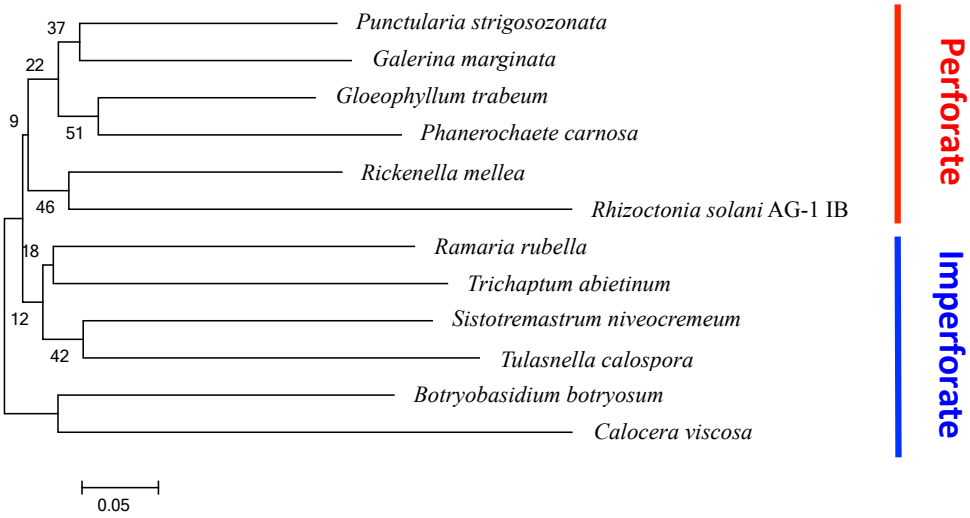



#6 ML

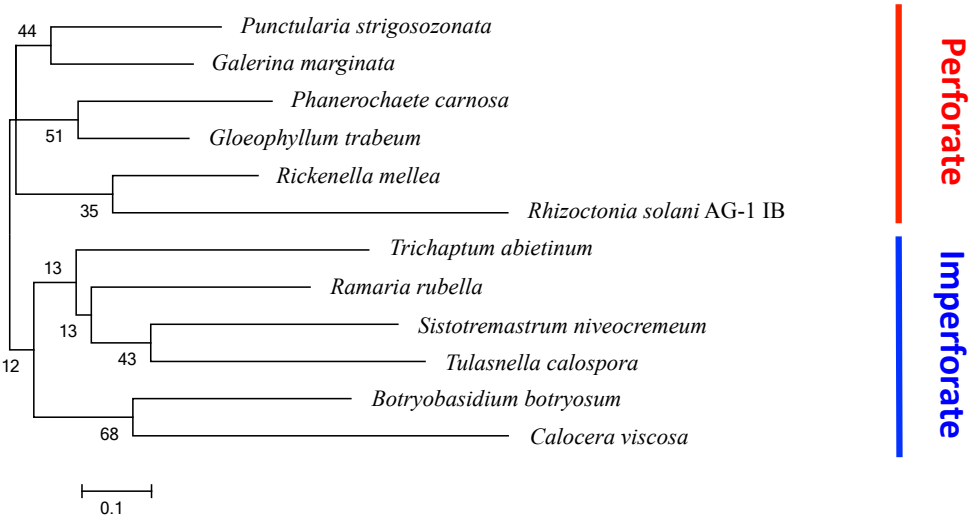



**Supplemental Figure 1.** (continued)

#7 NJ

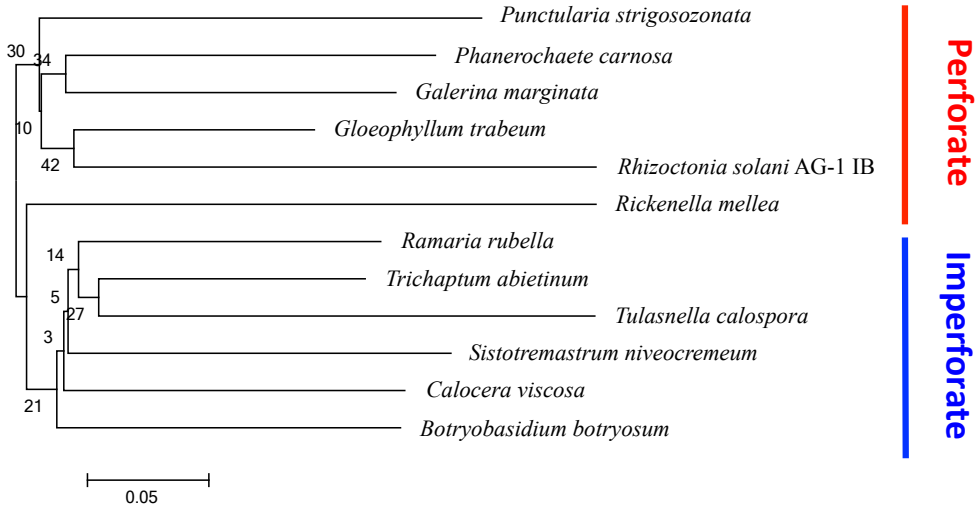



#7 ML

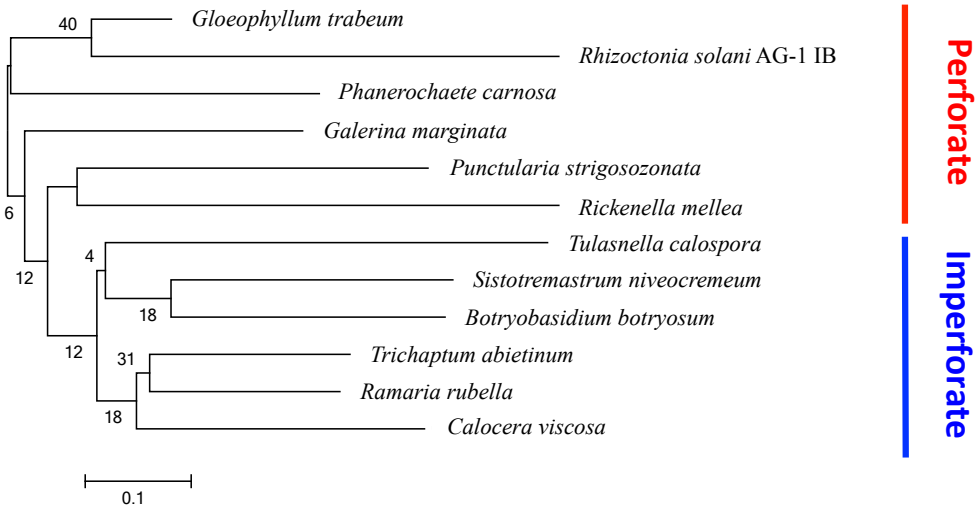



**Supplemental Figure 1.** (continued)

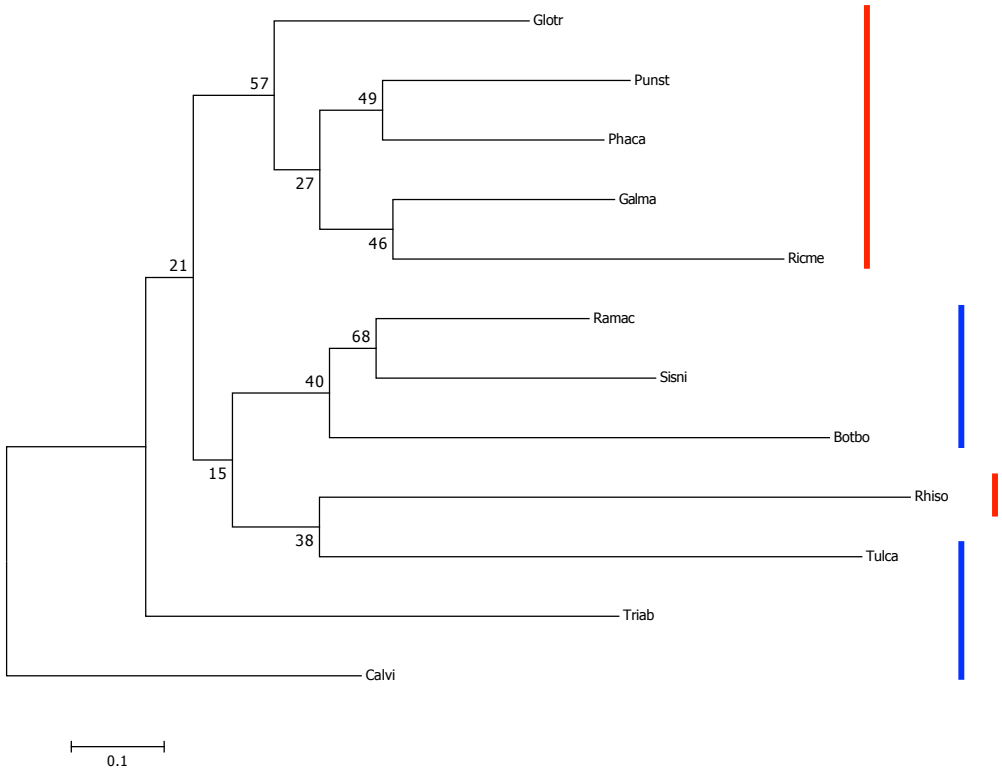
#8 NJ



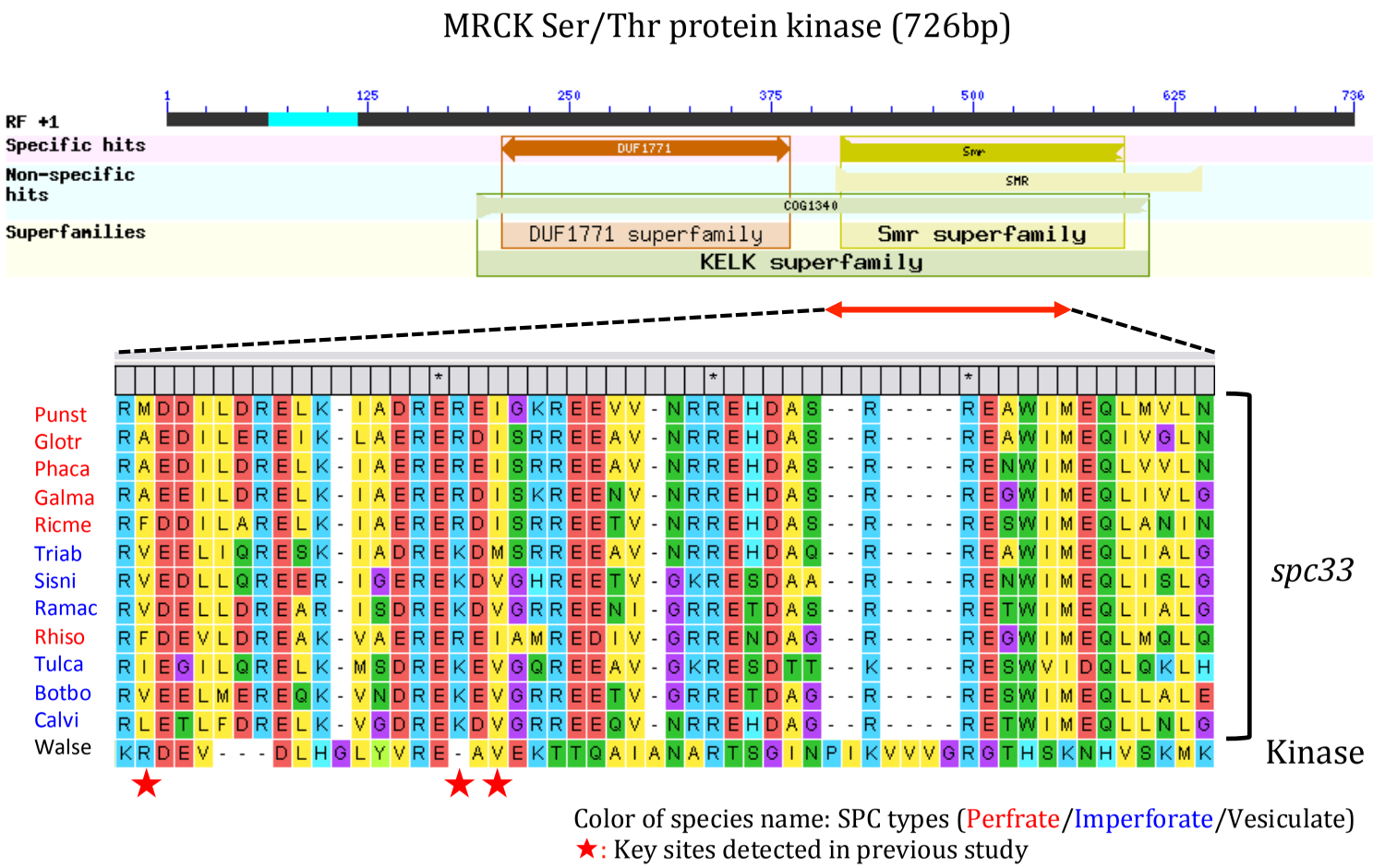

#8 ML




**Supplemental Figure 1.** (continued)




**Supplemental Figure 2.** Phylogenetic gene tree of *spc33* with partial-deletion option. All positions with less than 65% site coverage were eliminated.




**Supplemental Figure 3.** Key sites in *spc33* aligned with the functional motif of MRCK Ser/Thr pritein kinase superfamily.

**Supplemental Table 1.** Homology of amino acid sequences of spc33 between Rhiso and other species. Perforate type species are colored in red.

| Species | Identity % | Identity sites |
|---|---|---|
| Ricme | 21.31% | 127 |
| Galma | 20.97% | 125 |
| Triab | 19.97% | 119 |
| Ramac | 18.96% | 113 |
| Tulca | 18.79% | 112 |
| Calvi | 18.79% | 112 |
| Glotr | 18.12% | 108 |
| Botbo | 17.79% | 106 |
| Phaca | 17.28% | 103 |
| Sisni | 14.26% | 85 |
| Punst | 13.59% | 81 |

**Supplemental Table 2.** Number of simple sequence repeat (SSR) in spc33-containing scaffolds detected by RepeatMasker 4.0.8.

| Species | No. of SSRs |
|---|---|
| Punst | 182 |
| Glotr | 272 |
| Phaca | 379 |
| Galma | 311 |
| Ricme | 17 |
| Triab | 106 |
| Sisni | 33 |
| Ramac | 24 |
| Rhiso | 6 |
| Tulca | 61 |
| Botbo | 68 |
| Calvi | 148 |
| Walse | 33 |