DOI: 10.6092/issn.2531 – 7342/6647

Who is out there? What are they doing? Application of metagenomics and metaproteomics to reveal soil functioning

Antonietta Mello[1], Elisa Zampieri[2]

[1]Istituto per la Protezione Sostenibile delle Piante, SS di Torino, CNR, Viale P.A. Mattioli 25, I-10125 Torino
[2]Dipartimento di Scienze Agrarie, Forestali e Alimentari (DISAFA), Università di Torino, Largo Paolo Braccini 2, I-10095 Grugliasco (TO);
CREA, Council for Agricultural Research and Economics, Rice Research Unit S.S. 11 to Torino, Km 2,5 13100, Vercelli

corresponding author: Antonietta Mello, antonietta.mello@ipsp.cnr.it

Abstract

In order to link microbial community composition to ecological processes happening in the brûlé, a metaproteomics analysis was applied to a brûlé previously characterized by metagenomics. The metagenomics data had showed a reduced fungal biodiversity, a dominance of *Tuber melanosporum* and a reduced presence both of ectomycorrhizal Basidiomycota and of bacteria belonged to *Pseudomonas* and Flavobacteriaceae inside the brûlé. By metaproteomics analysis, the identified proteins revealed which biological processes were more represented or only present in the brûlé, and, among them, processes related to multiple stresses were identified in herbaceous plants. This study demonstrates that combining the data of metagenomics and metaproteomics gives the opportunity to potentially reveal the functioning of any environment.

Keywords: truffle-ground soil; Tuber melanosporum;fungi; bacteria; herbaceous plants

Riassunto

L'obiettivo di questo lavoro è stato quello di applicare lo strumento della metaproteomica a suoli di tartufaia già approfonditamente caratterizzati tramite approcci di metagenomica. Per fare questo è stato impiegato un database creato ad hoc a partire dagli organismi fungini, batterici e dalle piante identificati negli studi di metagenomica. Attraverso l'identificazione delle proteine estratte dal suolo della tartufaia si è cercato di chiarire quali processi metabolici fossero ivi attivi e quali fossero specificatamente presenti nel pianello di *Tuber melanosporum*, cioè in un'area caratterizzata da scarsa vegetazione attorno alla sua pianta ospite. Gli studi di metagenomica avevano messo in evidenza una minore biodiversità dei funghi nel pianello e che *T. melanosporum* era il fungo dominante, mentre i funghi ectomicorrizici appartenenti al phylum dei Basidiomycota diminuivano, suggerendo competizione con il tartufo. All'interno del pianello risultavano anche scarsamente presenti batteri appartenenti sia al genere *Pseudomonas* che alla famiglia delle Flavobacteriaceae. Con l'applicazione della metaproteomica è stato possibile individuare quali processi biologici fossero maggiormente presenti nel pianello, o esclusivi. I processi legati a multipli tipi di stress erano principalmente

DOI: 10.6092/issn.2531 – 7342/6647

rappresentati, e presenti soprattutto nelle piante erbacee. Questo studio ha permesso quindi di svelare il funzionamento di un ambiente così particolare come il pianello, dove il tartufo nero è il protagonista.

Parole chiave: suolo di tartufaia; Tuber melanosporum; funghi; batteri; piante erbacee

Metagenomics is the genomic analysis of the microorganisms contained in environmental samples such as ocean, soil and gut (Venter *et al*., 2004, Vogel *et al.*, 2009, Feltman, 2013). It is based on direct extraction of DNA of microorganisms from any environment, in this way avoiding their cultivation through the classical microbiology methods. The transition from classical microbiology to modern metagenomics studies has required the development of new technologies, such as the high-throughput sequencing technologies and the development of new computational methods to collect and extract information from complex datasets (Escobar-Zepeda *et al*., 2015). The species richness and the evenness are the two attributes that describe a microbial community although the use of diversity indexes is a better way to quantify and compare microbial diversity among samples. According to Escobar-Zepeda *et al.* (2015), the microbial diversity can be studied using two different approaches: (1) Amplicon sequencing or (2) Shotgun metagenomics. In the first approach, taxonomical informative markers such as 16S rRNA gene for prokaryotes and intergenic transcribed spacers (ITS) or the large ribosomal subunit (LSU) gene for eukaryotes are amplified from DNA extracted from communities. Since the amplicon sequence analysis relies on just one gene, rather than on all the genes from the community, is more correctly defined "metaprofiling" instead of metagenomics. In the second approach, large fragments or even complete genomes from a community can be reconstructed through the characterization of several coding and non-coding sequences that can be used as phylogenetic markers. Metaprofiling has been widely used to answer the predominant question in microbial ecology: "Who is out there? "

A study-case among the various studied environmental samples is represented by the "brûlé", which is an area of the truffle-grounds characterized by scarce vegetation around the host plant of the black truffle *Tuber melanosporum*. Analyses of ITS and 16S of the rRNA genes directly amplified from the soil of a brûlé, in a French truffle-ground, have revealed its bacterial and fungal communities composition (Napoli *et al.*, 2010; Mello *et al.*, 2011; Mello *et al.*, 2013; Mello *et al.*, 2015). Clear differences were shown between the fungal communities present inside and outside the brûlé, together with lower fungal biodiversity inside the brûlé where *T. melanosporum* was the dominant fungus whereas the other ectomycorrhizal fungi decreased in abundance, indicating that *T. melanosporum* competes with them (Napoli *et al.*, 2010; Mello *et al.*, 2011).

The patchy herbaceous plants in the brûlé were extensively colonized by arbuscular mycorrhizal fungi, as were the plants outside the brûlé. Examination of this fungal diversity in the soil showed reduced species richness inside the brûlé, compared with that outside the brûlé. Furthermore, members of Diversispora, Acaulospora, and Archaeospora were only found in the brûlé, in roots or soil, revealing clear differences between inside and outside the brûlé (Mello *et al.*, 2015). Also bacterial communities living in the brûlé were investigated. According to Mello *et al.* (2013), Firmicutes (*e.g.*, *Bacillus*), several genera of Actinobacteria, and a few Cyanobacteria were more abundant inside the brûlé, whereas *Pseudomonas* and several genera within the family Flavobacteriaceae were more abundant outside the brûlé. Taking together, these metagenomics studies clearly revealed the fungal and bacterial community composition inside and outside the brûlé (Table 1), answering to the question: Who is there? But in microbial ecology it is necessary answering also to a second question: What are they doing?

In order to link microbial community composition to ecological processes happening in the brûlé, Zampieri *et al*. (2016) applied a metaproteomics analysis to the brûlé previously characterized by metagenomics.

DOI: 10.6092/issn.2531 – 7342/6647

Tab.1. Metagenomics data concerning the bacterial and the fungal communities present in the truffle-ground analysed by Napoli *et al*. (2010); Mello *et al.* (2011); Mello *et al.* (2013); Mello *et al*. (2015).
Tab.1 Dati di metagenomica riguardanti le comunità batteriche e fungine presenti nella tartufaia analizzata da Napoli et al. (2010); Mello et al. (2011); Mello et al. (2013); Mello et al. (2015).

Bacteria	Fungi
Actinomyces	*Acaulospora* sp.
Arthrobacter	*Acremonium* sp.
Bacillus	*Amanita* sp.
Chryseobacterium	*Aporospora terricola*
Flavobacterium	*Archaespora trappei*
Frankia	*Arthrographis* sp.
Geobacillus	*Ascobolus* sp.
Geodermatophilus	*Aspergillus insuetus*
Massilia	*Aspergillus ustus*
Microbacterium	*Beauveria felina*
Micromonospora	*Bionectria ochroleuca*
Mycobacterium	*Cladophora* sp.
Myxococcus	*Chaetomium globosum*
Nitrospira	*Claroideoglomus* sp.
Pedobacter	*Coniosporium* sp.
Prochlorococcus	*Cordyceps bassiana*
Pseudomocardia	*Cryptococcus aerius*
Pseudomonas	*Curvularia inaequalis*
Riemerella	*Cylindrocarpon* sp.
Rubrobacter	*Didymella bryoniae*
Saccharothrix	*Diversispora* sp.
Streptosporangium	*Epicoccum* sp.
	Funneliformis caledonius
	Funneliformis mosseae
	Fusarium oxysporum
	Fusarium solani
	Genea
	Geoglossum
	Geosmithia sp.
	Gibberella moniliformis
	Glomus indicum
	Glomus sp.
	Helminthosporium solani
	Hymenogaster sp.
	Hypocrea koningii
	Inocybe sp.
	Leptosphaerulina sp.
	Lewia infectoria

DOI: 10.6092/issn.2531 – 7342/6647

	Lophiostoma sp.
	Microdochium sp.
	Mortierella alpina
	Myrothecium cinctum
	Naucoria (Alnicola) sp.
	Nectria sp.
	Neonectria radicicola
	Paraglomus majewskii
	Penicillium canescens
	Periconia macrospinosa
	Peziza sp.
	Phaeosphaeria sp.
	Phoma sp.
	Pluteus sp.
	Podospora sp.
	Preussia sp.
	Psathyrella
	Pulvinula constellatio
	Pyrenochaeta sp.
	Rhizophagus irregularis
	Rhizophydium sp.
	Sclerocystis sp.
	Scleroderma sp.
	Scutellospora sp.
	Septoglomus sp.
	Spizellomyces sp.
	Tetracladium sp.
	Tomentella sp.
	Trichoderma sp.
	Tricholoma sp.
	Tuber magnatum
	Tuber melanosporum
	Tuber rufum
	Tuber sp. (*T. borchii* complex)
	Vermispora sp.
	Verrucaria sp.
	Verticillium catenulatum
	Verticillium dahliae
	Xerocomus rubellus

Metaproteomics has been recently applied to different environments, i.e., soil, sediments, marine, and freshwater systems, proving to be a powerful tool to describe metabolic processes active in these environments (Siggins *et al.*, 2012). It is the study of all the proteins expressed by the organisms within an ecosystem at a specific time (Bastida *et al.*, 2014). It has the potentiality to provide a link between

DOI: 10.6092/issn.2531 – 7342/6647

phylogenetic microbial community data and functional ones (Bastida and Jemlich, 2016; Keblinger *et al.*, 2016; Wang *et al.*, 2016), but it is negatively influenced by different factors (e.g., soil heterogeneity, high microbial diversity, incomplete metagenomic information, etc.) (Keblinger *et al.*, 2016 and references inside; Wang *et al.*, 2016). The extraction of proteins from soils is a challenge because of organic compounds, such as complex carbohydrates, lipids and phenolic compounds, humic acids, and inorganic compounds, such as silt and clay minerals (Keblinger *et al.*, 2016). Another obstacle in metaproteomics analysis could be the protein identification, which depends on the features (design, capacity, quality) of the database chosen for the analysis (Wang *et al.*, 2016). To overcome these difficulties, Zampieri and colleagues (2016) used three different protein extractions methods to increase the yield of proteins from the studied brûlé and they built a homemade database based on the metagenomics data previously obtained from the same truffle-ground. In this study, the authors managed to extract 638 proteins, and in particular, 411 and 309 proteins were specific for inside and outside the brûlé, respectively. The organismal classification demonstrated that bacterial proteins were more abundant than those of Eukaryota, both inside and outside the brûlé, and surprisingly that proteins of Viridiplantae were more inside than outside (Zampieri *et al.*, 2016) (Fig. 1).

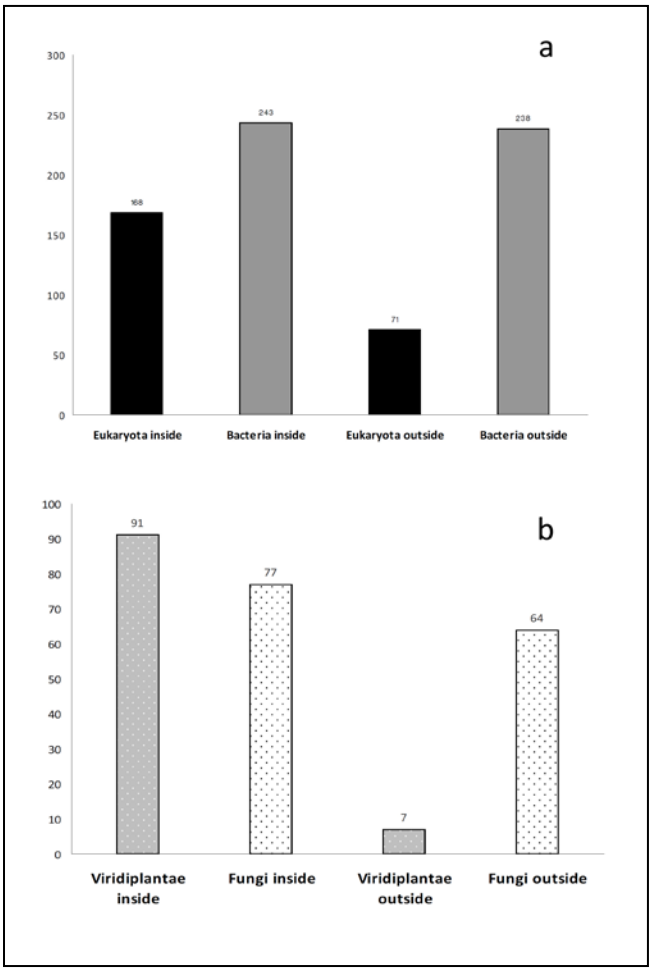

Fig. 1. Organismal classification of the identified proteins inside and outside the brûlé at superkingdom level (a) and at fungal and viridiplanta level (b) on the base of the metaproteomics data of Zampieri *et al.* (2016).
Fig. 1. Classificazione degli organismi delle proteine identificate dentro e fuori il pianello a livello di superregno (a) e a funghi e di piante (b) sulla base dei dati di metproteomica di Zampieri et al. (2016).

DOI: 10.6092/issn.2531 – 7342/6647

The identified proteins were characterized on the basis of Gene Ontology that allows to define the biological pathways in which they are involved, as well as their specific molecular functions. Moreover, the authors identified over-represented biological processes and molecular functions by using the Fisher's Exact Test: in particular, 57 processes were over-represented inside the brûlé and, among them, 14 were only specific inside, and related to stress, glycolysis and the Krebs cycle, sulphur metabolism, etc. Interestingly, the processes related to response to some kinds of stress (osmotic and salt), stimuli (abiotic and temperature), inorganic substances and metal ions (cadmium) were present only inside the brûlé, demonstrating how the two environments were different from a metabolic point of view, and the stress proteins were identified in plants, fungi and bacteria, suggesting that the organisms living in the brûlé face stress experience (Zampieri *et al*., 2016). Moving to *T. melanosporum* proteins, the search was done against the *T. melanosporum* genome database from MycorWeb (version of February 2015; 12826 sequences (http://mycor.nancy.inra.fr/IMGC/TuberGenome/download.php?select= fast). Overall, 265 *T. melanosporum* proteins were identified, whose 148 were specific inside the brûlé and 92 outside. Although the majority of *T. melanosporum* truffles are usually collected inside the brûlé, the presence of black truffle proteins outside the brûlé was expected, in agreement with the metagenomics data, namely nucleotidic sequences of *T. melanosporum*. *T. melanosporum* proteins as heat shock proteins and co-chaperonine, laccase and tyrosinase were identified only inside the brûlé, but the differences between inside and outside were not significant on the base of the Fisher Exact Test (Zampieri *et al*., 2016). This finding may suggest that the truffle proteins are not involved in the phenomenon of brûlé, i.e., the decrease of herbaceous plants, which are on the contrary affected by volatile organic compounds, delivered by *T. melanosporum* (Splivallo *et al.,* 2007, 2011; Streiblová *et al.,* 2012). Truffle volatiles could be then responsible of the stress and defence response in the organisms living inside the brûlé, principally the herbaceous plants where chaperonine proteins, superoxide dismutase, heat shock proteins, late embryogenesis abundant protein were over-represented (Zampieri *et al.,* 2016). Even if the biodiversity is reduced where the vegetation is scanty, the brûlé is an active environment.

This study demonstrates that combining the data of metagenomics and metaproteomics we have nowadays the opportunity to potentially reveal the functioning of any environment. The year of the soil has been just celebrated and new frontiers are open to disclose not only "Who there is" but also "What they are doing".

References

Bastida F., Hernández T., García C. (2014). Metaproteomics of soils from semiarid environment: functional and phylogenetic information obtained with different protein extraction methods, Journal of Proteomics, 101, 31-42.

Bastida F., Jehmlich N. (2016). It's all about functionality: How can metaproteomics help us to discuss the attributes of ecological relevance in soil?, Journal of Proteomics, 144, 159-161.

Escobar-Zepeda A., Vera-Ponce de León A., Sanchez-Flores A. (2015). The Road to Metagenomics: From Microbiology to DNA Sequencing Technologies and Bioinformatics, Frontiers in Genetics, 6, 348.

Feltman R. (2013). The Gut's Microbiome Changes Rapidly with Diet, Scientific American, 14.

Keiblinger K.M., Fuchs S., Zechmeister-Boltenstern S., Riedel K. (2016). Soil and leaf litter metaproteomics-a brief guideline from sampling to understanding, FEMS Microbiology Ecology, 92, fiw180.

Mello A., Napoli C., Murat C., Morin E., Marceddu G., Bonfante P. (2011). ITS-1 versus ITS-2 pyrosequencing: a comparison of fungal populations in truffle-grounds, Mycologia, 103, 1184-1193.

Mello A., Ding G.C., Piceno Y.M., Napoli C., Tom L.M., DeSantis T.Z., Andersen G.L., Smalla K., Bonfante P. (2013). Truffle brûlés have an impact on the diversity of soil bacterial communities, PLos One, 8, e61945.

DOI: 10.6092/issn.2531 – 7342/6647

Mello A., Lumini E., Napoli C., Bianciotto V., Bonfante P. (2015). Arbuscular mycorrhizal fungal diversity in the *Tuber melanosporum* brûlé, Fungal Biology, 19, 518-527.

Napoli C., Mello A., Borra A., Vizzini A., Sourzat P., Bonfante P. (2010). *Tuber melanosporum*, when dominant, affects fungal dynamics in truffle grounds, New Phytologist, 85, 237-247.

Siggins A., Gunnigle E., Abram F. (2012). Exploring mixed microbial community functioning: recent advances in metaproteomics, FEMS Microbiology Ecology, 80, 265-280.

Splivallo R., Novero M., Bertea C.M., Bossi S., Bonfante P. (2007). Truffle volatiles inhibit growth and induce an oxidative burst in *Arabidopsis thaliana*, New Phytologist, 175, 3417-3424.

Splivallo R., Ottonello S., Mello A., Karlovsky P. (2011). Truffle volatiles: from chemical ecology to aroma biosynthesis, New Phytologist, 189, 688-699.

Streiblová E., Gryndlerová H., Gryndler M. (2012). Truffle brûlé: an efficient fungal life strategy, FEMS Microbiology Ecology, 80, 1-8.

Venter J.C., Remington K., Heidelberg J.F., Halpern A.L., Rusch D., Eisen J.A., Wu D., Paulsen I., Nelson K.E., Nelson W., Fouts D.E., Levy S., Knap A.H., Lomas M.W., Nealson K., White O., Peterson J., Hoffman J., Parsons R., Baden-Tillson H., Pfannkoch C., Rogers Y.H., Smith H.O. (2004). Environmental genome shotgun sequencing of the Sargasso Sea, Science, 304, 66-74.

Vogel T.M., Hirsch P.R., Simonet P., Jansson J.K., Tiedje J.M., van Elsas J. D., Nalin R., Philippot L., Bailey M.J. (2009). Advantages of the metagenomic approach for soil exploration: reply from Vogel et al. Nature Reviews Microbiology, 7, 756-757.

Wang D.Z., Kong L.F., Li Y.Y., Xie Z.X. (2016). Environmental Microbial Community Proteomics: Status, Challenges and Perspectives, International Journal of Molecular Sciences, 17, E1275.

Zampieri E., Chiapello M., Daghino S., Bonfante P., Mello A. (2016). Soil metaproteomics reveals an inter-kingdom stress response to the presence of black truffles, Scientific Reports, 6, 25773.