ACTA SCIENTIFIC MICROBIOLOGY (ISSN: 2581-3226)

Volume 3 Issue 3 March 2020

Research Article

Metagenomic Study of the Taxonomic Profile of Rhizobacterial Communities in Soils Contaminated with Mercury

Marina Robas Mora[1]*, Agustín Probanza[2] and Pedro Antonio Jiménez Gómez[1]

[1]Microbiology Area, Pharmaceutical and Health Sciences Department, Faculty of Pharmacy, Universidad San Pablo CEU, Boadilla Del Monte, Madrid, Spain

[2]Plant Physiology Area, Pharmaceutical and Health Sciences Department, Faculty of Pharmacy, Universidad San Pablo CEU, Boadilla Del Monte, Madrid, Spain

***Corresponding Author:** Marina Robas Mora, Microbiology Area, Pharmaceutical and Health Sciences Department, Faculty of Pharmacy, Universidad San Pablo CEU, Boadilla Del Monte, Madrid, Spain.

DOI: 10.31080/ASMI.2020.03.0528

Received: January 18, 2020

Published: February 19, 2020

Abstract

Soil is a complex ecosystem whose homeostasis is affected by the presence of contaminants such as mercury. Knowing the effects of this heavy metal on edaphic microbial communities can help to establish soil quality bioindicators. For the present study, rhizospheric soils have been selected from three indigenous plants of the Almadén mining district (Spain): R. induratus (A), R. bucephalophorus (B) and *Avena sativa* (C), as well as bulk soil samples in plots subject to different mercury concentrations. Rhizospheres contain a huge heterogeneity of scarce microorganisms that are difficult to identify, whose functional and regulator role in their communities, is unknown. Metagenomic techniques allow to know the structure, diversity and abundance of microorganisms that make up these communities, as well as the effects that contaminants may have on their bacterial members, and the natural role played by the native plants that harbor them. The results on how mercury and the plant effect affect the taxonomic composition and microbial diversity, show a reduction in soil heterogeneity in the presence of mercury, and a partial shield of this effect by *Avena sativa* (C) in soils with high mercury concentration. This finding opens expectations to the selection of PGPR strains for future phytoremediation trials.

Keywords: Metagenomics, soil microbiome, BBH, LCA, Shannon index (H´), Simpson index (DSi).

Originality-Significance statement: The originality of this research relies on the fact that there is no scientific evidence of how different edaphic mercury concentrations select certain strains within a complex microbial community. Knowing the effects that this type of heavy metals have on the edaphic communities, as well as the selection that the plants make of the rhizospheric microorganisms opens doors to the selection of plant growth promoting strains for further uses in phytoremediation. Taking into account these results, in studies that are being carried out at present, we have selected 24 bacterial strains, all isolated from the rhizosphere of the plants tested. Metagenomics studies (NGS) are being carried out to know the genetic diversity and to check the existence of interesting genes from a biotechnological point of view, in view of the potential restoration of mining soils affected by this heavy metal.

Introduction

Mercury (Hg) pollution is a serious problem given its environmental repercussions [1] which may ultimately be transferred to the trophic chain, potentially having adverse effects on human health [2-5]. Although some studies have examined certain aspects

regarding biological taxonomic diversity in areas exposed to this contaminant to date, studies have yet to analyze overall taxonomic heterogeneity of such ecosystems as a good method to understand the factors that regulate the organization of the microbial communities. Gaining further knowledge regarding the microbial heterogeneity, will help in the discovery of underlying ecological processes [6].

In order to understand the hierarchy and composition of the microbial communities, it is 69 essential to explain the driving forces behind the selection of abundant and minority microorganisms, especially when the systems receive selective pressure from a contaminant [7].

Nevertheless, microorganisms live as part of complex microbioms. Thus, the main objective of microbial ecology is to understand the structure and function of these communities, as well as the interactions between microorganisms and their environment in diverse niches. The difficulty in cultivating a huge proportion of microorganisms of a microbial community in addition to huge functional redundancy in most soil microbiomes has limited the precise understanding of these.

To overcome this problem, recently, studies have examined the resilience capacity of microbial communities under stress conditions caused by industrial contaminants using metagenomic tools [8]. This requires technologies offering huge data sets which may be difficult to faithfully obtain, analyze and interpret [7]. The current rapid development of molecular (DNA-based) methods that facilitate deciphering microbiomes with respect to key functions will enable the development of improved criteria by which molecular information can be tuned to yield molecular markers of soil life support functions [9].

To delve further into this issue, NGS (Next Generation Sequencing) techniques permit high performance when obtaining sequencing data that reveals the biological heterogeneity beyond the results obtained from cultures [10]. For its part, quantitative PCR (qPCR) monitors the abundance of soil indicator organisms or genes. On top of that, the activity patterns of the respective microbes can be assessed by reverse transcription (RT)-qPCR [9]. Thus, the approach to the metagenomics concept must be oriented towards studying microbial communities so as to understand how they function and how their members interact within their niches.

Currently, many studies compare the populational structures and the phylogenetic diversity obtained from the gene sequencing coding for 16S rRNA gene. Metadata, in an environmental context, of the metagenomics samples, is especially important in comparative analyses, as it allows for the study of the effect of the habitat on the community structure and function. The application of this type of selected macro-ecological concepts, like the "functional response groups" concept for analyzing the soil microbiome will result in identifying groups of soil bacteria that respond similarly to challenges under comparable conditions Unfortunately, there are few studies devoted to this end using soil samples that are contaminated with heavy metals [11].

A key objective is to identify microbial groups that are responsible for specific characteristics of certain ecosystems. In this sense, it may be interesting the selection of bioindicators that report on soil quality. Such indicators should ideally describe organisms with key functions in the system, or with key regulatory/connecting roles (so-called keystone species) [9], in our case, for example, microorganisms with a special capacity to resist or buffer heavy metal contamination. This could become a useful tool for accessing the high biodiversity of the environmental samples, either through the use of a gene marker, such as 16S rRNA gene or through the use of random sequencing [12].

Despite all of this, our understanding of soil continues to be limited in many ways. The imperfect tools to describe microbial communities and limited possibilities to assign traits to community members make it difficult to link microbes to functions [13]. This is the reason why Jansson and Hofmockel [14], consider that the next frontier lies in understanding the metaphenome, this means, the product of the combined genetic potential of the microbiome and available resources.

To delve further into this issue, NGS (Next Generation Sequencing) techniques permit high performance when obtaining sequencing data that reveals the biological heterogeneity beyond the results obtained from cultures [10]. For its part, quantitative PCR (qPCR) monitors the abundance of soil indicator organisms or genes. On top of that, the activity patterns of the respective microbes can be assessed by reverse transcription (RT)-qPCR [9]. Thus, the approach to the metagenomics concept must be oriented towards studying microbial communities so as to understand how they function and how their members interact within their niches.

Currently, many studies compare the populational structures and the phylogenetic diversity obtained from the gene sequencing coding for 16S rRNA gene. Metadata, in an environmental context, of the metagenomics samples, is especially important in comparative analyses, as it allows for the study of the effect of the habitat on the community structure and function. The application of this type of selected macro-ecological concepts, like the "functional response groups" concept for analyzing the soil microbiome will result in identifying groups of soil bacteria that respond similarly to challenges under comparable conditions Unfortunately, there are few studies devoted to this end using soil samples that are contaminated with heavy metals [11].

A key objective is to identify microbial groups that are responsible for specific characteristics of certain ecosystems. In this sense, it may be interesting the selection of bioindicators that report on soil quality. Such indicators should ideally describe organisms with key functions in the system, or with key regulatory/connecting roles (so-called keystone species) [9], in our case, for example, microorganisms with a special capacity to resist or buffer heavy metal contamination. This could become a useful tool for accessing the high biodiversity of the environmental samples, either through the use of a gene marker, such as 16S rRNA gene or through the use of random sequencing [12].

Despite all of this, our understanding of soil continues to be limited in many ways. The imperfect tools to describe microbial communities and limited possibilities to assign traits to community members make it difficult to link microbes to functions [13]. This is the reason why Jansson and Hofmockel [14], consider that the next frontier lies in understanding the metaphenome, this means, the product of the combined genetic potential of the microbiome and available resources.

The concept of specific diversity in the ecology of communities and the diversity indixes continue to be widely used in ecology studies. Specific diversity is an emerging property in biological communities that is related to the variety existing in these communities. This attribute is the expression of two components: number of species present in the community (species richness) and distribution of the abundance between the species making up the community (equitability). Some of the most widely used diversity indices are the Simpson index (DSi) and the Shannon-Weaver index (H').

This work attempts to analyze the heterogeneity and abundance of the best represented taxonomic groups of rhizosphere microbial soil communities of three native species: *Rumex induratus* (Boiss and Reut), *Rumex bucephalophorus* (L.) and *Avena sativa* (L.) in two homologous biotopes differentiated by an unequal concentration of Hg salts. Likewise, it is necessary to determine the effect of the presence of Hg on the composition of said communities and to determine the diversity as assessed via the calculation of indices as widely described in the literature.

Materials and Methods

This study was carried out with samples obtained in the mining district of Almadén, Ciudad Real (Spain). Two experimental plots were used: Plot M, classified as a zone of high Hg concentration, with 1,710 mg/kg total Hg, 0.609 mg/Kg soluble Hg and 7.3 mg/Kg exchangeable Hg. Plot R, corresponds to a plot with a lower Hg concentration, specifically, with 122.4 mg/kg total Hg, < 0.02 mg/kg soluble Hg and 1.20 mg/kg exchangeable Hg [15].

Soil samples were collected from the following plant species: *Rumex induratus* (Boiss and Reut). (A), *Rumex bucephalaphorus* L. (B), *Avena sativa* L. (C), in order to obtain rhizosphere soil. For each sample, the soil from five sub-samples was combined and homogenized in order to create a replica; thus, the total of replicas was three per plant, per sampling zone. As a trial control, a bulk soil sample (SL) was taken from each sampling plot, analyzed in triplicate. In order to obtain soil and plant samples, the procedure from Ruiz-Palomino., *et al.* was followed. Throughout this work, the following terms were used: Plot M samples: A-Hg; B-Hg, C-Hg and SL-Hg; Plot R samples: A-neg; B-neg, C-neg and SL-neg.

To avoid obtaining abnormal data for the physiological variables of microbiological activity due to seasonal situations, the sampling corresponds to the maximums of biological activity for this altitude and latitude, those of this zone of the central Iberian Peninsula (Ruiz-Palomino., *et al.* 2005).

DNA was obtained and purified from the soil samples using a "DNeasy Blood and Tissue kit" (Qiagen©) following manufacturer instructions. Purified DNA was quantified using Picogreen based on 40 pg. The preparation of the libraries was carried out by the Genomics Unit of the Parque Científico de Madrid.

The initial PCR conditions implied an initial phase of denaturalization of the DNA strand carried out at 94 ºC for 5 minutes, then

27 consistent cycles were carried out in three phases per cycle: denaturalization at 94 ºC for 30 seconds, annealing phase at 52 ºC for 30 seconds and an action phase of the polymerase at 72 ºC for 45 seconds. After the 27 PCR cycles, an elongation phase was conducted at 72 ºC for 5 minutes and a final conservation phase at 4 ºC.

For this amplification of 27 cycles of PCR, the polymerase "Q5 Hot Start High-Fidelity"© (New England Biolabs) was used. 100 nM of the following oligonucleotides were added [16] for the amplification of the V3-V4 regions of the 16S rRNA gene: 16S rRNA gene PCR Amplicon Forward Primer = 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG 16S rRNA gene PCR Amplicon Reverse Primer = 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC

The analysis corresponding to the 16S rRNA gene amplicons of the V3 and V4 regions were sequenced using Illumina© (MiSeqrun PE 2x300) technology with a standard pattern for quantification. Quality control was carried out using the Fast QC tool: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Readings were pre-processed in order to obtain quality sequences, thus ensuring that the taxonomic assignment was sufficiently rigorous.

Amplifications of the bacteria detected in the samples were identified based on the BLAST results. The taxonomic assignment was carried out using Best Blast Hit (BBH) and Lowest Common Ancestor (LCA) algorithms. Readings were analyzed following a procedure of advanced methods of parallelization given the possibilities offered by Amazon Web Services (AWS). Illumina© readings were paired readings. The FLASH program was used for the assembly (fusion) of the readings of each pair.

The computer operations were carried out by the company "Era7 Bioinformatics"®, using MG7 method. Readings were assigned a node of the taxonomic tree based on the similarity of the sequence with 16S rRNA genes extracted from the "RNAcentral" database (http://rnacentral.org/) which includes rRNAs from a wide set of important databases such as SILVA, GreenGenes, RDP, RefSeq and ENA (MG7) available at the following URL: https://era7bioinformatics.com/en/page.cfm?id=464and title=microbiomes:-Mg7

In this work, the "OTU" (Operational Taxonomic Unit) was used at a given significance level. The statistical analysis of the data for the creation of heat maps, hieratic ordering of the diversity in clusters graphed in dendrograms, the diversity correlation and the study of partial least squares regression (PLS regression) was carried out using the Metagenassist program [17]. Finally, in order to calculate the alpha diversity, the Simpson [18] and Shannon-Weaver [19] indices were used.

Results

Organizing the results of the relative taxonomic abundance requires the application of mathematical algorithms. In figure 1, the results of the taxonomic distribution accumulated in percentage are presented, applying the LCA algorithm. Thus, it is seen that more than 90% of all of the heterogeneity includes the presence of the following groups: *Proteobacteria, Actinobacteria, Cyanobacteria*, *Acidobacteria* and Bacteroidetes, *Firmicutesand Verrumicrobiota*. Comparing the results of the soils with a greater concentration of Hg with those having lower concentrations, it may be observed that the communities change substantially in terms of composition in these seven best represented taxa.

The results obtained when applying the BBH algorithm are similar (Figure 2). Thus, the fact that both algorithms reveal the same information suggests the high quality of the processed sample, given that the number of well assigned readings and the total number of readings are quite similar.

Figure 1 and 2: Taxonomic distribution accumulated in percentage, applying the LCA (Figure 1) and BBH (Figure 2) algorithms.

Continuing with the application of the LCA algorithm, we then grouped the soils considering only the thirty taxa at the level of the best represented "order" or "family" for all soils by direct taxonomic assignment (Figure 3). The presence of members of the *Pseudomomas* genre in B- Hg is relevant, given that it accounted for more than 25% of the abundance of the sample. It is also noteworthy that there was a large abundance of the *Rhizobiales* order (approximately 5%). Similarly, there was a noteworthy presence of *Microcystis spp.* in the SL-Hg soil, which although found in all of the samples (5% A-Hg, 3% B-Hg and 4% C-Hg), made up over 55% of the free soil with a high concentration of Hg Applying the BBH algorithm, we find that the taxonomic heterogeneity appears to decrease in all cases as the concentration of Hg in the soil increases. So, in Figure 4, we see that the plant effect is not capable of correcting this factor. Thus, while A-Hg and C-Hg samples reveal a certain parallelism, SL-Hg sample is quite different based on the oversupply of *Microcystis* spp. and B-Hg has a high abundance of members of the *Pseudomonas* genre (28%) (Figure 4).

Figure 3 and 4: Distribution of the thirty best represented taxa by direct taxonomic assignment, applying LCA (Figure 3) and BBH (Figure 4) algorithm.

At an organizational level, several groupings may be established with regards to inter-species compatibility. Therefore, the results of the readings and the heterogeneity of all samples were analyzed collectively, attempting to analyze the correlation between taxa. Said analyses have been included in graphic form in figure 5.

Figure 5: Heatmap comparing samples heterogenity with the taxonomical species level identification.

Analyzing this information, we see the segregation of three large clusters, which may be simultaneously divided into six "sub-clusters". In the first of these, *Enterobacteriales* and *Pseudomonadales* were collectively segregated between other groups. This is also the case in the B-Hg sample of Figure 3.

On the other hand, in the next cluster, *Actinomycetales* are separated from the *Clostridiales*, represented in the following "sub-cluster". The same results are seen in a homologous manner in Figures 1-2 in which the totality of the samples was analyzed, highlighting that these groups do not appear in the same samples. Similarly, there is a high degree of compatibility between the *Acidobacteriales* and *Sphingomonadales*, which appear homologous in the sample of plant "C", as graphed in Figure 4.

The Partial Least Squares Regression (PLS regression) was found within the area of the principal components regression. The components obtaining the highest values best explained the model. In all analyses in which component 1 participated (19.2% of the accumulated variance) segregation was observed between the soils having greater concentrations of Hg, as compared to those with lower concentrations. This is seen in the projection in which components 1 and 2 participate (Figure 6).

Also, the analysis resulting from the combination of components 1, 2 and 3 via PLS regression, considering the participation of the three components, make up 47.5% of the variance. Once again, an effective segregation of the soils is observed based on the contaminant effect (Figure 7).

Figure 6 and 7: 2D (Figure 6) and 3D (Figure 7) PCA resulting from the projection of components 1, 2, 3 by PLS "Partial least squares regression".

Finally, in order to analyze the biological diversity of the samples, calculations were made with the two most widely used indices of microbiological ecology for this purpose. This information, which is included in table 1, reveals that except for the soils under the influence of plant "C", the diversity values are clearly higher in soils having a low selective pressure of Hg. Nevertheless, it is seen that the diversity of the soils is high, including the soils under strong Hg selective pressure. Similarly, we find that the most notable abundance is combined in a very reduced number of groups that are quite well represented in all of the samples.

	Shannon-Weaver (H') diversity Index	Simpson (D) diversity Index
A-Hg	4,617	0,021
A-neg	5,151	0,011
B-Hg	3,864	0,095
B-neg	4,838	0,021
C-Hg	5,215	0,018
C-neg	5,015	0,025

Table 1: Diversity indeces for each sample.

Discussion

Soil is the land ecosystem having the greatest biodiversity [20]. Bacterial populations may diversify via the phenomena described as "phase shifting", by which phenotypic and genotypic modifications occur within a population, thereby forming sub- populations. This process has been related to the adaptation to diverse environmental conditions, as well as to harsh changes in ecosystems (Davison., *et al*. 2004) offering the different sub-populations an advantage in order to survive in a greater quantity of habitats and to take on distinct changes [21].

The identification of the 16S rRNA gene as a *"*key stone*"* for the assessment of the microbial heterogeneity and evolution [22] and finally, via the analysis of the sequences of the 5S and 16S rRNA gene of different environments [23], managed to discover the heterogeneity of microorganisms without the need to culture them. These works revealed that the vast majority of microbial biodiversity had not been detected when the study methods were exclusively based on the culturing of microorganisms. Currently, it is estimated that 95 - 99% of the microbial community present in the environment is not really accessible via traditional culture methods. Analyses of community metagenomics allows for the possibility of approaching a more integral interpretation of how ecosystems respond to environmental stress.

Analyzing the results on the relative abundance of OTUs requires both the estimation of the best blast hits (BBH), as well as inferring the results of grouped readings with the LCA statistic [8].

Janssen [24], analyzing different metagenomics studies, suggested that the most abundant soil groups are *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Verrucomicrobia*, *Baceroidetes* [25] and *Firmicutes* [26], which make up a mean of 92% of the soil libraries. Rocha., *et al*. [27] added *Plantomycetes* to this list. These data coincide with those obtained in our study.

The members of the *Acidobacteria* phylum make up a mean of 20%. Similarly, *Verrucomicrobia* makes up a mean of 7% of the soil bacterial communities [24]. Both *Phyla* possess a very low number of referenced species and practically all of the information known about them comes from basic studies of non-cultured microorganisms. The *Verrumicrobia* taxa, in 2005, had ten well-defined species and *Acidobacteria* had barely three [28]. Thus, our knowledge of

these soil taxa is even more limited [29]. Although there are representatives of these Phyla that have been cultured [30-32], none have been cultured from soil samples.

Similarly, it has been found that in soils from this study, the participation of members of the *Cyanobacteria* taxon were very relevant. Their abundance has been previously noted in aquatic ecosystems but in soils, not in an abundant manner. So, few taxonomic groups make up the most significant abundance of a system and many unknown taxonomic groups of limited representation are not abundant [24].

Considering the results of abundance while comparing soils of greater or less Hg presence, the relevance of the *Pseudomonas* genus is evident in these soil types. Traditionally, the following genre have been considered to be very well represented: *Agrobacterium, Alcaligenes*, *Arthrobacter*, *Bacillus*, *Flavobacterium, Micromonospora, Nocardia*, *Pseudomonas* and *Streptomyces* in the cultivable portion of the soil [33]. However, the study of the 16S rRNA genes in soils has allowed for a more direct census of the soil bacteria, without the limitations inherent in studies based on cultures. This suggests that the members of nine genres represent only 2.5 to 3.2% of the soil bacteria. Regardless, the role of *Pseudomonas spp*. is noteworthy, as it contributes with a participation of approximately 1.6% in soils of a very different nature and it is one of the best represented genera as well as one that widely participates in the functionality of these ecosystems [34]. The strains of the *Pseudomonas* genre are capable of processing, integrating and reacting to a wide variety of changing environmental conditions and they reveal a strong capacity to react to physiochemical and biological signals. Strains that are capable of achieving resistance to heavy metals, organic solvents and detergents have been found, permitting them to exploit a wide range of carbon sources as nutrients, and to colonize environments and niches that are difficult to be colonized by other microorganisms. Therefore, it is not surprising that the bacteria of the Pseudomonas genre are considered to be a paradigm of metabolic versatility and key microorganisms in the recycling of organic material in aerobic compartments of ecosystems, playing an essential role in the improvement and maintenance of environmental quality. This may be an explanation for their participation, even more noteworthy, in soils sampled for this work that are contaminated with Hg in which they make up 9.4%.

Along these lines, it is interesting to note the participation of species of the *Microcystis* genre, belonging to the *Cyanobacteria*. *Microcystis spp* phylum, has preference for eutrophic environments [35]. It is not unusual, therefore, that their presence in soils contaminated with Hg exceeds 16% of the microbial abundance. On the other hand, this may be a potential threat, given that these soils may be shared as a reserve for these species, which, under favorable conditions, may form large superficial blooms. Furthermore, they may produce neurotoxins and hepatotoxins, such as microcystin [36] and cyano-peptolins [37].

Gram negative bacteria and, especially the alpha-proteobacteria, tend to have higher growth rates than the Gram positives, especially in environments with a high concentration of nutrients. Therefore, the evolution of the heterogeneity and abundance of the Gram negative bacteria in a community may be a good indicator [38]. On the other hand, the *Bacillus* and *Clostridium* genera have been considered common members of the soil bacterial community, being null or scarcely detected in this study. The *Bacilli* and *Clostridia* classes of the *Firmicutes* phylum make up some 214 genera, including *Bacillus spp.* and *Clostridium spp.* It is estimated that they contribute to only a mean of 2%.

It is possible that the members of this group are insufficiently represented in the libraries because the cells or spores may be difficult to lyse, not being detected therefore, in the analyses based on the PCR that is the base of the extraction of DNA from soil. Regardless of the methodology chosen, a critical step in the success of the metagenomics is the extraction of DNA [39]. Until evidence of bias, the members of this group should be considered relatively minor components of the soil bacterial communities. This does not contradict with any local or sporadically abundant cases [24].

Many authors postulate that data regarding a bacterial community may be used as a measurable indicator of the environmental health and potential production [40]. However, it is necessary to assess the relationship existing between the bacterial composition of the soil communities based on geographic location the influence of pH [41], ground temperature (Fiere and Jackson, 2006), the impact of human activity as well as the accumulation of heavy metals amongst other factors. Similarly, the C/N ratio may explain the distribution of such relevant taxa as *Proteobacteria*, *Actinobacteria*, *Acidobacteria* and *Firmicutes* It is estimated that this relation, in addition to the availability of phosphorus and aluminum, may explain up to 15% of the abundance of the most relevant taxa of a community As occurs with the members of the *Gaiellaceae* family, whose abundance is correlated with the C/N ratio (Hermans *et al*., 2017)

converting all of these taxa into potential bio-indicators of soil health, as seen in our results, in which a greater abundance of *Gaiellaocculata* was found in soils having a low concentration of Hg.

On the other hand, it is also useful to consider the relative abundance of the *Archaea* group, especially in the *Korarqueota* division which, until now, only includes ribosomal sequences with no cultivable species These results lead us to believe that, although it is estimated that their soil participation is low, [24] the importance of the *Archaea* domain and its biological participation in the microbial regulation of soils with Hg are still to be determined. A high level of non-cultivable bacteria have been found in soils with nothing being known of their functional role [42,43]. It is necessary to continue to research strategies of metatranscriptomics and new methods to further our knowledge of the rhizosphere, especially because the complexity of soil environments, where metagenome analysis include relic DNA extracted from dead and dormant cells [44,45] and DNA that is trapped in biofilms [46]. Even viable cells that are actively growing only regulate gene expression as needed, and not all genes are expressed at any given time.

Similarly, when considering the behavior of a community based on the effect of a contaminant, it may be seen that a temporary decrease in heterogeneity α and a high rate of rotation for heterogeneity β suggest that the determinant processes are the main drivers of the succession of the abundant sub-community. However, the great richness of the accumulative species indicates that the stochastic processes led to the succession of minority and lesser-known sub- communities [7]. The abundant bacteria contribute to the primary functions of the degradation of the contaminants, but the minority bacteria provide a substantial fraction of auxiliary functions, indicating a segregation of functions that may be the result of the heterogeneity. The main forces (that is, stochastic or deterministic processes) leading to the microbial succession may be dependent on the members of the community of low or high abundance in temporary microcosms with contaminants. Therefore, the abundance of minority sub-communities is essential in processes of success of the community structure [7] and it is necessary to develop new techniques to permit a greater knowledge of these still little known fractions. The study of the total real heterogeneity of the microbial community found in an ecosystem, and the description of its metabolic capacities, through the analysis of metagenomics (analyzed based on the understanding of the analyzed environmental changes) permits a better understanding of the roles played by microbiota in the maintenance and balance of an ecosystem [47]. Overall understanding of the ecological relations established between the soil microbiota and its environment is a huge challenge, considering the complex and very diverse nature of this community and the multiple variations of its structure between distinct micro-habitats. Massive analyses of biodiversity situate us at the beginning of this task [48]. Thus, many authors agree that the current challenge is to go beyond predictive understanding of gene function based on the genome/metagenome to understanding of actual functions carried out by the soil microbiome in situ (Jansson and Hofmockel, 2018).

Upon analyzing the information appearing in Figure 5, it may be interesting to examine the great biodiversity of these soils. The capacity of some bacteria to produce substances that permit them to compete against other microbial groups, such as the production of antibiotics, has been well-documented [49]. Likewise, we know mechanisms by which microorganisms are capable of stimulating or inhibiting the genetic expression in bacteria, favoring or hindering their genetic expression to activate their metabolism in processes regulated by "quorum sensing- quorum quenching". These are some of the mechanisms by which specific groups of bacteria are found to be statistically exclusive. So, expanding upon the knowledge of the results, such as the inclusion in the graph of the correlation of abundance between taxonomic groups, may allow us to advance in the identification of mechanisms and molecules regulating the compatibility and conglomerate of bacteria with biotechnological purposes.

The rhizosphere produces a characteristic change in the distribution and activity of the microorganisms associated with the roots, as compared to the bulk soil [50]. This is the so-called "rhizophere effect". The composition of the rhizophere populations depends upon the composition of the radicular exudate, as well as the plant species, root type, plant age, phenological state, as well as the type and history of the soil. As was the case with most of other authors, in our study, we observed a domination of Gram negative bacilli, including the species of the *Pseudomonas* genera [34].

Despite different approaches, similar principles have been found to govern the heterogeneity patterns of natural environments. This is the classic principle of "competitive exclusion" [51,52].

In line with García-Salamanca., *et al*. [53], the taxonomic and functional structures are influenced by biotic and abiotic factors, including physico-chemical characteristics, the soil composition and the presence or absence of plant coverage. These authors suggest that the selective pressure of the roots produce modifications through the contribution of exudates that act to attract or repel different members of the microbial community. Similarly, the biomass of arbon and nitrogen also directly correlates with the pH of the microbial community. The richness of a community is moderately affected by the presence of heavy metals [54], although this impact is not equal in culturable and non-culturable fractions [55].

Studies on heterogeneity based on rRNA have permitted more detailed analyses. Some authors have found a certain "shielding" effect that moderates the impact of contamination on microbial communities [54]. This effect may be asymmetric in different clusters of the microbial soil community [55].

The partial least squares regression (PLS regression) models are especially suitable when the predictors matrix has more variables than observations and when there is multicollinearity between the values of X [56].

Thus, in Figure 6 and Figure 7, it is seen that while the plant effect and the soil composition determine the typology of the community, it is the presence of Hg that establishes the definitive form of the microbial heterogeneity.

The concept of specific diversity in community ecology has been intensely debated by ecologists, with semantic, conceptual and technical issues resulting from its use [57]. Specific diversity is related with the expression of two components. The first of these is the number of species present in the community, and it is referred to as species richness. The second component is the equitability, and it describes how the abundance is distributed between the species making up the community.

The diversity indices include, in one single value, the specific richness and the equitability. The Shannon-Weaver (H') and Simpson (DSi) indices are widely used to assess microbial diversity of microbial communities in metagenomic studies [8].

Despite the fact that the majority of the communities, including microbial ones, are highly diverse [58], it is also well-documented that the presence of a contaminant reduces the biological heterogeneity. So, the values of microbial diversity measured with the Shannon index are significantly reduced due to the effect of the contamination [8] in studies of the resilience of the microbial communities under situations of industrial contamination. In this way, in the presence of a contaminant, many of the taxa living in a microbial community appear to be ecologically equivalent, as they also are in terms of resource requirements [58]. In this work, we have verified that the microbial diversity is inversely proportional to the concentration f Hg in the rhizosphere samples, with no major reductions. Therefore, unlike the postulate of Patel [8], this idea goes against that expected according to natural selection, especially under conditions of strong selective pressure. Thus, we may be facing what has been referred to as "the paradox of diversity" [59]. The diversity of the microorganisms may represent the capacity of the ground to handle disruptions [60-66]. According to this principle, in this work, considering the microbial number and heterogeneity existing in soils, it is understood that this appears to be an environment that is suitable for this type of microorganisms, which are adapted. This is due to the fact that bacteria are very adaptable, both physiologically as well as genetically, in the face of variation of environmental conditions: horizontal and vertical gene transfers, high mutation rate, phase variation, etc. It may be inferred that the contamination effect that takes place in this study zone has a marked chronic nature, as long as it has been produced over the past 2000 years. Similarly, considering the values recorded in the Shannon-Weaver index (H') in soils influenced by the *Avena sativa* (L.) effect, it is seen that the plant clearly favors diversity. Thus, the high density and diversity that characterizes the soil microbiota has allowed for adaptation to changes in the environmental conditions through adjustments in the activity rates, biomass and/or community structure [9].

Simpson's diversity index [18] is one of the parameters that allows us to measure he richness of organisms. It is also referred to, in the literature, as the index of diversity of species or the index of dominance. In ecology, it is also used to quantify the biodiversity of a habitat.

As in the studies by Ladygina and Hedlund (2010) and later, by Golebiewski., *et al*. (2014) our results allow us to verify the shielding effect that took place, leading to an eventual reduction in the diversity of the rhizophere soil communities based on the effect of heavy metal contaminants. Similarly, the diversity suffered a dramatic reduction in the bulk soil samples having a high concentration of Hg.

Conclusions

The analysis of the microbial soil communities through the use of the LCA and BBH algorithms show that the best represented taxa are *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Verrucomicrobia*, *Bacteroidetes* and *Firmicutes.* Dislike other authors that frequently identify *Bacillus* and *Clostridium* genera as well as other Gram positive bacteria in soil samples, in this work it has been found that their abundance is quite reduced. On the other hand, the participation of gram negative bacteria dominates, even in free soils. Of special note is the abundance of *Pseudomonas spp.*

Bulk soils heterogeneity and abundance are influenced by the distinct mercury concentrations found in the experimental plots. High concentration of Hg favors a high prevalence of species of the *Microcystis* genre. The multivariate analysis of the partial least squares regression permits segregation of the soils based on their microbial communities, highlighting the importance of Hg pollution as a conditioning element for the heterogeneity and abundance of its members.

In the rhizospheres, there is a huge heterogeneity of microorganisms that are difficult to identify and quite scarce, whose functional role in their communities, as well as potential regulator role, is unknown. Thus, it may be deduced that they accumulate a very high percentage of genes (eventually even unknown) and may act as a gene reservoir of molecules, with potential biotechnological uses. Therefore, in future studies, it may be useful to analyze their potential as bioindicators of the quality and evolution of the soil in different processes.

The Shannon (H´) and Simpson (DSi) indices reveal very high diversity levels in rhizophere communities. However, the bulk soils suffer from a reduction in diversity when exposed to high concentrations of Hg. This noteworthy finding may be related (in environments altered by the presence of Hg) to the fact that the roots of the studied plants are capable of partially shielding the toxic effects of Hg in the microbial communities. The notably protective function of *Avena sativa* (L.) plant on the soil microbiota suggests the need for future studies on phytorhizo- mediation with mercury-tolerant PGPR bacteria.

Acknowledgements

This work was supported by the research Project FUSP-BS-PPC01-2014, founded by Fundación Universitaria San Pablo CEU and Banco de Santander. This work is also linked to the Project AGL2013-40758-R, founded by Ministerio de Economía y Competitividad, Gobierno de España.

Bibliography

1. Lloret Quirante J., *et al*. "Distrito minero de Almadén. Sostenibilidad del cierre de una actividad milenaria". *Journal of Experimental Botany* 65 (2015): 1399-1409.

2. Martínez A., *et al*. "Sampling high to extremely high Hg concentrations at the Cerco de Almadenejos, Almadén mining district (Spain): The old metallurgical precinct (1794 to 1861 AD) and surrounding areas". *Journal of Geochemical Exploration* 109 (2011): 70-77.

3. Higueras P., *et al*. "Intraplate mafic magmatism, degasification, and deposition of mercury: The giant Almadén mercury deposit (Spain) revisited". *Ore Geology Reviews* 51 (2013): 93-102.

4. Das S., *et al*. "Genetic basis and importance of metal resistant genes in bacteria for bioremediation of contaminated environments with toxic metal pollutants". *Applied Microbiology and Biotechnology* 100 (2016): 2967-2984.

5. Fernández-Naranjo N., *et al*. "Sediment characterization of an artificial reservoir impacted by mining activity in the Almadén mercury district (Spain)". Conference: IX Congreso Geológico de España 16.1 (2016): 137-140.

6. Goberna M., *et al*. "Opposing phylogenetic diversity gradients of plant and soil bacterial communities". *Proceeding Biology Sciences* 283.1825 (2016): 2015-3003.

7. Jiao M., *et al*. "Resilience associated with mental health problems among methadone maintenance treatment patients in Guangzhou, China". *AIDS Care* 29.5 (2017): 660-665.

8. Patel V., *et al*. "Response and resilience of soil microbial communities inhabiting in edible oil stress/contamination from industrial estates". *BMC Microbiology* 22.16 (2016): 50.

9. Schloter M., *et al*. "Microbial indicators for soil quality". *Biology and Fertility of Soils* 54.1 (2018): 1-10.

10. Nacke H., *et al*. "Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories". *FEMS Microbiology Ecology* 78.1 (2011): 188-201.

11. Gołębiewski M., *et al*. "16S rDNA pyrosequencing analysis of bacterial community in heavy metals polluted soils". *Microbial Ecology* 67.3 (2014): 635-647.

12. Guo W., *et al*. "Nonparametric regularized regression for phenotype- associated taxa selection and network construction with metagenomic count data". *Journal of Computational Biology* 23.11 (2016): 877-890.

13. Baldrian P. "The known and the unknown in soil microbial ecology". *FEMS Microbiology Ecology* 95.2 (2019).

14. Jansson JK and Hofmockel KS. "The soil microbiome — from metagenomics tometaphenomics". *Current Opinion in Microbiology* 43 (2018): 162-168.

15. Moreno-Jiménez E., *et al*. "Using Mediterranean shrubs for the phytoremediation of a soil impacted by pyritic wastes in Southern Spain. A field experiment". *Journal of Environmental Management* 92 (2011): 1584-1590.

16. Ashelford KE., *et al*. "New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras". *Applied and Environmental Microbiology* 72.9 (2006): 5734-5741.

17. Arndt D., *et al*. "METAGENassist: A comprehensive web server for comparative metagenomics". *Nucleic Acids Research* 40 (2012): 80-95.

18. Simpson EH. "Measurement of Diversity". *Nature* 163 (1949): 688-688.

19. Shannon CE. "A mathematical theory of communication". *The Bell System Technical Journal* (1948): 623-656.

20. Germida JJ and Siciliano SD. "Taxonomic diversity of bacteria associated with the roots of modern, recent and ancient wheat cultivars". *Biology and Fertility of Soils* 33 (2001): 410-415.

21. Dybvig K. "DNA rearrangements and phenotypic switching in prokaryotes". *Molecular Microbiology* 10.3 (1993): 465-471.

22. Woese CR. "Bacterial evolution". *Clinical Microbiology Reviews* 51.2 (1987): 221-271.

23. Pace NR., *et al*. "The analysis of natural microbial populations by ribosomal RNA sequences". *Advances in Microbial Ecology* 9 (1986): 1-55.

24. Janssen PH. "Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and rRNA genes". *Applied and Environmental Microbiology* 72.3 (2006): 1719-1728.

25. Buckley DH and Schmidt TM. "Diversity and dynamics of microbial communities in soils from agro-ecosystems". *Environmental Microbiology* 5 (2003): 441-452.

26. Felske A., *et al*. "Response of a soil bacterial community to grassland succession as monitored by 16S rRNA levels of the predominant ribotypes". *Applied and Environmental Microbiology* 66 (2000): 3998-4003.

27. Da Rocha UN., *et al*. "Exploration of hitherto-uncultured bacteria from the rhizosphere". *FEMS Microbiology Ecology* 69.3 (2009): 313-328.

28. Garrity GM., *et al*. "Taxonomic outline of the prokaryotes, release 5.0". (2004).

29. Zul D., *et al*. "Effects of plant biomass, plant diversity, and water content on bacterial communities in soil lysimeters: implications for the determinantsof bacterial diversity". *Applied and Environmental Microbiology* 73 (2007): 6916-6929.

30. Janssen PH., *et al*. "Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, ctinobacteria, Proteobacteria, and Verrucomicrobia". *Applied and Environmental Microbiology* 68 (2002): 2391-2396.

31. Sait M., *et al*. "Cultivation of globallydistributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys". *Environmental Microbiology* 4 (2002): 654-666.

32. Davis KER., *et al*. "Effects of growth medium, inoculum size, and incubation time on the culturability and isolation of soil bacteria". *Applied and Environmental Microbiology* 1 (2005): 826-834.

33. Alexander M. "Introduction to soil microbiology". 2nd ed. John Wiley and Sons, New York NY (1977).

34. Berg G and Smalla K. "Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere". *FEMS Microbiology Ecology* 68.1 (2009): 1-13.

35. Xie L., *et al*. "The low TN:TP ratio, a cause or a result of Microcystis blooms?". *Water Research* 37.9 (2003): 2073-2080.

36. Le T. "Mictocystis: Toxic Blue-Green Algae. Office of Environmental Health Hazard Assessment". USA (2009).

37. Gademann K., *et al*. "Multiple toxin production in the cyanobacterium microcystis: isolation of the toxic protease inhibitor cyanopeptolin 1020". *Journal of Natural Products* 73.5 (2010): 980-984.

38. Zelles L., *et al*. "Discrimination of microbial diversity by fatty acid profiles of phospholipids and lipopolysaccharides in differently cultivated soils". *Microbial Population Dynamic Plant and Soil* 170.1 (1992): 115-122.

39. Singh BK., *et al*. "Influence of grass species in soil type on rhizosphere microbial community structure in grassland soils". *Applied Soil Ecology* 36 (2009): 147-155.

40. Lau SCK., *et al*. "Genome sequence of the pink-pigmented marine bacterium Loktanella hongkongensis type strain (UST950701-009PT), a representative of the Roseobacter group". *Standards in Genomic Sciences* 10 (2015): 51.

41. Lauber LC., *et al*. "The influence of soil properties on the structure of bacterial and fungal communities across land-use types". *Soil Biology and Biochemistry* 9.40 (2008): 2407-2415.

42. Zhang HB., *et al*. "Bacterial diversity at different depths in lead-zinc mine tailings as revealed by 16S rRNA gene libraries". *Journal of Microbiology* 45.6 (2007): 479-484.

43. Van Elsas JD., *et al*. "A procedure for the metagenomics exploration of disease-suppressive soils". *Journal of Microbiology Methods* 75.3 (2008): 515-522.

44. Carini P., *et al*. "Relic DNA is abundant in soil and obscures estimates of soil microbial diversity". *Nature Microbiology* 19 (2016):16242.

45. Emerson JB., *et al*. "Schrödinger's microbes: tools for distinguishing the living from the dead in microbial ecosystems". *Microbiome* 5 (2017): 86.

46. Flemming HC and Wingender J. "The biofilm matrix". *Nature Reviews Microbiology* 8 (2010): 623-633.

47. Green ST and Rubin EM. "Metagenomics: DNA sequencing of environmental samples". *Nature Reviews Genetics* 6 (2005): 805-814.

48. Falkowski PG and Colomban V. "Shotgun Sequencing in the Sea: A Blast from the Past?". *Science* 304.5667 (2004): 58-60.

49. Robas Mora M., *et al*. "Effect of the Type of Vitis vinifera Cultivation in the Cenophenoresistome and Metabolic Profiling (CLPP) of Edaphic Bacterial Communities". *Journal of Agriculture, Science and Technology* 7.8 (2018).

50. Haichar FZ., *et al*. "Plant host habitat and root exudates shape soil bacterial community structure". *The ISME Journal* 2.12 (2008): 1221-1230.

51. Gause GF. "The struggle for existence.a classic in mathematical biology and ecology". Georgiĭ Frantsevich) 1910-1986 Baltimore, The Williams and Wilkins company (1934).

52. Hardin G. "The competitive exclusion principle". *Science* 131.3409 (1960): 1292-1297.

53. García-Salamanca A., *et al*. "Bacterial diversity in the rhizosphere of maize and the surrounding carbonate-rich bulk soil". *Microbial Biotechnology* 6.1 (2013): 36-44.

54. Lorenz N., *et al*. "Response of microbial activity and microbial community composition in soils to long-term arsenic and cadmium exposure". *Soil Biology and Biochemistry* 38.6 (2006): 1430-1437.

55. Ellis RJ., *et al*. "Cultivation-dependent and independent approaches for determining bacterial diversity in heavy-metal-contaminated soil". *Applied and Environmental Microbiology* 69 (2003): 3223-3230.

56. Tenenhaus M. "Component-based structural equation modelling". *Total Quality Managementand Business Excellence* 19.7 (2008): 8.

57. Hurlbert SH. "The nonconcept of Species Diversity: A Critique and Alternative Parameters". *Ecology* 52 (1971): 577-586.

58. Rappé MS and Giovannoni SJ. "The uncultured microbial majority". *Annual Review of Microbiology* 57 (2003): 369-394.

59. Tilman D. "Causes, consequences and ethics of biodiversity". *Nature* 405 (2000): 208-211.

60. Bardgett RD. "Causes and consequences of biological diversity in soil". *Zoology* 105.4 (2002): 367-374.

61. Daniel R. "The metagenomics of soil". *Nature Review Microbiology* 3 (2005): 470-478.

62. Lozupone CA and Knight R. "Species divergence and the measurement of microbial diversity". *FEMS Microbiology Reviews* 32.4 (2008): 557-578.

63. Woese CR and Fox GE. "Evolution Phylogenetic structure of the prokaryotic domain: The primary kingdoms". *Proceedings of the National Academy of Sciences of the United States of America* 74.11 (1977): 5088-5090.

64. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

65. https://era7bioinformatics.com/en/page.cfm?id=464 and title=microbiomes:-Mg7

66. http://rnacentral.org/expert-database