Bastida, F., Torres, I.F., Moreno, J.L., Baldrian, P., Ondoño, S., Ruiz-Navarro, A., Hernández, T., **Richnow, H.H.**, **Starke, R.**, García, C., **Jehmlich, N.** (2016):
The active microbial diversity drives ecosystem multifunctionality and is physiologically related to carbon availability in Mediterranean semi-arid soils
Mol. Ecol. **25** (18), 4660 – 4673

http://dx.doi.org/10.1111/mec.13783

Received Date : 21-Apr-2016

Revised Date : 07-Jul-2016

Accepted Date : 21-Jul-2016

Article type : Original Article

TITLE PAGE

The active microbial diversity drives ecosystem multifunctionality and is physiologically related to carbon-availability in Mediterranean semiarid soils

Felipe Bastida[1*], Irene F. Torres[1], José L. Moreno[1], Petr Baldrian[2], Sara Ondoño[1], Antonio Ruiz-Navarro[1], Teresa Hernández[1], Hans H. Richnow[3], Robert Starke[4], Carlos García[1], & Nico Jehmlich[4]

[1]CEBAS-CSIC. Department of Soil and Water Conservation. Campus Universitario de Espinardo, 30100, Murcia, Spain.

[2]Laboratory of Environmental Microbiology, Institute of Microbiology of the CAS, Vídeňská 1083, Praha 4 14220, Czech Republic.

[3]Helmholtz-Centre for Environmental Research - UFZ, Department of Isotope Biogeochemistry, Permoserstr. 15, 04318 Leipzig, Germany

[4]Helmholtz-Centre for Environmental Research - UFZ, Department of Molecular Systems Biology, Permoserstr. 15, 04318 Leipzig, Germany

*Corresponding author: fbastida@cebas.csic.es

Phone: +34 968396106; Fax: +34968396213

Running Head: Active diversity in semiarid soils.

Competing interests: The authors declare no competing interests.

The active microbial diversity drives ecosystem multifunctionality and is physiologically related to carbon-availability in Mediterranean semiarid soils

Abstract

Biogeochemical processes and ecosystemic functions are mostly driven by soil microbial communities. However, most methods focus on evaluating the total microbial community and fail to discriminate its active fraction which is linked to soil functionality. Precisely, the activity of the microbial community is strongly limited by the availibility of organic carbon (C) in soils under arid and semiarid climate. Here, we provide a complementary genomic and metaproteomic approach to investigate the relationships betweeen the diversity of the total community, the active diversity and ecosystem functionality across a dissolved organic carbon (DOC) gradient in South-East Spain.

DOC correlated with the ecosystem multifunctionality index composed by soil respiration, enzyme activities (urease, alkaline phosphatase and β-glucosidase) and microbial biomass (phospholipid fatty acids, PLFA). This study highlights that the active diversity (determined by metaprotoemics) but not the diversity of the whole microbial community (evaluated by amplicon gene sequencing) is related to the availability of organic C and it is also connected to the ecosystem multifunctionality index.

We reveal that DOC shapes the activities of bacterial and fungal populations in Mediterranean semiarid soils and determines the compartmentalization of functional niches. For instance, *Rhizobales* thrived at high-DOC sites probably fueled by metabolism of one-C compounds. Moreover, the analysis of proteins involved in the transport and metabolism of

carbohydrates revealed that *Ascomycota* and *Basidiomycota* occupied different nutritional niches. The functional mechanisms for niche specialization were not constant across the DOC gradient.

Introduction

Ecosystemic functions and biogeochemical processes are majorly driven by soil microbial communities. However, most methods focus on evaluating total microbial community and fail to determine its active fraction (Blagodatskaya and Kuzyakov, 2013). Indeed, the active microorganisms compose only about 0.1-2% of the total microbial biomass in soil (Blagodatskaya & Kuzyakov, 2013). The assessment has ecological consequences because the attempt to link the microbial community to ecosystem functionality can be conceptually limited by the fact that only a minor proportion of microorganisms is actually active in soils (Blagodatskaya & Kuzyakov, 2013) and mediate ecosystem processes.

Probably, the absence of adequate approaches has limited our knowledge about the functionalities of microbial communities. In the recent years, the development of metaproteomics has expanded this field and allows the understanding of the activity and function of microbial populations (Verberkmoes *et al.,* 2009). Here, we provide a complementary approach to describe the diversity and composition of the total and the active microbial communites accomplished by amplicon gene sequencing and metaproteomics. The identification of the repertoire of proteins within environmental samples (Hettich *et al*., 2013), allows the identification of the key drivers, soil processes, and cellular functionalities that actually regulate microbial adaptations to the environment, as well as providing a link to phylogeny (Schneider *et al*., 2012; Keiblinger *et al*., 2012). Given the link to functionality, metaproteomics can be used to estimate the active diversity of soil and its relationship to ecosystem functionality. Furthermore, as a step forward approach in microbial ecology, deeper insights will be obtained by protein abundance and phylogenetic analysis of proteins involved in translation and ribosome structures.

The activity of microbial populations and the microbial-mediated ecosystem processes are often limited by the availability of organic C in arid and semiarid ecosystems (García *et al.,* 1994; Bastida *et al.,* 2006). As highlighted by Guige *et al.* (2015), dissolved organic C (DOC) is an indicator of C-availability that contains quickly available energy sources for microbial activity (Bowen *et al.,* 2009) and is strongly linked to the activity of soil microbial community. This pool of organic C is related to to plant cover and vegetation type (Kalbitz *et al.,* 2000; Sanderman and Amundson, 2008), mineralogy (Kaiser et al., 2001) and climate (Camino-Serrano *et al.,* 2014). However, DOC as a factor shaping the diversity, composition and functionalities of soil microbial communities has been rarely investigated in depth (Guigue *et al*., 2015). Moreover, the cellular and physiological responses of microbial populations to the availability of organic C are still obscure. These responses can be the basis of ecosystems functionality and, for instance, can exert control of organic matter decomposition and the cycling of elements in semiarid areas.

In this study, the aim is to explore comparatively the links between the diversity and composition of the total and the active community, and ecosystem functional indicators across DOC gradients in semiarid Mediterraean soils. We hypothesize that the active diversity will be more related to ecosystem functionality in semiarid soils than genomic diversity. In addition, we expect that niche adaptations of particular populations to organic C availability can be explained by changes in cellular physiology that, in turn, will explain variations in their abundance and changes in the diversity of the active microbial community. Overall, our study allows a better understanding of the functional adaptations of microbial populations to C availability in soil and the relationship to the functionality of semiarid ecosystems.

Material and methods

Soil sampling and sites decription

Soil sampling was carried out in April 2014. The surveyed soils correspond to a dryland region (Mediterranean semiarid climate) in South-East Spain. A total of 60 samples were collected that correspond to 20 environmental areas (3 replicates per area). For each replicate, six different sub-samples of soil (20 cm depth) were collected and subsequently mixed. Soils were selected with the purpose of establishing a marked gradient of organic matter availability. The selected soils correspond to typical habitats in dryland regions, ranging from bare, pre-desertic soils developed in marsh substrates with intense erosion and degradation to highly productive transects with dense forest cover (Table 1). The vegetation cover differed also between sites in terms of abundance and typology, from xerophytic shrubs and grasslands to forest areas dominated by *Pinus* sp. and *Quercus* sp. We are aware that abiotic factors (lithology, microclimate, altitude, etc.) and biotic characteristics (e.g., vegetation) may exert control of belowground microbial processes. These differences in abiotic and biotic factors generate a gradient of organic C availability and energy (García *et al*., 1994; Bastida *et al*., 2006). Indeed, among the other chemical parameters analyzed (total organic C, total N, inorganic C, water-soluble N, and pH), DOC was the only variable that formed part of the linear regression models that significantly explained bacterial biomass, fungal biomass, bacterial diversity, and CO_2-emissions.

Chemical soil analysis, respiration, phospholipid fatty acid analysis (PLFAs) and enzyme activities

A characterization of the soils, including chemical indicator, the localization, plant cover, and dominant species was performed (Table 1). The dissolved organic C (DOC) of the soil was extracted with distilled water (1:5, w:v) by shaking for 2 h, followed by centrifugation at 13000 rpm for 15 min and filtration. The analysis of the C content in the extracts was performed in a C analyzer for liquid samples (Multi N/C 3100, Analytik Jena). The total nitrogen content (N) and total organic C were determined using an Elemental Analyzer (C/N

Flash EA 112 Series-Leco Truspec). As aqueous solution 1:5 (w:v) was used to measure pH in a pH meter (Crison mod.2001, Barcelona, Spain). Soil respiration, enzyme activities (β-glucosidase, phosphatase and urease) and PLFAs were analysed as described previously (Bastida *et al*., 2015a) (Supplementary Information).

Protein extraction from soil and mass spectrometric analysis

Protein extraction was performed according to the method described by Chourey et al. (2010), which has been found suitable for semiarid soils (Bastida et al., 2014). The cell lysis and disruption of soil aggregates were performed by boiling at 100ºC for 10 min in sodium dodecyl sulfate (SDS) buffer. The concentration and purification steps were performed using trichloroacetic acid (TCA) and three acetone washing steps. Protein pellets were resuspended in SDS lysis buffer, containing 4% SDS, 0.1 mM dithiothreitol, and 100 mM Tris HCl, and incubated for 5 min at 95°C. The resuspended proteins were loaded on SDS gels (4% stacking gel, 12% separating gel) and run into the separating gel for 10 min. After electrophoresis, the gels were stained with colloidal Coomassie brilliant blue (Bastida *et al*., 2016). The gel area containing the protein mixture of each sample was sliced into one piece. Gel pieces were destained and subsequently digested using 100 ng of trypsin (Sigma, Munich, Germany), overnight at 37 ºC. The peptide mixture was extracted twice with 100% acetonitrile and concentrated by vacuum centrifugation for 15 min. Tryptic peptides were reconstituted in 0.1% formic acid prior LC-MS measurement (Bastida *et al*., 2016).

The peptide lysates were separated on a UHPLC system (Ultimate 3000, Dionex/Thermo Fisher Scientific, Idstein, Germany). Five µL samples were first loaded for 5 min on the precolumn (µ-precolumn, Acclaim PepMap, 75 µm inner diameter, 2 cm, C18, Thermo Scientific) at 4% mobile phase B (80% acetonitrile in nanopure water with 0.08% formic acid), 96% mobile phase A (nanopure water with 0.1% formic acid), then eluted from the analytical column (PepMap Acclaim C18 LC Column, 25 cm, 3 µm particle size, Thermo Scientific) over a 120 min linear gradient of mobile phase B (4-55% B). Mass spectrometry

was performed on an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) with a TriVersa NanoMate (Advion, Ltd., Harlow, UK) source in LC chip coupling mode. The MS was set on top speed with a cycle time of 3 s, using the Orbitrap analyzer for MS and MS/MS scans with higher energy collision dissoziation (HCD) fragmentation at normalized collision energy of 28%. MS scans were measured at a resolution of 120,000 in the scan range of 400-1,600 *m/z*. MS ion count target was set to 4×10^5 at an injection time of 60 ms. Ions for MS/MS scans were isolated in the quadrupole with an isolation window of 2 Da and were measured with a resolution of 15,000 in the scan range of 350-1400 *m/z* in the Orbitrap. The dynamic exclusion duration was set to 30 s with a 10 ppm tolerance around the selected precursor and its isotopes. Automatic gain control target was set to 5×10^4 with an injection time of 120 ms.

Proteome Discoverer (v2.0, Thermo Scientific) was used for protein identification and the acquired MS/MS spectra were searched with Mascot (v2.4.1) against the NCBInr database (08/24/2014, containing 48,094,830 sequences). Enzyme specificity was selected to trypsin with up to two missed cleavages allowed using 10 ppm peptide ion tolerance and 0.05 Da MS/MS tolerances. Oxidation (methionine) and carbamylation (lysine and arginine) were selected as variable modifications and carbamidomethylation (cysteine) as a static modification. Only peptides with a false discovery rate (FDR) <0.01 calculated by Percolator (Käll et al., 2007) and peptide rank = 1 were considered as identified. The abundance of one detected protein was quantified using the average abundance of the top-3 peptide assigned to this protein. The abundance of all detected proteins in each sample was normalized to the abundance of a given protein. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaino et al., 2013) with the dataset identifier PXD003572.

The “PROteomics results Pruning & Homology group ANotation Engine” (PROPHANE) (http://www.prophane.de) was applied to assign proteins to their phylogenetic and functional origin (Schneider *et al*., 2011). The abundance of functional proteins was retrieved attending

to the area under the curve (AUC) of peptides. The relative phylogenetic contribution for each functional protein was analyzed using Unipept (http://unipept.ugent.be/) (Mesuere *et al*., 2015). The diversity of the active microbial community was calculated by the Shannon-Wiener index (Shannon and Weaver, 1949) based on NSAF values at genus level. A total of 7280 protein groups and 27245 peptides were identified in the metaproteome of the 60 soil samples.

DNA extraction and amplification

DNA was extracted from 500 mg of freeze-dried soil of each sample, using a Fast DNA Spin Kit for soil and the FastPrep® Instrument (MP Biomedicals, Santa Ana, CA). The V4 region of bacterial 16S rRNA gene was amplified using the barcoded primers 515F and 806R (Argonne National Laboratory) as described previously (Caporaso *et al*., 2012). PCR amplification of the fungal ITS2 region from DNA was performed using barcoded gITS7 and ITS4 (Ihrmark *et al*., 2012) in three PCR reactions per sample as described previously (Zifcakova *et al*., 2016). Further details on sequencing and data processing are provided in Supplementary Information.

Data analysis and assessment of ecosystem multifunctionality

Multifunctionality represents the provision of multiple ecosystem processes and services (Bradford *et al*., 2014; Delgado-Baquerizo *et al*., 2016). To obtain a quantitative multifunctionality index for each site, a common procedure consists of the standardization of individual functions using the Z-score transformation (Delgado-Baquerizo *et al*., 2016). The standardized rates of ecosystem functions are then averaged to obtain a multifunctionality index (Wagg *et al*., 2014; Delgado-Baquerizo *et al*., 2016; Lefcheck *et al*., 2015). The multifunctionality index was calculated from five parameters organic matter mineralization (soil respiration); cycling of C (β-glucosidase activity), P (phosphatase activity) and N

(urease activity); and microbial biomass (the sum of bacterial and fungal PLFAs). Soil enzyme activities have been widely used as indicators of soil quality (Bastida *et al*., 2006). Overall, this multifunctionality index (MF) is related to the provision of services mediated by soil microbial communities including biogeochemical cycling of elements, organic matter decomposition and biomass reservoir.

The Shannon-Wiener index of diversity was calculated for: i) bacterial diversity based on 16S rRNA gene sequencing; ii) fungal diversity based on ITS sequencing, and iii) diversity of the active community based on the taxonomic analysis of proteins. A Z-score transformation and averaged calculation was applied to obtain a joint genomic diversity index that included both bacterial and fungal diversities.

Spearman correlations between DOC content and the abundance of microbial populations (estimated by genomics and proteomics) and metabolic functions were determined. Similarly, correlations between DOC content, diversity and the multifunctionality index or their composite variables (soil respiration, β-glucosidase, phosphatase, urease and microbial biomass) were determined (Table 2; Fig 1; Supplementary Fig S1).

The relative abundances given by genomics and metaproteomics were classified by the use of percentiles at 20%. The statistical analyses were performed using IBM-SPSS Statistics (v22.0). The figures were produced using R software v.3.1.3.

Results

Relationships between DOC content, microbial diversity and functional ecosystemic indicators

The content of soil organic C ranged between 0.25 and 20%; total N between 0.04 and 1.07%; and the pH between 7.6 and 8.7. The DOC content ranged between 65 and 1700 mg C kg^{-1} soil (Fig. 1A; Table 1). The DOC content was not significantly correlated with the genomic diversity (Fig. 1B). The DOC content and the diversity of the active community were statistically correlated (r = 0.37, P = 0.004) (Fig. 1C; Table 2).

The ecosystem multifunctionality index correlated with DOC content (r = 0.86, P < 0.001) (Fig. 1D; Table 2). The multifunctionality index did not correlate with genomic diversity (r= 0.12, P =0.64) but did so with active diversity (r = 0.50, P < 0.001) (Fig. 1E and Fig. 1F; Table 2). Singular activity indicators (respiration, β-glucosidase, phosphatase, urease) were correlated significantly with the active diversity but not with genomic diversity (Table 2; Supporting Information Fig. S1).

Statistical descriptors of the linear and non-linear regressions are observed in Table S1 (Supporting Information Table S1). Linear or logarithm models showed the highest R^2 values (P < 0.05) as observed in Figure 1. The unique exception was the relationship between the active diversity and ecosystem multifunctionality that was slightly better described by the quadratic model (R^2= 0.28, P < 0.001) than by the linear model (R^2= 0.26, P < 0.001) (Fig. 1F; Table S1).

Composition of total vs active bacterial communities

In total, 957709 bacterial sequences were obtained by 16S rRNA gene sequencing after quality trimming and removal of chimera sequences. The clustering at a 97% similarity threshold yielded 27897 OTUs. Metaproteomics revealed 3082 non-redundant microbial proteins.

Overall, 16S rRNA gene sequencing revealed higher abundances of *Acidobacteria* and *Actinobacteria* than the data provided by metaproteomics (MP) and the phylogenetic analysis of proteins involved in ribosome structures and translation (RP). In contrast, the abundance of *Cyanobacteria* was higher in MP than in 16S rRNA gene. A DOC-dependent gradient was observed for *Cyanobacteria*, with the relative abundance of their 16S, MP and RP all significantly decreasing with the increase in WSC (Fig. 2).

Regardless of the approach, *Proteobacteria* was the most abundant bacterial phylum accounting for up to 60% of the bacterial proteins (Fig. 2). The low abundance of *Planctomycetes* was noticeable when studied by 16S rRNA gene sequencing. However, the

activity of this phylum (as revealed by MP and RP) was much higher and reached up to 30%, in the case of RP, in some sites.

Regarding the 16S rRNA gene analysis, the relative abundances of *Actinobacteria* and *Proteobacteria* increased in parallel to the increase of DOC, as demonstrated by the positive correlations with DOC of each phylum ($P < 0.01$) (Fig. 3). Conversely, the relative abundances of *Acidobacteria*, *Bacteroidetes*, *Cyanobacteria*, *Firmicutes*, *Gemmatimonadetes*, *Chloroflexi*, and *Verrucomicrobia* decreased when DOC increased ($P < 0.01$). *Planctomycetes* did correlate significantly with DOC ($r = 0.23$, $P = 0.082$). The relative abundances of the dominant bacterial orders followed a trend explained by the DOC gradients and correlation coefficients (Supporting Information Figs. S2 and S3).

Composition of total vs active fungal communities

In total, 456985 fungal sequences were obtained by ITS gene sequencing after quality trimming and removal of chimera sequences. The clustering at a 97% similarity threshold yielded 12406 OTUS. Fungal communities were dominated by *Ascomycota* and *Basidiomycota*, as depicted by ITS sequencing, metaproteomics, and ribosomal proteins (Fig. 2). Nevertheless, ITS sequencing was able to identify some minor phyla, such as *Glomeromycota* (up to 3.5% of the sequences) or *Chytridiomycota* (up to 0.64% of the sequences), whose proteins were not observed. Ribosomal and translation fungal proteins were observed throughout the DOC range with the exception of the lowest values, which points to an active fungal community in the presence of plant inputs. The abundance of RP proteins from *Basidiomycota* increased across the DOC range and exceeded the abundance of ascomycotal proteins at the highest DOC values.

With respect to the ITS analysis, the relative abundances of *Ascomycota*, *Glomeromycota*, and *Chytridiomycota* significantly decreased when DOC increased while the *Basidiomycota* increased (Fig. 3). The relative abundances of the dominant fungal orders estimated by ITS sequencing also significantly responded to the DOC content (Supporting Information S3 and S4).

The ratio between bacterial and fungal proteins was lower in soils with the highest amount of DOC, as a consequence of the increase in the amount of fungal proteins. This ratio ranged from 42 to 54 at the DOC content < 640 mg C kg^{-1}, but decreased to 11 in the soils with the highest DOC content (640-1600 mg C kg^{-1} soil). The abundances of proteins from *Basidiomycota* were positively correlated with DOC ($P < 0.01$) (Fig. 3).

Metabolic proteins in DOC gradients and phylogeny assignment

The abundances of proteins involved in chromatin, replication, transcription, nucleotide metabolism, translation, and signal transduction significantly increased with the soil DOC content (Fig. 3; Supporting Information Fig. S5). By contrast, negative response was found for proteins involved in carbohydrate metabolism, metabolism of coenzymes, inorganic ion transport, intracellular trafficking, and lipid transport and metabolism (Fig 3; Supporting Information Fig. S5).

The results revealed that *Proteobacteria* were the dominant phylum. In this group, the most abundant proteins were those related to cellular membranes and DNA, RNA and protein synthesis and metabolism. These proteins were more abundant in soil samples with high DOC content (Fig. 4). It is noteworthy that proteins involved in the transport and metabolism of carbohydrates in *Proteobacteria* were underrepresented in soils with little DOC (Fig. 4).

Planctomycetes showed the highest abundance of proteins involved in translation and protein modification in soils with high DOC contents. *Cyanobacteria* expressed a higher abundance of proteins in soils with the lowest level of available C (Fig. 4). For *Rhizobiales*, membrane and cell envelope proteins were more abundant in soils of low DOC content. Contrarily, proteins involved in transcription and translation were found only in soils with high DOC content. Similarly, *Planctomycetales* had a high abundance of proteins related to transcription, translation, and signal transduction in soils with high DOC content, while proteins related to the cell membranes of *Planctomycetales* were not identified.

Functional proteins in DOC gradients and phylogeny assignment

The abundance of RuBisCO decreased as DOC increased (Fig. 3; Supporting Information Fig. S6). The phylogenetic analysis of the cyanobacterial peptides of RuBisCO revealed *Oscillatoriales* and *Nostocales* as the main origins of this enzyme (Supplementary Fig. S7). The relative abundance of laccases increased with soil DOC (Fig. 3; Supplementary Fig. S6).

Similarly, cellulase, laccases, and xylanase proteins were more abundant overall in soils with high DOC content and a similar pattern was observed for proteins involved in the metabolism of less-complex sugars (i.e. galactosidases and glycosyl hydrolases) (Fig. 3; Fig. S6). Fungal peptides from cellulases derived from *Sebacinales* (*Basidiomycota*) and *Helotiales* (*Ascomycota*) were identified (Supporting Information Fig. S7). In the case of xylanase, besides fungal peptides, *Micromonosporales* bacteria appeared as the main producer of this enzyme. *Agaricales* were the main producers of laccases (Fig. S7).

Peptides from methanol dehydrogenase and methylene-tetrahydrofolate dehydrogenase (One-C metabolism enzymes) were identified and their abundance increased with the DOC content ($P < 0.01$) (Fig. S7).

Discussion

The relationships between microbial diversity, DOC and ecosystem functionality

Diversity of bacteria and fungi did not correlate with the DOC content. This may suggest that semiarid soils contain a high proportion of dormant or inactive microorganisms that do not require DOC-derived energy for metabolic processes. Similar results were previously observed in soils with differing contents of organic C (Hirsch *et al*., 2007; Bastida *et al*., 2013) and it has been argued that functional redundancy of microbial communities can explain the absence of a relationship between their diversity and the C content (Wertz *et al*., 2007).

The diversity of housekeeping genes has been proposed recently as a robust phylogenetic marker of diversity (Kostantinidis *et al*., 2006; Roux *et al*., 2011). However, the phylogenetic classification of proteins (and not genes) and the computation of diversity indices have not been performed so far in ecosystem gradients. In contrast to DNA diversity, the active diversity (analyzed by metaproteomics) was related to DOC content. The responsiveness of the active microbial diversity can be perceived as an ecological adaptation to C availability in soil and may be highlighted as a critical attribute of the ecosystems functionality. In fact, this active diversity (but not the DNA diversity) also significantly correlated with the ecosystem multifunctionality index. This observation might lead to fundamental considerations: i) the diversity of the active microbial community can be a more straightforward indicator of nutrient cycling and CO_2 emissions in semiarid ecosystems; and ii) the active diversity – as an attribute of the microbial community – can be taken into account when assessing future impacts of global change. Recently, several studies have shown that genome diversity was connected to multifunctionality of ecosystems (Wagg *et al*., 2014; Delgado-Baquerizo *et al*., 2016) and concluded that greater diversity would promote the co-metabolic degradation of plant inputs in soil and hence increase the functionalities of soil microbial community (Delgado-Baquerizo *et al*., 2016). Here, we provided evidence that the active diversity is more related to ecosystem functionality than genomic diversity. It is worth to highlight that this assessment is valid for both the multifunctionality index (Fig. 1) and singular indicators (i.e. respiration, enzyme activities or microbial biomass; Fig. S1). Overall, the linear-type relationship (although not very different with logarithm and quadratic models) between DOC content and the ecosystem multifunctionality could indicate that higher active diversity increases the rates of ecosystem processes in soil.

Bacterial adaptations in DOC gradients

Soil C sources are considered key ecological drivers of microbial community dynamics (Drenovsky *et al*., 2004; de Vries *et al*., 2012). *Proteobacteria* was a dominant taxon and its abundance increased with the soil DOC content. Nevertheless, the magnitude of the

response to the DOC gradient was more marked when using the entire metaproteome or ribosomal proteins than with 16S rRNA gene sequencing. This finding points to a high stimulation of this phylum when DOC increased. Flynn *et al*. (2010) suggested that an organism that carries an “excess” of ribosomes will be able to respond rapidly in terms of growth to beneficial conditions and this might be a feature of the copiotrophic lifestyle of the *Proteobacteria*.

Proteobacteria and *Actinobacteria* have been proposed as copiotrophs and it is assumed that they are able to assimilate easily available organic compounds and to grow rapidly (Fierer *et al*., 2006; Eilers *et al*., 2010). Lauro *et al*. (2009) indicated that, in marine environments, copiotrophs are characterized by a lower abundance of lipid transport and metabolism proteins. Our gradient study may extend these functional insights and explain the metabolic plasticity of particular clades. Specifically, the relative abundance of proteins involved in the metabolism and transport of lipids in *Proteobacteria* was lower at high-DOC sites than in the low DOC soils, which is in agreement with Lauro *et al*. (2009). However, the increased abundance of proteins involved in lipid metabolism in soils of low DOC could explain the phylogenetic exclusion of other clades by *Proteobacteria* and the overall dominance of *Proteobacteria* in soils, even in those with a low DOC content. The increase of lipid metabolism in *Proteobacteria* might be related to the degradation of the cell wall and membranes from other clades, giving *Proteobacteria* a strong competitive advantage (Goldfarb *et al*., 2011).

Although this is the first evidence of metaproteome relationships to copiotrophy-oligotrophy on soil gradients, other authors observed that the abundance of genes involved in DNA and RNA replication processes increased as soil N augmented and this feature was more common in copiotrophic than in oligotrophic bacteria (Fierer *et al*., 2007). However, contrarily to *Proteobacteria*, *Planctomycetes* and *Bacteroidetes* showed a scarcity of cell wall proteins that may be indicative of a non-competitive use of resources for development. Indeed, *Bacteroidetes* have been proposed as a copiotrophic phylum (Fierer *et al*., 2007). Nevertheless, the negative correlation between DOC and the genomic abundance of this

group and the absence of correlation with the abundance of proteins from this clade do not support such copiotrophic behavior. Other groups also decreased with DOC, which points to oligotrophic behavior in terms of 16S rRNA gene abundance. For instance, *Firmicutes* would be placed in this group due to the harsh environmental conditions and their capacity to form endospores (Tardy *et al*., 2015) which would explain the increased activity of this phylum at higher DOC content, i.e., improved environmental conditions. *Acidobacteria* have been observed to be abundant in rhizosphere soils, but a large decrease in *Acidobacteria* has been shown also after the addition of C substrates (Cleveland *et al*., 2007). In our study, *Acidobacteria* also showed an oligotrophic lifestyle (negative correlation with DOC), as observed by other authors (Trivedi *et al*., 2013; Leff *et al*., 2015).

Noteworthy, the activity of *Planctomycetes* was surprisingly high in comparison to their abundance indicated by genomics. This underexplored group has emerged also as a highly active microbial group in soil restoration studies (Bastida *et al*., 2015a). The ubiquitous activity of *Planctomycetes* in the soils analyzed here remains a mystery. Lauro *et al.* (2009) highlighted that *Plancomycetes* possess features of both copiotrophs (large genomes) and oligotrophs (a single copy of rRNA operon); this ecological evolution may provide them with sucess in terms of activity across gradients of C availability. This flexible lifestyle might be coupled to the capacity for sensing the environment, by the production of proteins involved in signal transduction (Lauro *et al*., 2009).

Given their capacity for C-fixation, *Cyanobacteria* constitute a key ecological phylum in drylands, which are usually nutrient-depleted (García *et al*., 1994). *Cyanobacteria* are known for their capacity to thrive in low-resource-availability soils and pre-desertic conditions (Belnap and Lange, 2003; Bastida *et al*., 2014b). Our results suggest that *Cyanobacteria* occupy the functional niche of plants regarding C-fixation in pre-desertic soils, the ones with the lowest DOC content (Bastida *et al*., 2015b). It is of note that these organisms constitute the exclusive C-fixers in these environments, as revealed by the increased abundance of cyanobacterial RuBisCO in soils with low DOC content. However, why the activity of *Cyanobacteria* was high in comparison to their overall abundance remains unclear.

Moreover, the translation and ribosomal cyanobacterial proteins were rare. Alltogether, these results could be explained by the large size of the *Cyanobacteria* phylum (Portillo *et al*., 2013), that probably makes a greater contribution to the amount of proteins in the microbial community, despite the fact that their translation and growth capacity is probably slower than for other clades.

In conjuntion with genomics, the identification of enzymes involved in the metabolism of one-C compounds provides insights into the ecology of some populations. The metabolism of plant cell wall compounds (such as lignin) can produce methanol through the action of pectin methyl esterases (Galbally and Kristine, 2002). This methanol can be used as a C and energy source by some microorganisms besides methanotrophs (Sy *et al*., 2005). The identification of the pyrrolo quinoline quinone-containing methanol dehydrogenase from several *Rhizobiales* indicates a niche specialization based on the metabolism of C-metabolism. Indeed, it has been suggested that the metabolism of one-C compounds is active during bacterial colonization of the phyllosphere (Delmotte *et al*., 2009) and rhizosphere (Knief *et al*., 2012).

Rhizobiales showed high capacity for microbial biomass development and activity, as indicated by high abundances of 16S rRNA gene sequences and cell envelope proteins across the gradient, which supports the competitiveness of this group, not only in association with N_2-fixing plant species. The functional analysis of *Rhizobiales* proteins revealed that the basal metabolism at low-DOC sites is maintained due to the expression of proteins involved in the transport and metabolism of carbohydrates and amino acids. Indeed, Delmotte *et al*. (2009) suggested that the abundance of carbohydrate transporters in the phyllosphere microbiome can be viewed as a mechanism of adaptation to the environment for some microbial groups. However, as copiotrophic organisms, *Rhizobiales* thrived much more at high-DOC sites and we cannot discount the idea that this activation was energetically fueled by the metabolism of one-C compounds.

Fungal adaptations in DOC gradients

Our knowledge of the fungal dynamics in response to environmental gradients is much scarcer than for bacterial communities (Liu *et al*., 2015) and this assessment is particulary true for arid and semiarid ecosystems. ITS sequencing allowed the identification of non-dominant fungal groups such as *Glomeromycota*, composed mainly of mycorrhizal fungi. The abundance of this group diminished with increasing DOC and this trend can be explained given the lesser ecological value of *Glomeromycota* to their hosts in habitats with sufficient nutritional resources (Johnson *et al*., 2010; Leff *et al*., 2015).

Metaproteomics provided a simultaneous quantification of bacterial and fungal abundances that rarely can be achieved by PCR-based methods. Zifcakova *et al*. (2016) indicated that bacterial biomass is quantitatively dominant in soil ecosystems and this assessment remains valid for the active fraction, given the higher amount of bacterial proteins relative to fungal proteins. However, although bacterial biomass is quantitatively dominant, fungi have been suggested to be key drivers of decomposition processes in forest environments (Baldrian *et al*., 2012). Indeed, the bacterial-to-fungal protein ratio was higher at sites where the DOC was 0 to 640 mg C kg^{-1}, but fell strongly at the sites with the higest DOC level (640-1600 mg C kg^{-1}). These sites are located at higher elevation and dominated by *Pinus* and *Quercus*, whose plant debris is chemically different to that in shrublands and grasslands. These areas can represent the climax soil under the Mediterranean climate and it has been indicated that, as an ecosystem undergoes sucession, the proportion of fungi rises (Baldrian *et al*., 2008).

Indeed, plant cover – particularly the presence of *Pinus* trees – has been found to raise the abundance of *Basidiomycota* (Tardy *et al*., 2015) and a similar trend was found here at forest sites with the highest DOC content and lignocellulosic compounds with a high C/N ratio (i.e. pine needles). Similarly, Schneider *et al*. (2012) found an increase of proteins of *Basidiomycota* when the C/N ratio of litter increases. The division dynamics within the kingdom *Fungi* are related to the type of plant inputs, with a high content of lignified compounds, and the existence of enzymatic mechanisms for the decomposition of such substrates (de Boer *et al*., 2005). The increased abundance and activity of *Basidiomycota*

across the organic C-availability gradient was made clear by the ITS, MP, and RP approaches, although the increase in RP proteins at the sites where the DOC was 640-1600 mg C kg^{-1} was more notable and, probably, indicative of the high activity of this fungal group. We found that the high translation capacity of the members of the *Basidiomycota* in soils with high DOC content is coupled to the production of extracellular enzymes for the decomposition of lignin and celluloses (Baldrian *et al*., 2008; Voriskova and Baldrian, 2013). Indeed, several potential extracellular enzymes involved in organic matter decomposition could be identified and related taxonomically to *Basidiomycota*: laccase was identified from *Agaricales* (*Basidiomycota*), and cellulases from *Helotiales* (*Ascomycota*) and *Sebacinales* (*Basidiomycota*). Moreover, when comparing the protein abundances involved in the transport and metabolism of carbohydrates, it is clear that *Ascomycota* and *Basidiomycota* occupy different "nutritional" niches: *Ascomycota* expressed these proteins in soils of moderate C content, while for *Basidiomycota* the protein levels were higher in soils with high DOC content - probably as a mechanism of plant cell-wall degradation.

It should be highlighted that this study was performed by searching MS/MS spectra against NCBI database but the use of specific soil metagenomes could have increased the success of protein identification. Nevertheless, we reveal comprehensive phylogenetic-functional adaptations of soil microbial communities to organic C availability in semiarid Mediterranean soils. Functional and phylogenetic analyses of proteins, as indicators of population activities complemented DNA-based community analyses and allowed a distinction of lifestyles that could not be achieved solely by the amplification of taxonomical genes. Importantly, the active diversity of the microbial community, estimated by metaproteomics, was more sensitive to organic C availability than DNA diversity and seemed a better predictive indicator of ecosystem functionality in the studied soils.

The active diversity is likely regulated through the plasticity of the physiological adaptations of populations which, in turn, modulates their absolute abundances. These physiological capacities of microbial populations are not constant across DOC gradients in semiarid soils.

Interestingly, functional mechanisms for nutritional-niche specialization were deciphered in both the fungal and bacterial kingdoms.

Acknowledgements

F. B thanks the Spanish Government for his “Ramón y Cajal” contract and FEDER funds (RYC-2012-10666). I.F. Torres acknowledges the Spanish Ministry for her FPI fellowship. The authors thank the Spanish Ministry for the CICYT projects AGL2010-16707 and AGL2014-54636-R. The authors are grateful to the Fundación Séneca (19896/GERM/15). F.B and N.J are grateful to the German Academic Exchange Service (DAAD) (PPP-Spain, 57049579) for their funded project. P. Baldrian was supported by the research concept of the Institute of Microbiology of the CAS (RVO61388971) and by C4Sys (LM2015055). N.J was partially funded by SFB Aquadiva. The authors are grateful for the use of the analytical facilities of the Centre for Chemical Microscopy (ProVIS) at the Helmholtz Centre for Environmental Research, which is supported by European Regional Development Funds (EFRE - Europe funds Saxony) and the Helmholtz Association. Benjamin Scheer (Department of Isotope Biogeochemistry) is acknowledged for the technical support for protein analysis in the ProVIS laboratory. Kathleen Eismann (Department of Molecular Systems Biology) is acknowledged for the technical support for proteome sample preparation.

Conflict of Interest

The authors declare no conflict of interest.

References

Baldrian P *et al.* (2008) Enzyme activities and microbial biomass in topsoil layer during spontaneous succession in spoil heaps after brown coal mining. *Soil Biology and Biochemistry,* **40**, 2107-2115.

Baldrian P *et al.* (2012) Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. The *ISME Journal,* **6**, 248-258.

Bardgett RD (2005) The biology of soil. A community and ecosystem approach. Oxford University Press, Oxford.

Bastida F, Moreno JL, Hernández T, García C (2006) Microbiological degradation index of soils in a semiarid climate. *Soil Biology and Biochemistry,* **38**, 3463-3473.

Bastida F, Hernández T, Albaladejo J, García C (2013) Phylogenetic and functional changes in the microbial community of long-term restored soils under semiarid climate. *Soil Biology and Biochemistry,* **65**, 12-21.

Bastida F *et al.* (2014a) Characterization of the microbial community in biological soil crusts dominated by *Fulgensia desertorum* (Tomin) Poelt and *Squamarina cartilaginea* (With.) P. James and in the underlying soil. *Soil Biology and Biochemistry,* **76**, 70-79.

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

Bastida F, Selevsek N, Torres IF, Hernández T, García C (2015a) Soil restoration with organic amendments: linking cellular functionality and ecosystem processes. *Scientific Reports,* **5**, 15550.

Bastida F *et al.* (2015b) Deforestation fosters bacterial diversity and the cyanobacterial community responsible for carbon fixation processes under semiarid climate: a metaproteomics study. *Applied Soil Ecology,* **93**, 65-67.

Bastida F *et al*. (2016) The ecological and physiological responses of the microbial community from a semiarid soil to hydrocarbon contamination and its bioremediation using compost amendment. *Journal of Proteomics*, **135**, 162-169.

Bernard L *et al.* (2007) Dynamics and identification of soil microbial populations actively assimilating carbon from C-13-labelled wheat residue as estimated by DNA- and RNA-SIP techniques. *Environmental Microbiology,* **9,** 752-764.

Belnap J, Lange OL (2003) Biological soil crusts: structure, function and management. Springer, Berlin. 503 p.

Blagodatskaya E. Kuzyakov Y (2013) Active microorganisms in soil: critical review of estimation criteria and approaches. *Soil Biology and Biochemistry,* **67**, 192-211.

Bowen SR, Gregorich EG, Hopkins DW. (2009). Biochemical properties and biodegradation of disolved organic matter from soils. *Biol Fertil. Soils* **45:**733-742.

Bradford MA. *et al.* (2014) Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition. *Proceedings of the National Academy of Sciences of the United States of America,* **111**, 14478-14486.

Camino-Serrano M, Gielen B, Luyssaert S *et al.* (2014) Linking variability in soil solution dissolved organic carbon to climate, soil type and vegetation type. *Global Biogeochemical Cycles*, **28**, 497-509.

Chourey K *et al.* (2010) Direct cellular lysis/protein extraction protocol for soil metaproteomics. *Journal of Proteome Research,* **9**, 6615-6622.

Cleveland CC, Nermegut DR, Schmidt SK, Townsend AR (2007) Increases in soil respiration following labile carbon additions linked to rapid shifts in soil microbial community composition. *Biogeochemistry,* **82**, 229-240.

de Boer W, Folman LB, Summerbell RC, Boddy L (2005) Living in a fungal world: impact of fungi on soil bacterial niche development. *FEMS Microbiology Reviews,* **29**, 795-811.

de Vries FT *et al.* (2012) Abiotic drivers and plant traits explain landscape scale patterns in soil microbial communities. *Ecology Letters,* **15**, 1230-1239.

Delgado-Baquerizo M *et al.* (2016) Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nature Communications,* **7**, 10541.

Delmotte N. *et al.* (2009) Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences of the United States of America,* **106**,16428-16433.

Drenovsky RE, Vo D, Graham KJ, Scow KM (2004) Soil water content and organic carbon availability are major determinants of soil microbial community composition. *Microbial Ecology,* **48**, 424-430.

Eilers KG, Lauber CL, Knight R, Fierer N (2010) Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. *Soil Biology and Biochemistry,* **42**, 896-903.

Fierer N, Bradford MA, Jackson RB (2007) Toward an ecological classification of soil bacteria. *Ecology,* **88**, 1354-1364.

Flynn KJ *et al.* (2010) Is the growth rate hypothesis applicable to microalgale? *Journal of Phycology,* **46**, 1-12.

García C, Hernández T, Costa F (1994) Microbial activity in soils under Mediterranean environmental conditions. *Soil Biology and Biochemistry,* **26**, 1185-1191.

Galbally IE, Kristine W (2002) The production of methanol by flowering plants and the global cycle of methanol. *Journal of Atmospheric Chemistry,* **43**, 195-229.

Goldfarb KC *et al.* (2011) Differential growth responses of soil bacterial taxa to carbon substrates of varying chemical recalcitrance. *Frontiers in Microbiology,* **2:**article94 1-10.

Guigue J *et al.* (2015) Water-extractable organic matter linked to soil physico-chemistry and microbiology at the regional scale. *Soil Biology and Biochemistry,* **84**, 158-167.

Hettich RL, Pan C, Chourey K, Giannone RJ (2013) Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. *Analytical Chemistry,* **85**, 4203-4214.

Hirsch PR *et al.* (2007) Starving the soil of plant inputs for 50 years reduces abundance but not diversity of soil bacterial communities. *Soil Biology and Biochemistry,* **41**, 2021-2024.

Hortal S *et al.* (2015) Benefactor and allelopathic shrub species have different effects on the soil microbial community along an environmental severity gradient. *Soil Biology and Biochemistry,* **88**, 48-57.

Johnson NC, Wilson GWT, Bowker MA, Wilson JA, Miller RM. (2010) Resource limitation is a driver of local adaptation in mycorrhizal symbioses. *Proceedings of the National Academy of Sciences of the United States of America,* **107**, 2093-2098.

Kaiser K, Guggenberger G, Zech W (2001) Isotopic fractionation of dissolved organic carbon in shallow forest soils as affected by sorption. *European Journal of Soil Science,* **52**, 585-597.

Kalbitz K, Geyer S, Gehre M (2000) Land use impacts on the isotopic signature (^{13}C, ^{14}C, ^{15}N) of water-soluble fulvic acids in a German fen area. *Soil Science,* **165**, 728-736.

Käll L, Canterbury JD, Weston J, Noble WS, MacCoss MJ (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nature Methods,* **4**, 923-925.

Keiblinger KM *et al.* (2012) Soil metaproteomics – Comparative evaluation of protein extraction protocols. *Soil Biology and Biochemistry,* **54**, 14-24.

Konstantinidis KT, Ramette A, Tiedje JM (2006) Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. *Applied and Environmental Microbiology,* **72**, 7286-7293.

Knief C *et al.* (2012) Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. The *ISME Journal,* **6**, 1378-1390.

Lauro FM *et al.* (2009) The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America,* **106**, 15527-15333.

Lefcheck JS *et al.* (2015) Biodiversity enhances ecosystem multifunctionality across throphic levels and habitats. *Nature Communications,* **6**, 6936.

Leff JW *et al.* (2015) Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academy of Sciences of the United States of America,* **112**, 10967-10972.

Liu J *et al.* (2015) Soil carbon content drives the biogeographical distribution of fungal communities in the black soil zone of northeast China. *Soil Biology and Biochemistry,* **83**, 29-39.

Mesuere B *et al.* (2015) The Unipept metaproteomics analysis pipeline. *Proteomics,* **15**. 1437-1442.

Odum EP (1969) The strategy of ecosystem development. *Science,* **164,** 262-270.

Portillo MC, Leff JW, Lauber CL, Fierer N (2013) Cell size distributions of soil bacterial and archaeal taxa. *Applied and Environmental Microbiology,* **79**, 7610-7617.

Roux S, Enault F, Bronner G, Debroas D (2011) Comparison of 16S rRNA and protein-coding genes as molecular markers for assessing microbial diversity (Bacteria and Archaea) in ecosystems. *FEMS Microbiol Ecology,* **78**, 617-628.

Sanderman J, Amundson R (2008) A comparative study of dissolved organic carbon transport and stabilization in California forest and grassland soils. *Biogeochemistry,* **92**, 41-59.

Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana.

Schneider T *et al.* (2011) Structure and function of the symbiosis partners of the lung lichen (*Lobaria pulmonaria* L. Hoffm.) analyzed by metaproteomics. *Proteomics,* **11**, 2752-2756.

Schneider T *et al.* (2012) Who is who in litter decomposition? Metaprotoemics reveals major microbial players and their biogeochemical functions. The *ISME Journal,* **6**, 1-14.

Sy A, Timmers ACJ, Knief C, Vorholt JA (2005) Methylotrophic metabolism is advantageous for *Methylobacterium extorquens* during colonization of *Medicago truncatula* under competitive conditions. *Applied and Environmental Microbiology,* **71**, 7245-7252.

Tardy V *et al.* (2015) Shifts in microbial diversity through land use intensity as drivers of carbon mineralization in soil. *Soil Biology and Biochemistry,* **90**, 204-213.

Trivedi P, Anderson IC, Singh BK (2013) Microbial modulators of soil carbon storage: integrating genomic and metabolic knowledge for gobal prediction. *Trends in Microbiology,* **21**, 641-651.

VerBerkmoes NC, Denef, VJ, Hettich RL, Banfield, JF (2009) Functional analysis of natural microbial consortia using community protoemics. *Nature Reviews in Microbiology,* **7**,196-205.

Vizcaino JA et al. (2013) The Proteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acid Research,* **41**, D1063-1069.

Voriskova J, Baldrian P (2013) Fungal community on decomposing leaf litter undergoes rapid successional changes. *The ISME J,* **7**, 477-486.

Wagg C, Franz Bender S, Widmer F, van der Heijden MGA (2014) Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proceedings of the National Academy of Sciences of the United States of America,* **111**, 5266-5270.

Wertz S *et al.* (2007) Decline of soil microbial diversity does not influence the resistance and resilience of key soil microbial functional groups following a model disturbance. *Environmental Microbiology,* **9**, 2211-2219.

Zifcakova L, Vetrovsky T, Howe A, Baldrian P (2016) Microbial activity in forest soil reflects the changes in ecosystem properties between summer and winter. *Environmental Microbiology,* **18**, 288-301.

Data Accessibility

All the 16S and ITS sequence data used in this study are deposited in the MG-RAST server under project 4686402.3 (http://metagenomics.anl.gov/linkin.cgi?project=16983). The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaino et al., 2013) with the dataset identifier PXD003572.

Supporting Information

Additional supporting information maybe found in the online version of this article.

Author Contributions

F.B. and J.L.M conceived the idea and experimental design and collected the samples. F.B., I.F.T., S.O., R.S., N.J. performed metaproteomic analysis. I.F.T, S.O. and F.B performed genomic and biochemical analysis. P.B provided the support for genomic analysis and bioinformatics. A.R.N and R.S. provided statistical support. R.S. and N.J were encharged on figure design. F.B. wrote the paper with substantial input from all the authors.

Figure Legends

Figure 1. The dissolved organic C (DOC) content in the studied gradient (A) and its relationships with diversity and ecosystem multifunctionality: DOC and genomic diversity (B); DOC and active diversity (C); DOC and ecosystem multifunctionality (D); Ecosystem multifunctionality and genomic diversity (E); and ecosystem multifunctionality and active

diversity (F). Regression equations are exclusively plotted if the Spearman correlation analysis between the variables resulted significant (P<0.05).

Figure 2. The composition of bacterial and fungal communities as evaluated by 16S rRNA gene sequencing and ITS sequencing (Genomics), metaproteomics and ribosomal proteins across the dissolved organic C (DOC)gradient. Regression curves are plotted for the most abundant groups.

Figure 3. Spearman correlation coefficients between dissolved organic C (DOC) and the relative abundance of populations analyzed by genomics and metaproteomics, functional proteins and cellular functions. Only significant correlation coefficients are showed (P<0.05).

Figure 4. Cellular functionalities of phyla (A) and orders (B) calculated by ranges of dissolved organic C (DOC) content. Data are presented as percentiles of DOC content. Cell cycle (proteins involved in

Table Legends

Table 1. Studied areas and soil characteristics

Table 2. Spearman correlation analysis between the dissolved organic C (DOC) content, diversity proxies, activity indicators, microbial biomass and the ecosystem multifunctionality index.

Table 1. Studied areas and soil characteristics[1]

Site	Coordinates	Height (m, asl[2])	Plant cover (%)	Plant types	TOC		TN		DOC	
Low productive, desertic soils					Mean	SD	Mean	SD	Mean	SD
Abanilla	38º 12' 16''N; 1º 5' 19''W	200	0	Bare	0.26	0.01	0.05	0.00	75.4	10.7
Gebas	37º 54' 38''N; 1º 24' 7''W	315	5	*Rosmarinus officinalis*	0.45	0.01	0.06	0.00	85.9	7.1
Mula	38º31'48''N; -1º41'38''W	266	15	*Thymus sp.*	0.65	0.13	0.08	0.02	108.7	2.6
Shrublands and Grasslands										
Abanilla	38º 12' 13''N; 1º 5' 18''W	194	60	*Xerophytiic shrub*	1.16	0.46	0.10	0.01	166.5	46.8
Gebas	37º 54' 15'' N; 1º 27' 9'' W	422	80	*Xerophytiic shrub*	3.26	0.40	0.29	0.10	346.4	26.4
Benablón	38º3'8''N; 1º53'48''W	791	80	*S.tenacissima, R. officinalis*	3.89	0.24	0.37	0.03	20.7	29.9
Benablón	38º3'8''N; 1º53'48''W	791	90	*Grassland*	4.80	0.18	0.42	0.02	476.9	104.0
Gebas	37º54'15''N; 1º24'38''W	369	40	*Pinus halepensis, Stipa tenacissima*	2.34	0.46	0.14	0.03	181.6	9.7
El Sabinar	38º1'52''N; 2º10'32''W	1207	85	*Thymus sp.*	2.39	0.05	0.25	0.01	90.7	14.7
Revolcadores	38º3'49''N; 2º17'48''W	1400	80	*Grassland*	3.23	0.29	0.26	0.02	125.4	12.2
Forest										
El Sabinar	38º1'52''N; 2º10'32''W	1207	100%*	*Juniperus sp., Thymus sp.*	9.88	0.23	0.65	0.03	495.5	59.9
Sª.Espuña	37º51'25''N; 1º28'13''W	405	100%*	*Pinus halepensis*	6.50	1.24	0.44	0.03	491.8	50.4
Sª.Espuña	37º51'13''N; 1º30'12''W	770	100%*	*Quercus coccifera*	5.84	1.04	0.46	0.05	844.1	182.0
Sª.Espuña	37º51'13''N; 1º30'12''W	770	100%*	*P.halepensis*	7.45	0.23	0.31	0.03	528.4	81.7
Sª.Espuña	37º51'43''; 1º32'21''W	990	100%*	*Juniperus sp.*	15.79	0.45	0.75	0.02	664.4	62.5
Sª.Espuña	37º52'4''N; 1º33'56''W	1360	100%*	*Quercus rotundifolia*	17.42	2.27	1.00	0.05	629.4	37.7
Sª.Espuña	37º52'4''N; 1º33'56''W	1360	100%*	*Pinus nigra*	10.80	1.76	0.59	0.02	431.9	38.0
Barranda	38º0'58''N; 2º2'27''W	993	100%*	*P.halepensis, R.officinalis, Asparagus horridus*	11.33	4.06	0.58	0.08	1040.5	114.5
Revolcadores	38º3'49''N; 2º17'48''W	1400	100%*	*Q.rotundifolia*	13.42	1.25	0.71	0.02	1596.8	137.1
Revolcadores	2º17'48''N; 2º17'48''W	1400	100%*	*Pinus pinaster, Quercus sp., A.horridus, Thymus sp.*	5.77	0.20	0.27	0.01	253.9	31.6

[1]TOC (Total organic C, %); TN (Total N, %), DOC (Dissolved organic C, mg C kg^{-1} soil). SD (Standard deviation of the mean); [2]Height (m), above the sea level (asl). *Sampled under tree canopy.

Table 2. Spearman correlation analysis between DOC, diversity proxies, activity indicators, microbial biomass and the ecosystem multifunctionality index.

	DOC	GD	BD	FD	AD	MF	SR	βGA	PA	UA
GD	0.096 (0.46)									
BD	0.31 (0.81)	0.76 (<0.001)								
FD	-0.71 (0.59)	0.79 (<0.001)	0.24 (0.069)							
AD	0.37 (0.004)	0.084 (0.35)	0.23 (0.084)	0.004 (0.98)						
MF	0.76 (<0.001)	0.125 (0.64)	0.10 (0.40)	0.11 (0.39)	0.50 (<0.001)					
SR	0.83 (<0.001)	0.008 (0.95)	0.061 (0.84)	0.002 (0.99)	0.54 (<0.001)	0.79 (<0.001)				
βGA	0,70 (<0.001)	0.17 (0.19)	0.13 (0.3)	0.14 (0.27)	0.49 (<0.001)	0.96 (<0.001)	0.72 (<0.001)			
PA	0.82 (<0.001)	0.12 (0.37)	0.15 (0.24)	0.045 (0.73)	0.35 (<0.001)	0.93 (<0.001)	0.74 (<0.001)	0.87 (<0.001)		
UA	0.38 (0.003)	-0.045 (0.73)	-0.11 (0.40)	0.024 (0.86)	0.31 (<0.001)	0.70 (<0.001)	0.46 (<0.001)	0.67 (<0.001)	0.59 (<0.001)	
MB	0.51 (<0.001)	0.38 (0.002)	0.31 (0.015)	0.33 (0.010)	0.47 (<0.001)	0.86 (<0.001)	0.56 (<0.001)	0.86 (<0.001)	0.71 (<0.001)	0.60 (<0.001)

DOC (Dissolved organic C); GD (Genomic diversity calculated as the averaged Z-scores of bacterial and fungal Shannon index by genomics); BD (bacterial genomic diversity); FD (fungal genomic diversity); AD (active diversity); MF (ecosystem multifunctionality index); SR (soil respiration); βGA (β-glucosidase activity); PA (phosphatase activity); UA (urease activity); MB (microbial biomass, sum of bacterial and fungal PLFAs). Correlation coefficients are followed by *P* values in parenthesis.

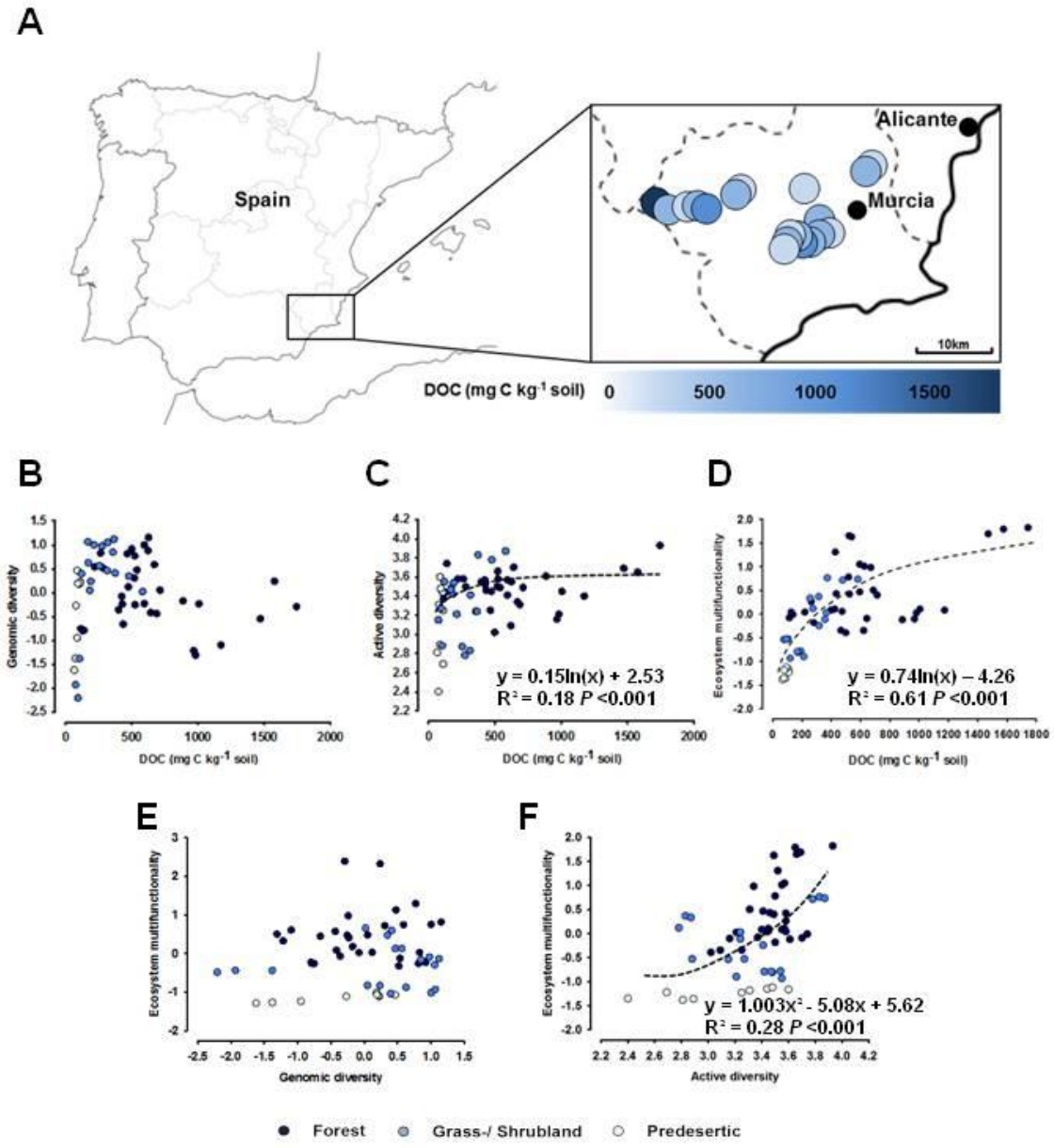

Figure 1. The dissolved organic C (DOC) content in the studied gradient (A) and its relationships with diversity and ecosystem multifunctionality: DOC and genomic diversity (B); DOC and active diversity (C); DOC and ecosystem multifunctionality (D); Ecosystem multifunctionality and genomic diversity (E); and ecosystem multifunctionality and active diversity (F). Regression equations are exclusively plotted if the Spearman correlation analysis between the variables resulted significant (P<0.05).

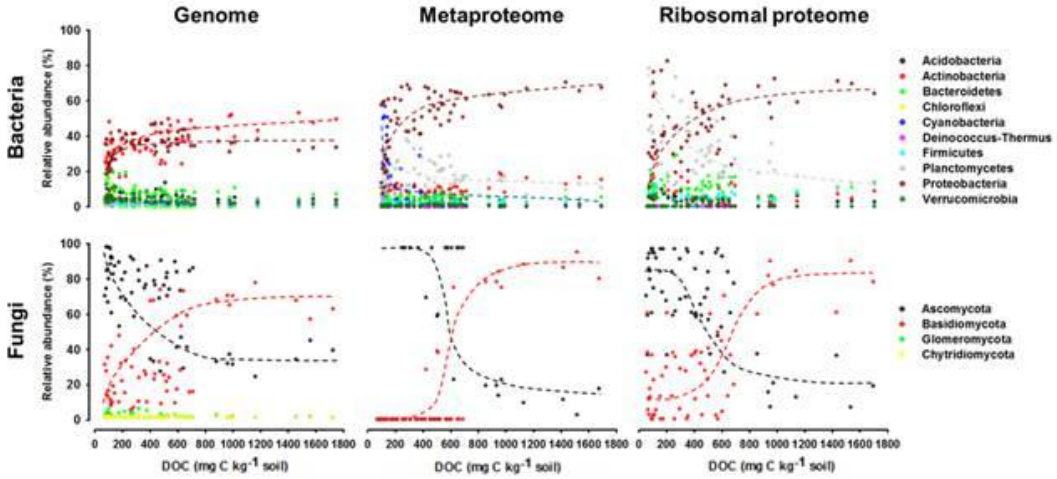

Figure 2. The composition of bacterial and fungal communities as evaluated by 16S rRNA gene sequencing and ITS sequencing (Genomics), metaproteomics and ribosomal proteins across the dissolved organic C (DOC) gradient. Regression curves are plotted for the most abundant groups.

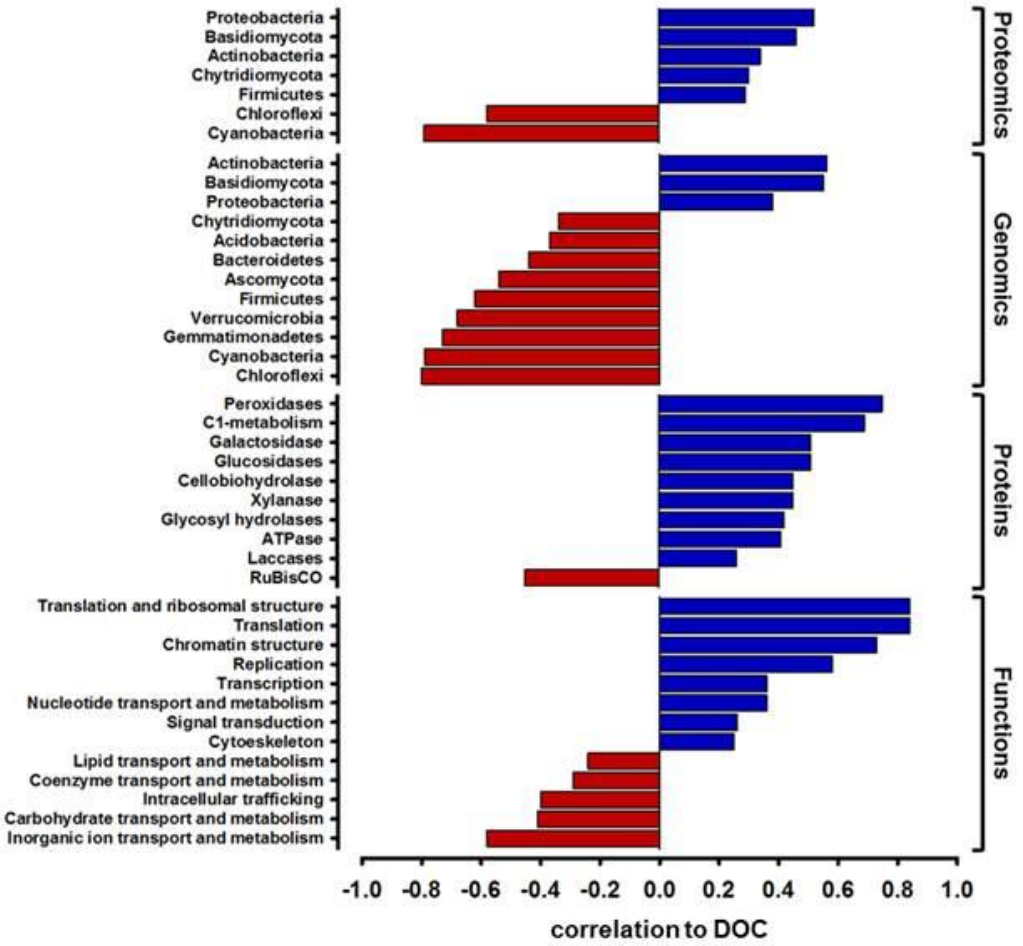

Figure 3. Spearman correlation coefficients between dissolved organic C (DOC) and the relative abundance of populations analyzed by genomics and metaproteomics, functional proteins and cellular functions. Only significant correlation coefficients are showed (P<0.05).

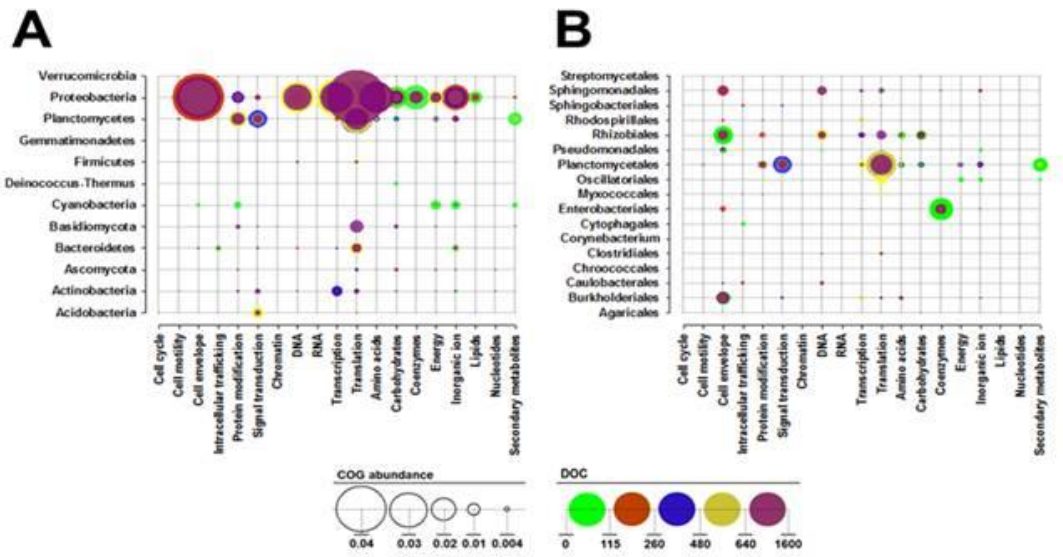

Figure 4. Cellular functionalities of phyla (A) and orders (B) calculated by ranges of dissolved organic C (DOC) content. Data are presented as percentiles of DOC content. Abbreviations: Cell cycle (Cell cycle control, cell division, chromosome partitioning); Cell envelope (Cell wall, membrane and envelope biogenesis); Intracellular trafficking (Intracellular trafficking, secretion, and vesicular transport); Protein modification (Postranslational modification, protein turnover, chaperones); Chromatin (Chromatin structure and dynamics); DNA (Replication, recombination and repair); RNA (RNA processing and modification); Translation (Translation, ribosomal structure and biogenesis); Amino acids (Amino acid transport and metabolism); Carbohydrate (Carbohydrate transport and metabolism); Energy (Energy production and conversion); Inorganic ion (Inorganic ion transport and metabolism); Lipids (Lipid transport and metabolism); Nucleotides (Nucleotide transport and metabolism); Secondary metabolites (Secondary metabolites biosynthesis, transport and catabolism).