Journal of Microbiological Methods 88 (2012) 224–228

Contents lists available at SciVerse ScienceDirect

Journal of Microbiological Methods

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

Comparative assessment of fungal cellobiohydrolase I richness and composition in cDNA generated using oligo(dT) primers or random hexamers

Carolyn F. Weber [1], Cheryl R. Kuske *

Los Alamos National Laboratory, Bioscience Division, Mail Stop 888, Los Alamos, NM 87545, United States

ARTICLE INFO

Article history:
Received 25 October 2011
Received in revised form 28 November 2011
Accepted 28 November 2011
Available online 8 December 2011

Keywords:
cDNA synthesis
Random hexamers
Oligo(dT)
Cellobiohydrolase I gene
Soil fungi

ABSTRACT

Understanding soil fungal distribution and activities, particularly at the level of gene expression, is important in unveiling mechanisms regulating their activities *in situ*. Recent identification of fungal genes involved in carbon cycling has provided the foundation for developing reverse-transcriptase PCR assays to monitor spatiotemporal gene expression patterns in soils and other complex microbial systems. The polyadenylated 3′ ends of eukaryotic mRNA transcripts enables the use of oligo(dT) primers for cDNA synthesis, but this can result in the overrepresentation of the 3′ end of transcripts in cDNA pools. In an effort to increase the uniformity of transcripts represented in cDNA pools, random hexamers have been used. The use of both priming methods is abundant in the literature, but we do not know how these methods perform relative to each other. We performed comparative richness and compositional analyses of the fungal glycosyl hydrolase family 7 cellobiohydrolase I gene *cbh*I amplified from soil cDNAs that had been generated using either oligo(dT) primers or random hexamers. Our results demonstrate that similar *cbh*I richness and composition were recovered using both approaches. Richness estimates and compositional profiles of *cbh*I sequence libraries generated from random hexamer-primed cDNA were more variable than from libraries generated from oligo(dT) primed cDNA. However, our overall results indicate that, on average, comparable richness and composition were recovered from soil cDNAs when either priming method was used.

Published by Elsevier B.V.

1. Introduction

Fungi often dominate soil microbial biomass where they play critical roles in biogeochemical cycling (Ehrlich, 2006), but we do not yet understand how fungal distribution and activities *in situ* affect these cycles. Increased information on genes involved in specific pathways (Gregoriev et al., 2011) and rapid development of molecular tools are enabling unprecedented progress in this area of research, particularly at the level of gene expression (Kellner et al., 2010). Increasingly, studies are being conducted that examine the richness and composition of expressed functional genes that encode key enzymes in carbon and nitrogen cycling (e.g. cellobiohydrolase, laccase, manganese peroxidase; Luis et al. 2005; Kellner et al. 2010, Courty et al. 2008; Edwards et al. 2011; Kellner and Zak 2009).

Recent studies have demonstrated that the gene encoding the catalytic subunit of cellobiohydrolase I, that cleaves cellobiose from the reducing and non-reducing ends of the cellulose molecule, can be specifically amplified by PCR from soil fungal communities. This gene serves as a suitable genetic marker for a phylogenetically broad subset of cellulolytic fungi (Edwards et al., 2008; Weber et al., 2011; Baldrian et al. 2011), and is a useful molecular target for assaying the response of cellulolytic fungi to various environmental perturbations (e.g. increasing atmospheric CO_2) that potentially affect carbon cycling in soil. In particular, monitoring the expression of this gene *in situ* will identify fungi that may be important in soil cellulose degradation. Understanding the limitations of methods used to monitor expression of this gene is necessary for making proper ecological interpretations from soil gene expression studies.

Two-step reverse transcription PCR is commonly used to assay target gene expression in environmental samples. Total RNA is first extracted and then reverse transcribed to complimentary (c) DNA after any desired selection for specific fractions of the RNA (e.g. mRNA). Specific genes of interest are then PCR-amplified from the cDNA for downstream sequencing, fingerprinting or other analyses. However, characteristics of cDNA pools may vary dramatically depending on how the reverse transcriptase reaction is primed. Two commonly used types of primers are oligo(dT) and random hexamers. Our study investigated the potential impacts of primer set choice on the richness and composition of an expressed fungal functional gene in soil.

For about 40 years, it has been known that eukaryotic mRNAs are polyadenylated at the 3′ end of transcripts (Sarkar, 1996) and this is hypothesized to contribute to their stability and enhance translation (Anderson, 2005; Guhaniyogi and Brewer, 2001, Sachs, 1998). Polyadenylation also has been discovered in eukaryotic rRNAs (Slomovic et

* Corresponding author. Tel.: +1 505 664 4800.
E-mail address: kuske@lanl.gov (C.R. Kuske).
[1] Present address: Idaho State University, Department of Biological Sciences, 921 South 8th Ave, Stop 8007, Pocatello, ID 83209.

0167-7012/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.mimet.2011.11.016

al., 2006; Kuai et al., 2004), some bacterial mRNAs (Gopalakrishna et al. 1981; Anderson, 2005), *Chlamydomonas reinhardtii* chlorplast RNAs (Komine et al., 2000), and noncoding RNAs (Li et al., 2002; Kushner, 2002). Nonetheless, pre-selecting for or targeting polyadenylated eukaryotic mRNA transcripts using oligo(dT) primers during reverse transcription have remained methods of choice for examining functional gene transcripts (i.e. Cloonan et al. 2008, Mortazavi et al., 2008) amidst total RNA, which is overwhelmingly dominated by ribosomal RNA (ca. 90%).

The use of oligo(dT) primers to synthesize cDNA can result in the overrepresentation of the 3′ end of transcripts in the cDNA pool (Mortazavi et al., 2008). The ability of the reverse transcriptase enzyme to transcribe cDNA over the entire length of a transcript is a function of the sequence characteristics and length. Consequently, not all transcripts will be represented equally well in the cDNA pool (Mortazavi et al., 2008). To circumvent this problem, many transcriptomic studies, utilizing both shotgun and target gene approaches, use random hexamers to generate cDNA, in attempts to capture greater representation of transcripts along their entire lengths (Mortazavi et al., 2008). A recent study in which transcriptomes were sequenced on the Illumina genoma analyzer platform noted that random hexamers introduced bias in the nucleotide composition at the beginning of the reads in the resulting transcriptome, impacting the uniformity of the reads (Hansen et al., 2010).

Many of the methods being used for examining gene expression in microbial communities were originally optimized for use in single model organisms in the laboratory and continue to evolve as technologies advance. Furthermore, we do not know how well these methods perform when applied to complex environments or if results obtained with different methods are comparable; the latter impacts our ability to compare across studies using different approaches. In particular, biases that may result from use of different reverse transcriptase PCR approaches for specific functional genes remain largely unknown. This combined with the increasing abundance of studies in the literature that assess the richness and composition of expressed target functional genes in soils using either random hexamers and oligo(dT) priming make it important to know how these two methods perform relative to each other.

We tested the hypothesis that use of random hexamers or oligo(dT) primers to synthesize cDNA would provide significantly different measures of richness and composition in libraries of expressed fungal cellobiohydrolase genes. To accomplish this, we extracted RNA from triplicate soil samples, synthesized cDNA from each RNA extract using either random hexamers or oligo(dT) primers and generated libraries of cellobiohydrolase I gene fragments. Richness estimates and compositional analyses of the two library types were compared.

2. Materials and methods

2.1. Study site, soil collection and experimental design

In April 2010, triplicate surface soil cores (0–2 cm) were collected at the U.S. Department of Energy FACE (Free-air Carbon Dioxide Enrichment) site in the Duke Forest (North Carolina) using 7.5 cm diameter aluminum core tubes. Samples were collected in an equilateral triangle pattern approximately 1 m apart. Samples were collected in a plot that had been exposed to ambient levels of CO_2 and had been fertilized with nitrogen since 2005 at a rate of 11.2 g N m^{-2} (manually applied pellet ammonium nitrate). Other physical and chemical characteristics of the site have been described previously (McCarthy et al. 2010; Weber et al., 2011; http:/face.env.duke.edu/fertilization.cfm). Immediately after collection, each of the three soil samples was homogenized in ziptop plastic bags. A representative subsample of each soil was collected in a separate 50 mL Falcon tube and then flash frozen in liquid nitrogen. Samples were stored on dry ice during transport back to the laboratory where they were stored at −70 to −80 °C until RNA was extracted.

2.2. RNA extractions and reverse-transcription

Prior to extracting RNA, samples were crushed under liquid nitrogen using a mortar and pestle. For each of the three soil samples, RNA was extracted from three 2.0 g-subsamples (nine total RNA extracts). RNA was extracted using the MoBio RNA Powersoil Total RNA Isolation Kit (MoBio Laboratories, Carlsbad, CA) according to the manufacturer's protocol. DNA was removed by treating the extracts with Turbo DNase (Ambion, Austin, TX) at 37 °C for one hour followed by inactivation using 0.2 volumes of DNase inactivation reagent (Ambion, Austin, TX) per extract. Samples were incubated with the inactivation reagent with occasional vortexing for two minutes. The inactivation reagent was pelleted via centrifugation at 8000×*g* for 1.5 min. Resulting DNA-free RNA was transferred to nuclease-free sterile microcentrifuge tubes. Concentrations and purity of RNA extracts were determined using a Nanodrop 2000c (Thermo Scientific, Wilmington, DE). Each of the extracts was adjusted to a concentration of ca. 60 ng μl^{-1}.

Soil cDNA was synthesized immediately from each of the nine RNA extractions. From each of the nine RNA extracts, two pools of single-stranded complimentary DNA (cDNA) were generated; one cDNA pool was generated by reverse transcribing RNA using an Oligo(dT)18 primer (OdT) and the other was generated using random hexamers (RH) from Roche Applied Science (Indianapolis, IN). For each reverse-transcriptase reaction, 3.5 μl of RNA and 1 μl of 20 μM oligo(dT)18 primer or 20 μM random hexamers were placed into 0.2 mL PCR tubes and incubated at 72 °C for 3 min and then 42 °C for 2 min in an Eppendorf Master Cycler Pro (Eppendorf North America, Hauppauge, NY). Immediately after incubation, 5.5 μl of master mix containing the following reagents (Clontech, Mountain View, CA) was added to each reaction (final concentrations listed): 1X First-Strand Buffer, 2.5 mM DTT, 1 mM dNTP, 0.25 μl RNase Inhibitor and 1 μl SMARTscribe reverse transcriptase. Reverse transcription was carried out at 42 °C for 90 min and terminated at 70 °C for 10 min in an Eppendorf Master Cycler Pro (Eppendorf North America, Hauppauge, NY). Each of the single-stranded cDNA products was diluted with 40 μl of 1X TE buffer (pH = 8.0; Roche Applied Sciences, Indianapolis, IN).

2.3. PCR amplification, cloning and sequencing of cbhI

From each of the 18-cDNA preparations, *cbh*I was PCR-amplified in triplicate 25 μl reactions using primers fungalcbhIF and fungalcbhIR (Edwards et al., 2008), reagents and thermal cycling conditions previously described (Weber et al., 2011). All PCR products were visualized on an ethidium bromide-stained 1% agarose gel (TBE). Triplicate products were pooled and purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) prior to cloning using the TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA). From each cloning reaction, 96 clones were manually selected. Bidirectional sequencing was completed using Sanger technology.

2.4. Sequence assembly, alignment and analysis

Bidirectional reads were assembled using Sequencher version 4.7 (Gene Codes Corporation, Ann Arbor, MI) and were translated using the Baylor College of Medicine Search Launcher (http://searchlauncher.bcm.tmc.edu/seq-util/seq-util.html). Sequences less than 470 nucleotides in length and those containing ambiguities were eliminated from the data set. Inferred amino acid sequences from all 18 libraries were aligned together using default parameters in Muscle 3.6 (Edgar, 2004). The alignment was imported into ARB (Pruesse et al., 2007) to generate a distance matrix based on 169 positions; positions where gaps occurred most often were not included in the distance matrix calculation. The