CHAPTER

6

Molecular Ecology

Sarah C. Watkinson
University of Oxford, Oxford, UK

THE APPLICATION OF DNA TECHNOLOGY IN THE ECOLOGY OF FUNGI

Introduction

In this chapter we give an introduction to the molecular methods based on DNA technology which are transforming our understanding of the diversity and activities of fungi in the environment. Being mostly microscopic and often growing within opaque substrates or the living tissues of other organisms, fungi are hard to observe in their natural habitats, except when conspicuously sporulating. In laboratory culture, some may reveal more informative characteristics, but readily-culturable fungi are now known to constitute only a minority of those present in most habitats. A high degree of taxonomic expertise is needed to identify even sporulating fungi by morphology alone, because sporophore form, on which identification is usually based, is typically convergent so that distantly-related species can look the same. Using DNA sequences to identify fungal species has several advantages over identification based on morphological features. The bases in DNA provide a large number of non-adaptive characters. Environmental sequences can be placed within a clade of near neighbours, giving the predictive value of natural classification. By analysing the fungal nucleic acids in environmental samples, we can discover what kinds of fungi are present, where they are, and what substances they consume and change. Rapid technical developments in recent years mean that many processes in sequencing and bioinformatics can be automated.

Generic primers to amplify sequences in the internal transcribed spacer (ITS) region of ribosomal DNA (Figure 6.1) were first developed in the mid-1990s. Highly conserved ribosome-encoding flanking sequences facilitated primer design. For example, basidiomycete DNA could be amplified selectively from the mixture of fungal DNA extracted from soil samples. Species-specific ITS sequences or their restriction fragments were separated using gel-based methods and bacterial cloning. The value of DNA-based identification in fungal ecology was soon evident. For example, the fungi in ectomycorrhizal roots (Chapter 7) could be identified by matching their DNA to readily-identifiable aboveground mushrooms. This showed that the ectomycorrhizal species most conspicuous by their fruiting bodies were not necessarily

http://dx.doi.org/10.1016/B978-0-12-382034-1.00006-2

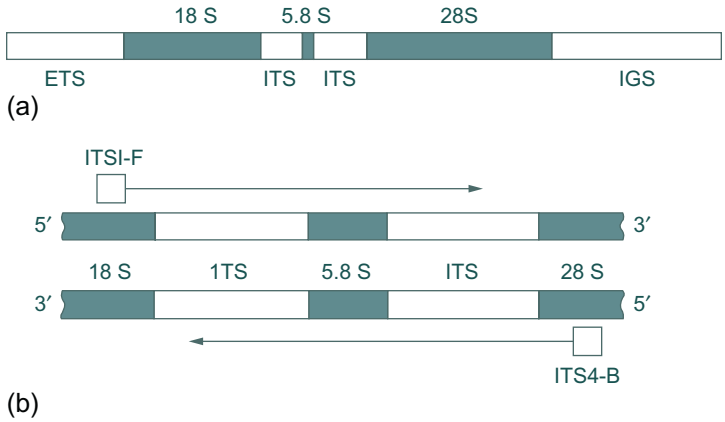

FIGURE 6.1 The utilisation of ribosomal DNA in fungal identification. (a) single copy of ribosomal DNA (rDNA), with coding and non-coding regions. An internally transcribed spacer (ITS) lies between the genes coding for 18S and 28S RNA, with that coding for 5.8S RNA embedded within it. Beyond the RNA genes are the externally transcribed spacer (ETS) and the intergenic spacer (IGS). (b) The complementary strands in the ITS region of DNA, with adjacent coding regions, showing the points at which the two primers, ITS1-F and ITS4-B, bind, and the direction of DNA replication towards the 3′ end of each strand.

the most abundant on roots, while some ectomycorrhizal fungi, although widespread on roots, were never recorded aboveground at all. The new molecular tools for investigating the diversity of underground fungi sparked an explosion of investigations in environmental fungal diversity. Not only soil fungi, but any fungal samples could be characterised in this way without the need for isolation into culture, including mycelial fragments, spores from air and water, and pathogenic fungi from tissues of diseased plant or animals. As fungal sequences accumulated in databases such as GenBank (www.ncbi.nlm.nih.gov/genbank/) sequence matching using the basic linear alignment search tool (BLAST) became the norm. Even when there is no exact match, BLAST searching, which provides an indication of the degree of similarity between sequences, may indicate the position of a fungal sequence in a phylogeny of relatives.

A new generation of automated high-throughput sequencing and bioinformatics technology has created a second revolution in molecular fungal ecology. Thousands, if not millions, of environmental sequences can now be separated and identified in parallel. High-throughput methods have largely replaced the labour-intensive method of bacterial cloning as a way of separating fungal ITS sequences in environmental samples, although the quick and inexpensive cloning and gel methods can still be useful in smaller studies (Figure 6.2).

Molecular fungal ecology has, however, not superseded long-established methods of microbiology. Direct observation and quantification at all scales, and culture on defined media, remain essential techniques in fungal ecology. However, these approaches are now generally combined with molecular methods in any investigation of the role of fungi in ecosystems.

FUNGAL DIVERSITY IN THE ENVIRONMENT

Community Analysis

The overall diversity of fungi present in an environmental sample can now be captured by sequencing the entire range of genetic material that it contains, an approach termed

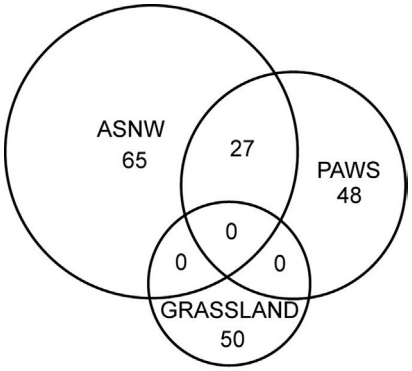

FIGURE 6.2 The use of small-scale sequencing and cloning methods to compare fungal community diversity under different vegetation types. The Venn diagram shows the numbers of unique fungal ITS types recovered from three main habitat types (ancient woodland, plantation, and adjacent grassland). Soil fungi from ancient semi-natural woodland exist in sites converted to non-native conifer plantations but not in grassland soil. *Source: Johnson et al. (2014).*

metagenomics. To obtain the **metagenome** of the sample, either DNA can be sampled, or alternatively RNA (which is then copied into the more stable cDNA for storage and sequencing). RNA sequences have the advantage of reflecting only the diversity and activities of fungi that are actively growing and synthesising new ribosomes and proteins, because RNA is generated only from genes being actively transcribed. In natural environments, the **metatranscriptome**, consisting of all the RNA present, reveals fungal genes expressed at the time of sampling. Sequencing systems such as 454-pyrosequencing, Illumina, and Ion Torrent generate sequences of all the nucleic acids present, in the form of short sequences of up to 400 bp. The mass of sequence data is analysed and interpreted using dedicated computational tools which are continually being developed and refined. The extensive replication achievable with such high-throughput 'massively parallel' sequencing approaches makes it possible to extract robust data, even in the face of the multiple sources of biological and technical variation encountered in any investigation of microbes in real environments. The metagenome and metatranscriptome can reveal both microbial community diversity and patterns of gene expression.

Even without assignment to named taxonomic groups, the metagenome provides a snapshot of diversity that can point to patterns in fungal communities. It may also discover new fungal taxa (Chapter 1). Examples of this approach include analyses of fungal diversity in soils in the United States under prairie grassland and in French mountain forest. The prairie soils, from different sites across Kansas, were reported to contain an astonishing fungal diversity at all sampled sites, with over a million types of fungi and little phylogenetic overlap between sites, a finding which must merit further examination in view of the extent and importance of this North American grassland biome. In a study of forest soils in France, analysis was targeted at Ascomycete and Basidiomycete fungi growing symbiotically on tree roots as ectomycorrhiza, by using primers selective for Dikarya and using the curated mycorrhiza-specific database UNITE as well as GenBank for sequence matching. Around 1000 separate operational taxonomic units, OTUs, (described in Chapter 1) were found in each 4 g soil sample, a much higher diversity than had been predicted from previous observations. Of these, 81% were Dikarya with agaricomycetes predominating, with relatively few ectomycorrhizal taxa accounting for most of the species diversity.

Methodology

The technical complexity of highly-automated high-throughput sequencing inevitably introduces multiple potential sources of error. Understanding the likely pitfalls at every stage of the procedure (Figure 6.3) requires a range of expertise, from field sampling strategy to computational analysis, and collaboration between experts in a range of fields is essential to ensure that sources of error are understood and controlled for as far as possible.

DNA is often extracted from samples of a few grammes at most, so it is a challenge to design a scale and pattern of sampling that reflects real fungal diversity in a large field habitat. There is likely to be variation in size and abundance of fungal individuals, in their patterns of distribution, and in the kinds of materials to be sampled. Some individual fungi, such as those forming extensive mycelial networks in forest soils, extend over many metres, while other species will be present as scattered microscopic spores. Moreover, samples from most natural habitats will vary in texture, composition, and density. The process of collecting samples may itself cause rapid and confounding changes in nucleic acid composition, for example, fast-growing saprotrophic species may multiply preferentially following accidental damage, causing nuclear multiplication and consequent over-representation of their DNA, so

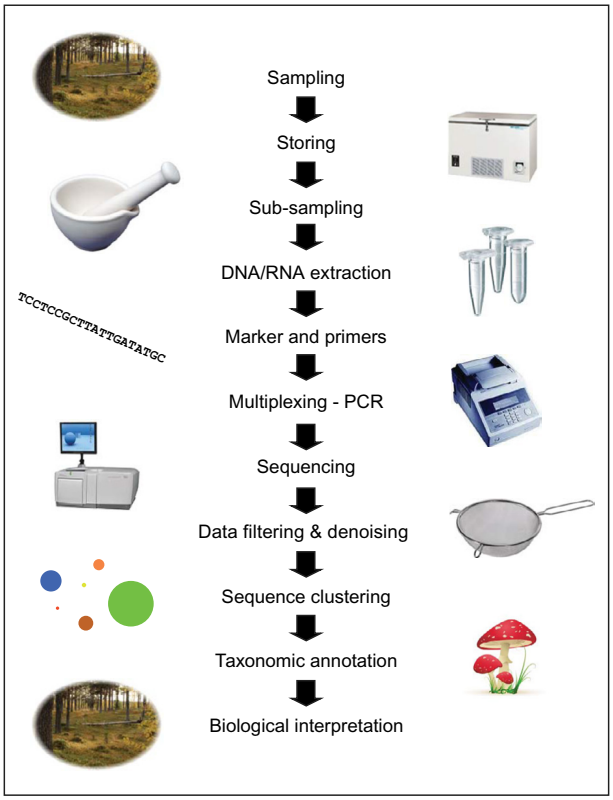

FIGURE 6.3 Overview of the steps involved in high-throughput sequencing of fungal communities. *Source: Lindahl et al. (2013).*

samples must be frozen or chemically preserved immediately on collection. If RNA is being sampled it must be frozen instantly to −80 °C and copied to the more stable *c*DNA for long-term storage. The efficiency of DNA extraction may vary between samples, and soils and plant material may contain inhibitors of downstream processes including the PCR.

Identification

The choice of sequence to be used for identification depends on the level of taxonomic discrimination required. The ITS region (Chapter 1, p. 5) is widely used to identify species and has been proposed as a universal 'barcode' locus. For higher-order taxa, the 28S subunit sequence is useful and the D1/D2 domains of the 28S characterise a taxonomic level approximating to the genus. Sequences in the long and short subunits – LSU and SSU—are used for the Glomeromycota, in which there are insufficient data available for other regions. These ITS sequences are combined with other conserved genes for phylogenetic species determination, with six gene regions generally used: 18S rRNA, 28S rRNA, 5.8S rRNA, elongation factor 1-(*EF1*), and two RNA polymerase II subunits (*RPB1* and *RPB2)*. The choice of primers inevitably introduces bias, and to overcome this it is desirable to use many different primers, and primers with degenerate positions, so that the resulting amplicons represent as many as possible of the fungi present in the sample. Artefacts can arise during nucleic acid amplification, including the formation of chimaeras, mispriming, and formation of heteroduplexes, which can give overestimates of the real sequence diversity. Sequences are aligned using automated bioinformatics programmes to generate clusters of taxonomically meaningful groups based on similarity. Where sequences differ by less than a predetermined maximum, usually 3%, they are considered to represent a single OTU. The details of alignment programmes are ever-changing, as multiple programmes are developed to cluster similar sequences into taxonomically meaningful groups. The computational analysis of large sets of environmental sequences is non-trivial, and the reader is referred to specialist sources for detailed up-to-date information.

Identifying Cryptic Species, Clones, and Genets

Phylogenetic determination of individuals from environmental sequences has led to the recognition that many fungi that were assigned to a single named species on morphological features, exist in nature as multiple types with important differences in ecological function. For example, mycorrhizal fungi that are indistinguishable on the basis of morphology may consist of phylogenetically distinct clades differing in host plant specificity, preferred soil type, or in other functional aspects of the symbiosis (Chapter 7). These variants are termed **cryptic species**.

In studying the biology of single species of fungi in natural environments, it is frequently important to be able to distinguish individual organisms from each other, and to characterise mycelium or spores as members of the same or separate clones (see Chapter 4 for a discussion of fungal clones and genets). Environmental samples can be assigned to an individual clone using hyper-variable sequences such as microsatellite DNA or single nucleotide polymorphisms (SNPs) as a fingerprint. This has many practical applications in mycology. By tracing the spatial distribution of clones of ectomycorrhizal fungi it is possible to determine the underground extent and potential connectivity of the common networks (p. 226) of

ectomycorrhizal basidiomycetes, some of which are capable of extending for several metres through the soil. Because they can act as channels carrying plant nutrients, water, and photosynthate between plants, their extent and host connections of these fungal channels may influence the structure of the aboveground ecosystem (see Chapter 7). Determining whether infective propagules are clonal or recombinant can provide valuable clues in the population genetics of pathogens (Chapter 8), pointing to the mode of spread of disease and hence the most effective targets for its control.

Diagnostic Methods for Known Species

Known species or strains of fungi can be quantified from environmental samples by **quantitative PCR**, in which the transcript level of species-specific sequences is assayed fluorimetrically. Both DNA and RNA (reverse-transcribed to cDNA) can be determined. This is a valuable diagnostic method in medicine and agriculture, being very sensitive and specific, and capable of being automated for the simultaneous analysis of hundreds of samples. Portable PCR machines enable on-site detection. Quantitative PCR has numerous applications in agriculture, including detecting and quantifying fungi present at very low levels. It can be used to detect crop and human pathogens, fungicide-resistant strains, and strains carrying biotechnologically-introduced traits. Imaging methods are used for diagnosing and visualising known species active *in situ*. Figure 6.4 shows the use of transformation with green fluorescence protein (GFP) to image active cells of the insect pathogen *Metarhizium* within the body of its mosquito larva host.

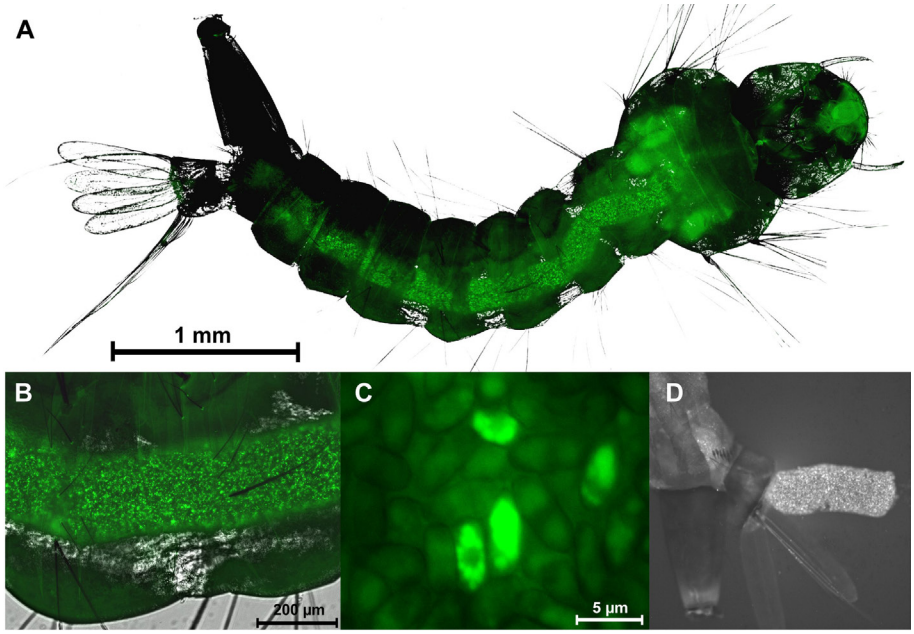

FIGURE 6.4 The insect pathogen *Metarhizium brunnei*, transformed with green fluorescence protein (GFP) and visualised by fluorescence microscopy within the body of its host, a larva of the mosquito *Aedes aegypti*. Species of *Metarhizium* are being investigated for biological control of mosquitoes that are vectors of human disease, including malaria, dengue, and yellow fever. (a) conidia in the gut lumen of the larva, (b, c) conidia expressing GFP in the gut lumen, and (d) active conidia in faecal pellets. *Source: Butt et al. (2013).*

Classification and Nomenclature of Environmental Fungal Sequences

Molecular environmental sampling is now widely and successfully used for fungal species discovery and identification. This has created a problem: the rate of discovery of new species from their sequences is now too fast for them all to be assimilated into the established system of classification and nomenclature. To quote David Hibbett (see Further Reading): 'Fungal taxonomy seeks to discover and describe all species of fungi, to classify them according to their phylogenetic relationships, and to provide tools for their identification… to enable communication about fungal diversity, there is a pressing need to develop classification systems based on environmental sequences'.

Currently, when a potentially new fungal species is discovered, it is described and named according to the International Code of Nomenclature for algae, fungi, and plants. The name is then added to Index Fungorum (http://www.indexfungorum.org/), an ongoing list derived from world mycological literature. This system provides a registry of new taxonomic discoveries, and a standardised nomenclature for fungal species, essential for scientific communication in biology. The current system of identification and nomenclature based on specimens – herbarium and culture collections curated by taxonomists – makes it possible to identify species and test prior taxonomic hypotheses. However, the majority of fungal species or OTUs found as environmental sequences have no known fruit bodies, and have not been cultured so cannot be deposited as voucher specimens in herbaria, or cultures in a culture collection. Fungal taxonomists are exploring ways of integrating sequence-based and specimen-based fungal taxonomy into a single internationally recognised system. It has been estimated that naming and classifying the estimated five million or more undescribed fungal species using traditional specimen-based methods would take over a thousand years. Some mycologists therefore proposed that environmental sequences shown to be valid indications of the existence of new fungal taxa should be named and accepted into modern systems of taxonomy. For this, it is necessary to develop a system of nomenclature that will link fungal sequences in GenBank with specimen-based fungal herbarium taxonomy. The 2011 review by Hibbett and colleagues (listed below under Further Reading) reported 91,225 fungal ITS sequences in GenBank. These could be sorted into 16,969 clusters of 93% similarity. Of these, 37% were known only from environmental DNA, and only 13% of fungal sequences in GenBank matched sequences from described specimens. However, the proportion of remaining sequences with no matches to herbarium specimens does not necessarily represent fungi new to science, because many well described species of fungi listed in *Index Fungorum* are still not represented in GenBank. Most of the herbarium voucher specimens at the Royal Botanic Gardens at Kew, UK, when submitted to GenBank BLAST searches, produced no matches, or incorrectly named ones. While automated sequence identification can replace taxonomic expertise, it introduces new potential sources of error. For example, public databases such as GenBank may contain naming errors and wrongly-attributed sequences, which may not be corrected because only the original depositor – who need not be a taxonomist – can alter the GenBank record. For this reason, curated quality-assessed sequence databases are maintained for some groups of fungi where reliable identification is critical for a particular scientific community, for example, in mycorrhizal research UNITE, (http://unite.zbi.ee), SILVA (http://www.arb-silva.de), MaarjAM (http://maarjam.botany.ut.ee), and for plant and human pathology, Fusarium-ID (http://isolate.fusariumdb.org/index.php), and the Aspergillus Genome Database (http://isolate.fusariumdb.org/index.php) in plant and human pathology.

Efforts are currently under way to incorporate ITS sequences from professionally identified and curated herbarium voucher specimens and isolates into a single publicly-accessible database. These 'species barcodes' will in future cross-reference environmental sequences into a quality-controlled system of classification and identification, and provide taxonomic authentication for fungal sequence accessions in GenBank. In this way, the majority of fungi that are known only as environmental sequences might eventually be linked into a specimen-based list, such as Index Fungorum, to provide a single integrated catalogue.

LINKING FUNGAL DIVERSITY AND ECOSYSTEM PROCESSES

Characterising fungal diversity is only the first step in investigating the importance of fungi in an ecosystem. The next is to link fungal diversity to ecosystem function. Molecular methods are providing keys to what has been termed the 'black box' of microbial ecosystem processes. To illustrate this, we describe how a combination of molecular and chemical approaches is revealing the key role of fungi in the maintenance of boreal forest, the northern hemisphere conifer-dominated biome which is a major carbon sink in the global carbon cycle. These sub-arctic forests grow in areas of low sunlight and high rainfall, and their highly-leached soils are poor in available nitrogen and minerals. Soils are typically highly stratified, with a carbon-rich layer of recently-fallen litter that becomes gradually comminuted by animals and saprotrophic microbes to fragments and soluble organic material in the underlying humic layers. The mycelium of ectomycorrhizal and saprotrophic fungi accumulates and redistributes scarce nitrogen and minerals, and thus plays a central role in plant nutrition, and in the turnover of carbon and nitrogen. An analysis of pine forests in Sweden has shown that both the diversity of soil fungi and their nutritional mode – saprotrophic or ectomycorrhizal – differ between soil horizons (Figure 6.5). From sequence analysis of soil samples, the rapidly-decomposing carbon-rich litter layer is occupied mainly by saprotrophic species. Ectomycorrhizal species (Chapter 7) dominate deeper soil layers and nitrogen isotope ratios indicate that these are responsible for most nitrogen mobilisation in the soil. Carbon isotope ratios show that plant-litter-derived carbon can persist for as long as 45 years in lower soil horizons, in the form of humic and phenolic materials and organic nitrogen compounds which break down extremely slowly. Many of these long-lived substances are now known to be formed by fungi, and to represent an important contribution to soil carbon sequestration.

The process of fungal plant litter decomposition is a pivotal process in carbon and nitrogen cycles of terrestrial ecosystems (Chapter 9). Patterns of gene expression in environmental samples can reveal the production of specific enzymes in natural habitats and identify the fungi producing them. For example, an investigation of spruce (*Picea abies*) forest in Bohemia analysed the metagenome and metatranscriptome in the surface litter layer of the soil. Sequences of the fungal wood-decomposing enzyme, cellobiohydrolase (Chapter 5) were identified and attributed to the producing fungal species. RNA analysis, combined with isolation, culture and biochemical experiments, demonstrated that relatively few key species, for example, species of the agaricomycete *Mycena* and ascomycetes of the Xylariales, are responsible for most of the litter decomposition, with transcripts of ascomycete cellobiase predominating in the litter layer **transcriptome** (see below).

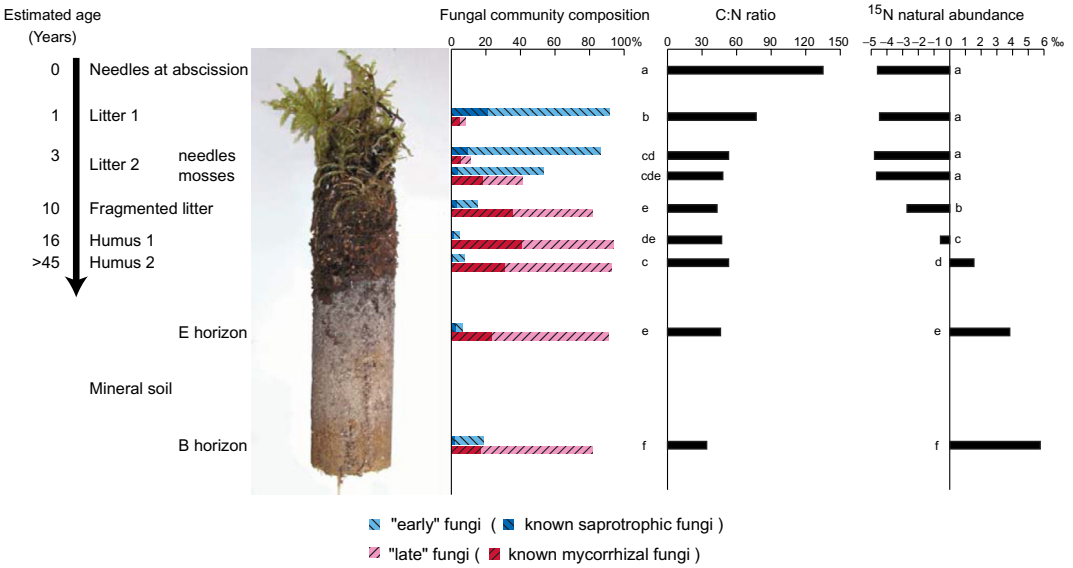

FIGURE 6.5 Fungal community composition, carbon:nitrogen (C:N) ratio and ^{15}N natural abundance throughout the upper soil profile in a Scandinavian *Pinus sylvestris* forest. Molecular methods for identification of fungi, combined with chemical analyses of carbon and nitrogen, indicated different roles of saprotrophic and ectomycorrhizal fungi in the carbon and nitrogen dynamics of separate soil horizons. *Source: Lindahl et al. (2007).*

A specific gene or genes may have a key role in the ecology of a species, for example, virulent strains of a pathogenic fungus may differ from avirulent ones by the possession of a single gene (Chapter 8). Such genetic determinants of pathogenicity are identified by gene replacement (Figure 6.6). In the rice blast disease caused by *Magnaporthe oryzae*, a functional signal transduction pathway is necessary for germinating conidia of the fungus to penetrate into the tissue of a rice leaf and establish infection (see Chapter 8 for details of this plant pathogen). Experimental mutant strains in which a key gene is disrupted cannot infect the leaf, but virulence can be restored by transformation to reintroduce an undamaged copy of the gene. Gene replacement proved that components of the fungal signal transduction pathway – receptors and protein kinases – were required for infection in rice blast, indicating that sensing and response was involved in the fungus's entry into the leaf tissue.

Stable isotope probing (SIP) is used to identify naturally-occurring fungi that break down specific substrates. An example of its use might be the search for a decomposer to break down a xenobiotic pollutant compound (Chapter 5, p. 184). A heavy isotope such as ^{13}C, ^{15}N, or ^{18}O is incorporated into the target material by chemical synthesis or metabolism. The labelled target substrate is placed in the environment to be sampled until it becomes naturally colonised by decomposers present among the pool of soil microbes, including fungi. The DNA of organisms assimilating the target substrate becomes marginally heavier, and can be separated on a density gradient. The DNA bands can then be probed both for taxonomically diagnostic sequences, and for genes encoding enzymes likely to be expressed in assimilating the target compound. Organisms identified by SIP as using a specific substrate can be further characterised by imaging techniques that are being progressively developed and refined.

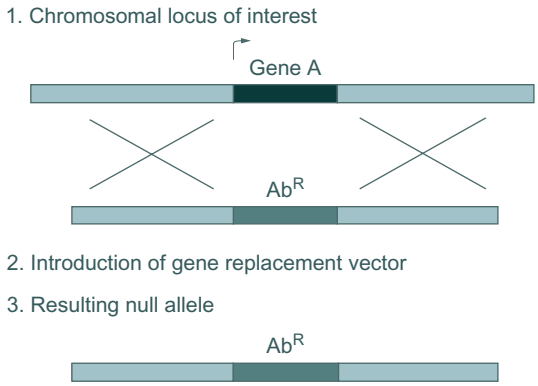

FIGURE 6.6 One-step gene replacement in fungi. In this process, a genomic clone spanning a gene of interest is first identified from a genomic library. A restriction fragment containing regions of at least one kilobase on either side of the gene of interest is cloned, and an antibiotic resistance gene cassette (Ab^r) is inserted into the space normally occupied by the protein-encoding part of the gene of interest. The antibiotic resistance gene cassette will contain a promoter and terminator sequence to allow high level expression in the fungus. The gene replacement vector is then introduced into the fungus by transformation. A double cross-over event can occur, allowing the selectable marker gene to replace the chromosomal copy of the gene. The resulting locus is known as the **null allele**, as it expresses none of the biological activity associated with the native gene. Targeted gene replacements provide a direct test of the role of specific genes in fungal pathogenicity. *Source: By Talbot N.J., reproduced from The Fungi, second edition, 2001.*

In the case of organisms for which taxonomically diagnostic sequences are known, fluorescent probes can be combined with micro autoradiography (fluorescent *in situ* hybridisation and micro autoradiography, FISH-MAR) to visualise individual cells, with a resolution of 0.5–2 μm – the range of diameters of hyphae and many spores. Raman spectroscopy can be used to enhance the resolution and power of cell imaging, by picking up the red shift in mass spectra of compounds that have incorporated a heavy atom. Instruments have been devised that image red shift effects in single cells. This isotope technique has the advantage that it labels cells as a whole, without being localised to a single marker compound such as DNA. In the future, ever higher resolution and more rapid procedures are expected to be developed to observe and identify microbial cells *in situ* and *in vivo*.

WHOLE-GENOME SEQUENCING AND COMPARATIVE GENOMICS

The whole genome sequence of a species contains the organism's entire nuclear genetic repertoire and is a valuable aid to knowledge of its ecology, evolution, and physiology.

The first whole fungal genome to be published was that of baker's yeast, *Saccharomyces cerevisiae* in 1996. Since then, improving sequencing technology has added scores of other species. A project under way at the U.S. Department of Energy Joint Genome Institute (JGI) aims to sequence the genomes of a thousand fungi: (http://genome.jgi.doe.gov/programs/fungi/1000fungalgenomes.jsf).

Such massive allocation of resources reflects recognition of the power of genomics for understanding and exploiting fungal biology. In the JGI Community Sequencing Program, mycologists from around the world are invited to make a case for the potential advantages of sequencing a given fungus, and species are selected on a peer-reviewed basis. The resulting genomic data, together with pipelines of computational procedures for their interpretation, are made available to the scientific community, providing a huge resource for investigating fundamental aspects of fungal biology. Once the whole genome sequence of a species has been assembled, it is scanned for open reading frames, sequences predicted to function as protein-encoding genes. These are annotated with their likely functions by matching to databases of known genes and proteins. For example, sequences may be identified as likely to encode enzymes associated with particular cell functions in nutrition and metabolism, or clusters of sequences may indicate that a species possesses the metabolic equipment to synthesise various classes of secondary metabolite.

Genome and Transcriptome

The genome sequence data alone show what genes a species possesses, but does not prove that they are functional. For this, the next step is to examine the **transcriptome**, the entire RNA produced by gene expression of a fungus. High-throughput methods, for example, Illumina and 454-pyrosequencing, show all the RNA (*c*DNA) sequences present and are used for large transcriptomic samples. Alternatively, where the whole genome sequence is known, microarrays can be used. This is a quicker and less expensive method for smaller samples, and has the advantage of giving immediate results. The DNA of all the genes annotated in the genome, or a selected subset, are arrayed on microscope slides, and transcripts of each gene are identified and quantified by hybridisation to its DNA template.

Changes in the transcriptome can show spatial and temporal patterns of gene expression associated with the functional adaptation of the organism, such as interactions with the environment and other organisms, developmental changes or nutrient acquisition. In this way, the fully-sequenced genome becomes a vast source of objectively-acquired information on cellular and metabolic aspects of niche adaptation. Figure 6.7 shows how the transcriptome of the timber dry rot fungus *Serpula lacrymans* changes according to whether it is growing on wood or on glucose as sole carbon/energy source. A wood substrate induces the expression of multiple carbohydrate-active enzymes as well as membrane transport proteins presumably active in enzyme secretion and nutrient uptake.

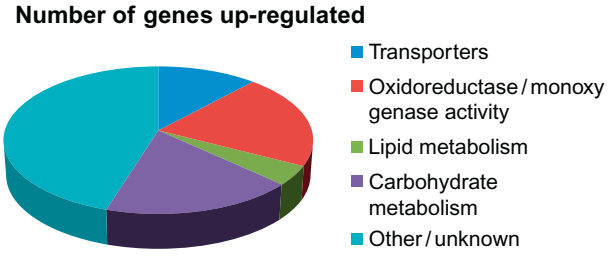

FIGURE 6.7 Functional characterisation of *Serpula lacrymans* transcripts with significantly increased gene expression when grown on wood compared with glucose-based medium, identified by microarray analysis. Expression of glycoside hydrolases involved in cellulose utilisation increased over a hundredfold in mycelium feeding on wood. *Source: Eastwood et al. (2011).*

As progressively larger sets of fungal genomes become available for comparison, it is becoming possible to distinguish a core fungal genome from genes unique to individual species. At the time of writing, a study of 33 basidiomycetes and 30 other fungi were found to possess a core fungal genome of around 5000 genes. However, as many as half of the proteins in Basidiomycota lacked homologues in other groups of fungi, and 23% were unique to a single organism, with slightly lower figures for unique proteins in Ascomycota.

Gene Families and Niche Adaptation

Comparative genomics of multiple species can reveal not only their phylogenetic relationships, but also patterns of past lineage divergences that underlie the unique niche adaptation of individual species. Comparing the genomes of a number of different species from a range of ecological niches reveals expansions and contractions of gene families associated with different life styles, providing a clue to proteins which might give a species a selective advantage in a particular ecological niche. The agaricomycetes found in forest habitats include species with a range of fairly distinct nutritional modes. Physiology and biochemistry distinguish brown rot species that feed on cellulose and hemicellulose from wood and other forms of plant remains, with little or no lignin decomposition, white rot species that also remove most of the lignin from wood using ligninolytic peroxidases, and others that acquire sugars from host photosynthesis by mycorrhizal symbiosis with tree roots (Ectomycorrhizal fungi, see Chapter 7). Brown rot decomposition dispenses with some or all of the ligninolytic peroxidases that enable white rot fungi to break down lignin, and depolymerize crystalline cellulose partly by means of the oxidative Fenton's process (described in Chapter 5). From species phylogeny, the brown rot mode of nutrition has evolved convergently from lignin-decomposing white rot in several separate lineages. Ectomycorrhizal nutrition also shows convergence, with separate phylogenetic origins (Figure 7.10).

Evolutionary adaptation to these different nutritional niches has been accompanied by expansions and contractions in gene families. In agaricomycete phylogeny, lineages whose modern representatives utilise wood by brown rot decay are nested within wider clades of fungi mainly represented by white rotting species. It is inferred that brown rot wood decomposition evolved repeatedly as a feeding strategy that conferred advantages in some niches. The phylogeny of gene families associated with breaking down wood polymers (oxidases, including peroxidases and mono-oxygenases, hydrolases including a range of endo- and exoglucanases) shows that evolution of brown rot species from white rot ancestry was accompanied by the loss of many oxidative enzymes as well as a reduction and refinement in the suite of hydrolases. A similar loss of degradative enzymes occurred alongside the switch to symbiosis in ectomycorrhizal species. As the genomes of more and more species become known, it is, however, becoming clear that agaricomycete modes of wood decay encompass a continuous spectrum. While some fungi such as *Postia placenta* and *Phanerochaete chrysosporium* can be assigned to the traditional groups of brown and white rot fungi respectively, other more recently sequenced species such as *Jaapia argillacea* and *Botryobasidium botryosum* share features of both.

Reconciling gene phylogeny with molecular clock data in a **chronogram** can show the approximate dates of phylogenetic divergences (Figure 7.11). The results seem to present a picture of changes in early forest ecosystems. It has been suggested that the accumulation of wood, which formed the coal in carboniferous geological strata, might have been brought to

an end when fungi evolved the capacity to feed on dead wood, returning its carbon to the atmosphere as carbon dioxide. Brown rot, which is the most common form of decay of conifer wood and leaves a lignin-rich residue, may have contributed to the development of the humus-rich, nitrogen poor acid soils of boreal forest.

Adaptation by Changes in Genome Architecture

Comparing whole genomes has led to greatly increased understanding of gene evolution, including the discovery of the importance of **horizontal gene transfer** in which genes have been directly transferred across fungal lineages, without sexual recombination. It is inferred when some genes, or a gene cluster, are found only in a single lineage nested within a phylogenetic tree. In fungi this appears to have been a key evolutionary process in opening up new niches and enabling adaptive radiation into new habitats. A whole set of genes encoding the enzymes and transport proteins required for nitrate utilisation, derived from several different ancestral lineages, may have enabled ascomycetes and basidiomycetes to colonise terrestrial soils in which nitrate is the main available form of nitrogen (Chapter 5). Genes that confer virulence may spread within a clade of fungi and generate new plant diseases. A recently emergent wheat disease is caused by *Septoria nodorum.* Its fully-sequenced genome revealed the presence of a gene for a protein toxin that acts as a virulence factor (i.e. necessary for virulence, see Chapter 8). This gene was found to be homologous with one encoding a protein toxin in an established wheat pathogen, *Pyrenophora tritici-repentis*, which causes yellow spot of wheat. The toxin gene appears to have been conveyed by horizontal transfer to the previously nonpathogenic *Septoria nodorum*, endowing this fungus with virulence against susceptible races of wheat. Horizontal transfer may involve not only small sequences but whole gene clusters. The recently sequenced genome of *Podospora anserina* contains a 23-gene cluster encoding the complete pathway for production of sterigmatocystin, a highly toxic aflatoxin precursor, which is presumed to have been horizontally transferred in its entirety from *Aspergillus nidulans*.

Transposable elements are autonomously-replicating, non-protein-encoding sections of the genome that can relocate around the genome (Chapter 4, p. 132). They have been recently recognised in fungi, where it is now believed that they may play an important part in the evolution of symbiosis. Transposable elements make up an exceptionally high proportion of the genome in some mycorrhizal species: around 60% in the truffle *Tuber melanosporum*, and 21–24% in *Laccaria bicolor*.

Transposable elements can cause changes in genome architecture and gene function. Genes can become silenced, or duplicated. Gene duplication allows a gene to acquire new functions by mutation, since one copy remains to maintain the original function, leaving the second copy free to mutate without compromising the organism's fitness. Inserted transposable elements may interfere with meiosis and thus remove sections of the genome from the possibility of recombination. Comparisons of symbiotic and non-symbiotic species of agaricomycetes has shown that transposable elements are usually more common in symbiotic fungi than their saprotrophic close relatives. Moreover, genes that have been identified as having crucial roles in the development of symbiotic tissues appear both to be associated with transposable elements, and also to lack orthologues in other organisms. Both these observations suggest the possibility that transposable elements have played a part in the genetic innovation underlying a move from saprotrophic to biotrophic nutrition.

Further Reading

General Works on Methodology

Grigoriev, I., 2013. Fungal genomics for energy and environment. In: Horwitz, B.A., Mukherjee, P.K., Mukherjee, M., Kubicek, C.P. (Eds.), Genomics of Soil- and Plant-Associated Fungi. Springer, Berlin Heidelberg, pp. 11–27.

Hirsch, P.R., Mauchline, T.H., Clark, I.M., 2010. Culture-independent molecular techniques for soil microbial ecology. Soil Biol. Biochem. 42, 878–887.

Lindahl, B.D., Nilsson, R.H., Tedersoo, L., Abarenkov, K., Carlsen, T., Kjøller, R., Kõljalg, U., Pennanen, T., Rosendahl, S., Stenlid, J., Kauserud, H., 2013. Fungal community analysis by high-throughput sequencing of amplified markers – a user's guide. New Phytol. 199, 288–299.

Lindahl, B.D., Kuske, C.R., 2014. Metagenomics for Study of Fungal Ecology. In: The Ecological Genomics of Fungi. John Wiley & Sons Inc, Hoboken, pp. 279–303.

Macleod, D., 2006. Principles of Gene Manipulation and Genomics, SB Primrose & RM Twyman. Blackwell Publishing, Oxford.

Martin, F., Cullen, D., Hibbett, D., Pisabarro, A., Spatafora, J.W., Baker, S.E., Grigoriev, I.V., 2011. Sequencing the fungal tree of life. New Phytol. 190, 818–821.

Neufeld, J.D., Wagner, M., Murrell, J.C., 2007. Who eats what, where and when? Isotope-labelling experiments are coming of age. ISME J. 1, 103–110.

Wellington, E.M.H., Berry, A., Krsek, M., 2003. Resolving functional diversity in relation to microbial community structure in soil: exploiting genomics and stable isotope probing. Curr. Opin. Microbiol. 6, 295–301.

Examples of Molecular Approaches to Investigating the Diversity and Ecology of Fungi

Baldrian, P., Kolařík, M., Štursová, M., Kopecký, J., Valášková, V., Větrovský, T., Žifčáková, L., Šnajdr, J., Rídl, J., Vlček, Č., 2011. Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. ISME J. 6, 248–258.

Buée, M., Reich, M., Murat, C., Morin, E., Nilsson, R.H., Uroz, S., Martin, F., 2009. 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. New Phytol. 184, 449–456.

Butt, T.M., Greenfield, B.P., Greig, C., Maffeis, T.G., Taylor, J.W., Piasecka, J., Dudley, E., Abdulla, A., Dubovskiy, I.M., Garrido-Jurado, I., 2013. Metarhizium anisopliae pathogenesis of mosquito larvae: a verdict of accidental death. PloS One 8, e81686.

Clemmensen, K.E., Bahr, A., Ovaskainen, O., Dahlberg, A., Ekblad, A., Wallander, H., Stenlid, J., Finlay, R.D., Wardle, D.A., Lindahl, B.D., 2013. Roots and associated fungi drive long-term carbon sequestration in boreal forest. Science 339, 1615–1618.

Cornell, M., 2007. Comparative genome analysis across a kingdom of eukaryotic organisms: specialization and diversification of the fungi. Genome Res. 17, 1809–1822.

Eastwood, D.C., et al., 2011. The Plant Cell Wall–Decomposing Machinery Underlies the Functional Diversity of Forest Fungi. Science 333, 762–765.

Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., Robeson, M., Edwards, R.A., Felts, B., Rayhawk, S., Knight, R., Rohwer, F., Jackson, R.B., 2007. Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. Appl. Environ. Microbiol. 73, 7059–7066.

Hess, J., Skrede, I., Wolfe, B., LaButti, K., Ohm, R.A., Grigoriev, I.V., Pringle, A., 2014. Transposable element dynamics among asymbiotic and ectomycorrhizal Amanita fungi. Genome Biol. Evol. 6 (7), 1564–1578.

Hibbett, D.S., Ohman, A., Glotzer, D., Nuhn, M., Kirk, P., Nilsson, R.H., 2011. Progress in molecular and morphological taxon discovery in Fungi and options for formal classification of environmental sequences. Fungal Biol. Rev. 25, 38–47.

Hibbett, D.S., Donoghue, M.J., 2001. Analysis of character correlations among wood decay mechanisms, mating systems, and substrate ranges in homobasidiomycetes. Syst. Biol. 50, 215–242.

Hibbett, D., Thorn, R., 2001. Basidiomycota: homobasidiomycetes. In: McLaughlin, D.J., McLaughlin, E.G., Lemke, P.A. (Eds.), The Mycota VII, Part B: Systematics and Evolution. Springer, Berlin, pp. 121–168.

James, T.Y., et al., 2006. Reconstructing the early evolution of fungi using a six-gene phylogeny. Nature 443, 818–822.

Johnson, J., Evans, C., Brown, N., Skeates, S., Watkinson, S., Bass, D., 2014. Molecular analysis shows that soil fungi from ancient semi-natural woodland exist in sites converted to non-native conifer plantations. Forestry 87, 705–717.

Kershaw, M.J., Talbot, N.J., 2009. Genome-wide functional analysis reveals that infection-associated fungal autophagy is necessary for rice blast disease. Proc. Natl. Acad. Sci. U.S.A. 106 (37), 15967–15972.

Kohler, A., Kuo, A., Nagy, L.G., Morin, E., Barry, K.W., Buscot, F., Canbäck, B., Choi, C., Cichocki, N., Clum, A., 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nat. Genet. 47, 410–415.

Lindahl, B.D., Ihrmark, K., Boberg, J., Trumbore, S.E., Högberg, P., Stenlid, J., Finlay, R.D., 2007. Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. New Phytol. 173, 611–620.

Martin, F., et al., 2008. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature 452, 88–92.

Martinez, D., Challacombe, J., Morgenstern, I., Hibbett, D., Schmoll, M., Kubicek, C.P., Ferreira, P., Ruiz-Duenas, F.J., Martinez, A.T., Kersten, P., 2009. Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion. Proc. Natl. Acad. Sci. 106, 1954–1959.

Martinez, D., Larrondo, L.F., Putnam, N., Gelpke, M.D.S., Huang, K., Chapman, J., Helfenbein, K.G., Ramaiya, P., Detter, J.C., Larimer, F., 2004. Genome sequence of the lignocellulose degrading fungus Phanerochaete chrysosporium strain RP78. Nat. Biotechnol. 22, 695–700.

Pickles, B.J., Genney, D.R., Potts, J.M., Lennon, J.J., Anderson, I.C., Alexander, I.J., 2010. Spatial and temporal ecology of Scots pine ectomycorrhizas. New Phytol. 186, 755–768.

Richards, T.A., 2011. Genome evolution: horizontal movements in the fungi. Curr. Biol. 21, R166–R168.

Richards, T.A., Talbot, N.J., 2013. Horizontal gene transfer in osmotrophs: playing with public goods. Nat. Rev. Microbiol. 11, 720–727.

Riley, R., et al., 2014. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. Proc. Natl. Acad. Sci. U.S.A. 111, 9923–9928.

Rosling, A., Timling, I., Taylor, D.L., 2013. Archaeorhizomycetes: patterns of distribution and abundance in soil. In: Horwitz, B.A., Mukherjee, P.K., Mukherjee, M., Kubicek, C.P. (Eds.), Genomics of Soil- and Plant-Associated Fungi. Springer, Berlin Heidelberg, pp. 333–349.

Skrede, I., Engh, I., Binder, M., Carlsen, T., Kauserud, H., Bendiksby, M., 2011. Evolutionary history of Serpulaceae (Basidiomycota): molecular phylogeny, historical biogeography and evidence for a single transition of nutritional mode. BMC Evol. Biol. 11, 230.

Slot, J.C., Hibbett, D.S., 2007. Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. PLoS One 2. e1097.

Talbot, J. M., Martin, F., Kohler, A., Henrissat, B., and Peay, K. G., 2015. Functional guild classification predicts the enzymatic role of fungi in litter and soil biogeochemistry. Soil Biol. Biochem. 88, 441– 456.

Tunlid, A., Rineau, F., Smits, M., Shah, F., Nicolas, C., Johansson, T., Persson, P., Martin, F., 2013. Genomics and spectroscopy provide novel insights into the mechanisms of litter decomposition and nitrogen assimilation by ectomycorrhizal fungi. Genom. Soil Plant Assoc. Fungi 191, 441–456.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., Dufresne, A., 2015. The importance of the microbiome of the plant holobiont. New Phytol. 206, 1196–1206.

van der Heijden, M.G.A., Martin, F.M., Selosse, M.-A., Sanders, I.R., 2015. Mycorrhizal ecology and evolution: the past, the present, and the future. New Phytol. 205, 1406–1423.