



Universiteit  
Leiden  
The Netherlands

## **Assemblage and functioning of bacterial communities in soil and rhizosphere**

Yan, Y.

### **Citation**

Yan, Y. (2016, June 8). *Assemblage and functioning of bacterial communities in soil and rhizosphere*. Retrieved from <https://hdl.handle.net/1887/40026>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/40026>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/40026> holds various files of this Leiden University dissertation.

**Author:** Yan Y.

**Title:** Assemblage and functioning of bacterial communities in soil and rhizosphere

**Issue Date:** 2016-06-08

## Chapter 5

# **Functional genes rather than taxonomical composition of the rhizosphere bacterial community determine plant biomass in *Jacobaea vulgaris***

Yan Yan, Eiko E. Kuramae, Mattias de Hollander, Johannes A. van Veen,  
Peter G.L. Klinkhamer

*Manuscript*

## Abstract

The relationship between plants and their surrounding microbiota belowground is complex and has been the focus of much research. In reality, the functionality of the microorganisms that are involved in plant-microbe interactions is still not well understood. We used 16S rRNA gene sequencing to determine the bacterial composition and a shotgun metagenomics approach to determine the functional traits of the rhizosphere microbiome. We examined the effects of the taxonomical composition and the functional traits of the bacterial community in the rhizosphere on plant biomass production. Both were significantly different among soils inoculated with different dilutions of the original bacterial community. Plant biomass production was on average the lowest and showed the highest variation after inoculation of the undiluted communities. A combination of unsupervised multivariate statistics and partial correlations showed that *Arthrobacter* was the taxonomical group that was most strongly related to plant biomass and that ‘transporters’ genes were the functional genes most strongly related to plant biomass. Both were positively correlated to plant biomass and positively correlated with each other. Specifically, the ‘monosaccharide transporters’ genes significantly positively correlated to plant biomass when all three dilutions samples were taken together, and this group of genes increased significantly upon dilutions in the rhizosphere. The frequency of ‘transporters’ genes was higher in *Arthrobacter* than in other components of the bacterial community. Partial correlation indicated that after taking the frequency of ‘transporters’ genes into account the correlation between *Arthrobacter* and plant biomass was no longer significant while after taking the frequency of *Arthrobacter* into account the correlation between ‘transporters’ genes and plant biomass was still highly significant. Although these results should be considered with caution this seems to suggest that functional genes rather than the taxonomical composition of the bacterial community of the rhizosphere determine plant biomass production.

## Keywords

Rhizosphere metagenome | Functional traits | Unsupervised multivariate analyses | Plant biomass

## 5.1. Introduction

The microbiome of the rhizosphere plays critical roles in the functioning of terrestrial ecosystems (Philippot et al 2013). The rhizomicrobiome drives and responds to the specificity of its environment, including host plant characteristics (Haichar et al 2008, Bulgarelli et al 2012), and factors such as pH, salinity, moisture and the availability of nutrients (Fierer and Jackson 2006, Logue and Lindstrom 2010, Nemergut et al 2010, Brockett et al 2012). The rhizomicrobiome plays a key role in plant development and the productivity of the aboveground vegetation (van der Heijden et al 1998, Wagg et al 2011).

Despite the general acceptance that plant roots select specific microbial species which directly or indirectly influence host plant physiology and development (Mendes et al 2011), the extent to which functional traits linked with the rhizosphere microbiome determine colonization and impact on the host plant remains largely unknown. Therefore, characterization of the functional traits of the rhizosphere microbiome is crucial for understanding the effect of soil-borne microbes on plant development.

Advanced shotgun metagenomics approaches offer promising tools to target the microbial genes related to host plant-microbe interactions and so the associated functions. Current studies using this approach that focus on describing the microbiome of humans or other mammal host revealed that the microbiome composition and functions are determinative for the physiology of the host (Turnbaugh et al 2006, Tremaroli and Backhed 2012). Transcription analyses of bacterial genes in the rhizosphere have mostly been performed on single rhizobacterial strains (Mark et al 2005, Matilla et al 2007, Dennis et al 2010). However, because of the complexity of the rhizomicrobiome and the inability to culture many microorganisms, comprehensive, overall, pictures of the microbial community and its functionality related to its link with host plant productivity in natural ecosystems are scarce (Ofek-Lalzar et al 2014, Bulgarelli et al 2015). Thus, in order to improve our understanding of the mechanisms of plant-microbe interactions, we need to characterize better the fundamental ecological processes that underlie the composition and the functionality of the rhizomicrobiome.

The major aim of this study was to acquire better understanding of the relationship between the rhizosphere microbiome and plant growth both at the

level of the taxonomical composition and at the level of the functional genes of the bacterial community. To establish differences in the microbial communities and differences in plant growth, serial dilutions of a soil suspension were prepared and the obtained inocula were, subsequently, re-inoculated into the original soil previously sterilized by  $\gamma$ -irradiation. After an incubation period, plants were potted in the soil samples. We used *Jacobaea vulgaris*, one of the most common weeds in the Netherlands. We used 16S rRNA gene sequencing to assess the composition of the bacterial community in the rhizosphere and a total DNA shotgun metagenomics approach to assess the potential microbiome functionality. In Chapter 4 we already showed that the selection of rhizosphere microbial communities from soil communities was strongly based on the functional traits of the selected microbes. So, here, we hypothesized that selection on the basis of particular functional traits will also have a strong impact on plant growth. We addressed three basic questions: 1) Is plant growth related to the taxonomical composition of bacterial communities in the rhizosphere? 2) Is plant growth related to the frequency of particular functional genes in the rhizosphere? 3) Is the taxonomical composition related to the frequency of particular functional genes and if so which of the two is most strongly related to plant growth?

## **5.2. Materials and methods**

### *5.2.1. Soil sampling and plant selection*

Thirty liters of soil were collected at a depth of 15 cm from a dune soil in Meijendel, The Netherlands. The soil had a sandy texture, an organic matter content of 9.1%, pH of 7.4 and the ammonium, nitrate and phosphorus content of 30.4 mg/kg, 2.2 mg/kg and 15.2 mg/kg respectively. The soil was sieved and homogenized and stored in 500 g aliquots in plastic bags. One bag of soil was kept separately to prepare the inoculum. The soil was sterilized by  $\gamma$ -irradiation ( $> 25$  kGray, Isotron, Ede, the Netherlands). Sterility was tested by spreading 0.5 g of the soil from the inoculum-bag onto TSA and PDA media. No bacterial and fungal growth was observed on agar plates after 6 days for 6 replicates. A subsample of the fresh soil was used to determine soil moisture (24 h, 105 °C). For the dilution treatments, a 10 % suspension of untreated soil in sterilized water ( $10^{-1}$ ) was sequentially diluted to obtain further dilutions of  $10^{-6}$  and  $10^{-9}$

and these were added to the sterilized soil. The  $10^{-1}$  suspension was considered to be the undiluted treatment.

*Jacobaea vulgaris* was used as study plant species. Seeds were collected in Meijendel (52°9'N, 4°22'E), The Netherlands. One seed was propagated by tissue culture. This genotype showed a strong negative feedback in the inoculated soil compared to growth in sterile soil in a previous study (Joosten et al 2009). Since tissue culture plants are more or less “sterile”, it was reasonable to use this “clean” plant for the experiments. After 8 weeks of incubation of the inoculated soils, at the moment that the regrown microbial communities reached similar abundances (Chapter 2), tissue culture plants were potted in 0.5 L pots containing the incubated soil. Samples were taken from the bulk soil at the moment of planting. Plants were grown randomly distributed in a climate room (relative humidity 70%, light 16h at 20 °C, dark 8h at 20 °C). Sterile demineralized water was given every two days with additions of 10 ml nutrient solution (Steiner 1968) once every two weeks, in order to avoid nutrient limitation to plant growth. After 6 weeks of plant growth, plants were harvested and gently shaken to remove the loosely adhering soil after which rhizosphere soil samples were collected by removing the remnant soil with a fine sterile brush. Soil samples were stored at -20 °C for further analysis. Harvested plant parts (shoots and roots) were freeze-dried at -80 °C for one week until constant weight. The design of the experiment included 3 dilutions, with 6 replicates each and duplicate samples per replicate for both the incubated soil and rhizosphere samples. Given that during plant growth the soil was only isolated by a layer of tin foil from the atmosphere, there is a possibility that this could constitute an unknown source of bacteria. However, we assumed that this would not have a major effect on our results as we know that the bulk soil had a full grown community of over  $10^9$  cells per gram of soil after the 8-week pre-incubation period in closed bags following inoculation of the (un-) diluted suspensions as found in Chapter 2.

### 5.2.2. Amplicon sequence analysis

The raw data was processed using the QIIME v.1.6.0 pipeline (Caporaso et al 2010). Low quality sequences below 150 bp in length or with an average quality score below 25 were removed. After denoising the sequences using

Denoiser 0.91 (Reeder and Knight 2010), and testing for chimeras using USEARCH (Edgar et al 2011), Operational Taxonomic Units (OTUs) were identified using the UCLUST 1.2.21 algorithm (Edgar 2010) with a phylotype defined at the 97% sequence similarity level. The resulting OTUs were aligned against the Ribosomal Database Project database (Cole et al 2009).

### *5.2.3. Metagenomics library preparation for DNA shotgun sequencing*

Shotgun metagenomic analyses were conducted on the soil DNA extracts (according to the supplier's manual (MO BIO Laboratories, Carlsbad, CA, USA) following the Illumina Pair-End Prep kit protocol with sequencing performed using 2×300 bp sequencing run on the Illumina Miseq2000 (Macrogen Inc. Company, South Korea).

Paired end reads were trimmed using Sickle (Joshi and Fass 2011) with a minimum PHRED score of 30 and at least 150 bp in length. Subsequently a co-assembly of all data was made with Spades 3.1.1 (Bankevich et al 2012) at different k-mer length of 31,91,101 and 121. Following the final assembly genes are predicted using Prodigal 2.61 (Hyatt et al 2010) and converted from GFF (General Feature Format) to GTF (General Transfer Format) using cufflinks 2.1.1 (Trapnell et al 2010). Per sample, reads were mapped to contigs using BamM 1.4.1 (Imelfort 2015), which uses BWA 0.7.12-r1039 (Li and Durbin 2009) and samtools 1.2 (Li et al 2009). Next the number of reads per sample mapping to genes was calculated using featureCounts (Liao et al 2014). To annotate the set of genes hmmsearch 3.0 (Finn et al 2015) was used to screen the FOAM (Prestat et al 2014) set of Hidden Markov Models (release 1.0). Scripts provided by FOAM were used to select the best hit to the database. For each gene the best KO hits were added to the count matrix of featureCounts as a single column. Thereafter, the KO column was aggregated using the Python Pandas library (McKinney 2015). Hits to multiple KO terms were split. Finally, for each FOAM level a count matrix was made. The whole analysis has been implemented in a Snakemake workflow (Koster and Rahmann 2012).

#### 5.2.4. Data analysis

Alpha diversity calculations were performed based on the rarefied OTU table to compare the diversity among samples at a given level of sampling effort (Hughes and Hellmann 2005). The OTU table was rarefied to 1,535 reads by “single rarefaction” QIIME script since this number was the lowest number of reads for all samples. Four undiluted  $10^{-1}$  samples were filtered out because of the very low number of reads. The average sequence reads from 3 sterilized controls were used as a baseline that was subtracted from the reads of all samples.

All statistical analyses were conducted using R and the vegan package (Dixon 2003). To assess whether manipulation of the bacterial community could explain changes in total plant biomass, ANOVA (False Discovery Rate-corrected) was determined across dilution groups. Data was transformed to fit normal distributions when needed. Unsupervised Principal Component Analysis (PCA) was applied by PAST (Hammer et al 2001). PCAs were performed to visualize the different dilution effects on both taxonomical profiles and functional traits based on normalized functional data. Each broad functional category could be divided in in a subset of functions based on the FOAM dataset. The weight of each taxonomical unit and each functional trait was assigned on the PC score, respectively. In this way the important species functional traits for the PCA separation were distinguished from all other functions on the basis of PC score. The unsupervised analyses were followed up by correlation analyses of the selected potentially important taxonomical units and functional traits with plant biomass. We then used partial correlations to identify the most important taxonomical units and functional traits that were related to plant biomass. As a last step we used again partial correlations with plant biomass to identify whether the taxonomical composition or functional traits were the most important to explain differences in plant biomass.

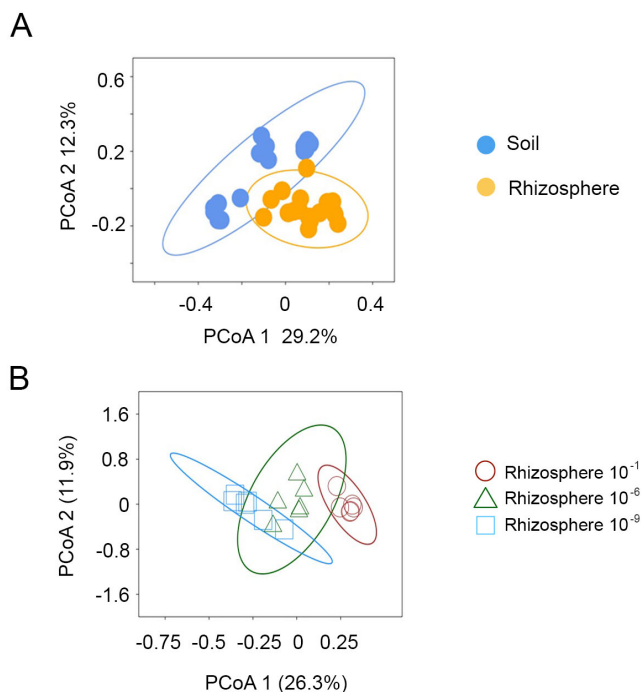
Network analysis was conducted based on the correlations between the selected functional traits and plant biomass of the undiluted  $10^{-1}$  rhizosphere samples. Significant correlations were identified based on P-values  $< 0.05$ , this corresponds to correlation coefficients  $> 0.5$  or  $< -0.5$ . The resulting correlation matrix was translated into an association network using Cytoscape 3.2.1 (Shannon et al 2003).

A heatmap was created using the relative abundance of the selected functional traits and classified by R package. The distance used were Pearson correlation for clustering the genes. Partial correlations were calculated by R and the 'ppcor' package.

### 5.3. Results

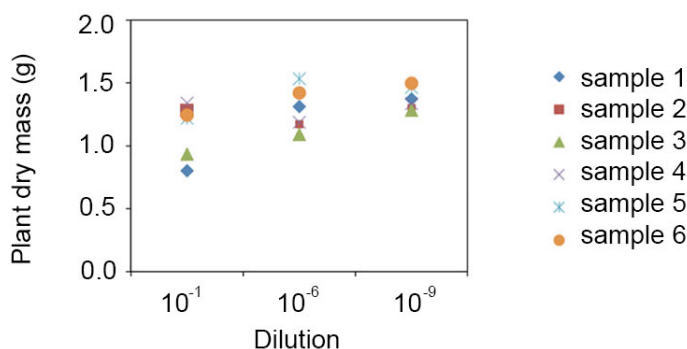
#### 5.3.1. Effect of bacterial community composition on plant biomass

Clearly, as was already demonstrated in Chapter 4, dilution and rhizosphere selection led to changes in diversity and structure of the bacterial communities (Fig. 5.1A). The taxonomical profile of rhizosphere samples showed a significant separation amongst three dilutions (Fig. 5.1B; ANOSIM,  $R = 0.49$ ,  $P < 0.01$ ) with PC1 and PC2 explained 26.3 % and 11.9 % of the observed variation, respectively.



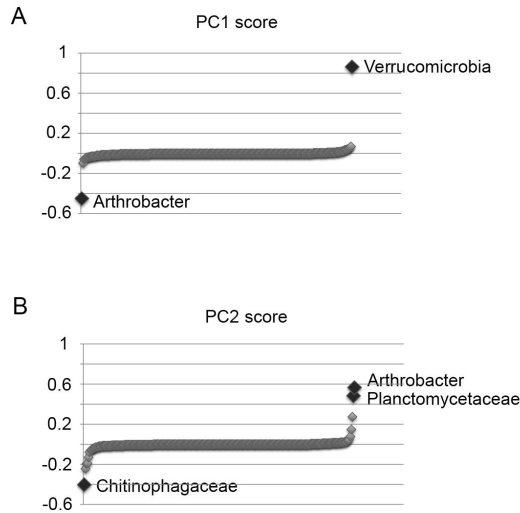
**Figure 5.1.** PCoA of Bray-curtis similarity matrix among samples using taxonomic profiles based on the relative abundance of OTUs. (A) Variation between samples of soil and rhizosphere. (B) Variation between dilutions of rhizosphere samples.

The plants grew significantly less well in the undiluted  $10^{-1}$  rhizosphere samples compared to the diluted samples (Fig. 5.2). Variation in plant biomass production among replicated samples differed for the different dilutions. In the undiluted  $10^{-1}$  rhizosphere samples, the largest variation in plant biomass amongst replicates was observed.



**Figure 5.2.** Effect of soil microbial communities on plant dry biomass (Mean dry weight,  $n=12$  per dilution).

To determine the species that were potentially responsible for the differences in plant biomass production, we determined Spearman's rank correlation between the PC1 score of the rhizosphere taxonomic profile and plant biomass of the undiluted  $10^{-1}$  samples. Interestingly, the PC1 and PC2 scores significantly correlated with plant biomass in soils inoculated with undiluted  $10^{-1}$  samples (PC1:  $n = 8$ ,  $R = -0.91$ ,  $P < 0.001$ ; PC2:  $R = 0.82$ ,  $P < 0.01$ ). To pre-select OTUs, we zoomed in on PC1 and PC2 of the taxonomical profile and selected species with scores  $< -0.3$  and  $> 0.3$ . This resulted in two species from PC1 and three species from PC2 (Fig. 5.3). One group of OTUs (*Arthrobacter*) overlapped so this resulted in four species in total (Fig. 5.4). *Arthrobacter* was negatively correlated with PC1 and positively with PC2, and as expected it showed a positive correlation with plant biomass (Fig. 5.4;  $n = 8$ ,  $R = 0.87$ ,  $P < 0.01$ ). *Planctomycetales* was positively correlated with PC2 and as expected it was also positively related to plant biomass ( $n = 8$ ,  $R = 0.79$ ,  $P < 0.01$ ). *Verrucomicrobia* and *Chitinophagaceae* that positively correlated to PC1 and PC2, respectively, were not significantly correlated to plant biomass, although the trends were in the direction as expected on basis of their PC scores (Fig. 5.4B).



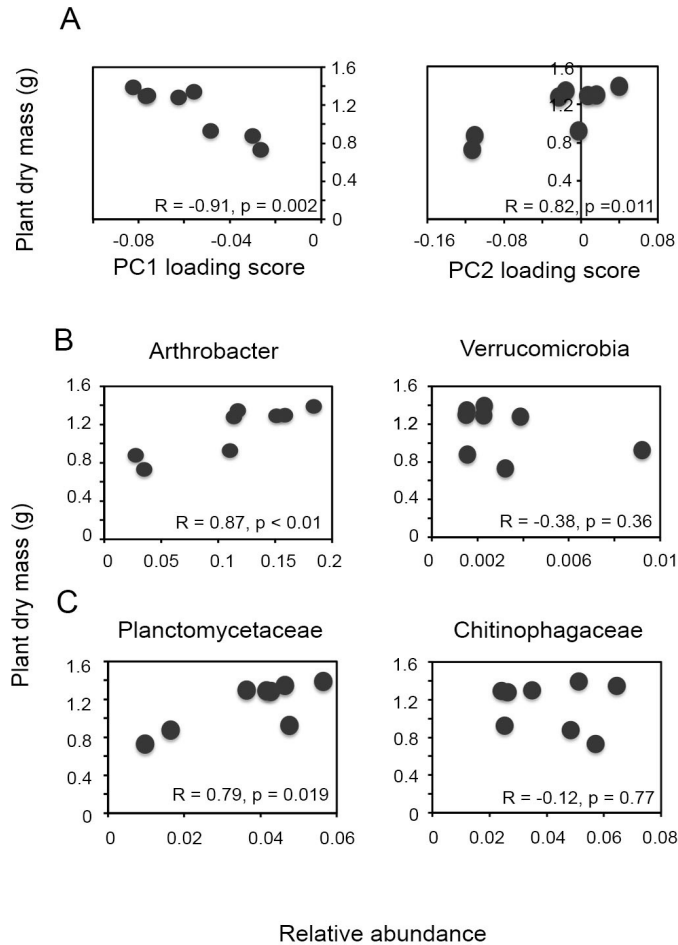
**Figure 5.3.** Loading plot of Principle Component Analysis (PC1) of the taxonomical profiles with red dots as important factors and gray dots as not important. Species names are indicated near each red dot.

To identify the potentially most important taxonomical unit, *i.e.* *Arthrobacter* or *Planctomycetaceae* we calculated the partial correlations with plant biomass. After taking *Arthrobacter* into account *Planctomycetaceae* was no longer correlated with plant biomass while after taking *Planctomycetaceae* into account *Arthrobacter* still significantly positively correlated to plant biomass ( $n = 8$ ,  $R_p = 0.80$ ,  $P < 0.05$ ; Table 5.1). Interestingly, we found in our previous paper (Chapter 4) that, indeed, *Arthrobacter* occurred at higher frequency in the rhizosphere compared to bulk soil.

**Table 5.1.** Partial correlation matrix between the two main species, *i.e.* *Arthrobacter* and *Planctomycetaceae*, controlling plant biomass.

| Variables studied              | Arthrobacter | Planctomycetaceae | Plant biomass 10 <sup>-1</sup> |
|--------------------------------|--------------|-------------------|--------------------------------|
| Arthrobacter                   | 1.00         | - 0.12 (0.79)     | 0.80* (0.03)                   |
| Planctomycetaceae              |              | 1.00              | 0.52 (0.23)                    |
| Plant biomass 10 <sup>-1</sup> |              |                   | 1.00                           |

Values indicate partial correlation coefficients ( $P$ -value) between two species; \* $P < 0.05$ .

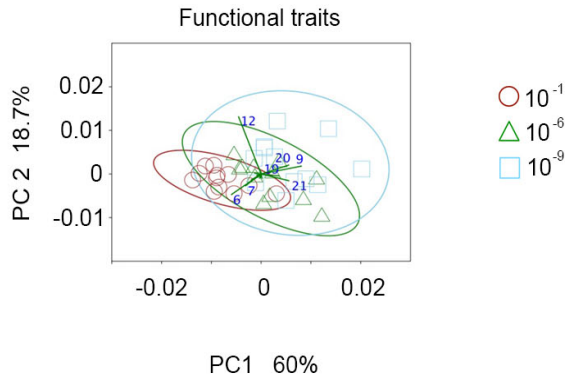


**Figure 5.4.** Interactions between species and their host plant. (A) Correlations between PC1 and PC2 of the taxonomical profiles in the rhizosphere and the plant biomass of the undiluted  $10^{-1}$  samples; (B) Correlations between plant biomass of the undiluted  $10^{-1}$  samples and the species that were selected from the PC1 score; (C) Correlations between plant biomass of the undiluted  $10^{-1}$  samples and the species that were selected from the PC2 score.

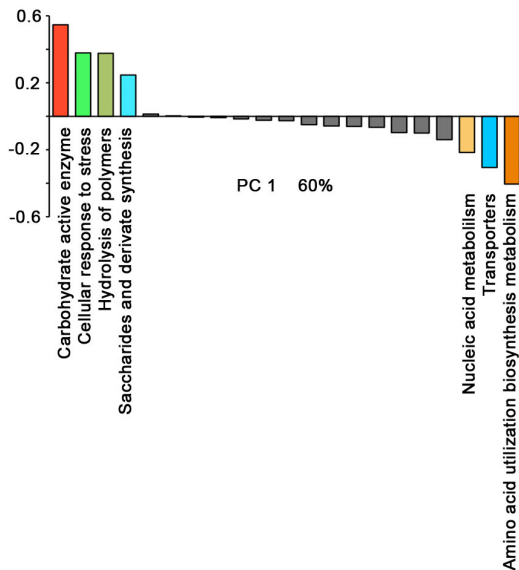
### 5.3.2. Differences in functional traits among the bacterial communities

To further assess the functional traits responsible for the discrimination amongst the dilutions of the rhizosphere samples, an unsupervised multivariate data analyses, Principal Component Analysis (PCA), was performed on the bacterial functional profile (Fig. 5.5). A PCA of the functional profile based on FOAM Dataset ‘level 1’ of the rhizosphere samples showed a significant separation amongst three dilutions (ANOSIM,  $R = 0.41$ ;  $P < 0.001$ ) with PC1 and PC2 explaining 60 % and 18.7 % of the observed variation, respectively.

Based on the PC1 score, the functional traits with scores < -0.3 and > 0.3 were selected. As a result, seven out of twenty-one functional traits ('level 1') were identified as significantly influenced by dilutions (Fig. 5.6).



**Figure 5.5.** Principle component analysis of functional traits of the rhizosphere samples (FOAM 'level 1'). The functions responsible for the PCA separation were indicated in the biplot. 6: amino acid utilization biosynthesis metabolism; 7: nucleic acid metabolism; 9: carbohydrate active enzyme; 12: transporters; 19: saccharide and derivate synthesis; 20: hydrolysis of polymers; 21: cellular response to stress.



**Figure 5.6.** Loading plot of Principle Component Analysis (PC1) of functional traits of the rhizosphere samples (FOAM 'level 1') with colored bars as the important factors and the gray bars as not important.

### 5.3.3. Functional traits related to plant biomass

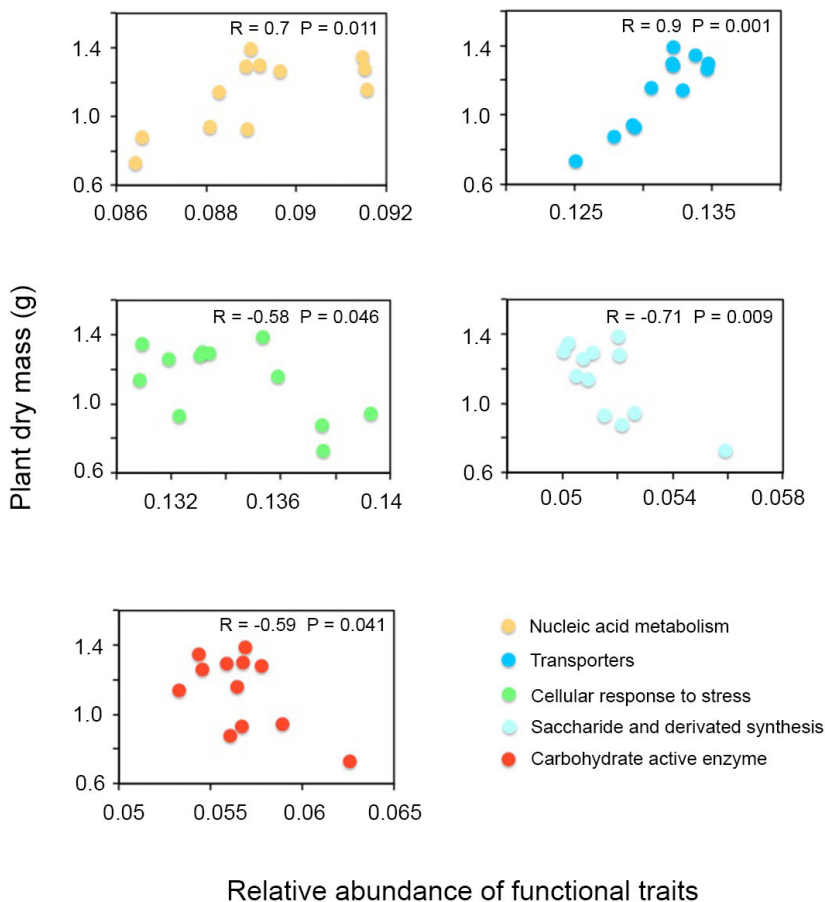
In order to predict the effect of the first two principal components of the functional profile on plant biomass, Spearman's rank correlations between the first two principal components and plant biomass of each dilution was performed (Table 5.2). For the undiluted  $10^{-1}$  samples, plant biomass was significantly, negatively, correlated with PC1 ( $n = 12$ ,  $R = -0.69$ ,  $P < 0.01$ ) and positively correlated with PC2 ( $n = 12$ ,  $R = 0.86$ ,  $P < 0.001$ ). The trends were the same for the diluted samples except for PC2 of the  $10^{-9}$  dilution (Table 5.2). Based on the biplot (Fig. 5.5), the important functions that were responsible for the differences in the PCA separation, *i.e.* 'level 1': 'amino acid utilization biosynthesis metabolism', 'nucleic acid metabolism', 'carbohydrate active enzyme', 'transporters', 'saccharide and derivate synthesis', 'hydrolysis of polymers' and 'cellular response to stress', were not correlated to plant biomass when all three dilutions are taken together. Thus, in order to examine which functional traits contributed to plant biomass, we focused on the undiluted,  $10^{-1}$ , samples where we observed the largest differences in the plant biomass production. Overall, five out of seven functional traits (at 'level 1') were, on basis of relative gene abundances, significantly correlated with plant biomass (Fig. 5.7).

**Table 5.2.** Spearman's correlation coefficients between PC1 loading of functional traits and plant dry biomass of each dilution.

| Dilution  | PC loading | Spearman's coefficient | P value  |
|-----------|------------|------------------------|----------|
| $10^{-1}$ | PC 1       | - 0.69                 | 0.013 ** |
| $10^{-6}$ | PC 1       | - 0.33                 | 0.291    |
| $10^{-9}$ | PC 1       | - 0.48                 | 0.118    |
| $10^{-1}$ | PC 2       | 0.86                   | 0.001*** |
| $10^{-6}$ | PC 2       | 0.52                   | 0.079    |
| $10^{-9}$ | PC 2       | - 0.28                 | 0.381    |

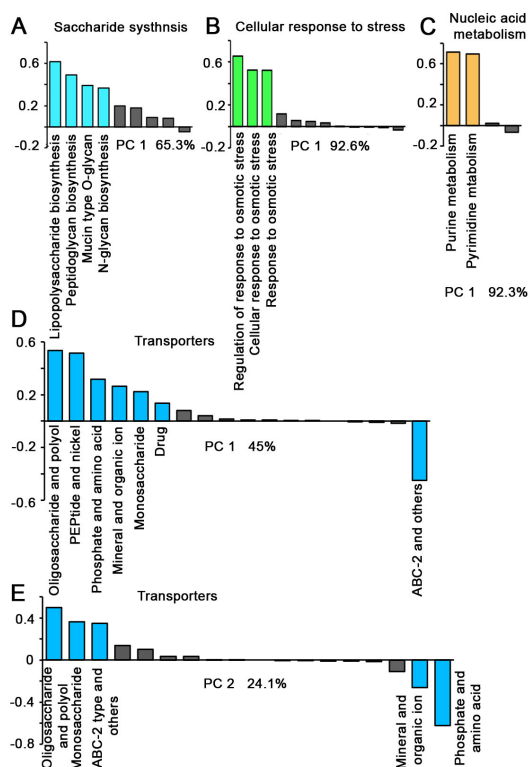
To pre-select functional traits at deeper levels (e.g. 'level 2' or 'level 3'), first we tested the five functional traits ('level 1'), and only for the four ones for which we found a significant correlation with plant biomass we zoomed in at deeper levels of particular functional traits. Correlations between each potential

functional trait and plant biomass were determined to generate a correlation network. For three of the functional traits, we zoomed in at ‘level 2’, for two others we could zoom in at ‘level 3’ (Fig. 5.8). The results of the network analysis indicated twelve functional traits belonging to four broad functional categories that were significantly correlated with plant biomass (Fig. 5.9). For one category, i.e. ‘carbohydrate active enzymes’, no deeper level function was correlated with plant biomass. Functional traits related to ‘transporters’ and ‘nucleic acid mechanism’ showed positive correlations with plant biomass. In contrast, functional traits that were related to ‘cellular response to stress’ and ‘saccharide and derivate synthesis’ were negatively correlated with plant biomass.



**Figure 5.7.** Correlations between the functional traits of the rhizosphere samples and the plant biomass of the undiluted  $10^{-1}$  samples. The colour of each dot indicates the functional categories of ‘level 1’ in the FOAM dataset to which the traits belong.

To further determine the above selected twelve functional genes (e.g. ‘level 2’ or ‘level 3’) that could determine the differences in plant biomass when all three dilutions were taken together, Spearman correlations were performed between these functional genes and plant biomass, respectively. ‘Monosaccharide transporters’ genes significantly positive correlated to plant biomass ( $n = 36$ ,  $R = 0.48$ ,  $P < 0.01$ ).

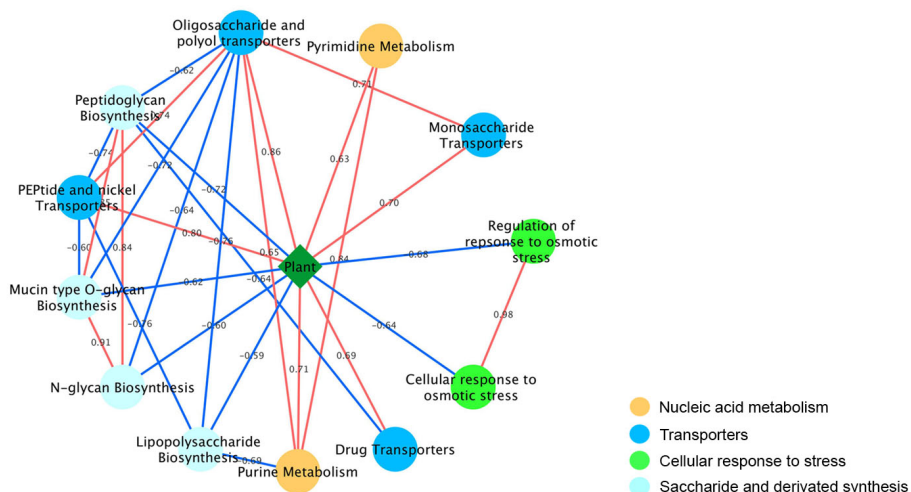


**Figure 5.8.** Loading plot of principle component analysis (PC1) of functional traits of the rhizosphere samples (FOAM ‘level 1’) with colored bars as important factors.

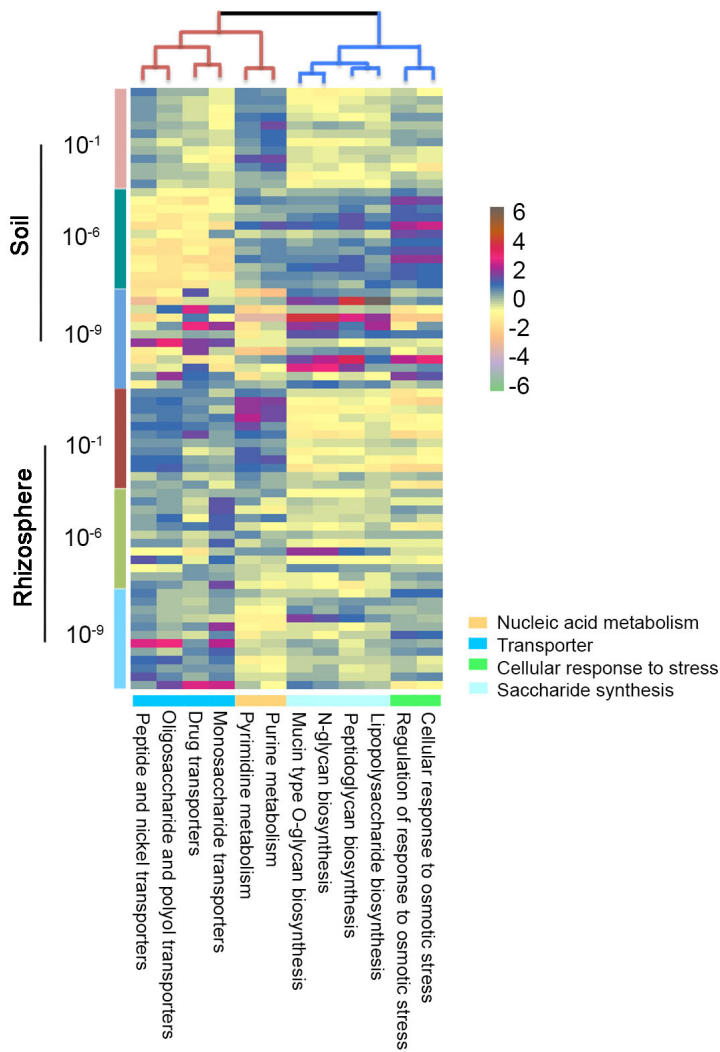
#### 5.3.4. Abundance of predicted functional traits in the rhizosphere and in the soil

We visualized the relative abundances of these twelve selected functional traits in the rhizosphere and in the bulk soil of the three dilutions in a heatmap (Fig. 5.10). The functional traits, which were positively correlated with plant biomass of the undiluted  $10^{-1}$  samples, *i.e.* genes related to ‘transporters’ and ‘nucleic acid metabolism’, clustered together and were over-represented in the rhizosphere compared to the soil samples (Fig. 5.9). This overrepresentation

was strongest for the ‘transporters’ genes. Therefore we analyzed this category in more detail. Permanova test yielded significant results for the interaction between dilutions and the presence of plants for the ‘transporters’ genes ( $F = 7.97$ ,  $P < 0.0001$ ). Visual inspection of the heatmap clearly showed effects of both dilutions and the presence of plants (Fig. 5.10). The strongest differences that were consistent with the Permanova test were between the soil samples and rhizosphere samples of  $10^{-6}$  dilution. Furthermore, Permanova tests showed that both dilutions and the presence of plants had a significant influence on the relative abundance of ‘transporters’ genes (‘level 1’), respectively (Table 5.3;  $F = 14.98$ ,  $P < 0.001$ ;  $F = 15.13$ ,  $P < 0.001$ ). The relative abundance of ‘ABC transporters’ genes in the rhizosphere, involved in the uptake of monosaccharides, oligosaccharide and other compounds, was significantly affected by dilutions ( $F = 14.98$ ,  $P < 0.001$ ;  $F = 10.59$ ,  $P < 0.001$ ;  $F = 11.72$ ,  $P < 0.001$ , for the three ‘transporters’ genes, respectively), and was higher in the rhizosphere than in the soil ( $F = 15.13$ ,  $P < 0.001$ ;  $F = 12.87$ ,  $P < 0.001$ ;  $F = 21.77$ ,  $P < 0.001$ ). Both dilutions and the presence of plants increased the relative abundance of ‘drug transporters’, respectively ( $F = 17.97$ ,  $P < 0.001$ ;  $F = 18.72$ ,  $P < 0.001$ ).



**Figure 5.9.** Network correlations of twelve functional traits (FOAM ‘level 2’ and ‘level 3’) with the plant biomass of the undiluted  $10^{-1}$  samples. Red lines indicate positive correlations, blue lines indicate negative correlations. The colour of each node indicates the functional categories of level 1 to which the traits belong.  $P < 0.05$ ,  $R > 0.05$  or  $R < -0.05$ .



**Figure 5.10.** Heatmap of relative abundance of twelve functional traits in each dilution of the soil and the rhizosphere. Red lines cluster functional traits positively correlated with plant biomass of undiluted  $10^{-1}$  samples, blue lines cluster functional traits negatively correlated with plant biomass of undiluted  $10^{-1}$  samples. The colour at the bottom of the heatmap profile indicates the functional categories of level 1 to which the traits belong.

**Table 5.3.** Two-way Permanova using Bray-Curtis similarity showing the effects of dilution and plant presence on functional traits.

| Functions                          |                                     |                                               | Factors        | Sum of<br>sqrs | df | Mean<br>Square | F     | p       |
|------------------------------------|-------------------------------------|-----------------------------------------------|----------------|----------------|----|----------------|-------|---------|
| Level 1                            | Level 2                             | Level 3                                       |                |                |    |                |       |         |
| Transporters                       |                                     |                                               | Plant presence | 0.019          | 1  | 0.019          | 15.13 | < 0.001 |
|                                    |                                     |                                               | Dilution       | 0.038          | 2  | 0.019          | 14.98 | < 0.001 |
|                                    |                                     |                                               | Interaction    | 0.020          | 2  | 0.010          | 7.97  | < 0.001 |
| Transporters                       | ABC<br>transporters                 | Monosaccharide<br>transporters                | Plant presence | 0.019          | 1  | 0.019          | 15.13 | < 0.001 |
|                                    |                                     |                                               | Dilution       | 0.038          | 2  | 0.019          | 14.98 | < 0.001 |
|                                    |                                     |                                               | Interaction    | 0.020          | 2  | 0.010          | 7.97  | < 0.001 |
| Transporters                       | ABC<br>transporters                 | Oligosaccharide<br>and polyol<br>transporters | Plant presence | 0.002          | 1  | 0.002          | 12.87 | < 0.001 |
|                                    |                                     |                                               | Dilution       | 0.003          | 2  | 0.002          | 10.59 | < 0.001 |
|                                    |                                     |                                               | Interaction    | 0.002          | 2  | 0.001          | 5.27  | < 0.001 |
| Transporters                       | ABC<br>transporters                 | Peptide and nickel<br>transporters            | Plant presence | 0.002          | 1  | 0.002          | 21.77 | < 0.001 |
|                                    |                                     |                                               | Dilution       | 0.003          | 2  | 0.001          | 11.72 | < 0.001 |
|                                    |                                     |                                               | Interaction    | 0.002          | 2  | 0.001          | 8.84  | < 0.001 |
| Transporters                       | Major<br>Facilitator<br>Superfamily | Drug transporters                             | Plant presence | 0.001          | 1  | 0.001          | 17.97 | < 0.001 |
|                                    |                                     |                                               | Dilution       | 0.002          | 2  | 0.001          | 18.72 | < 0.001 |
|                                    |                                     |                                               | Interaction    | 0.001          | 2  | 0.001          | 13.28 | < 0.001 |
| Saccharides and derivate synthesis |                                     |                                               | Plant presence | 0.001          | 1  | 0.001          | 0.75  | 0.52    |
|                                    |                                     |                                               | Dilution       | 0.005          | 2  | 0.002          | 6.08  | < 0.001 |
|                                    |                                     |                                               | Interaction    | 0.001          | 2  | 0.001          | 1.23  | 0.30    |
| Cellular response to stress        |                                     |                                               | Plant presence | 0.001          | 1  | 0.001          | 1.022 | 0.317   |
|                                    |                                     |                                               | Dilution       | 0.021          | 2  | 0.010          | 44.84 | < 0.001 |
|                                    |                                     |                                               | Interaction    | 0.002          | 2  | 0.001          | 4.62  | 0.012   |

Functional traits related to ‘saccharides and derivate synthesis’, which were negatively correlated with plant biomass of the undiluted  $10^{-1}$  samples, were over-represented in the bulk soil compared to the rhizosphere samples (Fig. 5.10). This is also true for the genes of ‘cellular response to stress’. Permanova test for the effects of dilutions and the presence of plants on the relative abundance of ‘cellular response to stress’ related genes resulted in a significant interaction ( $F = 4.62$ ,  $P = 0.01$ ). Furthermore, Permanova test for the effect of dilutions and the presence of plants on the relative abundance of ‘saccharides and derivate synthesis’ resulted in significant results for dilutions (Table 5.1;  $F = 6.08$ ,  $P < 0.001$ ), but not for the presence of plants. However, dilutions also had a significant influence on the relative abundance of ‘cellular response to stress’ related genes ( $F = 44.84$ ,  $P < 0.001$ ), but not for the presence of plant.

The over representation in the rhizosphere was the strongest for the ‘transporters’ genes, the group of genes that also had the highest correlation with plant biomass ( $n = 12$ ,  $R = 0.90$ ,  $p < 0.001$ ). To further analyze which of the four groups of functional genes (‘level 1’) potentially was the most important one to explain variation in plant biomass we calculated partial correlations with plant biomass (Table 5.4). After taking the ‘transporters’ genes into account none of the other groups of functional genes was significantly correlated with plant biomass while after taking the other groups into account in each case ‘transporters’ genes were significantly correlated with plant biomass.

**Table 5.4.** Partial correlation matrix between ‘transporters’ and other functional traits controlling plant biomass production.

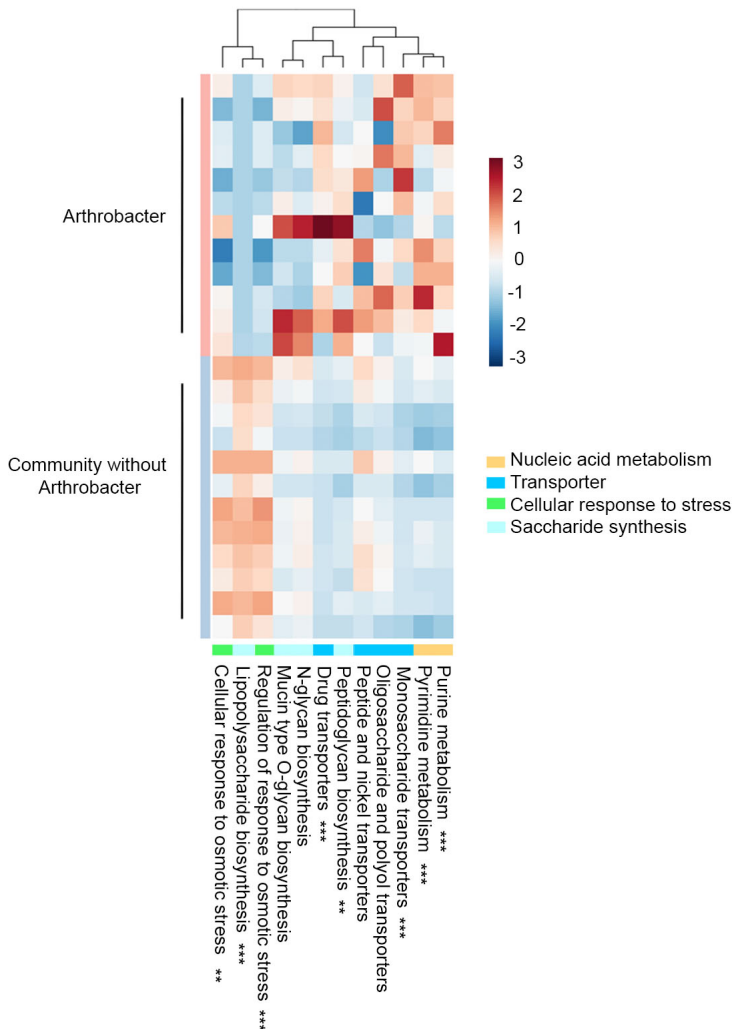
| Functional traits studied          | Nucleic acid metabolism           | Transporters   | Plant biomass 10 <sup>-1</sup> |
|------------------------------------|-----------------------------------|----------------|--------------------------------|
| Nucleic acid metabolism            | 1.00                              | 0.29 (0.530)   | 0.15 (0.750)                   |
| Transporters                       |                                   | 1.00           | 0.84*(0.020)                   |
| Plant biomass 10 <sup>-1</sup>     |                                   |                | 1.00                           |
|                                    | Carbohydrate Active enzyme        | Transporters   | Plant biomass 10 <sup>-1</sup> |
| Carbohydrate Active enzyme         | 1.00                              | -0.46 (0.300)  | 0.17 (0.710)                   |
| Transporters                       |                                   | 1.00           | 0.90**(0.006)                  |
| Plant biomass 10 <sup>-1</sup>     |                                   |                | 1.00                           |
|                                    | Saccharide and derivate synthesis | Transporters   | Plant biomass 10 <sup>-1</sup> |
| Saccharides and derivate synthesis | 1.00                              | -0.82* (0.020) | 0.58 (0.170)                   |
| Transporters                       |                                   | 1.00           | 0.91**(0.004)                  |
| Plant biomass 10 <sup>-1</sup>     |                                   |                | 1.00                           |
|                                    | Cellular response to stress       | Transporters   | Plant biomass 10 <sup>-1</sup> |
| Cellular response to stress        | 1.00                              | -0.62 (0.140)  | 0.36 (0.43)                    |
| Transporters                       |                                   | 1.00           | 0.91**(0.005)                  |
| Plant biomass 10 <sup>-1</sup>     |                                   |                | 1.00                           |

Values in the table indicate partial correlation coefficients (*P*-value) between two functional traits; \**P* < 0.05, \*\**P* < 0.01.

5.3.5. *A combined analysis of the effects of taxonomical composition and functional traits on plant biomass.*

We first analyzed the relative frequency of the twelve selected functional genes for *Arthrobacter* and the remainder of the bacterial community (Fig. 5.11). A heatmap revealed that functional traits belonging to ‘transporters’, ‘nucleic acid

metabolism', and most genes of 'saccharides and derivate synthesis' clustered together. Specifically, genes of 'purine metabolism', 'pyrimidine metabolism', 'drug transporters', 'monosaccharide transporters' and 'peptidoglycan biosynthesis' were significantly enriched in the community with *Arthrobacter*, while 'cellular response to osmotic stress', 'regulation of response to osmotic stress' and 'lipopolysaccharide biosynthesis' were clustered together and were significantly enriched in bacterial community without *Arthrobacter*.



**Figure 5.11.** Heatmap of relative abundance of the twelve functional traits of the undiluted  $10^{-1}$  rhizosphere samples for communities with and without *Arthrobacter*. The colour at the bottom of heatmap profile indicates the functional traits at 'level 1' (ANOVA: \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ ).

Because the frequency of ‘transporters’ genes is higher in the community with *Arthrobacter*, the correlation of ‘transporters’ genes with plant biomass may be indirect through *Arthrobacter* or vice versa. We therefore calculated partial correlations. After taking the frequency of ‘transporters’ genes into account, the correlation of *Arthrobacter* with plant biomass is no longer significant. While after taking the effect of *Arthrobacter* into account, the correlation of ‘transporters’ genes with plant biomass is still significant (Table 5.5;  $n = 12$ ,  $R_p = 0.79$ ,  $P = 0.036$ ). This suggests the ‘transporters’ genes are more important to explain differences in plant biomass than *Arthrobacter*.

**Table 5.5.** Partial correlation matrix between *Arthrobacter* and transporters controlling plant biomass production.

| Variables studied              | Transporters | Arthrobacter | Plant biomass 10 <sup>-1</sup> |
|--------------------------------|--------------|--------------|--------------------------------|
| Transporters                   | 1.00         | 0.05 (0.916) | 0.79* (0.036)                  |
| <i>Arthrobacter</i>            |              | 1.00         | 0.50 (0.258)                   |
| Plant biomass 10 <sup>-1</sup> |              |              | 1.00                           |

Values in the table indicate partial correlation coefficients ( $P$ -value) within *Arthrobacter* and transporters; \* $P < 0.05$ .

5.4. Discussion

In Chapter 2, we demonstrated that the dilution procedure changes the diversity and structure of the bacterial community after regrown in the soil and in Chapter 4 that further selection proceeds in the rhizosphere, largely on the basis of functional traits. This may imply that functional traits that are selected in the rhizosphere may also have a strong influence on plant growth. As we already showed in Chapter 4 that the selection of functional traits is not randomly associated with taxonomic selection, the functional selection as observed in the rhizosphere is, of course, intimately associated with selection of bacterial species. Therefore, here, we also assessed the taxonomic relationship between plant biomass and the bacterial community composition.

Indeed, we demonstrated that manipulation of the bacterial community by the dilution approach, affected plant biomass production. Plants gained the lowest biomass in soils inoculated with the lower dilutions, i.e. the more diverse rhizosphere communities. Recent studies on ‘plant-soil feedback’ have shown that rhizomicrobiome could directly or indirectly influence the composition and productivity (i.e. biomass) of plant communities (van der Heijden 2008, Joosten et al 2009, van Elsas et al 2012). Moreover, reduction of abundant and/or rare species by manipulation of microbial community could promote plant growth (Hol et al 2010). Hence, microbial community composition belowground has been identified as predictor of fitness of the aboveground vegetation (van der Heijden et al 2008, Lau and Lennon 2011, Wagg et al 2011). In our study, we detected two OTUs (Fig. 5.4), which were actually significantly related to plant biomass and, thus, being potential candidates to explain the observed differences by unsupervised multivariate analysis. The first taxonomical unit was *Arthrobacter* that is known to promote plant growth (Dimkpa et al 2009). *Arthrobacter* is typically found in soil and several species of *Arthrobacter* have been described as plant growth promoter (Gusain et al 2015, Ullah and Bano 2015). Analysis of the wheat rhizosphere using 16S rRNA gene sequencing revealed that *Arthrobacter* belonged to the group of rhizobacteria (Tahir et al 2015). The other group of bacteria that also showed a positive correlation with plant biomass was *Planctomycetaceae*. This genus is known to include typical rhizosphere species (Tesfaye et al 2003), but from literature it is not known if it includes growth-promoting species. Partial correlation analysis identified *Arthrobacter* as the most important one to explain differences in plant biomass (Table 5.1). Identifying genera that promote or inhibit plant growth gives us little information on the mechanisms causing these effects. It is therefore also of great interest to study, in addition to the taxonomic composition, the relationship between plant growth and the functional genes of the bacteria from the rhizosphere community.

This study showed the power of the metagenomics approach in combination with an unsupervised Principal Component Analysis to predict plant biomass production in relation to the functional traits of the rhizomicrobiome. Interestingly, as was shown in the heatmap and Permanova test, the functional traits that were positively correlated with plant biomass were over-represented in the rhizosphere compared to the bulk soil, which suggests that plants selected beneficial bacterial activities surrounding their roots. This

particular bacterial functionality may lead to plant growth promotion. In contrast, the functional traits that were negatively correlated with plant biomass were more abundant in the soil than in the rhizosphere, suggesting plants selected against such functions leading to over-representation in the bulk soil compared to the rhizosphere.

More precisely, a group of ‘ABC-type transporters’ of peptides, oligosaccharides and drugs and the uptake and release of many different compounds were over-represented in the rhizosphere compared to the bulk soil. This observation is consistent with the fact that numerous genes for ‘membrane transporters’ systems have been reported as enriched in the rhizosphere (Mendes et al 2014). Another group of functional genes that was over-represented in the rhizosphere compared to the bulk soil is linked to ‘nucleic acid metabolism’. Given that the category ‘nucleic acid metabolism’ involves several interconnected pathways, and may be indicative of cellular growth processes, this suggests higher bacterial growth and activity in the rhizosphere than in soil. This may presumably result in increased plant biomass production for instance by protection against pathogens or by increasing nutrient acquisition for the host.

In this study we observed not only positive but also negative correlations between functional traits of the rhizomicrobiome and plant biomass. As mentioned above, in contrast to the functional genes that were positively related to plant biomass, the ones that were negatively correlated with plant biomass were over-represented in the soil compared to the rhizosphere. This would suggest that plants selected against such genes in the rhizosphere. If, for example, plants create a less stressful environment for bacteria by rhizodeposition this would cause a less stressful environment in the rhizosphere compared to the soil and consequently to an under representation of these genes in the rhizosphere as compared to the soil. At the same time, if plants are growing well, they would produce more roots and more developed rhizosphere and thus to a less stressful environment for the bacteria, leading to a negative correlation between plant biomass and the density of stress genes. We should be careful however with such an interpretation because one of the negative correlations was between plant biomass and a group of osmotic stress genes, i.e. ‘cellular response to osmotic stress’ and ‘regulation of response to osmotic stress’. As larger plants take up more water this could create a more stress full

environment to the microbes and so one could expect a positive relationship between plant biomass and osmotic stress genes. Because plants were watered every two days during plant growth we may have created an environment that did not have moisture stress, nor for plants and nor for microbes.

We also found a negative relationship between genes related to 'saccharides and derivate synthesis'. It could be that this negative correlation is due to the fact that the plants that grow better provide more carbohydrates and saccharides to the microbes so that biosynthesis of these products is repressed in their rhizosphere. If the above reasoning is correct we would also expect that the rhizosphere provide an environment where these genes are repressed compared to the bulk soil. However, for this group of genes we did not find an under- or over representation in the rhizosphere compared to the bulk soil.

Partial correlation analyses identified 'transporters' genes as the most important ones to potentially explain the observed differences in plant biomass. Likewise we identified *Arthrobacter* as the most important taxonomical unit in this respect. Well-known activities of *Arthrobacter* are degradation of pollutants in the rhizosphere (Khan et al 2009), production of auxin that might stimulate nutrient uptake (Tsavkelova et al 2006) and production of indole-3-acetic acid (IAA) (Sziderics et al 2007). Because the frequency of 'transporters' genes was relatively high in *Arthrobacter* compared to the rest of the bacterial community we used partial correlation to test for the relative importance of the two for plant growth. This analysis suggests that the frequency of 'transporters' genes is the most important factor and that plants select for favorable functions rather than species to benefit their growth. Thus, the high abundance of 'transporters' genes in *Arthrobacter* may lead to the high abundance of *Arthrobacter* in the rhizosphere. Yet, we should treat these results with caution because our analyses may be biased by the fact that the two methods may have different sensitivity. Furthermore, the genes of 'monosaccharide transporters' ('level 3') significantly correlated positively to plant biomass when all three dilutions were taken together, and this group of genes increased significantly upon dilutions in the rhizosphere samples, indicating that this group of genes most likely explained the plant biomass differences in dilutions. Because the 'transporters' genes were enriched in the rhizosphere compared to the soil we can conclude that plants positively affect bacterial species with such genes and this may in part explain the positive correlation with plant growth. At the same

time bacterial species with these ‘transporters’ genes may stimulate plant growth, which would contribute to the positive correlation. Whether either one of the two or both explanations are true cannot be concluded from our experiments with certainty. Additional experiments are needed that use the microbial communities from soil of pots with different plant growth to inoculate sterile soil again and measure plant growth for a second and following generations in order to achieve maximum enrichment of the most responsible genes. Clearly, it is not very likely that there is a single function that determines the variation in plant biomass. The problem in the analyses of this type of studies is the fact that there are many potential factors and a limited number of replicates. We therefore used an unsupervised method to make an unbiased pre-selection of potential taxonomical units and groups of functional genes. This improves the statistical power but comes at the cost of many important factors going unnoticed.

Nevertheless, our study provides a comprehensive framework for understanding the mechanism of plant-microbe interactions. The latter is not merely of scientific interest but is also useful for the development of sustainable crop production systems, e.g. by application of beneficial soil microbial communities or species to optimize crop yields.

## 5.5. References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS *et al* (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **19**: 455-477.
- Brockett BFT, Prescott CE, Grayston SJ (2012). Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biology & Biochemistry* **44**: 9-20.
- Bulgarelli D, Rott M, Schlaeppi K, van Themaat EVL, Ahmadinejad N, Assenza F *et al* (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* **488**: 91-95.
- Bulgarelli D, Garrido-Oter R, Munch PC, Weiman A, Droge J, Pan Y *et al* (2015). Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell Host & Microbe* **17**: 392-403.
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Meth* **7**: 335-336.
- Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al* (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* **37**: D141-D145.
- Dennis PG, Miller AJ, Hirsch PR (2010). Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *Fems Microbiology Ecology* **72**: 313-327.
- Dimkpa C, Weinand T, Asch F (2009). Plant-rhizobacteria interactions alleviate abiotic stress conditions. *Plant Cell Environ* **32**: 1682-1694.
- Dixon P (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14**: 927-930.
- Edgar RC (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-2461.
- Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194-2200.
- Fierer N, Jackson RB (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 626-631.
- Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F *et al* (2015). HMMER web server: 2015 update. *Nucleic Acids Res* **43**: W30-38.
- Gusain YS, Kamal R, Mehta CM, Singh US, Sharma AK (2015). Phosphate solubilizing and indole-3-acetic acid producing bacteria from the soil of Garhwal Himalaya aimed to improve the growth of rice. *Journal of Environmental Biology* **36**: 301-307.
- Haichar FE, Marol C, Berge O, Rangel-Castro JI, Prosser JI, Balesdent J *et al* (2008). Plant host habitat and root exudates shape soil bacterial community structure. *ISME J* **2**: 1221-1230.
- Hammer et al HD, Ryan PD. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4**: **9**. pp.
- Hol WHG, de Boer W, Termorshuizen AJ, Meyer KM, Schneider JHM, van Dam NM *et al* (2010). Reduction of rare soil microbes modifies plant-herbivore interactions. *Ecology Letters* **13**: 292-301.
- Hughes JB, Hellmann JJ (2005). The application of rarefaction techniques to molecular inventories of microbial diversity. *Environmental Microbiology* **397**: 292-308.
- Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bmc Bioinformatics* **11**.
- Imelfort M, Woodcroft, B., Parks, D. (2015). Ecogenomics/BamM [WWW Document]. GitHub. URL <https://github.com/Ecogenomics/BamM> (accessed 7.21.15).

- Joosten L, Mulder PPJ, Klinkhamer PGL, van Veen JA (2009). Soil-borne microorganisms and soil-type affect pyrrolizidine alkaloids in *Jacobaea vulgaris*. *Plant and Soil* **325**: 133-143.
- Joshi NA, Fass JN (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available at <https://github.com/najoshi/sickle>.
- Khan MS, Zaidi A, Wani PA, Oves M (2009). Role of plant growth promoting rhizobacteria in the remediation of metal contaminated soils. *Environmental Chemistry Letters* **7**: 1-19.
- Koster J, Rahmann S (2012). Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* **28**: 2520-2522.
- Lau JA, Lennon JT (2011). Evolutionary ecology of plant-microbe interactions: soil microbial structure alters selection on plant traits. *New Phytologist* **192**: 215-224.
- Li H, Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754-1760.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.
- Liao Y, Smyth GK, Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923-930.
- Logue JB, Lindstrom ES (2010). Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints. *ISME J* **4**: 729-738.
- Mark GL, Dow JM, Kiely PD, Higgins H, Haynes J, Bayse C *et al* (2005). Transcriptome profiling of bacterial responses to root exudates identifies genes involved in microbe-plant interactions. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 17454-17459.
- Matilla MA, Espinosa-Urgel M, Rodriguez-Herva JJ, Ramos JL, Ramos-Gonzalez MI (2007). Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome Biology* **8**.
- McKinney W (2015). Python Data Analysis Library — pandas: Python Data Analysis Library [WWW Document]. URL <http://pandas.pydata.org/> (accessed 7.22.15).
- Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM (2014). Taxonomical and functional microbial community selection in soybean rhizosphere. *ISME J* **8**: 1577-1587.
- Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332**: 1097-1100.
- Nemergut DR, Cleveland CC, Wieder WR, Washenberger CL, Townsend AR (2010). Plot-scale manipulations of organic matter inputs to soils correlate with shifts in microbial community composition in a lowland tropical rain forest. *Soil Biology & Biochemistry* **42**: 2153-2160.
- Ofek-Lalzar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, Minz D (2014). Niche and host-associated functional signatures of the root surface microbiome. *Nature Communications* **5**.
- Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH (2013). Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology* **11**: 789-799.
- Prestat E, David MM, Hultman J, Tas N, Lamendella R, Dvornik J *et al* (2014). FOAM (Functional Ontology Assignments for Metagenomes): a Hidden Markov Model (HMM) database with environmental focus. *Nucleic Acids Res* **42**: e145.
- Reeder J, Knight R (2010). Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nature Methods* **7**: 668-669.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D *et al* (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**: 2498-2504.
- Steiner A (1968). Soilless culture. Proceedings of the 6th colloquium of the International Potash Institute, Florence, Italy. International Potash Institute, Bern, Switzerland, pp 324-341.

- Sziderics AH, Rasche F, Trognitz F, Sessitsch A, Wilhelm E (2007). Bacterial endophytes contribute to abiotic stress adaptation in pepper plants (*Capsicum annuum* L.). *Canadian Journal of Microbiology* **53**: 1195-1202.
- Tahir M, Mirza MS, Hameed S, Dimitrov MR, Smidt H (2015). Cultivation-Based and Molecular Assessment of Bacterial Diversity in the Rhizosphere of Wheat under Different Crop Rotations. *Plos One* **10**.
- Tesfaye M, Dufault NS, Dornbusch MR, Allan DL, Vance CP, Samac DA (2003). Influence of enhanced malate dehydrogenase expression by alfalfa on diversity of rhizobacteria and soil nutrient availability. *Soil Biology & Biochemistry* **35**: 1103-1113.
- Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ *et al* (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28**: 511-U174.
- Tremaroli V, Backhed F (2012). Functional interactions between the gut microbiota and host metabolism. *Nature* **489**: 242-249.
- Tsavkelova EA, Klimova SY, Cherdyntseva TA, Netrusov AI (2006). Microbial producers of plant growth Stimulators and their practical use: A review. *Applied Biochemistry and Microbiology* **42**: 117-126.
- Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**: 1027-1031.
- Ullah S, Bano A (2015). Isolation of plant-growth-promoting rhizobacteria from rhizospheric soil of halophytes and their impact on maize (*Zea mays* L.) under induced soil salinity. *Canadian Journal of Microbiology* **61**: 307-313.
- van der Heijden MGA, Bardgett RD, van Straalen NM (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters* **11**: 296-310.
- van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T *et al* (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**: 69-72.
- van der Heijden MGA (2008). The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems (vol 11, pg 296, 2008). *Ecology Letters* **11**: 651-651.
- van Elsas JD, Chiurazzi M, Mallon CA, Elhottova D, Kristufek V, Salles JF (2012). Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 1159-1164.
- Wagg C, Jansa J, Schmid B, van der Heijden MGA (2011). Belowground biodiversity effects of plant symbionts support aboveground productivity. *Ecology Letters* **14**: 1001-1009.

