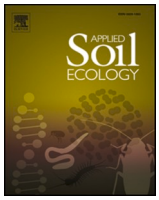
Applied Soil Ecology 173 (2022) 104388

Contents lists available at ScienceDirect

Applied Soil Ecology

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



Short communication

Complementary protein extraction methods increase the identification of the Park Grass Experiment metaproteome

Gerry A. Quinn [a,c,*], Alyaa Abdelhameed [a,b], Ibrahim M. Banat [c], Daniel Berrar [d], Stefan H. Doerr [e], Ed Dudley [a], Lewis W. Francis [a], Salvatore A. Gazze [a], Ingrid Hallin [f], G. Peter Matthews [f], Martin T. Swain [g], W. Richard Whalley [h], Geertje van Keulen [a]

[a] Institute of Life Science, Medical School, Swansea University, Wales, UK
[b] Department of Biotechnology, University of Diyala, Iraq
[c] Institute of Biomedical Sciences, Ulster University, N. Ireland, UK
[d] Data Science Laboratory, Tokyo Institute of Technology, Japan
[e] Department of Geography, College of Science, Swansea University, Wales, UK
[f] School of Geography, Earth and Environmental Sciences, University of Plymouth, UK
[g] IBERS, Aberystwyth University, Wales, UK
[h] Rothamsted Research, Hertfordshire, UK

ARTICLE INFO

Keywords:
Temperate-grasslands
Soil-Metaproteome
Protein-extraction
Biogeochemical-cycles
Regulation

ABSTRACT

Although the Park Grass Experiment is an important international reference soil for temperate grasslands, it still lacks the direct extraction of its metaproteome. The identification of these proteins can be crucial to our understanding of soil ecology and major biogeochemical processes. However, the extraction of protein from soil is a technically fraught process due to difficulties with co-extraction of humic material and lack of compatible databases to identify proteins. To address these issues, we combined two protein extraction techniques on Park Grass experiment soil, one based on humic acid removal, namely a modified freeze-dry, heat/thaw/phenol/chloroform (HTPC) method and another which co-extracts humic material, namely an established surfactant method. A broad range of proteins were identified by matching the mass spectra of extracted soil proteins against a tailored Park Grass proteome database. These were mainly in the categories of "protein metabolism", "membrane transport", "carbohydrate metabolism", "respiration" "ribosomal and nitrogen cycle" proteins, enabling reconstitution of specific processes in grassland soil. Protein annotation using NCBI and EBI databases inferred that the Park Grass soil is dominated by *Proteobacteria*, *Actinobacteria*, *Acidobacteria* and *Firmicutes* at phylum level and *Bradyrhizobium*, *Rhizobium*, *Acidobacteria*, *Streptomyces* and *Pseudolabrys* at genus level. Further functional enrichment analysis enabled us to connect protein identities to regulatory and signalling networks of key biogeochemical cycles, notably the nitrogen cycle. The newly identified Park Grass metaproteome thus provides a baseline on which future targeted studies of important soil processes and their control can be built.

1. Introduction

Park Grass Experiment (PGE) is an international reference soil with an extensive metadata set that encompasses many soil and temperate grass characteristics including hay yield, soil ecology and nutrient levels (Delmont et al., 2012; Silvertown et al., 2006). Although recent metagenomic studies have led to improved understanding of this soil's ecosystem, its metaproteome has still not been directly extracted (Delmont et al., 2015, 2012, 2011). These soil proteins play an integral role in facilitating many of the Earth's biogeochemical processes, however, their extraction can be problematic due to bonding to clay particles, co-extraction of humic materials and the poor availability of matching protein databases for soil protein identification (Qian and Hettich, 2017).

Indeed, it has taken many years to overcome some of these technical problems. Even now, the numbers of protein identifications per soil generally range from hundreds up to the low thousands (Abiraami et al., 2019). Many physico-chemical methods are used to extract soil proteins;

* Corresponding author at: Institute of Biomedical Sciences, Ulster University, Coleraine, N. Ireland, UK.
E-mail address: g.quinn@ulster.ac.uk (G.A. Quinn).

https://doi.org/10.1016/j.apsoil.2022.104388
Received 27 December 2021; Accepted 10 January 2022
Available online 21 January 2022

however, the most effective techniques are based on phenol/chloroform and surfactant extraction. These techniques are also Complementary since the phenol/chloroform method removes humic acids whilst the surfactant method does not (Bastida et al., 2014; Chourey et al., 2010; Greenfield et al., 2018; Heyer et al., 2019).

Another obstacle in the identification of soil proteins is the accurate matching of their mass spectra to a metagenomic or metaproteomic database. Many databases were originally derived from clinical or laboratory samples making it harder to achieve relevant matches to environmental proteins. However this has improved in recent years with the creation of environmentally oriented databases such as Park Grass (Callister et al., 2018; Delmont et al., 2011; Tringe et al., 2005).

To address these issues in soil protein identification, we have applied two compatible soil protein extraction techniques, namely a surfactant technique and an amended freeze dry heat/thaw/phenol/chloroform (HTPC) method, to extract a broad range of proteins and identify active biogeochemical processes and their control mechanisms in Park Grass reference soil. The results of this study could be used as a baseline for further functional ecological and climate change impact studies in temperate grassland soils.

2. Results and discussion

2.1. Analysis of park grass soil proteins at peptide level

Park Grass Experiment (PGE) soil was randomly sampled from the A horizon in untreated plot 3D, followed by determination of its physicochemical characteristics (Table S1), The extracted metaproteome of Park Grass revealed the identities of a total of 1266 proteins: 715 in the surfactant extract, 635 in the modified HTPC extract, with a shared identity of 84 proteins (Figs. S1, S2, Supplementary data S1a, S1b, S2 and Supplementary information S1a, S1b, S1c) using a tailored annotated soil proteome database (Supplementary information S1a).

The number of these proteins identified compared favourably with previous proteomic studies from a variety of soils, allowing for variation in physicochemical characteristics (Abiraami et al., 2019; Keiblinger et al., 2012). Our results also indicated that each extraction method enriched a different set of proteins, which we examined further by comparing peptide chemistries and (predicted) protein functions for each extraction method.

The differences in these two methods may explain the observed differences in protein identifications. Humic acids are reported to interfere with detection of the proteins by mass spectrometry (Qian and Hettich, 2017), however, their removal runs run the risk of excluding some proteins from identification. Therefore, we specifically selected the surfactant and modified HTPC methods over others tested because they were compatible per se with this soil (yielding quantifiable protein at gel or mass spectrometry level) (Supplementary Fig. S3) and were complementary with respect to removal or inclusion of humic acids.

Most importantly, our use of a compatible soil database which incorporated Park Grass metagenome was key to successful protein identification (Delmont et al., 2012, 2011; Tartaglia et al., 2020). Interestingly, the soil proteins matched to many soil databases from around the world (Delmont et al., 2012, 2011; Tartaglia et al., 2020; Tringe et al., 2005). However, it is difficult to assess the wider applicability of this database to other soils since our experiments were confined to Park Grass Experiment soil.

2.2. Differences in peptides extracted by each technique

When we compared the peptide chemistries of the two protein extractions we found that the HTPC extracts contained longer peptides than surfactant extracts (Supplementary Fig. S4a). Isoelectric point analysis indicated that the HTPC extract contained more acidic peptides in the pH region 3.66–4.75 and 5.21–6.90, and the surfactant extract contained more peptides at higher pIs in the range of 6.90–9.29 and 9.29–12.0 (Supplementary Fig. S4b and Data S2). Analysis of peptide amino acid "groups" (i.e. small, hydrophobic, hydrophilic, acidic, basic) demonstrated that surfactant-extracted peptides contained more aliphatic (13.1%) and basic amino acids (7.8%) whilst peptides from HTPC extracts contained more aromatic (31.8%) and acidic (18.9%) amino acids (Table S2). It may be that the longer peptide lengths in HTPC extract arose from reduced protein cleavage (by trypsin) caused by a greater number of acidic bases (Siepen et al., 2007) or a lower number of lysine or arginine residues (both observed in HTPC peptides). The more hydrophobic character of HTPC peptides may have been due to the enrichment of hydrophobic proteins in the phenol layer during phenol/chloroform partitioning. These observations suggested that the extraction methods yielded a slightly different proteome based on intrinsic protein properties, which could be exploited in further (targeted) studies.

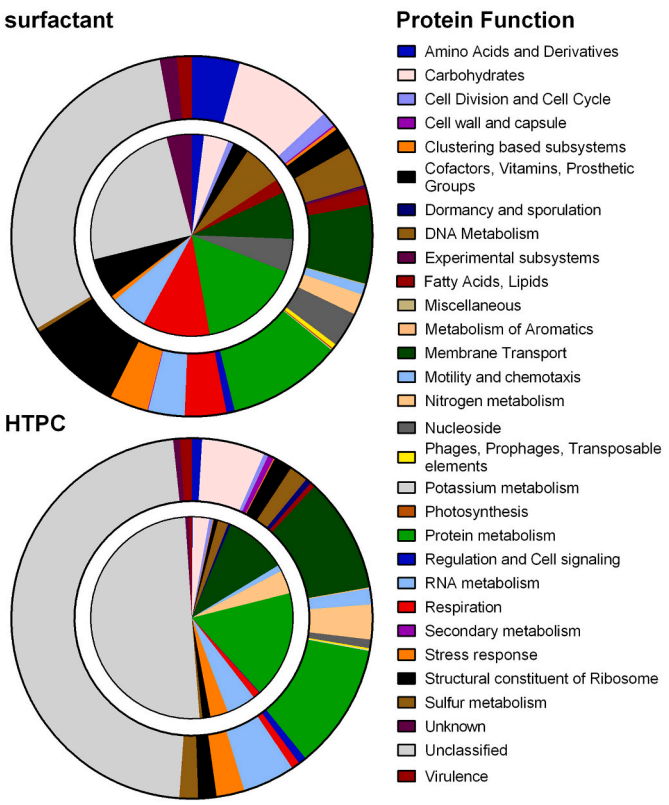


Fig. 1. Functional categories of Park Grass Experiment soil proteins extracted by surfactant and HTPC methods. Proteins quantified by normalised values (NSAF) based on the identification of 715 proteins by surfactant extraction and 635 proteins by HTPC extraction. Segments of the outer rings represent low stringency protein identification; inner segments represent high stringency protein identification. Data based on the homogenised Park Grass soil sample extracted in triplicate by HTPC and surfactant protein extraction methods. Annotation based on level 1 SEED categorisation by MG-RAST.

2.3. Difference in functional protein categories extracted by each technique

The most abundant functional protein categories identified in our soil protein extraction using low stringency soil protein identification criteria were protein metabolism (10.77%), membrane transport (8.71%), carbohydrates (7.23%), structural constituents of ribosomes (5.14%) and RNA metabolism (3.99%) (Fig. 1, Supplementary Table S3 and Data S3a).

High stringency protein analysis revealed the most abundant protein categories were protein metabolism (16.7%), membrane transport (8.77%), respiration (5.96%), RNA metabolism (5.32%), DNA

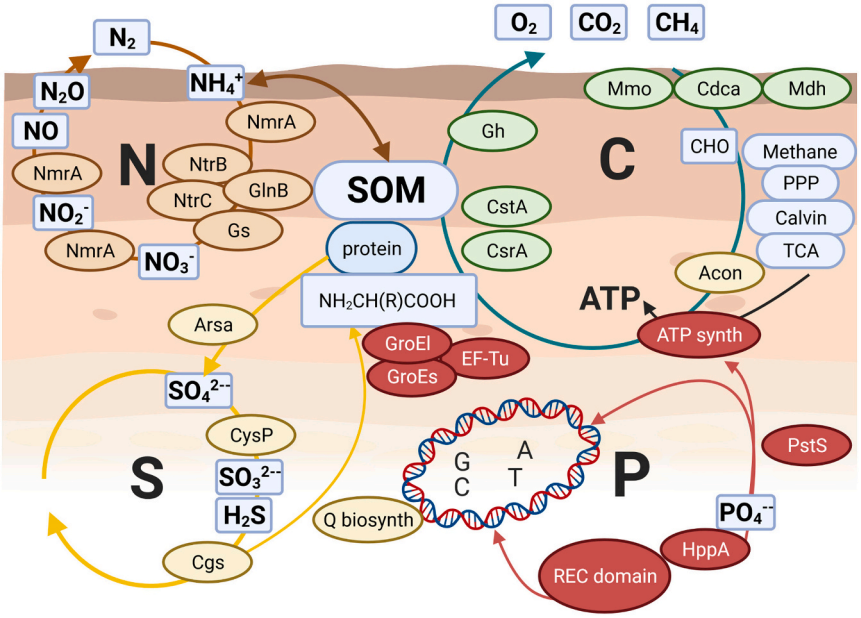


Fig. 2. Snapshot of proteins involved in key biogeochemical cycles in Park Grass Experiment soil.
Abbreviations: methanol dehydrogenase (Mdh), methane monooxygenase (Mmo), carbon monoxide dehydrogenase (Codh), carbonic anhydrase/carbonate dehydratase (Cdca), glycosyl hydrolase/glycosidase (Gh), carbon storage regulator protein (CsrA), carbon starvation protein (CstA), nitrate ABC transporter permease (NtrB), nitrogen regulatory protein I (NtrC), nitrogen regulation protein Pii-1 (GlnB), glutamate synthase (Gs), nitrogen metabolite repression protein A (NmrA), ATP synthase, phosphate ABC transporter substrate-binding protein (PstS), protein chaperonin (GroEL protein), glyceraldehyde-3-phosphate dehydrogenase (Gapdh), elongation factor Tu (Ef-Tu), GTPses, response regulator with CheY-like receiver, AAA-type ATPase, and DNA-binding domain (REC domain), arylsulfatase (Arsa), cystathionine gamma-synthase (Cgs), sulfate ABC transporter periplasmic sulfate-binding protein (CysP), aconitate hydratase (Acon), queuosine biosynthesis protein (Q biosynth).

metabolism (4.06%), structural constituents of ribosomes (3.90%) and carbohydrates (3.30%) (Fig. 1, Table S3, Supplementary Data S3b). Analysis of the differences in soil extraction methods by Bland-Altman plots indicated that the surfactant and HTPC methods identified different groups of proteins (Fig. S5). There were significantly more proteins associated with ribosomal- and respiration-protein categories identified in the surfactant extract in both the low and high stringency protein identifications (Fig. S5 and Table S4). In contrast, the modified HTPC method extracted significantly more "unclassified" proteins (protein whose function was shared in several categories) in both high and low stringency identification, and more proteins associated with nitrogen metabolism in the high stringency identification (Supplementary Fig. S5, Table S4, Data S3c and S3d). Furthermore, protein enrichment analysis of individual proteins using GO ontology revealed that the surfactant protocol extracted more ribosomal- and nucleotide/nucleoside related proteins, while the HTPC extracts were more enriched in nitrogen-metabolism related proteins (Supplementary Table S5, Data S4a, S4b and S4c).

Again, the functional differences in the proteins extracted by these two methods could be related to the chemistry of phenol/chloroform which is known to partition nucleotides, carbohydrates and partially denatured associated proteins into the upper polar layer of the separation process. Since this layer is removed in the separation process to reduce humic acid contamination, it would deplete nucleic acids, carbohydrates, and associated proteinaceous material while enriching for phenol-soluble hydrophobic proteins. In contrast, the surfactant protein extraction method dissolves both hydrophobic and hydrophilic proteins but it might not solubilize very hydrophobic regions that remain tightly folded inside the protein unless acted upon by a strong chaotropic agent like phenol. Similar observations have previously been documented by the group of Bastida who found that different proteomes correlated to different extraction methods (Bastida et al., 2018).

Surprisingly, the modified HTPC extraction method also enriched proteins associated with the nitrogen cycle. There may be several possible explanations for this. Although nitrogen cycle proteins are generally cytoplasmic and soluble, nitrogen-fixing cells have very thick peptidoglycan layers to protect proteins from atmospheric oxygen (and Gram-negative bacteria have additional thick polysaccharide layers). It may be that the modified HTPC method, which includes freeze/thaw steps, is more lytic to these resistive compartments than the surfactant method. An alternative possibility is that the HTPC method may preserve multi-protein conglomerates better than surfactant-based extractions (i.e., less sample loss). These findings may be related to the predominance of the genus *Bradyrhizobia* (which plays an important role in nitrogen fixation) in HTPC extracts identified at a ratio of 8:1 HTPC/surfactant. These results suggest that a preferential extraction method could be adopted for studies focussing on specific aspects of soil ecosystem functions.

Although we assume that many of the proteins identified in the Park Grass soil are from the soil microbiome, they can also be present in the soil as extracellular enzymes or proteins or they could be derived from dead or decaying material. However, the most abundant protein categories identified in our Park Grass metaproteome are similar to those found in actively dividing bacteria namely protein synthesis (including ribosomal proteins), energy metabolism and proteins with binding function (Ishihama et al., 2008).

2.4. Identification of proteins involved in key biogeochemical cycles

We used protein enrichment analysis to create a snapshot of proteins we identified in Park Grass soil that are involved in key biogeochemical processes (Fig. 2, Supplementary Data S5a-S5d).

Surprisingly, we identified both biogeochemical process proteins as well as proteins involved in regulatory and signalling networks. This was exemplified by the identification of nitrate ABC transporter permease (NtrB), nitrogen regulatory protein I (NtrC), nitrogen regulation protein PII -1 (GlnB), glutamate synthase (Gs) and nitrogen metabolite repression protein A (NmrA) in the nitrogen cycle. These results may coincide with a rise in species richness (flora) that was noted in Park Grass which was linked to a decrease in nitrogen input (Blake et al., 1999; Storkey et al., 2015).

We also identified phosphate ABC transporter substrate-binding protein (PstS). This is typically induced upon phosphate limitation which has been reported in Park Grass soil (McDowell et al., 2002).

We also identified sulfur cycle proteins including arylsulfatase (Arsa), cystathionine gamma-synthase (Cgs), sulfate ABC transporter periplasmic sulfate-binding proteins (Sbp and CysP). A very sharp local and national decline in the deposition of sulfur from the atmosphere since the 1980s has resulted in a decrease in soluble sulfate in the surface soil at Park Grass (Blake et al., 1999). As Sbp and CysP expression is repressed in the presence of external sulfur sources (Aguilar-Barajas et al., 2011) these findings indicate low sulfate and/or cysteine (bio)

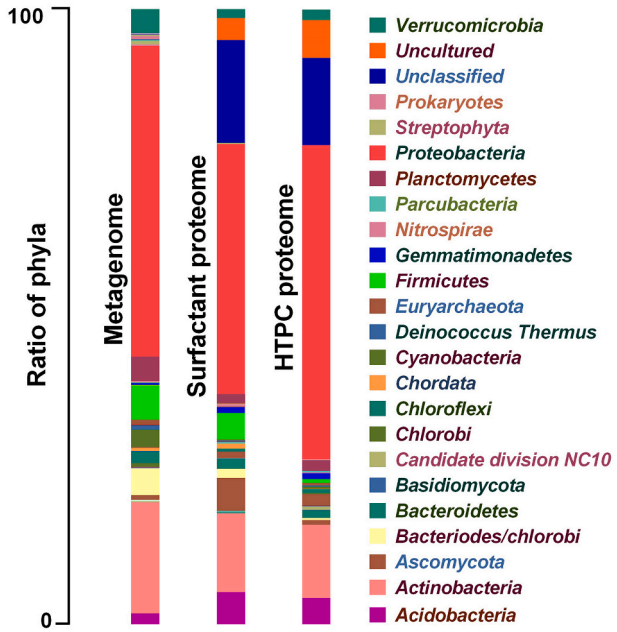


Fig. 3. Community structure of Park Grass inferred from metaproteomes compared to a previous metagenomic study. Soil proteins annotated with phylum and genus information from NCBI and EBI. Phyla ratios derived from 976,268 sequences from the original soil metagenome (metasoil F1, Feb 2009) and 1488 normalised weighted spectra for both the HTPC and surfactant methods. Only the most abundant phyla are displayed. The proteomic data is based on the homogenised Park Grass soil sample extracted in triplicate by the HTPC and surfactant protein extraction methods.

availability in Park Grass soil.

2.5. Inferring the community structure from Park Grass metaproteome

The annotation of the Park Grass soil metaproteome was also used to infer its community structure through the NCBI and EBI databases. The proteome-derived community structures were compared at phylum and genus levels to previous community structure based on metagenomic analysis (Delmont et al., 2011) (Fig. 3).

The phyla inferred from the annotation of the metaproteomes were similar to previous metagenomic analysis (Delmont et al., 2012, 2011) being dominated by *Proteobacteria* (46.31%), unclassified bacteria (16.44%), *Actinobacteria* (12.43%), and *Acidobacteria* (4.80%) (Supplementary Table S6a and Supplementary Data S6a-S6c). The HTPC extract displayed a significantly higher diversity ($p = 0.0002$) (Supplementary information Table S6b). This was confirmed by ANOSIM analysis which revealed a low correlation of the outputs between the two protein extraction methods ($R = -0.03$) (Supplementary Information Table S6c). This data suggests that each extraction protocol might have been more effective on different groups of bacteria. Analysis of the microbiome at the genus level revealed that Park Grass soil was dominated by rhizome-associated bacteria *Bradyrhizobium* (7.56%) followed by *Rhizobium* (2.91%), uncultured bacteria (2.88%), *Acidobacteria* (2.53%), *Streptomyces* (2.09%), *Pseudolabrys* (1.84%), *Neorhizobium* (1.53%) and *Candidatus Entotheonella* (1.51%) (Supplementary Fig. S6 and Data S6b).

The differences between our two protein extractions were more apparent at genus level especially for bacteria associated with the rhizosphere. *Bradyrhizobia* (a slow growing bacteria), was by far the most abundant genus identified primarily in HTPC extracts. However, *Rhizobium* (a faster growing bacteria), was primarily found in surfactant extracts. Other notable differences in soil microbiome extraction ratios were *Polaromonas* which was in greater abundance in surfactant extracts and *Methylobacterium* which was identified in greater abundance in HTPC extracts. Interestingly, *Bradyrhizobia* are important in a wide range of biogeochemical functions especially nitrogen fixation.

3. Conclusions

This soil metaproteomic study of Park Grass Experiment soil combined a modified phenol/chloroform method and an established surfactant method together with a compatible database to identify a broad range of proteins that might not have been apparent using a single extraction method. Further protein enrichment analyses linked active biochemical pathway players with signalling and regulatory networks of key biogeochemical cycles, which in turn, corresponded to nutrient availability in Park Grass soil, directly connecting our metaproteome to the Park Grass metadata repository. This metaproteome study may thus provide a basis for future targeted studies of soil ecosystem functions and their control processes.

Supplementary information is available for this paper. Methods for this manuscript can be found in supplementary methods, additional data in supplementary data and supplementary information.

All other data supporting the findings of this study are available within the article and its supplementary information files. All supplementary materials are publicly available at https://doi.org/10.1016/j.apsoil.2022.104388.

Declaration of competing interest

We have no conflict of interest to declare.

Acknowledgements

We thank Swansea University staff including Alun Davies, Kathryn Sinclair, Penny Diffley and Dr. Alex Griffiths-Harold, staff at Biological Mass Spectrometry Facility, Manchester University including Dr. David Knight and Dr. Julian Selley, and staff at The Ruđer Bošković Institute, Croatia, Dr. Dušica Vujaklija, Hrvoje Dagelić and Mario Strelar.

Ethics approval and consent

The soil sampling and proteomic extraction permissions were obtained in writing from Rothamsted Research (agreement ID 13231MTA272-3).

Funding

This work was supported by the Natural Environment Research Council [grant number NE/K004638/1 for G van Keulen and NE/K004212/1 for G. Peter Matthews].

Author contributions

Soil collection by RA, GVK, GQ and technical staff. Soil preparation by GQ in consultation with SHD, PM, GVK, IH, AG, LF, AA. Basic soil tests by GQ and Forestry commission. Soil proteomics by GQ and ED. Bioinformatics by ED, MS, GQ, AG, AA and DB. Coding by GQ and MS. Mass spectrometry by Manchester University Biological Mass Spectrometry Core Facility. Statistical consulting by DB. Manuscript writing, revision, editing and suggestions by all authors.

References

Abiraami, T.V., Singh, S., Nain, L., 2019. Soil metaproteomics as a tool for monitoring functional microbial communities: promises and challenges. Rev. Environ. Sci. Biotechnol. https://doi.org/10.1007/s11157-019-09519-8.

Aguilar-Barajas, E., Díaz-Pérez, C., Ramírez-Díaz, M.I., Riveros-Rosas, H., Cervantes, C., 2011. Bacterial transport of sulfate, molybdate, and related oxyanions. BioMetals 24, 687–707. https://doi.org/10.1007/s10534-011-9421-x.

Bastida, F., Hernandez, T., Garcia, C., 2014. Metaproteomics of soils from semiarid environment: functional and phylogenetic information obtained with different

protein extraction methods. J. Proteome 101, 31–42. https://doi.org/10.1016/j.jprot.2014.02.006.

Bastida, F., Jehmlich, N., Torres, I.F., Garcia, C., 2018. The extracellular metaproteome of soils under semiarid climate: a methodological comparison of extraction buffers. Sci. Total Environ. 619–620, 707–711. https://doi.org/10.1016/j.scitotenv.2017.11.134.

Blake, L., Goulding, K.W.T., Mott, C.J.B., Johnston, A.E., 1999. Changes in soil chemistry accompanying acidification over more than 100 years under woodland and grass at Rothamsted Experimental Station, UK. Eur. J. Soil Sci. 50, 401–412. https://doi.org/10.1046/j.1365-2389.1999.00253.x.

Callister, S.J., Fillmore, T.L., Nicora, C.D., Shaw, J.B., Purvine, S.O., Orton, D.J., White, R.A., Moore, R.J., Burnet, M.C., Nakayasu, E.S., Payne, S.H., Jansson, J.K., Paša-Tolić, L., 2018. Addressing the challenge of soil metaproteome complexity by improving metaproteome depth of coverage through two-dimensional liquid chromatography. Soil Biol. Biochem. 125, 290–299. https://doi.org/10.1016/j.soilbio.2018.07.018.

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. https://doi.org/10.1021/pr100787q.

Delmont, T.O., Robe, P., Cecillon, S., Clark, I.M., Constancias, F., Simonet, P., Hirsch, P. R., Vogel, T.M., 2011. Accessing the soil metagenome for studies of microbial diversity. Appl. Environ. Microbiol. 77, 1315–1324. https://doi.org/10.1128/AEM.01526-10.

Delmont, T.O., Prestat, E., Keegan, K.P., Faubladier, M., Robe, P., Clark, I.M., Pelletier, E., Hirsch, P.R., Meyer, F., Gilbert, J.A., Le Paslier, D., Simonet, P., Vogel, T.M., 2012. Structure, fluctuation and magnitude of a natural grassland soil metagenome. ISME J. 6, 1677–1687. https://doi.org/10.1038/ismej.2011.197.

Delmont, T.O., Eren, A.M., Maccario, L., Prestat, E., Esen, Ö.C., Pelletier, E., Le Paslier, D., Simonet, P., Vogel, T.M., 2015. Reconstructing rare soil microbial genomes using in situ enrichments and metagenomics. Front. Microbiol. 6 https://doi.org/10.3389/fmicb.2015.00358, 358-358.

Greenfield, L.M., Hill, P.W., Paterson, E., Baggs, E.M., Jones, D.L., 2018. Methodological bias associated with soluble protein recovery from soil. Sci. Rep. 8, 11186. https://doi.org/10.1038/s41598-018-29559-4.

Heyer, R., Schallert, K., Büdel, A., Zoun, R., Dorl, S., Behne, A., Kohrs, F., Püttker, S., Siewert, C., Muth, T., Saake, G., Reichl, U., Benndorf, D., 2019. A robust and universal metaproteomics workflow for research studies and routine diagnostics within 24 h using phenol extraction, FASP digest, and the MetaProteomeAnalyzer. Front. Microbiol. 10, 1883. https://doi.org/10.3389/fmicb.2019.01883.

Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F.U., Kerner, M.J., Frishman, D., 2008. Protein abundance profiling of the Escherichia coli cytosol. BMC Genomics 9, 102. https://doi.org/10.1186/1471-2164-9-102.

Keiblinger, K.M., Wilhartitz, I.C., Schneider, T., Roschitzki, B., Schmid, E., Eberl, L., Riedel, K., Zechmeister-Boltenstern, S., 2012. Soil metaproteomics - comparative evaluation of protein extraction protocols. Soil Biol. Biochem. 54, 14–24. https://doi.org/10.1016/j.soilbio.2012.05.014.

McDowell, R.W., Brookes, P.C., Maheu, N., Poulton, P.R., Johnston, A.E., Sharpley, A.N., 2002. The effect of soil acidity on potentially mobile phosphorus in a grassland soil. J. Agric. Sci. 139, 27–36. https://doi.org/10.1017/S0021859602002307.

Qian, C., Hettich, R.L., 2017. Optimized extraction method to remove humic acid interferences from soil samples prior to microbial proteome measurements. J. Proteome Res. 16, 2537–2546. https://doi.org/10.1021/acs.jproteome.7b00103.

Siepen, J.A., Keevil, E.-J., Knight, D., Hubbard, S.J., 2007. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. J. Proteome Res. 6, 399–408. https://doi.org/10.1021/pr060507u.

Silvertown, Poulton P., Johnstone, E., Edwards, G., Heard, M., Biss, P.M., 2006. The park grass experiment 1856–2006: its contribution to ecology. J. Ecol. 94, 801–814. https://doi.org/10.1111/j.1365-2745.2006.01145.x.

Storkey, J., Macdonald, A.J., Poulton, P.R., Scott, T., Köhler, I.H., Schnyder, H., Goulding, K.W.T., Crawley, M.J., 2015. Grassland biodiversity bounces back from long-term nitrogen addition. Nature 528, 401–404. https://doi.org/10.1038/nature16444.

Tartaglia, M., Bastida, F., Sciarrillo, R., Guarino, C., 2020. Soil metaproteomics for the study of the relationships between microorganisms and plants: a review of extraction protocols and ecological insights. Int. J. Mol. Sci. 21, 8455. https://doi.org/10.3390/ijms21228455.

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., Podar, M., Short, J.M., Mathur, E.J., Detter, J.C., Bork, P., Hugenholtz, P., Rubin, E. M., 2005. Comparative metagenomics of microbial communities. Science 308, 554–557. https://doi.org/10.1126/science.1107851.