European Journal of Soil Biology 77 (2016) 53–59

Contents lists available at ScienceDirect

European Journal of Soil Biology

journal homepage: http://www.elsevier.com/locate/ejsobi

Original article

A new primer set for DNA metabarcoding of soil Metazoa

E. Capra [a, *], R. Giannico [b], M. Montagna [c], F. Turri [a], P. Cremonesi [a], F. Strozzi [b], P. Leone [a], G. Gandini [d], F. Pizzi [a]

[a] Istituto di Biologia e Biotecnologia Agraria, Consiglio Nazionale delle Ricerche, via Einstein, 26900 Lodi, Italy
[b] Parco Tecnologico Padano, via Einstein, 26900 Lodi, Italy
[c] Università degli Studi di Milano, Dipartimento di Scienze Agrarie e Ambientali - Produzione, Territorio, Agroenergia, via Celoria 2, 20133 Milano, Italy
[d] Università degli Studi di Milano, Dipartimento di Medicina Veterinaria, via Celoria 10, 20133 Milano, Italy

ARTICLE INFO

Article history:
Received 19 February 2016
Received in revised form
4 October 2016
Accepted 14 October 2016

Handling editor: Christoph Tebbe

Keywords:
Metazoa
Soil
18S rRNA
Metagenomics

ABSTRACT

Reliable evaluations of soil biodiversity represent a key factor in understanding ecosystem services. To date, species-discriminating barcodes efficiently describe bacterial and fungal communities associated with environmental samples, whereas investigations of soil microfauna are often hampered by the lack of a marker region encompassing the taxonomic range of soil organisms. Two new PCR primer sets targeting the V4-V5 and V5-V7 variable regions of the ribosomal 18S RNA (18S rRNA) were designed to be specific for metazoans metabarcoding and capable of detecting the majority of their lineages. *In silico* and *in vivo* assays on four soil typologies were carried out to compare the newly developed primer sets with a selection of primers targeting the homologous gene, which were previously used to assess soil metazoan biodiversity. The new primer sets, both on the basis of the *in silico* and *in vivo* comparisons, were very selective and consistent when analysing metazoan biodiversity across the tested soil typologies. On the basis of the coverage index and taxonomic resolution, the new primers targeting the ribosomal 18S RNA outperformed the other primers, and they represent a promising tool for assessing soil metazoan biodiversity through metabarcoding approaches.

1. Introduction

The soil ecosystem hosts high numbers of organisms [15] and represents a major reservoir of biodiversity. In this complex system, invertebrates are key components both in terms of abundance and in providing ecosystem services, such as actively contributing to soil formation and cycling of elements [18]. Reliable approaches to measure and evaluate biodiversity of soil organisms represent promising tools for a plethora of purposes, encompassing conservation biology investigations, biomonitoring programs, and evaluation of soil alterations. Traditional approaches to investigate biodiversity of soil organisms rely on morphological identification of taxa; the disadvantages of this procedure are the high level of expertise and the time effort required [30]. In recent years, the increasing use of DNA markers in species identification [14,23,25], has changed the morphology-based perspective in biodiversity studies. The metabarcoding approach is the leading candidate [1,3], to investigate organism diversity and to provide ecological information [11,33]. Briefly, the metabarcoding approach consists of detection and identification of taxa from complex environmental matrices (e.g., soil, soil litter, canopy, and water) starting from DNA extracted from bulk samples. Regardless of the high-throughput sequencing platforms adopted, molecular markers and primers play a crucial role in the recovery of a high proportion of taxa diversity. For some metazoan groups, the most commonly used marker is a portion of the mitochondrial Cytochrome Oxidase subunit I (COI), characterized by a relatively high nucleotide substitution rate [24] which allows for discrimination of organisms at the species level. However, because of these features, it is difficult to identify regions conserved across the main animal lineages to design primers suitable for DNA metabarcoding. Nevertheless, primer sequences amplifying a 5′ fragment of COI in metazoans [10] were successfully adapted to high-throughput sequencing technologies to characterize the metazoan biodiversity of bulk samples, including arthropods [33], and coral fish gut content [19].

The use of markers more conserved than the COI gene represents a possible solution to obtain universal PCR primers for metazoans. The small ribosomal subunit 18S RNA (18S rRNA),

* Corresponding author.
E-mail address: e.capra@ibba.cnr.it (E. Capra).

http://dx.doi.org/10.1016/j.ejsobi.2016.10.005
1164-5563/

which possesses a lower nucleotide substitution rate compared to the COI gene, represents an alternative to mitochondrial markers. Nevertheless, a possible flaw of 18S rRNA is low discriminating capability in distinguishing among closely related species. 18S rRNA primers were successfully used in biodiversity surveys of marine ecosystems, characterized mainly by polychaetes [8,11], and of soil ecosystems characterized by nematodes and arthropods [26,32].

The type of sample can greatly influence the choice of the region (COI, 18S rRNA) for metazoan metabarcoding. Recently, 18S rRNA metabarcoding was successfully used to characterize the metazoan communities in leaf-litter and soil. Because of the inability of COI to amplify from soil samples, >99% of returned Operational Taxonomic Units (OTUs) were bacterial, and the use of COI was limited in evaluation of Metazoa present in Malaise traps and Canopy fogging[32]. Similarly, the characterization of communities associated with *Zostera marina* seagrass meadows, using both mitochondrial and nuclear ribosomal genetic markers (COI and 18S rRNA) revealed that the highest proportion of OTUs were non-Metazoa in the case of COI gene, whereas a large number of OTUs were identified as metazoans for 18S rRNA. Furthermore, the18S gene uncovered higher numbers of common species as well as a greater number of overall species when compared to COI [8]. Notwithstanding, the DNA metabarcoding of soil with primers targeting the 18S rRNA resulted in a high proportion of OTUs returned as non-metazoan hits [8,32]. Thus, currently, a set of primers targeting 18S rRNA specific for metazoans is still lacking. Until a few years ago, to study biodiversity associated with environmental samples, most investigators preferred a 454 pyrosequencing platform, as it supplies reads longer than other technologies such as Illumina sequencing. New Illumina platforms and strategies (*i.e.,* Miseq instrument), in combination with 2x300 bp paired-end, allowed sequencing of amplicons with length comparable to that previously obtained by 454 pyrosequencing, with the advantage of higher quality and lower costs per base [2,6,17]. Taking advantage of these developments, Illumina platforms have been recently adopted to characterize soil biodiversity of different taxonomic groups across the whole tree of life, including fungi [29], plants, metazoans, protozoans, and bacteria [12].

The aim of this study was to develop new sets of primers on the 18S rRNA capable of encompassing the wide taxonomic diversity of metazoans inhabiting the soil ecosystem. The new primer sets, suited for Illumina platforms, were compared with previously published primers on four soil typologies.

2. Materials and methods

2.1. Primer design and experimental setup

Two new sets of primers, targeting the 18S rRNA gene, were *in silico* designed and tested to evaluate their efficiency in recovering soil metazoan diversity. Primers were designed in regions showing variation between metazoan and non-metazoan groups, using Fungi as a reference. The subset included different metazoan taxa: Artropoda (Accession number AY703484, EU432215, AY859604), Annelida (Accession number AF411895, AJ272183), Nematoda (Accession number AY284671, AY284591), Mollusca (Accession number EF489341, FJ917212), Rotifera (Accession number DQ297695, AJ487049), Tardigrada (Accession number FJ435736, DQ839607), Cordata (Accession number AB211066), Platyhelminthes (Accession number AJ270162); and Fungi: Ascomycota (Accession number AY838789, AF258606), Basidiomycota (Accession number AB021676). Small subunit sequences (SSU) from the NCBI database were aligned with BioEdit [13]. The two primer sets, M1041F, M1648R (hereafter Meta1) and M620F, M1260R (hereafter Meta2), were designed using the Primer3 software [31]. These primers were compared with previously published primers targeting the 18S rRNA gene: SSU_FO4, SSU_R22 (hereafter Meta3) [11], and #1, #2_RC (hereafter Meta4) [21] (Table 1).

The performances of *in silico* primers were evaluated on SSU rRNA SILVA database, by using Test Prime 1.0, allowing for one mismatch (www.arb-silva.de/search/testprime; [16]). Finally, the selected primers (i.e., the newly developed and those previously published) were experimentally compared through metabarcoding of four soil typologies. Three technical replicates for each soil typology and the primer pair were carried out, resulting in a total of 48 samples.

2.2. Sampling and DNA extraction

A total of four soil samples were collected in Northern Italy from cornfield (C), woodland (W), grazeland (G), and heavily grazed pasture (H) after removing vegetation cover. Each sample consisted of soil core of 5 cm diameter and 20 cm of length. After homogenization, the samples were kept at −20 °C. DNA was extracted from soil (0.5 g) using Nucleo Spin Soil kit Macherey-Nagel (Düren Germany).

2.3. DNA metabarcoding protocol: PCR amplicons generation, library preparation, Miseq run

DNA was amplified with locus specific primers added to Illumina overhang adapter sequences. Forward overhang: 5′ TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[locus specific target primer], Reverse overhang: 5′ GTCTCGTGGGCTCGGAGATGT GTATAAGAGACAG-[locus specific target primer]. The 18S rRNA gene was amplified with the two new primer sets Meta1 and Meta2, and with two previously published primers pairs Meta3 and Meta4. PCRs were performed in 50 μl volumes using Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific, MA USA) following manufacturer instructions, with 0.2 μl of each primer (100 μM), and 2 μl of genomic DNA (5 ng/μl). For 18S rRNA amplification, the cycling conditions were 30 s at 98 °C, followed by 25 cycles of 1 min at 98 °C, 30 s at 58 °C, 1 min at 72 °C, and finally 7 min at 72 °C. Finally, amplicons were cleaned-up with Agencourt AMPure XP (Beckman, Coulter Brea, CA) and the sizes were checked with a 2100 Bioanalyser Instrument (Agilent Technologies, Santa Clara, CA). Libraries were prepared with a second stage PCR using Nextera XT Index 1 Primers (N7XX) and Nextera XT Index 2 Primers (S5XX) from the Nextera XT Index kit (FC-131-1001 or FC-131-1002), following 16S Metagenomic Sequencing Library Preparation protocol (http://www.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf) with some modifications. The libraries obtained were quantified by Real Time PCR with KAPA Library Quantification Kits (Kapa Biosystems, Inc. MA, United States) pooled in equimolar proportions and sequenced, MiSeq Reagent kit v3 (600 cycles), with paired-end reads of 300 bp.

2.4. Bioinformatics data processing and analyses

Illumina raw reads were trimmed using Trimmomatic v0.32 [4] requiring a minimum base quality of 20 (Phred scale) and a minimum read length of 36 nucleotides. Only trimmed reads were included in downstream analysis. For long amplicons with non-overlapping paired ends, the first paired read was concatenated to the reverse complement of second paired end read, separating them with a single N base. For partial overlapping reads, the consensus was performed using fastq-join tool (code.google.com/p/