Bastida, F., Crowther, T.W., Prieto, I., Routh, D., García, C., **Jehmlich, N.** (2018):
Climate shapes the protein abundance of dominant soil bacteria
Sci. Total Environ. **640–641** , 18 - 21

http://dx.doi.org/10.1016/j.scitotenv.2018.05.288

Title Page

Climate shapes the protein abundance of dominant soil bacteria

Felipe Bastida[1], Tom W. Crowther[2], Iván Prieto[1], Devin Routh[2], Carlos García[1], Nico Jehmlich[3]

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

[2]Institute of Integrative Biology, ETH Zürich, Univeritätstrasse 16, 8006 Zürich, Switzerland

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

Corresponding author:

Dr. Felipe Bastida

CEBAS-CSIC

Campus Universitario de Espinardo

CP 30100 PO Box 164, Murcia, Spain

Phone: +34 968396106 | Fax: +34 968396213 | E-mail: fbastida@cebas.csic.es

Type of submission: Short Communication

Keywords: soil microbial community; metaproteomics; climate change; microbial diversity

Climate shapes the protein abundance of dominant soil bacteria

Abstract

Sensitive models of climate change impacts would require a better integration of multi-omics approaches that connect the abundance and activity of microbial populations. Here, we show that climate is a fundamental driver of the protein abundance of Actinobacteria, Planctomycetes and Proteobacteria, supporting the hypothesis that metabolic activity of some dominant phyla may be closely linked to climate. These results may improve our capacity to construct microbial models that better predict the impact of climate change in ecosystem processes.

1.Introduction

Soil microorganisms mineralize the soil organic matter thus playing a pivotal role in nutrient biogeochemical cycles. Because they also produce gaseous metabolites with potential greenhouse effects, they are also able to modify both the local and global climate (Bradford *et al.,* 2016). In recent years, considerable efforts have been made to understand the factors governing soil microbial community composition through DNA sequencing approaches (Bastida *et al.,* 2016; Delgado-Baquerizo *et al.*, 2018). However, soil microbes are often not active and a large percentage of soil DNA can belong to non-living cells (Carini *et al.,* 2016). Consequently, conventional sequencing approaches (i.e. 16S rRNA gene amplicons) can be limited in their capacity to predict the connections between microbial populations and ecosystem functioning. While genomic information provides a wealth of important information about the potential molecular machinery that might be employed for life processes, proteins are the direct catalyzers of cellular and environmental processes (Hettich *et al.,* 2013).

As a result, soil metaproteomics (the direct identification of proteins in soil) has been proposed as a promising new approach for the evaluation of the active component of the soil microbiome at local scales (Hultman *et al.,* 2015; Bastida *et al.,* 2016; 2017). However, until now, metaproteomics have yet to be used to examine how the activity of soil microbes varies across broad spatial scales. Here, we use a comprehensive cross site investigation of soil metaproteomes in order to examine the abiotic factors that determine the global variability in the protein abundance of dominant bacterial phyla. Given that climate regulates the metabolic activity of soil bacteria and climate events can cause changes in the composition of soil microbial communities (Bell *et al*., 2014; Evans and Wallenstein, 2014), we hypothesize that climate factors will shape more the protein abundance of bacterial phyla than their abundance evaluated by phylogenetic gene markers.

2. Materials and methods

In order to span a large environmental gradient with strong climatic differences, soils from boreal, temperate and semiarid ecosystems were collected within a five-year period from 2010 to 2015 (Table S1, Supplementary Information). Soils were sampled from Long-term Ecological Research Stations in North-America and South-Europe. These soils are included in the studies of Crowther *et al*. (2014) and Bastida *et al*. (2016, 2017). Soils were collected using a sampling design entender to maximize variation in climatic conditions and biome types. Sites included: Bonanza Creek, Alaska 64.85 °N, -147.84 °E (BNZ); Coweeta LTER, North Carolina 35 °N, -83.5 °E (CWT); Hubbard Brook Experimental Forest, New Hampshire 43.94 °N, -71.75 °E (HBR); Konza Prairie Biological Station, Kansas 39.09 °N, -96.57 °E (KNZf); pre-desertic semiarid grassland, Murcia 37° 54' N, 1° 24' °W (GEB); pre-desertic semiarid forest Murcia 37° 54' N, 1° 24' °W (GEP); Mediterranean forest Albacete 38° 22' N, 2° 20' °W (ALB); and Mediterranean forest, Granada 38° 0' N, 2° 2' °W (MOJ). Three soil samples were collected in each of the sites. Each sample was a replicate. Soil samples were sieved (2 mm) and immediately preserved at -20ºC. The following parameters were determined for each site: content of clay, silt, salt, organic C, and total N, C-to-N ratio, pH, soil moisture, mean-annual temperature (MAT) and mean-annual precipitation (MAP) (Supplementary Information). The ratio between MAT and MAP was utilized as an aridity indicator: the higher the ratio, the higher the aridity. Soil organic C ranged between 0.5 and 15.5%; total N between 0.05 and 0.84%; pH between 4.2 and 8.6 (Table S1) and C/N ratio between 9 and 20. MAT ranged between -2.94ºC and 18ºC. MAP varied between 260 mm to 1400 mm. The bacterial community composition (16S rRNA gene amplicon sequencing) of studied samples was obtained from the abovementioned studies and compared to metaproteome data obtained here. Proteins were extracted and processed as described elsewhere (Chourey *et al.,* 2011; Supplementary

Information). Genomics and metaproteomics were analyzed in each replicate and utilized as individual replicate in the statistical analyses. The mass spectrometry data have been deposited in the PRIDE partner repository with the dataset PXD003572, PXD005447 and PXD009773.

Pearson correlation coefficients between the abundance of each phylum given by metaproteomics and that retrieved by 16S rRNA sequencing were obtained (Table S2, Supporting Information). These results revealed signficant correlation coefficients in the case of Acidobacteria, Actinobacteria and Cyanobacteria, but not in the case of Bacteroidetes, Firmicutes, Planctomycetes and Proteobacteria.

Since not all variables were linearly related to soil abiotic variables, we first selected the best fitting approach (linear, quadratic, exponential or inverse fit) for each abiotic variable based on an Akaike information criterion (AIC) approach (Burnham and Anderson, 2002). AIC is an estimator of the relative quality of statistical models for a given dataset and provides a statistical method to select the best model among different ones. To do this, for each bacterial phyla or cellular functionality separately and each abiotic variable as the non-dependent variable we constructed all four possible models (linear, quadratic, exponential or inverse fit) and selected the best fitting approach using the AIC (Burnham and Anderson, 2002). The best fitting approach for each individual abiotic variable was selected on the basis of the lowest AIC for each set of four models containing this variable. Once the best fitting approach for each abiotic variable was selected, we tested which abiotic variable best explained the abundance of phyla through 16S rRNA sequencing or metaproteomics, as well as the abundance of proteins involved in cellular functionalities. To do this, we constructed all possible models with each abiotic variable as the non-dependent variable (12 models in total for the abundance of each bacterial phylum studied by sequencing and metaproteomics, and for the abundance of proteins involved in cellular functionality). The abiotic variable that best explained the abundance of bacterial phyla through 16S rRNA gene sequencing or metaproteomics, and the abundance of proteins involved in different cellular functionalities was selected on the basis of the model with the lowest AIC. We then calculated the adjusted R^2 and *P* values and extracted the regression equation coefficients for each best univariate model. Statistical analyses were done using the R software (version 3.3.2, R Development Core Team, 2016).

3.Results and discussion

Metaproteomics revealed that the abundance of the dominant bacterial phyla was best explained by pH and MAT/MAP (Table 1). In contrast, climatic factors were not selected in the univariate models that explained the abundance of bacterial phyla, studied by 16S rRNA gene sequencing. Previous studies based on 16S rRNA gene sequencing have highlighted the paramount role of pH on shaping the composition and diversity of soil microbial community (Fierer and Jackson 2006), as we also observed here (Fig. 1). The protein abundance of Acidobacteria and that of Actinobacteria and Cyanobacteria were contrarily shaped by pH.

There is a need for accurately forecasting the ecological consequences of global change in soil microbial communities (Delgado-Baquerizo *et al.*, 2018a). The protein abundance of Actinobacteria, Planctomycetes and Proteobacteria, as well as the abundance of proteins involved in the production and conversion of energy, were best explained by MAT/MAP (Table 1). Proteobacterial protein content was negatively related to MAT/MAP and was higher in temperate and boreal sites (lower MAT/MAP) than in semiarid sites (higher MAT/MAP), while actinobacterial protein abundance was favored in arid environments (Fig. 1). This phylum has been suggested to be adapted to harsh conditions such soil drying through their peptidoglycan layer (Battistuzzi and Hedges, 2009). Indeed, a sequencing study revealed that Actinobacteria can outcompete other dominant groups such Acidobacteria under arid conditions (Delgado-Baquerizo et al., 2018b). Considering the proposed univariate models, the increase in MAT/MAP, as predicted by climate change models in many areas of boreal, temperate and semiarid ecosystems (IPCC 2013), will likely enhance the protein abundance of Actinobacteria and Planctomycetes, and decrease that of Proteobacteria. Moreover, the abundance of proteins related to cellular energetic processes (among them the F-type H+-transporting ATPase was dominant) increased linearly with MAT/MAP (Table 1). These findings suggest an acceleration of energetic metabolic processes in soil (i.e. greater soil organic matter mineralization) induced by global warming (Bond-Lamberty and Thomson, 2010) and which is in line with the patterns of carbon losses occurring at a global scale (Crowther *et al.,* 2016).

In despite the limitations of metaproteomics due to the absence of genome databases (Starke et al., 2017), the complicated protein extraction due to their interaction with humic substances and soil particles (Giagnoni et al., 2011), and the reduced number of analyzed samples, climatic conditions and biome types, our study provides the first tentative distribution of soil microbial proteins at broad spatial scales. As revealed by the univariate models, climate is a fundamental determinant of the protein abundance of Actinobacteria, Planctomycetes and Proteobacteria, yet not their genomic abundance, suggesting that their metabolic activity could be more closely linked to climate than community composition. Our findings have potential implications within a global change scenario and suggest that the metabolic activity of such bacterial phyla, as well as the amount of proteins involved in cellular energetic processes, could be sensitive to the forecasted increase in aridity and warming (Huang *et al*., 2016). These initial results are the first in what is likely to be a blossoming field in microbial ecology, which may transform our capacity to construct microbially-explicit biogeochemical models that better predict the cellular functionalities of soil microbial communities under current and future climate scenarios.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

Authors thank the Spanish Ministry for the project AGL2014-54636-R and Séneca Foundation for the project (19896/GERM/15). Kathleen Eismann is acknowledged for her help in preparing metaproteome samples for LC-MS analysis. Noah Fierer is acknowledged for sharing the sample set of this study.

References

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, 4660-4673.

Bastida, F., Torres, I.F., Andrés-Abellán, M., Baldrian, P., López-Mondéjar, R., Vetrovsky, T., Richnow, H.H., Starke, R., Ondoño, S., García, C., López-Serrano, F.R., Jehmlich, N., 2017. Differential sensitivity of total and active soil microbial communities to drought and forest management. Glob. Change Biol. 23, 4185-4203.

Battistuzzi, F.U., Hedges, S.B., 2009. A major clade of Prokaryotes with ancient adaptations to life on land. Mol. Biol. Evol. 26, 335-343.

Bell, C.W., Tissue, D.T., Loik, M.E., Wallenstein, M.D., Acosta-Martínez, V., Erickson, RA., Zak, J.C., 2014. Soil microbial and nutrient responses to 7 years of seasonally altered precipitation in a Chihuahuan Desert grassland. Glob. Change Biol. 20, 1657-1673.

Bond-Lamberty, D., Thomson, A., 2010. Temperature-associated increases in the global soil respiration record. Nature 464, 579-583.

Bradford, M.A., Wieder, W.R., Bonan, G.B., Fierer, N., Raymond, P.A., Crowther, T.W., 2016. Managing uncertainty in soil carbon feedbacks to climate change. Nat. Clim. Chang. 6, 751-758.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd edn. Springer- Verlag, New York, NY, USA

Carini, P., Marsden, P.J., Leff, J.W., Morgan, E.E., Strickland, M.S., Fierer, N., 2017. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. Nat. Microbiol. 2, 16242.

Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., Brodie, E.L., Hettich, R.L., 2010. Direct cellular lysis/protein extraction protocol for soil metaproteomics. J. Proteome Res. 9, 6615-6622.

Crowther, T.W., Maynard, D.S., Leff, J.W., Oldfield, E.E., McCulley, R.L., Fierer, N., Bradford, M.A., *2014*. Predicting the responsiveness of soil biodiversity to deforestation: a cross-biome study. Glob. Change Biol. 20, 2983-2994.

Crowther, T.W., Todd-Brown, K.E.O., Rowe, C.W., Wieder, W.R., Carey, J.C., Machmuller, M.B., *et al*., 2016. Quantifying global soil carbon losses in response to warming. Nature 540, 104-108.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D.J., Bardgett, R.D., *et al*. (2018a). A global atlas of the dominant bacterial found in soil. Science 359, 320-325.

Delgado-Baquerizo, M., Reith, F., Dennis, P.G., Hamonts, K., Powell, J.R., Young, A., Singh, B.K., Bissett, A., 2018b. Ecological drivers of soil microbial diversity and soil biological networks in the Southern Hemisphere. Ecology 99, 583-596

Evans, S.E., Wallenstein, M.D., 2014. Climate change alters ecological strategies of soil bacteria. Ecol. Lett. 17, 155-164.

Fierer, N., Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. Proc. Natl. Acad. Sci. USA 103, 626-631.

Giagnoni, L., Magherini, F., Landi, L., Taghavi, S., Modesti, A., Bini, L., Nannipieri, P., van der Lelie, D., Renella, G., 2011. Extraction of microbial proteome from soil: potential and limitations assessed through a model study. Eur. J. Soil Sci., 62, 74-81.

Hettich, R.L, Pan, C., Chourey, K., Giannone, R.J., 2013. Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. Anal. Chem. 85, 4203-4214.

Huang, J., Yu, H., Guan, X., Wang, G., Guo, R., 2016. Accelerated dryland expansion under climate change. Nat. Clim. Change. 6, 166-172.

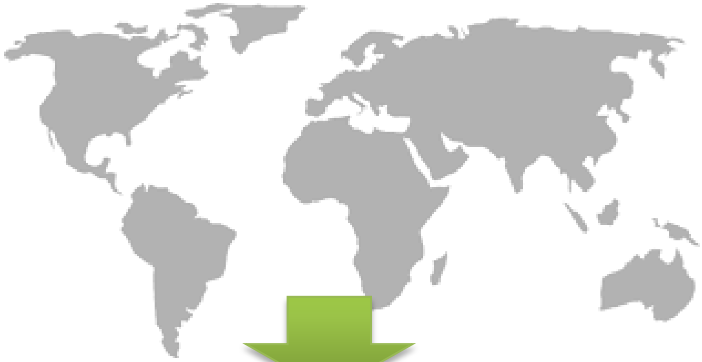
Hultman, J., Waldrop, M.P., Mackelprang, R., David, M.M., McFarland, J., Blazewicz, S.J., Harde, J., Turetsky, M.R., McGuire, A.D., Shah, M.B., VerBerkmoes, N.C., Lee, L.H., Mavrommatis, K., Jansson, J.K., 2015. Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. Nature 521, 208-212.

IPCC., 2013. Intergovernmental panel on climate change. Retrieved from http://www.ipcc.ch/.

Starke, R., Bastida, F., Abadía, J., García, C., Nicolás, E., Jehmlich, N., 2017. Ecological and functional adaptations to water management in a semiarid agroecosystem: a soil metaproteomics approach. Sci. Rep., 7, 10221.

Figure 1. The relative abundance of bacterial phyla studied by 16S rRNA amplicon sequencing (A) and metaproteomics (B), and the abundance of proteins involved in functional processes (C). Figure 1A is based on earlier data from Crowther et al. (2014) and Bastida et al. (2016, 2017).

*Graphical Abstract



Soil from different ecosystems

16S rRNA gene sequencing and Metaproteomics

Univariate models

Environmental and climate drivers of the abundance of bacterial phyla

***Highlights (for review : 3 to 5 bullet points (maximum 85 characters including spaces per bullet point)**

-Soil microorganisms play a pivotal role in biogeochemical cycles.

-Soil proteins from boreal, temperate and semiarid ecosystems were extracted.

-Abiotic variables that explained protein abundance were evaluated.

-Soil protein content of some bacteria phyla was linked to climate indicators.

Climate shapes the protein abundance of dominant soil bacteria

Abstract

Sensitive models of climate change impacts would require a better integration of multi-omics approaches that connect the abundance and activity of microbial populations. Here, we show that climate is a fundamental driver of the protein abundance of Actinobacteria, Planctomycetes and Proteobacteria, supporting the hypothesis that metabolic activity of some dominant phyla may be closely linked to climate. These results may improve our capacity to construct microbial models that better predict the impact of climate change in ecosystem processes.

1.Introduction

Soil microorganisms mineralize the soil organic matter thus playing a pivotal role in nutrient biogeochemical cycles. Because they also produce gaseous metabolites with potential greenhouse effects, they are also able to modify both the local and global climate (Bradford *et al.,* 2016). In recent years, considerable efforts have been made to understand the factors governing soil microbial community composition through DNA sequencing approaches (Bastida *et al.,* 2016; Delgado-Baquerizo *et al.*, 2018). However, soil microbes are often not active and a large percentage of soil DNA can belong to non-living cells (Carini *et al.,* 2016). Consequently, conventional sequencing approaches (i.e. 16S rRNA gene amplicons) can be limited in their capacity to predict the connections between microbial populations and ecosystem functioning. While genomic information provides a wealth of important information about the potential molecular machinery that might be employed for life processes, proteins are the direct catalyzers of cellular and environmental processes (Hettich *et al.,* 2013).

As a result, soil metaproteomics (the direct identification of proteins in soil) has been proposed as a promising new approach for the evaluation of the active component of the soil microbiome at local scales (Hultman *et al.,* 2015; Bastida *et al.,* 2016; 2017). However, until now, metaproteomics have yet to be used to examine how the activity of soil microbes varies across broad spatial scales. Here, we use a comprehensive cross site investigation of soil metaproteomes in order to examine the abiotic factors that determine the global variability in the protein abundance of dominant bacterial phyla. Given that climate regulates the metabolic activity of soil bacteria and climate events can cause changes in the composition of soil microbial communities (Bell *et al*., 2014; Evans and Wallenstein, 2014), we hypothesize that climate factors will shape more the protein abundance of bacterial phyla than their abundance evaluated by phylogenetic gene markers.

2. Materials and methods

In order to span a large environmental gradient with strong climatic differences, soils from boreal, temperate and semiarid ecosystems were collected within a five-year period from 2010 to 2015 (Table S1, Supplementary Information). Soils were sampled from Long-term Ecological Research Stations in North-America and South-Europe. These soils are included in the studies of Crowther *et al*. (2014) and Bastida *et al*. (2016, 2017). Soils were collected using a sampling design entender to maximize variation in climatic conditions and biome types. Sites included: Bonanza Creek, Alaska 64.85 °N, -147.84 °E (BNZ); Coweeta LTER, North Carolina 35 °N, -83.5 °E (CWT); Hubbard Brook Experimental Forest, New Hampshire 43.94 °N, -71.75 °E (HBR); Konza Prairie Biological Station, Kansas 39.09 °N, -96.57 °E (KNZf); pre-desertic semiarid grassland, Murcia 37° 54’ N, 1° 24’ °W (GEB); pre-desertic semiarid forest Murcia 37° 54’ N, 1° 24’ °W (GEP); Mediterranean forest Albacete 38° 22’ N, 2° 20’ °W (ALB); and Mediterranean forest, Granada 38° 0’ N, 2° 2’ °W (MOJ). Three soil samples were collected in each of the sites. Each sample was a replicate. Soil samples were sieved (2 mm) and immediately preserved at -20ºC. The following parameters were determined for each site: content of clay, silt, salt, organic C, and total N, C-to-N ratio, pH, soil moisture, mean-annual temperature (MAT) and mean-annual precipitation (MAP) (Supplementary Information). The ratio between MAT and MAP was utilized as an aridity indicator: the higher the ratio, the higher the aridity. Soil organic C ranged between 0.5 and 15.5%; total N between 0.05 and 0.84%; pH between 4.2 and 8.6 (Table S1) and C/N ratio between 9 and 20. MAT ranged between -2.94ºC and 18ºC. MAP varied between 260 mm to 1400 mm. The bacterial community composition (16S rRNA gene amplicon sequencing) of studied samples was obtained from the abovementioned studies and compared to metaproteome data obtained here. Proteins were extracted and processed as described elsewhere (Chourey *et al.,* 2011; Supplementary

Information). Genomics and metaproteomics were analyzed in each replicate and utilized as individual replicate in the statistical analyses. The mass spectrometry data have been deposited in the PRIDE partner repository with the dataset PXD003572, PXD005447 and PXD009773.

Pearson correlation coefficients between the abundance of each phylum given by metaproteomics and that retrieved by 16S rRNA sequencing were obtained (Table S2, Supporting Information). These results revealed signficant correlation coefficients in the case of Acidobacteria, Actinobacteria and Cyanobacteria, but not in the case of Bacteroidetes, Firmicutes, Planctomycetes and Proteobacteria.

Since not all variables were linearly related to soil abiotic variables, we first selected the best fitting approach (linear, quadratic, exponential or inverse fit) for each abiotic variable based on an Akaike information criterion (AIC) approach (Burnham and Anderson, 2002). AIC is an estimator of the relative quality of statistical models for a given dataset and provides a statistical method to select the best model among different ones. To do this, for each bacterial phyla or cellular functionality separately and each abiotic variable as the non-dependent variable we constructed all four possible models (linear, quadratic, exponential or inverse fit) and selected the best fitting approach using the AIC (Burnham and Anderson, 2002). The best fitting approach for each individual abiotic variable was selected on the basis of the lowest AIC for each set of four models containing this variable. Once the best fitting approach for each abiotic variable was selected, we tested which abiotic variable best explained the abundance of phyla through 16S rRNA sequencing or metaproteomics, as well as the abundance of proteins involved in cellular functionalities. To do this, we constructed all possible models with each abiotic variable as the non-dependent variable (12 models in total for the abundance of each bacterial phylum studied by sequencing and metaproteomics, and for the abundance of proteins involved in cellular functionality). The abiotic variable that best explained the abundance of bacterial phyla through 16S rRNA gene sequencing or metaproteomics, and the abundance of proteins involved in different cellular functionalities was selected on the basis of the model with the lowest AIC. We then calculated the adjusted R^2 and *P* values and extracted the regression equation coefficients for each best univariate model. Statistical analyses were done using the R software (version 3.3.2, R Development Core Team, 2016).

3.Results and discussion

Metaproteomics revealed that the abundance of the dominant bacterial phyla was best explained by pH and MAT/MAP (Table 1). In contrast, climatic factors were not selected in the univariate models that explained the abundance of bacterial phyla, studied by 16S rRNA gene sequencing. Previous studies based on 16S rRNA gene sequencing have highlighted the paramount role of pH on shaping the composition and diversity of soil microbial community (Fierer and Jackson 2006), as we also observed here (Fig. 1). The protein abundance of Acidobacteria and that of Actinobacteria and Cyanobacteria were contrarily shaped by pH.

There is a need for accurately forecasting the ecological consequences of global change in soil microbial communities (Delgado-Baquerizo *et al.*, 2018a). The protein abundance of Actinobacteria, Planctomycetes and Proteobacteria, as well as the abundance of proteins involved in the production and conversion of energy, were best explained by MAT/MAP (Table 1). Proteobacterial protein content was negatively related to MAT/MAP and was higher in temperate and boreal sites (lower MAT/MAP) than in semiarid sites (higher MAT/MAP), while actinobacterial protein abundance was favored in arid environments (Fig. 1). This phylum has been suggested to be adapted to harsh conditions such soil drying through their peptidoglycan layer (Battistuzzi and Hedges, 2009). Indeed, a sequencing study revealed that Actinobacteria can outcompete other dominant groups such Acidobacteria under arid conditions (Delgado-Baquerizo et al., 2018b). Considering the proposed univariate models, the increase in MAT/MAP, as predicted by climate change models in many areas of boreal, temperate and semiarid ecosystems (IPCC 2013), will likely enhance the protein abundance of Actinobacteria and Planctomycetes, and decrease that of Proteobacteria. Moreover, the abundance of proteins related to cellular energetic processes (among them the F-type H+-transporting ATPase was dominant) increased linearly with MAT/MAP (Table 1). These findings suggest an acceleration of energetic metabolic processes in soil (i.e. greater soil organic matter mineralization) induced by global warming (Bond-Lamberty and Thomson, 2010) and which is in line with the patterns of carbon losses occurring at a global scale (Crowther *et al.,* 2016).

In despite the limitations of metaproteomics due to the absence of genome databases (Starke et al., 2017), the complicated protein extraction due to their interaction with humic substances and soil particles (Giagnoni et al., 2011), and the reduced number of analyzed samples, climatic conditions and biome types, our study provides the first tentative distribution of soil microbial proteins at broad spatial scales. As revealed by the univariate models, climate is a fundamental determinant of the protein abundance of Actinobacteria, Planctomycetes and Proteobacteria, yet not their genomic abundance, suggesting that their metabolic activity could be more closely linked to climate than community composition. Our findings have potential implications within a global change scenario and suggest that the metabolic activity of such bacterial phyla, as well as the amount of proteins involved in cellular energetic processes, could be sensitive to the forecasted increase in aridity and warming (Huang *et al*., 2016). These initial results are the first in what is likely to be a blossoming field in microbial ecology, which may transform our capacity to construct microbially-explicit biogeochemical models that better predict the cellular functionalities of soil microbial communities under current and future climate scenarios.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

Authors thank the Spanish Ministry for the project AGL2014-54636-R and Séneca Foundation for the project (19896/GERM/15). Kathleen Eismann is acknowledged for her help in preparing metaproteome samples for LC-MS analysis. Noah Fierer is acknowledged for sharing the sample set of this study.

References

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, 4660-4673.

Bastida, F., Torres, I.F., Andrés-Abellán, M., Baldrian, P., López-Mondéjar, R., Vetrovsky, T., Richnow, H.H., Starke, R., Ondoño, S., García, C., López-Serrano, F.R., Jehmlich, N., 2017. Differential sensitivity of total and active soil microbial communities to drought and forest management. Glob. Change Biol. 23, 4185-4203.

Battistuzzi, F.U., Hedges, S.B., 2009. A major clade of Prokaryotes with ancient adaptations to life on land. Mol. Biol. Evol. 26, 335-343.

Bell, C.W., Tissue, D.T., Loik, M.E., Wallenstein, M.D., Acosta-Martínez, V., Erickson, RA., Zak, J.C., 2014. Soil microbial and nutrient responses to 7 years of seasonally altered precipitation in a Chihuahuan Desert grassland. Glob. Change Biol. 20, 1657-1673.

Bond-Lamberty, D., Thomson, A., 2010. Temperature-associated increases in the global soil respiration record. Nature 464, 579-583.

Bradford, M.A., Wieder, W.R., Bonan, G.B., Fierer, N., Raymond, P.A., Crowther, T.W., 2016. Managing uncertainty in soil carbon feedbacks to climate change. Nat. Clim. Chang. 6, 751-758.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd edn. Springer- Verlag, New York, NY, USA

Carini, P., Marsden, P.J., Leff, J.W., Morgan, E.E., Strickland, M.S., Fierer, N., 2017. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. Nat. Microbiol. 2, 16242.

Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., Brodie, E.L., Hettich, R.L., 2010. Direct cellular lysis/protein extraction protocol for soil metaproteomics. J. Proteome Res. 9, 6615-6622.

Crowther, T.W., Maynard, D.S., Leff, J.W., Oldfield, E.E., McCulley, R.L., Fierer, N., Bradford, M.A., *2014*. Predicting the responsiveness of soil biodiversity to deforestation: a cross-biome study. Glob. Change Biol. 20, 2983-2994.

Crowther, T.W., Todd-Brown, K.E.O., Rowe, C.W., Wieder, W.R., Carey, J.C., Machmuller, M.B., *et al*., 2016. Quantifying global soil carbon losses in response to warming. Nature 540, 104-108.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D.J., Bardgett, R.D., *et al*. (2018a). A global atlas of the dominant bacterial found in soil. Science 359, 320-325.

Delgado-Baquerizo, M., Reith, F., Dennis, P.G., Hamonts, K., Powell, J.R., Young, A., Singh, B.K., Bissett, A., 2018b. Ecological drivers of soil microbial diversity and soil biological networks in the Southern Hemisphere. Ecology 99, 583-596

Evans, S.E., Wallenstein, M.D., 2014. Climate change alters ecological strategies of soil bacteria. Ecol. Lett. 17, 155-164.

Fierer, N., Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. Proc. Natl. Acad. Sci. USA 103, 626-631.

Giagnoni, L., Magherini, F., Landi, L., Taghavi, S., Modesti, A., Bini, L., Nannipieri, P., van der Lelie, D., Renella, G., 2011. Extraction of microbial proteome from soil: potential and limitations assessed through a model study. Eur. J. Soil Sci., 62, 74-81.

Hettich, R.L, Pan, C., Chourey, K., Giannone, R.J., 2013. Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. Anal. Chem. 85, 4203-4214.

Huang, J., Yu, H., Guan, X., Wang, G., Guo, R., 2016. Accelerated dryland expansion under climate change. Nat. Clim. Change. 6, 166-172.

Hultman, J., Waldrop, M.P., Mackelprang, R., David, M.M., McFarland, J., Blazewicz, S.J., Harde, J., Turetsky, M.R., McGuire, A.D., Shah, M.B., VerBerkmoes, N.C., Lee, L.H., Mavrommatis, K., Jansson, J.K., 2015. Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. Nature 521, 208-212.

IPCC., 2013. Intergovernmental panel on climate change. Retrieved from http://www.ipcc.ch/.

Starke, R., Bastida, F., Abadía, J., García, C., Nicolás, E., Jehmlich, N., 2017. Ecological and functional adaptations to water management in a semiarid agroecosystem: a soil metaproteomics approach. Sci. Rep., 7, 10221.

Figure 1. The relative abundance of bacterial phyla studied by 16S rRNA amplicon sequencing (A) and metaproteomics (B), and the abundance of proteins involved in functional processes (C). Figure 1A is based on earlier data from Crowther et al. (2014) and Bastida et al. (2016, 2017).

Table 1. Best univariate models explaining the abundance of bacterial populations through 16S rRNA gene sequencing and metaproteomics, and the abundance of proteins involved in cellular functionality

	Dependent variable	Non-dependent Variable	Model type	AIC	Adjusted R^2	P value	a	b	c
Genomics	Acidobacteria	pH	Linear	156.03	0.82	<0.001	84.77	-9.04	
	Actinobacteria	pH	Exponential	69.79	0.67	<0.001	0.14	0.60	
	Bacteroidetes	C/N	Linear	171.73	0.17	0.0014	16.19	-0.50	
	Cyanobacteria	C	Inverse fit	98.44	0.88	<0.001	-0.48	5.12	
	Firmicutes	pH	Exponential	41.43	0.68	<0.001	9.84	-0.14	
	Planctomycetes	Silt	Linear	52.22	0.34	<0.001	4.25	-0.023	
	Proteobacteria	C/N	Linear	172.82	0.29	0.0014	17.65	0.69	
Metaproteomics	Acidobacteria	pH	Linear	259.08	0.64	<0.001	23.09	-2.35	
	Actinobacteria	MAT/MAP	Quadratic	168.59	0.17	0.003	6.77	1.79	-0.22
	Bacteroidetes	pH	Linear	128.46	0.50	<0.001	-4.09	1.13	
	Cyanobacteria	pH	Exponential	21.60	0.62	<0.001	0.14	0.60	
	Firmicutes	Clay	Inverse fit	103.56	0.10	0.049	0.84	4.02	
	Planctomycetes	MAT/MAP	Linear	191.97	0.42	<0.001	6.36	1.68	
	Proteobacteria	MAT/MAP	Linear	219.13	0.64	<0.001	68.97	-4.05	
	Carbohydrate metabolism	Silt	Linear	124.67	0.22	0.0053	6.76	-0.055	
	Energy production and conversion	MAT/MAP	Linear	131.18	0.59	<0.001	6.36	0.85	
	Transcription	C/N	Quadratic	121.14	0.39	<0.001	-10.19	1.69	-0.043
	Translation-ribosome	C/N	Linear	136.36	0.29	0.0014	-3.01	0.37	

Results show the best model among all possible models with each abiotic variable (lowest Akaike information criteria, AIC, see Supplementary material). For all variables, ΔAIC> 2 for the best fit model. Equation parameters (a,b,c) are given for linear ($y = a + bx$), quadratic ($y= a +bx + cx^2$), inverse fit ($y= a + (1/x)$) and exponential ($y=a*e^{bx}$) models. MAT/MAP is the ratio between mean annual temperature (MAT) and mean annual precipitation (MAP) and C/N is the total carbon:nitrogen ratio. The dependent variables were the relative abundance of bacterial phyla studied by 16S rRNA amplicon sequencing and metaproteomics.

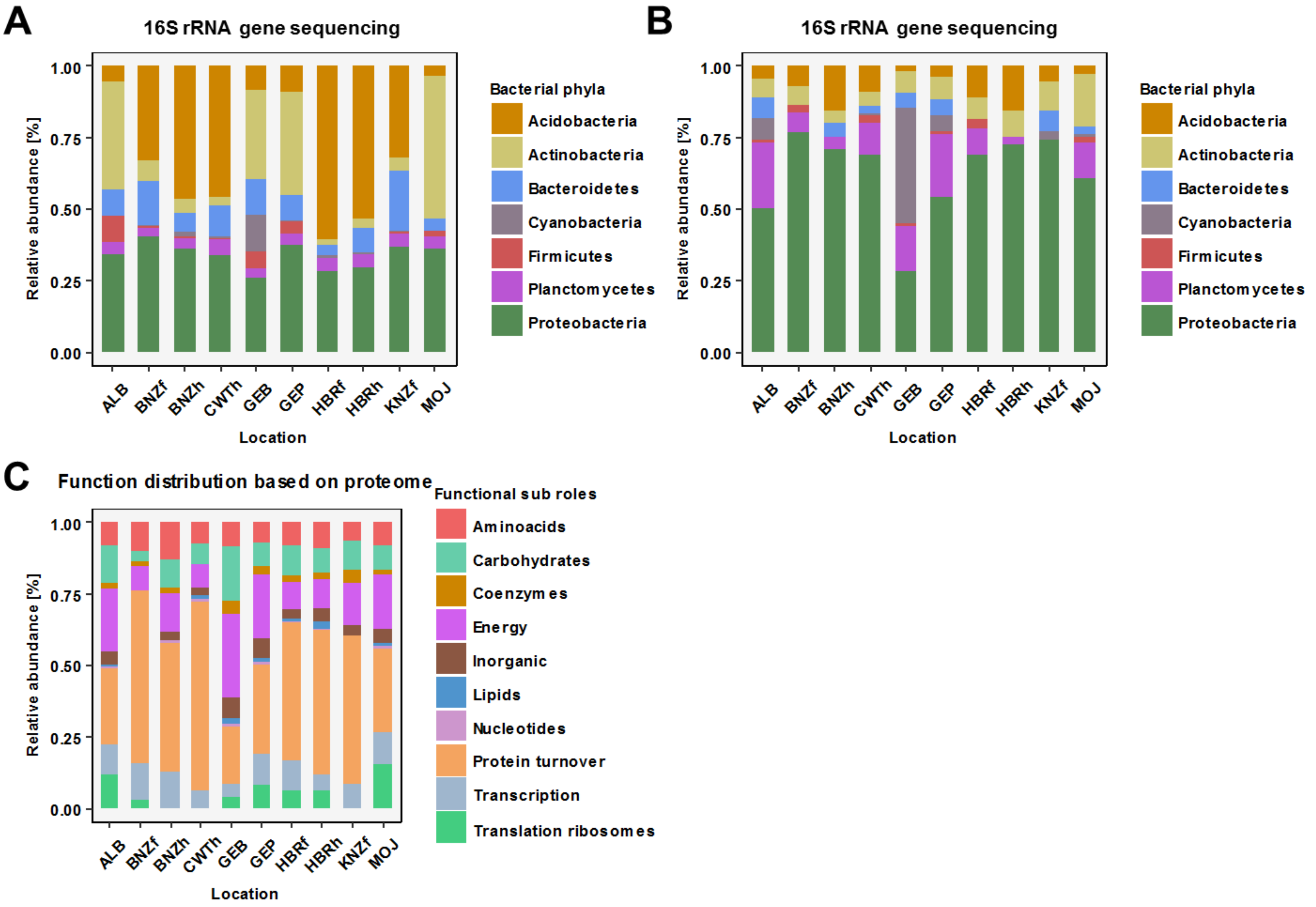

A
16S rRNA gene sequencing
Relative abundance [%]
1.00
0.75
0.50
0.25
0.00
ALB
BNZf
BNZh
CWTh
GEB
GEP
HBRf
HBRh
KNZf
MOJ
Location
Bacterial phyla
Acidobacteria
Actinobacteria
Bacteroidetes
Cyanobacteria
Firmicutes
Planctomycetes
Proteobacteria
B
16S rRNA gene sequencing
Relative abundance [%]
1.00
0.75
0.50
0.25
0.00
ALB
BNZf
BNZh
CWTh
GEB
GEP
HBRf
HBRh
KNZf
MOJ
Location
Bacterial phyla
Acidobacteria
Actinobacteria
Bacteroidetes
Cyanobacteria
Firmicutes
Planctomycetes
Proteobacteria
C
Function distribution based on proteome
Relative abundance [%]
1.00
0.75
0.50
0.25
0.00
ALB
BNZf
BNZh
CWTh
GEB
GEP
HBRf
HBRh
KNZf
MOJ
Location
Functional sub roles
Aminoacids
Carbohydrates
Coenzymes
Energy
Inorganic
Lipids
Nucleotides
Protein turnover
Transcription
Translation ribosomes

Figure 1