available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/funeco

Commentary

Prospects and challenges for fungal metatranscriptomics of complex communities

ABSTRACT

The ability to extract and purify messenger RNA directly from plants, decomposing organic matter and soil, followed by high-throughput sequencing of the pool of expressed genes, has spawned the emerging research area of metatranscriptomics. Each metatranscriptome provides a snapshot of the composition and relative abundance of actively transcribed genes, and thus provides an assessment of the interactions between soil microorganisms and plants, and collective microbial metabolic processes in many environments. We highlight current approaches for analysis of fungal transcriptome and metatranscriptome datasets across a gradient of community complexity, and note benefits and pitfalls associated with those approaches. We discuss knowledge gaps that limit our current ability to interpret metatranscriptome datasets and suggest future research directions that will require concerted efforts within the scientific community.

Introduction

Fungi have significant ecological roles as components of complex microbial communities in many diverse environments – including soil, marine and freshwater habitats, animal and insect digestive systems, and within plant and animal hosts. Perhaps most well known for their contributions to terrestrial ecosystems, fungi are the major decomposers of plant and soil organic matter and form critical nutritional linkages with plants through mycorrhizal, pathogenic and endophytic associations. Playing central roles in nutrient and mineral availability and mobilization, fungi represent some of the most functionally diverse organisms on Earth.

Molecular approaches based on the ribosomal RNA operon or single enzyme-coding genes have been used to assess taxonomic diversity and track fungal community composition. Reverse transcriptase PCR surveys using single enzyme-encoding genes (*i.e.* cellobiohydrolase; Weber et al., 2012) have identified specific transcripts in fungal communities that suggest ecological roles for fungal populations. However, single-gene studies are limited by primer selectivity, inability to capture the entire community, and in scope, failing to assess related functions or metabolic processes. Collectively surveying for all transcribed enzyme-coding genes in an environmental sample would improve our understanding of the metabolic diversity, activities, and community interactions among fungal species and in their associations with plants, bacteria, and other Eukarya.

The ability to extract and purify RNA directly from plants, decomposing organic matter, and soil, followed by high-throughput sequencing of the pool of expressed genes, is an emerging capability through which we may assess the plant–fungal interactions and fungal metabolism in soils. This approach, termed metatranscriptomics, generates a snapshot of the composition and relative abundance of actively transcribed genes at a single time point. Comparison of metatranscriptomes over time, and across environmental variables or gradients, provides information on the collective, interactive metabolism as the community responds to changing environmental conditions. Metatranscriptomic surveys have the potential to be particularly powerful when combined with parallel surveys, such as target-gene sequencing, metagenomics, meta-metabolomics and meta-proteomics. Partnered with information on soil geochemistry, plant species or nutritional status, or other experimental factors, these surveys become part of a comprehensive ecosystem study. Transcriptome and metatranscriptome surveys also facilitate discovery of novel enzymes and processes with importance to bioenergy, medical, agricultural, industrial and ecosystem climate response applications.

To efficiently survey the expressed gene pool, one must first obtain a mRNA sample from which DNA, ribosomal RNA, and impurities that inhibit downstream sample processing have been removed. This has been a significant technical challenge, especially in environmental samples where transcripts are low in abundance relative to the rRNA pool and where co-extracted inhibiting compounds are abundant. Recent advances in this area include rigorous extraction to remove RNAase enzymes and co-extracting contaminants, use of polyA-enrichment (where only the Eukarya are the target of investigation), and removal of rRNA using affinity reagents. Studies have now demonstrated successful extraction and purification of fungal mRNA from soil, plant roots, and decomposing leaf litter (see for examples, Griffiths et al., 2000; Liao et al., 2014; Weber et al., 2012).

Fungal transcriptomics – from single cultures to complex soil communities

Transcriptomes from single fungal cultures have provided key reference material for genome annotation and allowed the study of active fungal metabolism (for examples, Hori et al., 2014; Tisserant et al., 2012). Despite these successes, there remain significant gaps in our ability to use transcriptome datasets in their entirety, and there is a significant challenge to glean biologically meaningful inferences from them. These challenges become exponentially more daunting when assessing metatranscriptomic datasets that represent pairs of organisms or more complex communities.

Fig 1 illustrates a gradient of fungal transcriptome complexity from a single cultured fungus (box A) to very complex natural soil systems (C). It notes several features of current metatranscriptome studies, including approaches for analysis and key metabolic questions being addressed. The simplest case is a single fungal transcriptome (Fig 1, box A). A single-culture transcriptome approach involves sequencing both the genome and the expressed genes (e.g. using RNA-seq). Transcripts are identified by mapping sequences to the annotated fungal genome. Genome annotation quality can be variable, however, with many gene functions unidentified (hypothetical) or identified based solely on sequence homology to other genomes. This has been shown to be an exceptionally useful approach to identify metabolic pathways that are active in the presence of different substrates (See for example Hori et al., 2014). However, single-organism transcriptome datasets can be difficult to interpret because of differences in transcript turn-over rates and translational efficiencies, and the limitations of using homology-based gene assignments to correctly predict function. Interpretation is greatly improved when transcriptomes are combined with high-quality, accurately annotated genome and proteome information (Hori et al., 2014).

Metatranscriptome surveys have the potential to identify the metabolic interactions occurring between two or more organisms (Fig 1, box B). An example with relevance to plant health is between a mycorrhizal fungus and its plant partner (for examples, Liao et al., 2014; Larsen et al., 2011). With more organisms in the mRNA mix, the sequence datasets become more complex, but if annotated genome sequences for both fungus and plant are available, the metabolic signaling, interactions and their regulation may be determined. A notable challenge when simultaneously investigating two or

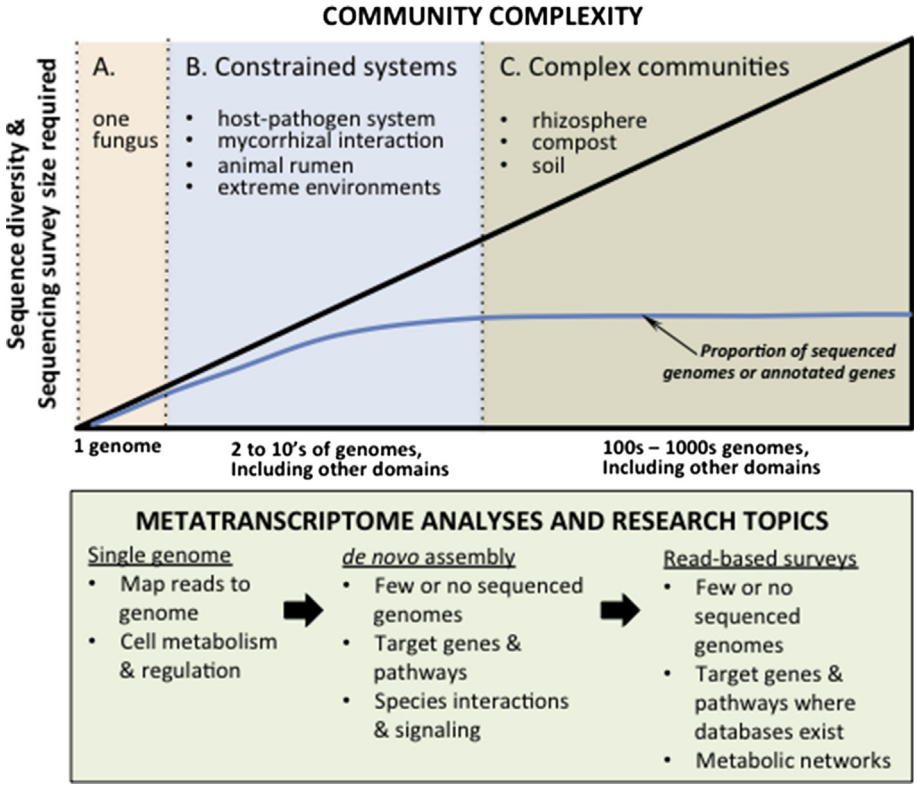

Fig 1 – The figure illustrates the impact of increasing fungal community complexity on our ability to assess and interpret metatranscriptome information. Panel A begins the continuum of complexity with transcriptome analysis of a single organism, where expressed genes may be mapped to a genome and where cell metabolism may be studied through transcriptome profiling. Panel B represents constrained systems where the focus is on a few co-habiting organisms that interact closely. For this level of metatransriptome study, *de novo* assembly or assessment of target genes and pathways that may influence the interaction are often the focus of study. Panel C represents complex communities such as the soil, where current analyses are often limited to read-based surveys.