NMR-environomics dataset

Deliverable D1.9

31 May 2022

Cristiana A.V. Torres[1], Marta Catalão[1], Filomena Freitas[1], Maria A.M. Reis[1]
Rafael Costa[2], Rui Oliveira[2]

[1]Biochemical Engineering Group, Chemistry Department, School of Sciences and Technology, NOVA University of Lisbon (FCT NOVA), Associação para a Inovação e Desenvolvimento da FCT NOVA

[2] Systems Biology & Engineering Group, Chemistry Department, School of Sciences and Technology, NOVA University of Lisbon (FCT NOVA), Associação para a Inovação e Desenvolvimento da FCT NOVA

PROMICOM

Harnessing the power of nature through PROductive MIcrobial CONsortia in biotechnology - measure, model, master

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101000733.

Prepared under contract from the European Commission

Grant agreement No. 101000733
EU Horizon 2020 Research and Innovation action

Project acronym:	**PROMICON**
Project full title:	**Harnessing the power of nature through PROductive MIcrobial CONsortia in biotechnology - measure, model, master**
Start of the project:	June 2021
Duration:	48 months
Project coordinator:	Dr. rer. nat. Jens O. Krömer Helmholtz Centre for Environmental Research - UFZ https://www.ufz.de/
Deliverable title:	Collection, pre-screening and characterization of natural microbiomes
Deliverable n°:	D1.9
Nature of the deliverable:	Report
Dissemination level:	Public
WP responsible:	WP1
Lead beneficiary:	NOVAID
Citation:	Torres, C.A.V., Catalão, M., Freitas, F., Reis, M.A.M., Costa, R., Oliveira, R. (2022). *NMR-environomics dataset*. Deliverable D1.9 EU Horizon 2020 PROMICON Project, Grant agreement No 101000733.
Due date of deliverable:	Month nº12
Actual submission date:	Month n°12

Deliverable status:

Version	*Status*	*Date*	*Author(s)*
1.0	Final	31 May 2022	NOVA.id

Table of contents

Preface

The main objectives of WP1 “Learning from nature – enabling technologies” is developing analytical tools and mathematical modelling approaches for the analysis of synthetic (bottom-up approach) and natural microbiomes (top-down approach). A key outcome is the establishment of standardized platforms for obtaining quantitative single cell data and connected coherent OMICS and Meta-OMICS data sets for complex microbiomes. In task 1.2 NOVA.id has developed a protocol for Proton nuclear magnetic resonance ((1)H-NMR) analysis for fast and reliable un-targeted exo-metabolomics in complex natural microbiome samples. This technique, based on the particular proton fingerprint of each molecule, does not require prior knowledge on the molecules present in the sample (untargeted method). It is thus particularly suitable for qualitative and quantitative analysis of complex natural microbiome samples containing a high number of unknown organic compounds. Spectral acquisition of microbiome samples was performed in a Bruker 600 MHz NMR spectrometer (available at NOVA.id facilities) and the respective protocol was optimized. Microbiome samples collected from natural habitats (WP2) were analysed by (1)H-NMR. Samples taken from the selection shake flask assays, resulting in time series data sets displaying how natural microbiomes change/interact with the environment (WP2), were also analysed by (1)H-NMR as a times series sequence. A total of 14 samples were analysed and included in the NMR-environomics data set annexed to this report in digital format. In the future, more samples of the microbiome selection assays will be analysed by (1)H-NMR (task 1.2). The (1)H-NMR data will support reverse metabolic function reconstruction, i.e. meta functional environomics (reverse reconstruction from the side of the envirome rather from the side of the genome) in task 1.3.

List of abbreviations

NMR	Nuclear magnetic resonance spectroscopy
PCA	Principal Component Analysis
COSY	Correlation Spectroscopy
NOSY	Nuclear Overhauser effect spectroscopy
ROESY	Ratating-frame NOSY
VOC	Volatile Organic Compounds
D2O	Deuterated water
(1)H-NMR	Proton nuclear magnetic resonance spectroscopy
1D NMR	1-dimensional NMR
2D NMR	2-dimensional NMR
PHA	Polyhydroxyalkanoates
EPS	Exopolysaccharides
VSS	Volatile Suspended Solids
Ac	Acetate

1 Background

Nuclear magnetic resonance (NMR) spectroscopy is arguably the most powerful tool for the study of molecular structures and interactions, and is increasingly being applied to environmental research, such as the study of wastewater (Maryam *et al.*, 2021). Particularly (1)H-Nuclear magnetic resonance ((1)H-NMR) spectroscopy is a powerful technique to analyse the composition of complex mixtures based on the particular proton fingerprint of each molecule. Among the NMR techniques (1)H-NMR is by far the most frequently adopted for exo-/endo-metabolomics. (1)H-NMR is particularly well established for analysis of biological fluids, cells, and tissue extracts (Duarte *et al.*, 2014). More recently complex wastewater samples containing an array of complex unknown organic compounds is being considered for (1)H-NMR analysis (Maryam et al., 2021). As illustrative example, Shumilina *et al.* (2020) applied 1-dimensional 600 MHz (1)H-NMR to identify and quantify volatile organic compounds in dairy industry waste samples. It was concluded that the NMR, chemical oxygen demand (COD) and total organic carbon (TOC) analyses complement each other providing a more accurate estimation of the organic matter content in wastewater.

The (1)H-NMR is a 1-dimensional (1D) spectra acquisition technique. Higher resolution 2-simensional (2D) (e.g. COSY, TOCSY and ROESY) greatly increases the dispersion of spectra peaks. Therefore, peaks that overlap significantly in the 1D spectrum may be much easier to resolve and assign in a 2D spectrum, and this has led to identification of metabolites that could be missed in 1D spectra (Bell *et al.*, 2014). This increased spectral dispersion is excellent for the study of complex mixtures such as unfractionated wastewater, as it helps overcome overlap by “spreading out” the signals over multiple dimensions even if the spectra acquisition takes a substantial amount of time (Maryam *et al.*, 2021).

While NMR can be used very effectively to pattern match and determine the identities of compounds, it also has the capability of quantifying the components present. For basic (1)H 1D NMR experiments, the major difference between quantitative and qualitative NMR is in the recovery (or relaxation) delay (Maryam *et al.*, 2021). (1)H-NMR spectra analysis has been approached mainly through multivariate statistical methods, such as principal components analysis (PCA), in which the compounds are not identified but their spectral patterns and intensities are compared to highlight relative differences between samples, hence avoiding issues related to spectral assignment and compound identification.

2 Materials and methods

2.1 Sample preparation for (1)H-NMR analysis

Sample preparation for (1)H-NMR analysis can essentially follow the minimal preparation protocol, or the freeze drying protocol. The minimal preparation protocol has typically lower resolution and a higher interference of water. The freeze drying protocol is more efficient at suppressing the water peak but loses some of the VOCs during the drying process. The freeze drying protocol was the one adopted in this report, as follows.

- The supernatant samples (5 mL) were freeze dried (ScanVac CoolSafeTM, LaboGene), at -100 ºC and 0.05 mbar, for 24 hours.
- The dried samples were dissolved in 1 mL of 0.1 M phosphate buffer (pH 7.2) solution. The solution was prepared by dissolving disodium hydrogen phosphate ($Na_2HPO_4 \cdot 12H_2O$) and sodium dihydrogen phosphate ($NaH_2PO_4 \cdot 2H_2O$) in heavy water (D2O)
- Then, 500 µL were placed into a 5 mm NMR tube.

- The NMR tubes were capped and stored at 4 °C until analysis.

2.2 (1)H-NMR spectra acquisition protocol

Spectral acquisition was performed in a Bruker 600 MHz NMR spectrometer (available at NOVA.id facilities). A NOESY-based pulse sequence with water suppression using presaturation was used, with the following parameters: spectral window of 20 ppm, with 64k points corresponding to 3.27 s acquisition time, 4 s relaxation time, 100 ms mixing time, and 25 °C working temperature. 128 scans were collected for each spectrum. For each sample both the 90º pulse and the water chemical shift center were determined. Shimming and tuning of the samples were performed automatically. The acquisition parameters were the same for all samples to obtain comparable results, including the receiver gain.

3 Results and discussion

3.1 (1)H-NMR analysis of natural microbiome samples

Samples produced in WP2 “Natural Consortia for Learning and Production” in Task 2.1 were selected for NMR analysis. NOVA.id has collected seven environmental samples originating from two different geographical areas, and encompassing distinct habitats, namely, forest soil, river sediments and plant roots (Table 1). The 7 microbiomes were characterized for their phenotypic and genotypic potential for the production of PHA and EPS and results presented in deliverable D2.1 “Collection, pre-screening and characterization of natural microbiomes”.

Table 1: Natural microbiomes samples: origin and collection date, and the pH and conductivity values of the aqueous suspensions of the processed samples.

Type of sample	Sample #	Sample origin	Collection date	pH	Conductivity (µS/cm)
Forest soil	**FF-01**	Mata do Buçaco (Aveiro)	13/11/2021	4.02	17.04
	FF-02	Mata dos Medos (Caparica, Setúbal)	05/11/2021	6.41	51.26
Sediment	**FF-03**	Tagus River Mouth (Porto Brandão, Setúbal)	19/10/2021	7.55	45.31
	FF-04	Tagus Estuarine Marshland (Corroios, Setúbal)	05/11/2021	7.52	36.63
Plant roots	***FF-05***	*Carpobrotus edulis* (Beach dune plant, Caparica)	05/11/2021	6.89	85.69
	FF-06	*Dicksonia antartica* (Fern, Mata do Buçaco, Aveiro)	13/11/2021	6.11	88.17
	FF-07	*Phaseolus lunatus* (Lima Bean, home vegetable garden, Aveiro)	13/11/2021	7.07	51.17

The microbiomes originating from Aveiro region (Mata do Buçaco soil, fern and lima bean roost) had lower mineral elements' contents (<100 ppm), with Na and K being the most relevant elements (53-65 ppm and 12-13 ppm, respectively). The same elements were found as the most relevant also in the soil and plant root samples collected at Caparica region (Mata dos Medos soil and beach dune plant roots). Significantly higher mineral elements' contents (>8000 ppm) were found for the river sediments samples (Tagus River Mouth and Tagus Estuarine Marshland sediments), with high contents of Na (6322-6377 ppm), Mg (1436-1525 ppm), K (399-869 ppm) and Ca (303-529 ppm). Traces of VFAs (lactate, acetic and isobutyric acid) and sugars (glucose, fructose, xylose and arabinose) were detected in the solid-free supernatant of the some of the samples.

The most abundant phyla in all the collected microbiomes were *Proteobacteria* (37.50 ± 12.75%), *Actinobacteria* (34.08 ± 15.47%), *Firmicutes* (10.57 ± 8.48%), and *Bacteroidetes* (3.55 ± 3.88%). The microbiomes originating from soil samples, despite their different geographical location (Caparica and Aveiro), had similar initial microbial composition, including members of the phyla *Acidobacteria* and *Verrucomicrobia*, for which no significant content was found in most of the other environmental samples, except for the plant roots of the fern and of lima bean.

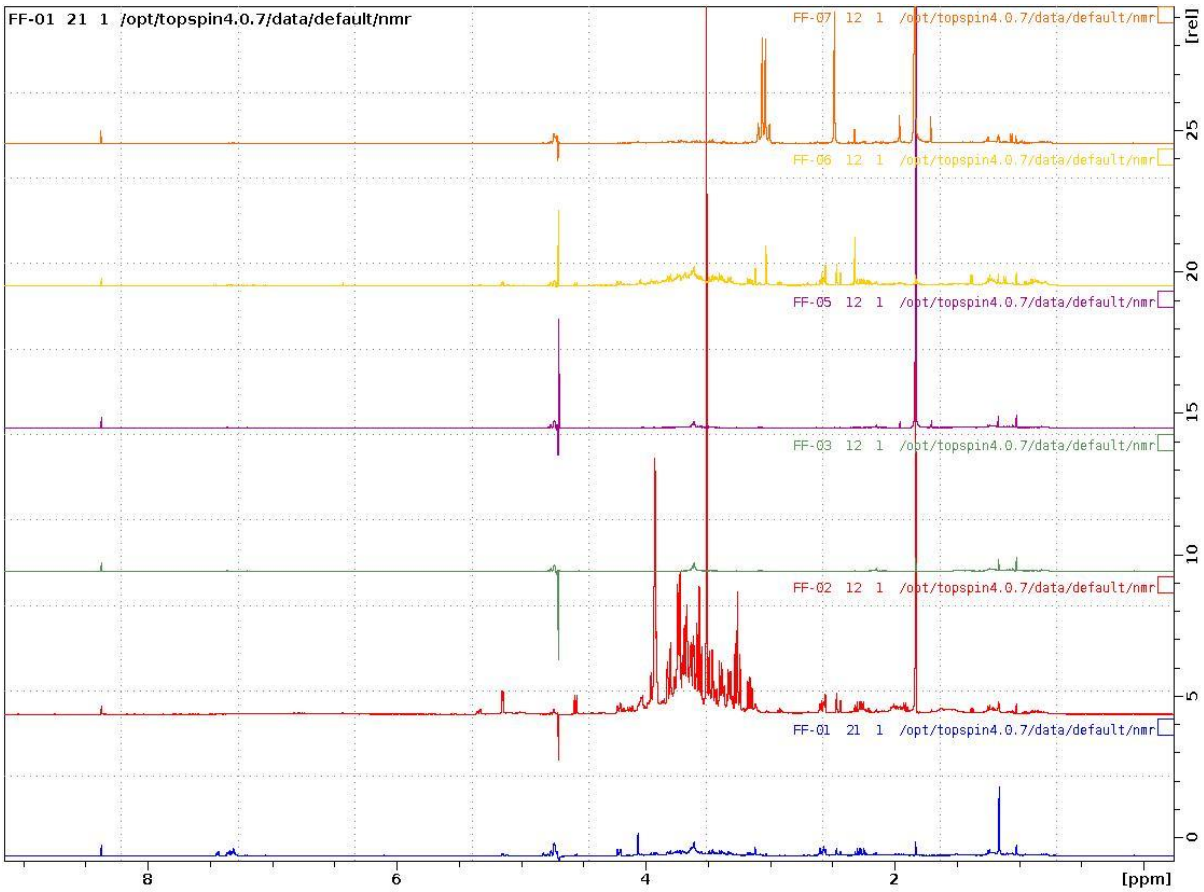

Figure 1. (1)H-NMR spectra of natural microbiome samples FF-01/FF-07 of Table 1. The different colours represent a different microbiome. Sample FF-04 could not be analysed due to the formation of a precipitate. Y-axis represents relative intensity. X-axis represents chemical shift (ppm).

The analysis of the natural microbiome samples by (1)H-NMR is shown in Figure 1. It is visually clear that the different microbiomes exhibit dissimilar spectra. Sample FF-03 (Sediment of Tagus River Mouth, Porto Brandão, Setúbal) appears to have a much lower organic compounds content in relation to the other samples. The sample FF-02 (Forest soil, Mata dos Medos, Caparica, Setúbal) shows a higher fragmentation suggesting a higher diversity of organic molecules in relation to the other samples. Also the intensity of peaks in the same region vary considerably from sample to sample. Sample FF-01 (Forest soil, Mata do Buçaco, Aveiro) despite having a similar microbial composition to sample FF-02 (Forest soil, Mata dos Medos, Caparica, Setúbal) displays aromatic peaks (around 7.4 ppm) not found in FF-02. Also the spectral region between 3.0-4.5 ppm in FF-02 denotes the presence of a higher concentration of carbohydrates in FF-02. The presence of carbohydrates can also be inferred in sample FF-06. Qualitatively the use of 1H-NMR profile is promising to characterize the microbiomes but further experiments will have to be performed to identify individual metabolites.

3.2 (1)H-NMR analysis of time series samples of microbiome selection assays

The natural microbiomes of Table 1 were subjected to a selective pressure, aiming to select for PHA-accumulating microorganisms as part of as part of deliverable D2.1 (WP2 “Natural Consortia for Learning and Production”, Task 2.1). The microbiomes obtained during the 10 days selection assays, were characterized for their phenotypic and genotypic potential for the production of PHA and EPS were also selected for (1)H-NMR analysis. Particularly, the 10 days selection assays for the Marshland sediments and the fern root microbiomes, were analysed by (1)H-NMR because these were the microbiomes that acquired higher PHA accumulation capacity. Some key properties of these samples are listed in Table 2.

Table 2: Main results obtained in the shake flask microbiome selection assays of samples listed in Table 1(*): TSS (g/L), VSS (g/L), overall acetate uptake (g/L), yield of VSS on a substrate basis $Y_{VSS/Ac}$ (g_{VSS}/g_{Ac}), specific acetate uptake rate for each feast-and-famine cycle ($g_{Ac}/(g_{VSS}.d)$); PHA_{max} (mg_{PHA}/g_{VSS}).

Microbiome	**TSS (g/L)**	**VSS (g/L)**	**Overall Ac uptake (g/L)**	**$Y_{VSS/Ac}$ (g_{VSS}/g_{Ac})**	**Specific Ac uptake rate (gAc/(g_{VSS}.d))**				**PHA_{max} (mg_{PHA}/g_{VSS})**
					1st cycle	**2nd cycle**	**3rd cycle**	**4th cycle**	*(day)*
Marshland sediments	39.05	12.19	62.30	0.20	0.228	0.620	0.616	0.446	46.7 *(5)*
Fern roots	29.99	17.66	43.43	0.41	0.000	0.424	0.514	0.275	11.2 *(3)*

(*)These data were generated in task 2.1 and is part of deliverable 2.1

The acetate and nitrogen consumption rates varied along the 4 cycles resulting in different final concentrations at the end of each cycle (Table 3). Samples of both Marshland sediments and Fern root microbiomes had always nitrogen present (e.g. nitrogen never depleted). On the other hand, the content of acetate in each sample was decreasing along the selection assay (for both microbiomes). The last two samples at the end of the end of the 4th cycle had none or almost no acetate. Such decrease in the acetate concentration towards then of the selection assay resulted from the microbiomes acclimatization to the experimental conditions. The observed acetate depletion in samples FF-10 and FF-11 led to the consumption of the previously stored PHA as alternative carbon source for energy and maintenance (Table 3).

Table 3: Nitrogen (N), Acetate(Ac), Volatile Suspended Solids (VSS) and Polyhydroxyalkanoates (PHA) concentrations at the end of each cycle of microbiome selection assays of Table 2

	Sample	t (d)	N (g/L)	Acetate (g/L)	VSS (g/L)	PHA (mg_{PHA}/g_{VSS})
Marshland sediments	**FF-08**	3	0.34	9.74	9.92	31.07
	FF-09	5	0.09	2.24	17.31	46.70
	FF-10	7	0.18	0.00	13.67	9.03
	FF-11	10	0.33	0.00	12.85	3.86
Fern root	**FF-12**	3	0.42	18.94	10.95	11.13
	FF-13	5	0.18	2.55	15.19	9.92
	FF-14	7	0.35	0.24	18.87	7.29
	FF-15	10	0.36	0.00	21.55	6.63

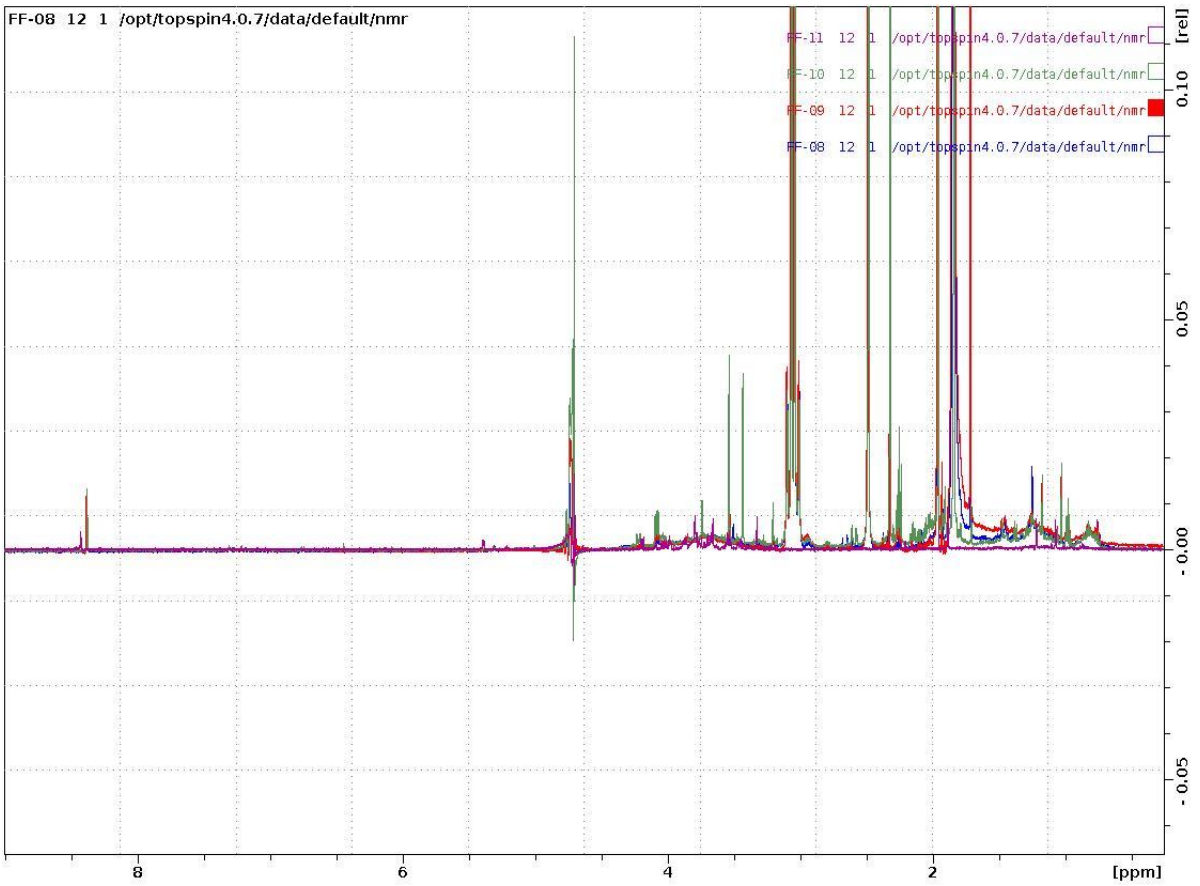

Figure 2. (1)H-NMR spectra of Marshland sediments selection assay (microbiome samples FF-08/FF-11). Sample FF-08 – 3 days selection. Sample FF-09 – 5 days selection. Sample FF-10 – 7 days selection. Sample FF-11 – 10 days selection. Y-axis represents relative intensity (rel). X-axis represents chemical shift (ppm)

The analysis of the Marshland sediments selected microbiomes by (1)H-NMR is shown in Figure 2. The analysis of the Fern root selected microbiomes by (1)H-NMR is shown in Figure 3. In both selection experiments it is visually clear that the spectra change over time. They show however a much higher degree of similarities when compared to the different microbiome sources of Figure 1. A large number of peaks change their intensities over time suggesting that their concentration changes over time, particularly in the methyl region corresponding to the acetate moiety (close to 2 ppm).. But also new peaks of new compounds appear as the selection process progresses.

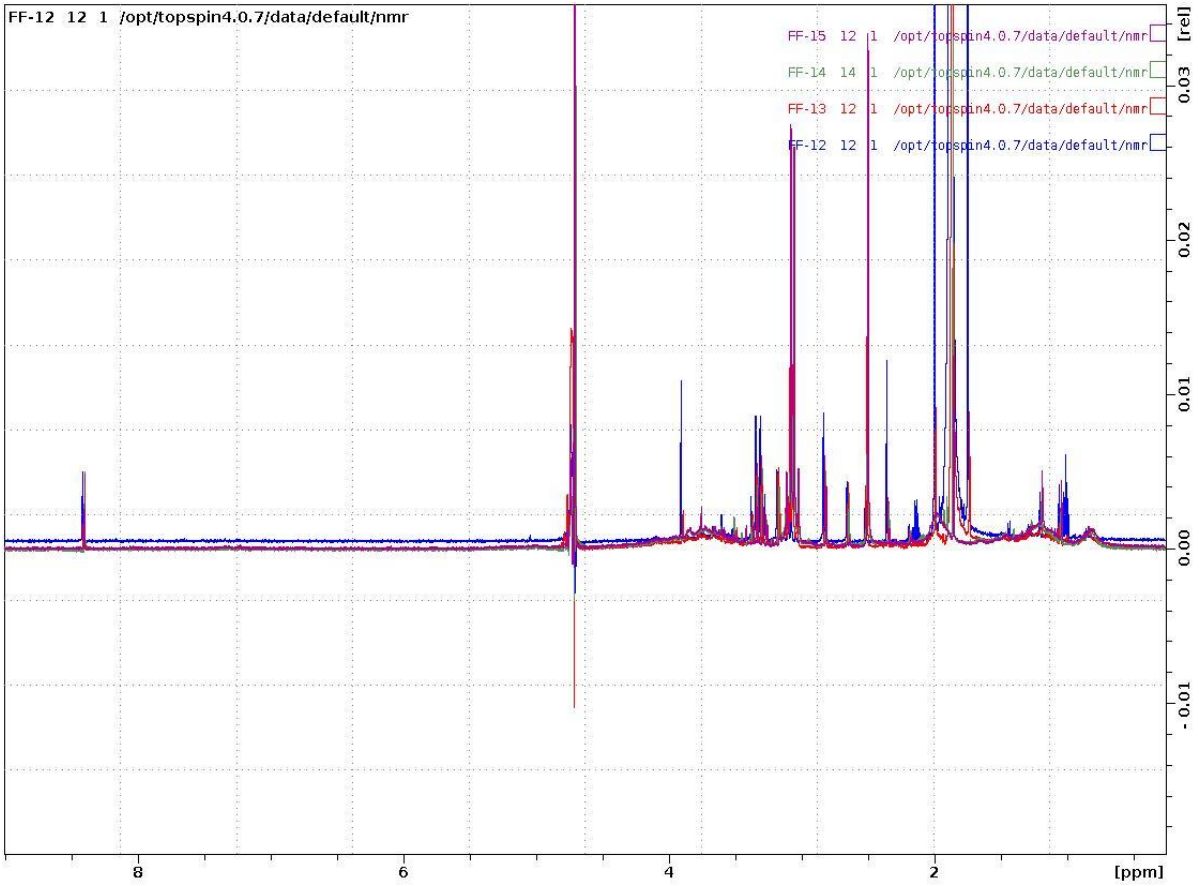

Figure 3. (1)H-NMR spectra of Fern root selection assay (microbiome samples FF-12/FF-15). Sample FF-12 – 3 days selection. Sample FF-13 – 5 days selection. Sample FF-14 – 7 days selection. Sample FF-15 – 10 days selection. Y-axis represents relative intensity (rel). X-axis represents chemical shift (ppm)

4 Concluding remarks

Despite the lower sensitivity when compared to Mass Spectrometry (MS), NMR has several advantages for metabolite profiling: it is nondestructive, requires minimal sample preparation, and provides specific structural information of unknown compounds. The latter is a key advantage for analysis of complex environmental samples with a large array of unknown compounds. Water, which is the largest constituent of microbiome samples, contributes a very large and broad signal to the (1)H-NMR spectrum. Not only can this overlap with and mask

signals of interest, it can also prevent optimization of receiver settings, leading to a lower signal-to-noise ratio, along with baseline roll and other distortions. The freeze drying sample preparation method adopted in this report has revealed itself very effective to suppress the water interference and increase spectra resolution.

Among the many different NMR techniques, (1)H-NMR affords the best sensitivity in comparison to other magnetic nuclei at natural abundance. The obtained results clearly show that the (1)H-NMR spectra provide very distinctive fingerprints of microbiome samples. Spectra of microbiomes originated from different sources are very dissimilar which opens the door for early-stage classification of microbiomes prior to selection. As more microbiomes are selected and analysed, multivariate data analysis such as Principal Component Analysis (PCA), should be applied to classify microbiomes in terms of phenotypic potential (PHA and EPS production capacity) in the end of the selection process. This would allow a high throughput pre-screening of microbiomes before entering the selection pipeline, which is a slow and costly process. The early stage detection of high potential natural microbiomes would be of high value for the top-down approach.

In the future, more samples of the microbiome selection assays will be analysed by (1)H-NMR (task 1.2). The (1)H-NMR data will support reverse metabolic function reconstruction, i.e. meta functional environomics (reverse reconstruction from the side of the envirome rather from the side of the genome) in task 1.3. The selected NOESY-based pulse sequence (4 s acquisition time, 1 s relaxation time, 100 ms mixing time, and 25 °C working temperature) generated spectra will be integrated with the Chenomx NMR Suite 7.1 software package. Reference spectra will be generated for each sample, which will allow to identify a large number of compounds in the samples and to obtain the respective concentrations.

6 References

Duarte T.M., Carinhas N., Silva A.C., Alves P.M., Teixeira A.P. (2014) 1H-NMR Protocol for Exometabolome Analysis of Cultured Mammalian Cells. In: Pörtner R. (eds) Animal Cell Biotechnology. Methods in Molecular Biology (Methods and Protocols), vol 1104. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-62703-733-4_16.

Shumilina, E., Johansen, T., Glasner, C. et al. Application of NMR Spectroscopy and Conventional Analytical Methods for the Assessment of Wastewater from Food Industry. Waste Biomass Valor 11, 1349–1357 (2020). https://doi.org/10.1007/s12649-018-0472-x

Maryam Tabatabaei, Daniel H. Lysak, Katelyn Downey, Flávio Vinicius Crizóstomo Kock, Xiang You, Rudraksha D. Majumdar, Andersson Barison, Luciano Morais Lião, Antonio Gilberto Ferreira, Venita Decker, Benjamin Goerling, Manfred Spraul, Markus Godejohann, Paul A. Helm, Sonya Kleywegt, Karl Jobst, Ronald Soong, Myrna J. Simpson, Andre J. Simpson (2021) NMR spectroscopy of wastewater: A review, case study, and future potential, Progress in Nuclear Magnetic Resonance Spectroscopy, Volumes 126–127, Pages 121-180, https://doi.org/10.1016/j.pnmrs.2021.08.001

N.G.A. Bell, L. Murray, M.C. Graham, D. Uhrín (2014) NMR methodology for complex mixture 'separation' Chem. Commun., 50, pp. 1694-1697, https://doi.org/10.1039/C3CC48907H

7 Supplementary information

The raw (1)H-NMR data generated in this report is available as supplementary file **NMR-envionomics-dataset (PROMICON Delivreable D1.9).zip**