



ANALYSIS OF THE BACTERIAL EPIPHYTIC MICROBIOTA OF OAK LEAF LETTUCE WITH 16S RIBOSOMAL RNA GENE ANALYSIS

Agnes Weiss^{1*}, Felix Scheller¹, Markus Oggenfuss², Fiona Walsh^{2,§}, Jürg E. Frey², David Drissner³, Herbert Schmidt¹

Address(es): Dipl.-Ing. Dr. Agnes Weiss

¹Department of Food Microbiology and Hygiene, Institute of Food Science and Biotechnology, University of Hohenheim, Garbenstrasse 28, 70599 Stuttgart, Germany, Tel.: +49 711 459 24255, Fax: +49 711 459 24199.

²Institute for Plant Production Sciences, and ³Institute for Food Sciences, Agroscope, Schloss 1, 8820 Wädenswil, Switzerland.

*Corresponding author: agnes.weiss@uni-hohenheim.de

doi: 10.15414/jmbfs.2015/16.5.3.271-276

ARTICLE INFO

Received 15. 5. 2015
Revised 12. 10. 2015
Accepted 14. 10. 2015
Published 1. 12. 2015

Regular article



ABSTRACT

The leaf microbiota has major influences on the quality of ready-to-eat lettuce. While studies investigating the epi- and endophytic microbiota of lettuce have been published, no protocols focusing only on the epiphytic microbiota exist. As the epiphytic microbiota may be especially influenced by technological steps in the production of ready-to-eat lettuce, an in-depth knowledge of these microorganisms is essential with regard to consumer safety and spoilage. Currently it is not clear to what extent results gained from single samples are representative of the community composition. A technique for the separation of bacterial cells from the leaf surface was applied to green oak leaf lettuce. The bacterial diversity was analysed in triplicate with high throughput Roche 454 sequencing of prokaryotic 16S rRNA genes to analyse the intra-sample variation. Sequence analysis revealed members of the phyla *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Gemmatimonadetes*, *Proteobacteria* and *Verrucomicrobia*, and of the candidate division WYO. The ten most abundant proteobacterial genera in all three samples were *Alkanindiges* (24.6%), *Pseudomonas* (11.3%), *Sphingomonas* (8.6%), *Janthinobacterium* (8.3%), *Acinetobacter* (4.3%), *Polaromonas* (1.3%), *Erwinia* (1.1%), and *Methylobacterium* (1.1%). The genera *Pedobacter* (2.5%) and *Hymenobacter* (1.4%) dominated the phylum *Bacteroidetes*. The intra-sample variation was less than 0.7% for seven of these most abundant genera with the exception of *Pseudomonas*, *Janthinobacterium* and *Alkanindiges*, where larger standard deviations were obtained. This low intra-sample variation demonstrates that the established technique based on oak leaf lettuce is suitable for the culture-independent analysis of the epiphytic bacterial microbiota of produce.

Keywords: Oak leaf lettuce; microbiota; biodiversity; next generation sequencing; *Alkanindiges*

INTRODUCTION

The microbial community of plant leaves comprises bacteria, yeast and fungi, which originate from diverse environmental sources such as soil, water, and air. The microbes exhibit commensal, mutualistic, or pathogenic characteristics towards the plant, whereby most of the interactions are commensal (Lindow and Brandl, 2003; Vorholt, 2012). Bacterial viable counts of up to 10⁷ colony forming units per gram of leaf have been described (Thunberg *et al.*, 2007; Wulfkuehler *et al.*, 2013 and 2014). According to Williams *et al.* (2013), only 1% to 10% of the total bacteria on leaves are cultivable. Thus, the results obtained by culture-dependent methods clearly differ from those determined by culture-independent sequence-based methods. Using culture-independent sequence-based methods the bacterial communities from several plants such as lettuce, spinach, rice, diverse tree species, and *Arabidopsis thaliana*, were investigated. These studies revealed that *Proteobacteria*, *Firmicutes*, *Actinobacteria* and *Bacteroidetes* were the most abundant phyla across plant bacterial communities (Knief *et al.*, 2012; Lopez-Velasco *et al.*, 2011; Rastogi *et al.*, 2012; Redford *et al.*, 2010; Reisberg *et al.*, 2012; Williams *et al.*, 2013). At the genus level, mainly *Pseudomonas*, *Sphingomonas*, *Methylobacterium*, *Bacillus*, *Massilia*, *Arthrobacter*, *Pantoea*, and *Erwinia* were found. *Sphingomonas* and *Methylobacterium* have been demonstrated as being the most abundant genera of the bacterial communities on leaves in a multitude of plant species including important crop plants like clover, soybean, tomato, rice, and several grassland species (Delmotte *et al.*, 2009; Enya *et al.*, 2007; Green, 2005; Kim *et al.*, 1998). Several techniques have been used to analyse community changes as a function of exogenous and endogenous factors such as geography, climate, radiation, and plant genetics (Finkel *et al.*, 2012; Hunter *et al.*, 2010; Izhaki *et al.*, 2013; Kadivar and Stapleton, 2003; Knief *et al.*, 2012; Rastogi *et al.*, 2012; Redford *et al.*, 2010; Vokou *et al.*, 2012).

In the present study, we developed an extensive protocol from sample preparation to next generation sequencing to assess the bacterial diversity on the surface of oak leaf lettuce. This cultivar was chosen as an example of a cultivar with a delicate leaf structure. A procedure for the separation of the bacterial cells from the leaf surfaces while retaining leaf integrity was established. A biodiversity analysis using Roche 454 sequencing of 16S rRNA gene amplicons was performed. This procedure was evaluated with DNA retrieved from a single green oak leaf lettuce head to determine the reproducibility of the method and to determine the degree of intra-sample variation within a single lettuce sample.

MATERIAL AND METHODS

Bacterial biomass harvest and DNA isolation

One head of conventionally grown green oak leaf lettuce (*Lactuca sativa* var. *crispa*) was obtained from a local retail store. The wrapper leaves were removed under sterile conditions by cutting with a scalpel. To gain a representative amount of biomass, ten middle-sized leaves from this lettuce head were used as described below. DNA was isolated from each of the leaves and the DNA preparations were subsequently combined to a single DNA solution. Bacterial biomass harvest was carried out as follows: the cut edges of ten middle-sized leaves (approximately 10 g of lettuce each) were sealed with commercial nail polish (Manhattan Lotus Effect, Coty Germany GmbH, Germany) to reduce chloroplast effusion. The leaves were transferred into sterile blender bags with lateral filters (BBAG-03, Gosselin SAS, France) and the appropriate nine-fold volume of sterile buffer (1% (w/v) TWEEN® 80 (Sigma-Aldrich Co., U.S.A.), 1% (w/v) buffered peptone water (Merck KGaA, Germany), 10 mM EDTA (Biomol GmbH, Germany); pH 6.7) was added. The leaves were completely covered by the buffer. The bags were then heat-sealed and placed into a Pulsifier® PUL 100E device (Microgen Bioproducts Ltd., UK), and treated for 45 s to detach microorganisms from the leaf surface. With this procedure, removal

of bacteria was highest and leaves remained largely intact (data not shown). Afterwards, the complete buffer of each leaf was divided into two equal volumes, which were transferred into two sterile 50 mL centrifuge tubes. The tubes were centrifuged at $8,000 \times g$ for 20 min at 4°C. Supernatants were discarded, and the four cell pellets of two leaves were pooled into 2 mL reaction tubes, resulting in five tubes, and stored at -20°C.

Genomic DNA was prepared from the five cell pellets modified from Wilson (1997). Each bacterial cell pellet was re-suspended in 567 µL lysozyme solution (3 mg/mL in TE-buffer (10 mM Tris, 1 mM EDTA; pH 8.0)) by vigorous blending. After the addition of cetyl trimethylammonium bromide (CTAB) and 0.7 M sodium chloride and incubation at 65°C for 10 min, the samples were cooled to room temperature, and 50 µL of ribonuclease A solution (3 mg/mL in TE-buffer) was added. The samples were mixed and incubated at 37°C for 2 h. The first extraction step was performed with 800 µL of a mixture of phenol:chloroform:isoamyl alcohol (25:24:1) (Carl Roth GmbH & Co. KG, Germany) by mixing and centrifugation in a bench top centrifuge ($17,950 \times g$, 4°C, 5 min). The upper aqueous phase was transferred into a fresh tube and 800 µL of chloroform:isoamyl alcohol (24:1) (Carl Roth GmbH & Co. KG) were added. The sample was mixed and centrifuged under the same conditions as described above. The aqueous phase was again transferred to a new tube and this step was repeated once. Finally, 600 µL of ice-cold isopropanol (-20°C; Merck KGaA) was added and the DNA was precipitated for at least 0.5 h at -20°C. The precipitate was centrifuged ($17,950 \times g$, 4°C, 15 min), the supernatant was discarded, and the pellet was washed with 1 mL of 70% (v/v) ethanol (-20°C). The dried pellet was re-suspended in 40 µL of DNase- and RNase-free sterile water (Gibco®, Thermo Fisher Scientific Inc.) and stored at 4°C overnight. The DNA was solubilised by heating at 65°C for 5 min. The solutions of the five parallel samples were pooled, dried with a SpeedVac™ Vacuum System (Thermo Fisher Scientific Inc.) at ambient temperature and re-suspended in 40 µL of DNase- and RNase-free sterile water. The DNA concentration and quality were determined with a NanoDrop™ ND-2000 device (PEQLAB Biotechnologie GmbH, Germany) and by agarose gel electrophoresis.

Amplification of 16S rRNA genes, library preparation and Roche 454 sequencing

The primers for the paired-end 16S rRNA gene community sequencing were 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') (Caporaso et al., 2012). Three samples of the DNA solution were used for Roche 454 sequencing. For each of the three samples, four different PCRs were performed to reduce the bias using dilutions of the purified DNA at concentrations of 10^1 , 10^2 , 10^3 , and 10^4 . Each 25 µL reaction contained 200 nM of each primer, 1 µL of template DNA as indicated above, 10 µL of 5 PRIME HotMasterMix (5 PRIME GmbH, U.S.A.), and 13 µL of PCR grade water. PCR cycle parameters consisted of an initial denaturation at 94°C for 3 min, then 35 cycles of denaturation at 94°C for 45 s, annealing at 50°C for 1 min and extension at 72°C for 1.5 min, followed by a final extension step at 72°C for 10 min, as described on the Earth Microbiome website (<http://www.earthmicrobiome.org/emp-standard-protocols/16S>). PCR products were obtained from template dilutions of 10^2 , 10^3 and 10^4 . These products were pooled for each sample separately (55 µL each), mixed with 275 µL of PB buffer (Qiagen GmbH, U.S.A.), and purified using a MinElute® PCR Purification Kit (Qiagen GmbH) according to the manufacturer's protocol. The amplicons were eluted with 20 µL of PCR grade water and quantified with a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific Inc.). The concentrations of the three samples were 50, 35 and 50 ng/µL, respectively. The following steps were performed for each sample independently: For the fragment end repair, 500 ng of purified amplicon was used and an adaptor ligation was carried out according to the Rapid Library Preparation Method Manual V. 3/2012 (F. Hoffmann-La Roche Ltd., Switzerland). The resulting sample (27 µL) was mixed with 3 µL of BlueJuice™ gel loading buffer (Thermo Fisher Scientific Inc.), run on a 1% agarose [TAE 1 ×] gel for 40 min at 4.5 V/cm. The ligation product was excised on a DarkReader® transilluminator (Clare Chemical Research, U.S.A.) and recovered using the MinElute® Gel Extraction Kit (Qiagen GmbH). Library quantitation was performed according to the Roche Manual with a QuantiFluor™ ST fluorimeter (Promega, U.S.A.). An aliquot of the DNA library was diluted to a working stock of 1×10^7 molecules/µL in TE buffer. An emulsion PCR was then performed with 11 µL of a working stock according to the emPCR Amplification Method Manual – Lib-L V. 3/2012 (F. Hoffmann-La Roche Ltd.). DNA library bead enrichment and sequencing primer annealing were performed according to the manufacturer's specifications using a REM-e module (F. Hoffmann-La Roche Ltd.) on a Biomek® 3000 robot (Beckman Coulter GmbH, U.S.A.). Finally, the parallel samples were independently sequenced on a 454 GS Junior System (F. Hoffmann-La Roche Ltd.) according to the Sequencing Method Manual V. 1/2013 (F. Hoffmann-La Roche Ltd.).

16S rRNA amplicon sequencing data analysis

Forward and reverse reads were joined using fastq-join from the ea-utils software package (Aronesty, 2011). The sequences were then quality-filtered to exclude sequences with ambiguous bases (N) and sequences containing base calls with less than 99% confidence (Phred score of < 20). The QIIME bioinformatics software suite (version 1.5.0) was used to subsample the resultant sequences (82,311, 99,654 and 95,216 sequences, respectively), to cluster those sequences into 97% identity operational taxonomic units (OTU) with UCLUST (Aronesty, 2011), to taxonomically classify each OTU with RDP classifier (Edgar 2010) at 97% sequence homology placement into the GreenGenes 12_10 database (DeSantis et al., 2006). All QIIME scripts were run using default parameters unless otherwise stated. The sequences were deposited in the NCBI sequence read archive (SRA) under the numbers SRR1523744 (sample 1), SRR1265099 (sample 2) and SRR1265426 (sample 3).

RESULTS AND DISCUSSION

16S rRNA gene amplicon sequencing results

In order to assess a method for the biodiversity analysis of leaf surface microbiota and to measure intra-sample variation, the total DNA of the leaf surface microbiota of a single lettuce head was isolated and quantified. A DNA concentration of 2,727.4 ng/µL with absorption ratios of A_{260}/A_{280} of 1.85 and A_{260}/A_{230} of 1.77 were obtained for the pooled DNA solution. Three samples of this DNA solution, called sample 1, sample 2 and sample 3, were analysed by 16S rRNA gene sequencing using a Roche 454 next generation sequencer as described above. The universal primers used have been described to cover the variable region 4 of the 16S rRNA gene of almost all bacterial phyla (Walters et al. 2011). The results and quality parameters of these runs are depicted in Table 1.

Table 1 16S rRNA gene amplicon sequencing results

| Sequence analysis parameters | Sample 1 | Sample 2 | Sample 3 |
|---|------------------------|------------------------|------------------------|
| Total sequences (OTUs) | 82,311 (1,355 OTUs) | 99,654 (1,667 OTUs) | 95,216 (1,364 OTUs) |
| Median length of sequences [bp] | 266 | 291 | 268 |
| Phred Score ± standard deviation | 36.73 ± 6.40 | 36.94 ± 6.12 | 37.07 ± 6.03 |
| Covered richness (Chao1) | 4644.0 | 5478.2 | 2370.2 |
| Phylogenetic distance (PD) estimator | 66.0 | 87.8 | 66.3 |
| Shannon's H index | 5.7 | 5.9 | 5.8 |
| Mitochondrial sequences (OTUs) | 129 (10 OTUs) | 192 (21 OTUs) | 169 (14 OTUs) |
| Chloroplast sequences (OTUs) | 10,992 (29 OTUs) | 9,853 (15 OTUs) | 15,468 (20 OTUs) |
| Bacterial 16S rRNA gene sequences (OTUs) | 71,190 (1,316 OTUs) | 89,609 (1,631 OTUs) | 79,579 (1,330 OTUs) |
| of these | | | |
| Abundant OTUs | 396 | 472 | 489 |
| Sequences in abundant OTUs | 69,284 | 87,954 | 77,947 |
| % of bacterial 16S rRNA gene sequences | 97.3 | 98.1 | 97.9 |
| % of bacterial 16S rRNA gene OTUs | 30.1 | 28.9 | 36.8 |

On average, 92,393 sequences were obtained per sample, with an average median length of 275 bp. Members of the domain *Archaea* were not detected, and the sequences assigned to eukaryotes originated from their mitochondria and chloroplasts (see Table 1). The covered richness expressed as the Chao1 index was rather high, reaching an average of 4164.1. For all samples the rarefaction curves showed asymptotic curve progression, indicating that a very high bacterial diversity was detected in this study (Figure 1). As expected, the total number of sequences included mitochondrial as well as chloroplast sequences (see Table 1), which were discarded from the subsequent analysis. Furthermore, sequences that could not be assigned to a phylum or a candidate division were excluded from the analysis. The mean number of bacterial 16S rRNA gene sequences was therefore 80,126. These data indicate a high level of taxonomic identification and classification with the method applied as well as that the applied database was very appropriate with respect to the analysed samples.

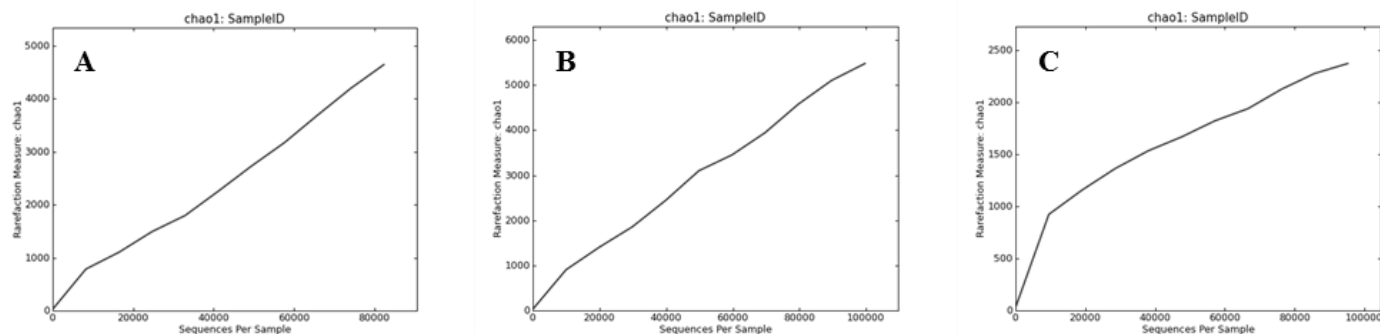


Figure 1 Rarefaction curves showing the covered richness expressed as the Chao1 index for the samples 1 (A), 2 (B) and 3 (C)

The data obtained by next generation sequencing were evaluated for the intra-sample variation. Approximately 450 OTUs per sample were detected. This demonstrates a diverse microbiota present on oak leaf lettuce, as the majority of sequences could be assigned to specific bacterial phyla. The number of sequences that remained unclassified or could only be assigned to the kingdom *Bacteria* strongly declined with increasing number of sequences per OTU in all samples. It has to be noted that diversity estimates are highly sensitive to the abundance of rare members of populations (White et al., 2010). Including OTUs with single reads into the evaluation, Kunin et al. (2010) found that sequencing errors inflated the actual diversity 100-fold, and therefore the present analysis was restricted to OTUs containing at least five sequences. Of the remaining OTUs on average six unclassified OTUs consisted of more than five sequences. For the kingdom *Bacteria* on average twelve OTUs with more than five sequences were found that could not be assigned to a further taxonomic level. These may represent hitherto unidentified bacteria, but might also be attributed to sequencing errors that were not removed from the dataset despite de-noising.

Phylogenetic analysis of abundant OTUs

The results gained from the three analysed samples were compared to determine the intra-sample variation. On average 452 OTUs contained at least five sequences, which corresponded to 32.0% of the bacterial 16S rRNA gene OTUs (see Table 1). Nevertheless, these OTUs contained on average 97.8% of the bacterial 16S rRNA gene sequences. Furthermore, on average 0.4% could be identified on the kingdom level as *Bacteria* only and could not be assigned to a specific phylum (Table 2).

Among the abundant OTUs sequences members of the phyla *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Gemmatimonadetes*, *Proteobacteria*, and *Verrucomicrobia* as well as the candidate division WYO were detected. As shown in Table 2, 89.9 ± 1.0% of the sequences belonged to the phylum *Proteobacteria*, followed by the phyla *Bacteroidetes* (7.1 ± 1.0%) and *Actinobacteria* (2.5 ± 0.4%). Only a minority of sequences was allocated to the remaining phyla. Within the *Proteobacteria*, on average 56.0% of the sequences were γ -*Proteobacteria*, 16.8% β -*Proteobacteria*, and 15.9% α -*Proteobacteria*, while only 0.1% of the sequences were δ -*Proteobacteria*. Only one OTU in sample 3 containing five sequences was allocated to ϵ -*Proteobacteria*, to the genus *Arcobacter*.

Table 2 Number of classified sequences allocated to abundant OTUs, respective number of OTUs and genera assigned to diverse phyla and their proportion within the of bacterial 16S rRNA gene sequences contained in abundant OTUs [%]

| Phylum | Sample 1 | | | | Sample 2 | | | | Sample 3 | | | |
|-------------------------|--------------------|----------------|--------|------|--------------------|----------------|--------|------|--------------------|----------------|--------|------|
| | Sequ. ^a | Number of OTUs | Genera | % | Sequ. ^a | Number of OTUs | Genera | % | Sequ. ^a | Number of OTUs | Genera | % |
| <i>Acidobacteria</i> | 6 | 1 | 1 | 0.0 | 0 | 0 | 0 | 0.0 | 5 | 1 | 0 | 0.0 |
| <i>Actinobacteria</i> | 1,430 | 32 | 9 | 2.1 | 2,472 | 38 | 9 | 2.8 | 1,974 | 39 | 11 | 2.5 |
| <i>Bacteroidetes</i> | 4,607 | 80 | 11 | 6.6 | 5,655 | 90 | 11 | 6.4 | 6,386 | 107 | 11 | 8.2 |
| <i>Firmicutes</i> | 157 | 8 | 3 | 0.2 | 135 | 5 | 3 | 0.2 | 237 | 15 | 4 | 0.3 |
| <i>Gemmatimonadetes</i> | 0 | 0 | 0 | 0.0 | 0 | 0 | 0 | 0.0 | 7 | 1 | 0 | 0.0 |
| <i>Proteobacteria</i> | 62,815 | 354 | 34 | 90.7 | 79,355 | 336 | 35 | 90.2 | 69,110 | 323 | 38 | 88.7 |
| <i>Verrucomicrobia</i> | 5 | 1 | 0 | 0.0 | 0 | 0 | 0 | 0.0 | 13 | 2 | 1 | 0.0 |
| WYO | 0 | 0 | 0 | 0.0 | 0 | 0 | 0 | 0.0 | 6 | 1 | 0 | 0.0 |
| identified | | | | 99.6 | | | | 99.6 | | | | 99.7 |
| <i>Bacteria</i> | 264 | 9 | - | 0.4 | 337 | 12 | - | 0.4 | 209 | 15 | - | 0.3 |

^a Sequ.: sequences.

The phyla detected in the current study match those reported for Romaine lettuce (Rastogi et al., 2012) and greenhouse-grown lettuce (Erlacher et al., 2014) for the total leaf microbiota of endophytic and epiphytic microorganisms, but, as expected, the composition differed. The majority of sequences were assigned to Gram-negative bacteria. This is in accordance with results of other authors (e.g., Rastogi et al., 2012; Vorholt, 2012), even though more members of the phylum *Firmicutes* were reported. The focus of this work was on the Gram-negative population, as several members are known to produce biofilms and to be responsible for spoilage. Nevertheless, the DNA isolation procedure applied in the present study was re-assessed with a panel of selected representative Gram-positive and -negative pure cultures. Furthermore, harsher DNA isolation procedures including bead-beating steps were evaluated for epiphytic phyllosphere community biomass samples, which yielded inferior DNA quality (data not shown). Cruaud et al. (2014) found in their comparison of a commercial DNA isolation kit with a conventional DNA isolation procedure for application in 454 pyrosequencing only low levels of dissimilarity in taxonomic affiliations. Furthermore, the primers applied in the current study should also have amplified Gram-positive sequences. The detected low numbers of Gram-positives could also be attributed to the leaf type, as the microbiota of lettuce was shown to depend on the cultivar (Hunter et al., 2010), the season (Williams et al., 2013) and the geographical location (Finkel et al., 2011; Redford et al., 2010).

In total, 58 genera were detected in both, sample 1 and 2, and 65 genera in sample 3 (Table 2). They contained on average 65.6% of the bacterial 16S rRNA gene sequences of abundant OTUs. Of these, only ten genera were assigned at least 1.0% of the sequences in one of the samples (see Table 3). Of these ten

genera, eight were identical in all three samples, and only *Methylobacterium* and *Polaromonas* were below 1.0% in one of the samples. For these ten most abundant genera, the number of sequences of the most abundant OTUs between the samples was comparable. For example, for the genus *Alkanindiges* two OTUs per sample contained the majority of the sequences. These were 8,665 and 8,610 sequences for sample 1, 10,357 and 9,396 sequences for sample 2 and 10,918 and 8,999 sequences for sample 3. The only exception was *Polaromonas*, where only two abundant OTUs were detected in sample 1, and one abundant OTU in each of the other two samples.

Of the *Proteobacteria* these main genera were the α -proteobacterial genera *Sphingomonas* and *Methylobacterium*, the β -proteobacterial genera *Janthinobacterium* and *Polaromonas*, as well as the γ -proteobacterial genera *Alkanindiges*, *Pseudomonas*, *Acinetobacter*, and *Erwinia*. To display the proteobacterial diversity, Figure 2 compares all proteobacterial genera that amounted at least 0.1% of the total number of sequences in the three samples.

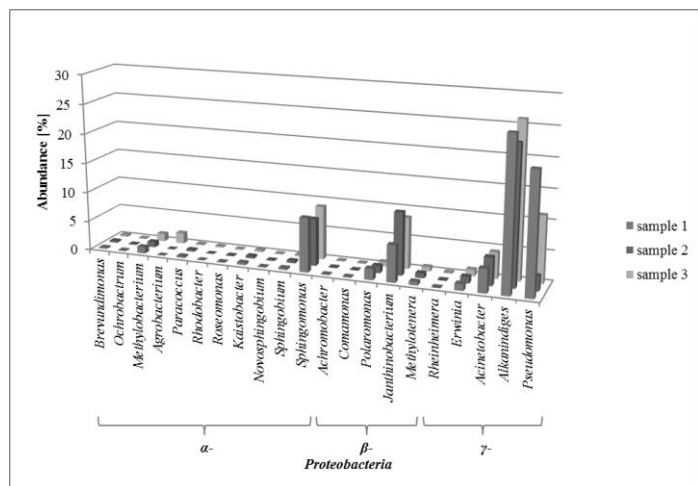


Figure 2 Overview of the genera of the phylum *Proteobacteria* with an abundance of at least 0.1% of the total sequences in one of the samples

Within the phylum *Bacteroidetes* the genera *Pedobacter* and *Hymenobacter* were most abundant (Table 3). As shown in Table 3, the standard deviations of the abundances were below 0.7% for seven of the ten genera. Larger standard deviations were calculated for *Pseudomonas*, *Janthinobacterium* and *Alkanindiges*. These results correlate well with higher numbers of sequences identified only at the respective family level (data not shown). For example, 20.5%, 2.6% and 11.2% of the sequences in abundant OTUs were allocated to the genus *Pseudomonas*. On the other hand, 3.5%, 22.4% and 9.6% were additionally allocated to the family *Pseudomonadaceae*, but no genus affiliation was possible.

Table 3 Number of classified sequences allocated to abundant OTUs, respective number of OTUs assigned to the ten most abundant genera and their proportion within the number of bacterial 16S rRNA gene sequences contained in abundant OTUs [%]

| Genus | Sample 1 | | | Sample 2 | | | Sample 3 | | | Mean | St. dev. |
|--------------------------|---------------------|------|------|---------------------|------|------|---------------------|------|------|------|----------|
| | Number of sequences | OTUs | % | Number of sequences | OTUs | % | Number of sequences | OTUs | % | | |
| <i>Alkanindiges</i> | 17,860 | 17 | 25.8 | 20,455 | 15 | 23.3 | 20,336 | 19 | 26.1 | 25.1 | 1.5 |
| <i>Pseudomonas</i> | 14,191 | 11 | 20.5 | 2,245 | 15 | 2.6 | 8,745 | 10 | 11.2 | 11.4 | 9.0 |
| <i>Sphingomonas</i> | 6,383 | 17 | 9.2 | 6,956 | 26 | 7.9 | 7,014 | 23 | 9.0 | 8.7 | 0.7 |
| <i>Janthinobacterium</i> | 4,344 | 5 | 6.3 | 9,235 | 7 | 10.5 | 6,691 | 9 | 8.6 | 8.5 | 2.1 |
| <i>Acinetobacter</i> | 2,795 | 15 | 4.0 | 4,136 | 16 | 4.7 | 3,459 | 17 | 4.4 | 4.4 | 0.4 |
| <i>Pedobacter</i> | 1,695 | 5 | 2.4 | 2,012 | 6 | 2.3 | 2,361 | 8 | 3.0 | 2.6 | 0.4 |
| <i>Hymenobacter</i> | 985 | 16 | 1.4 | 1,092 | 17 | 1.2 | 1,321 | 26 | 1.7 | 1.4 | 0.3 |
| <i>Polaromonas</i> | 1,306 | 2 | 1.9 | 1,064 | 1 | 1.2 | 560 | 1 | 0.7 | 1.3 | 0.6 |
| <i>Erwinia</i> | 760 | 2 | 1.1 | 967 | 2 | 1.1 | 852 | 2 | 1.1 | 1.1 | 0.0 |
| <i>Methylobacterium</i> | 782 | 7 | 1.1 | 764 | 7 | 0.9 | 1,044 | 10 | 1.3 | 1.1 | 0.2 |

The most abundant bacterial genus was *Alkanindiges*, a γ -*Proteobacterium*, which comprised on average $25.1 \pm 1.5\%$ of the bacterial 16S rRNA gene sequences in abundant OTUs. No species affiliation within this genus was possible in all samples. This genus was already reported to be an important member of the leaf microbiota of Romaine lettuce (Rastogi et al., 2012), and in the former study its presence correlated positively with that of *Xanthomonas campestris* pv. *vitians*. Contrarily, although *Alkanindiges* was the most abundant genus in the present study, no *Xanthomonas* sequences were detected. On the other hand, *Alkanindiges* and *Acinetobacter* were reported as indicators for healthy plants (Erlacher et al., 2014), and it should be studied further if the difference in *Alkanindiges* abundance could be attributed to the lettuce cultivar and growth conditions. The only genus of the family *Enterobacteriaceae* among the ten most abundant genera was *Erwinia*. *Erwinia* may cause plant disease, be

saprophytic or epiphytic, depending on the species (Hauben and Swings, 2005). It is commonly found on lettuce (Rastogi et al., 2012), and the presence of *Erwinia* on Romaine lettuce was hypothesized to correlate with a decreased survival of *E. coli* O157:H7 (Williams et al., 2013). *Escherichia* was only rarely detected on Romaine lettuce (Williams et al., 2013), which is also the case in the current study. Each sample contained only one OTU of *Escherichia*, with 25, 1 and 27 sequences, respectively.

The 1,192, 1,254 and 1,142 sequences allocated to rare OTUs amounted to 1.5% of the total bacterial 16S rRNA gene sequences on average. Table 4 gives an overview of the assignment of rare OTUs to the diverse phyla and candidate divisions. On average, only 744 (0.9%) single reads were found.

Table 4 Number of classified sequences allocated to rare OTUs, respective number of OTUs and genera assigned to diverse phyla and candidate divisions

| Phylum | Sample 1 | | | Sample 2 | | | Sample 3 | | |
|-------------------------|---------------------|------|--------|---------------------|------|--------|---------------------|------|--------|
| | Number of Sequences | OTUs | Genera | Number of Sequences | OTUs | Genera | Number of Sequences | OTUs | Genera |
| <i>Acidobacteria</i> | 13 | 10 | 1 | 21 | 17 | 1 | 18 | 12 | 1 |
| <i>Actinobacteria</i> | 95 | 70 | 11 | 116 | 89 | 17 | 92 | 58 | 14 |
| <i>Armatimonadetes</i> | 3 | 1 | 0 | 1 | 1 | 1 | 4 | 2 | 0 |
| <i>Bacteroidetes</i> | 158 | 117 | 12 | 183 | 128 | 13 | 210 | 115 | 15 |
| BRC1 ^a | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| <i>Chlamydiae</i> | 0 | 0 | 0 | 2 | 2 | 1 | 0 | 0 | 0 |
| <i>Chlorobi</i> | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 2 | 0 |
| <i>Chloroflexi</i> | 5 | 4 | 0 | 6 | 6 | 0 | 5 | 4 | 0 |
| <i>Firmicutes</i> | 26 | 21 | 9 | 70 | 52 | 10 | 42 | 25 | 11 |
| <i>Gemmatimonadetes</i> | 9 | 8 | 1 | 7 | 7 | 0 | 14 | 12 | 0 |
| <i>Nitrospirae</i> | 1 | 1 | 1 | 5 | 4 | 1 | 2 | 2 | 1 |
| <i>Planctomycetes</i> | 12 | 10 | 1 | 11 | 11 | 3 | 16 | 11 | 1 |
| <i>Proteobacteria</i> | 664 | 510 | 38 | 802 | 631 | 42 | 714 | 433 | 40 |
| <i>Tenericutes</i> | 5 | 2 | 0 | 4 | 2 | 0 | 1 | 1 | 0 |
| <i>Thermi</i> | 4 | 2 | 2 | 3 | 3 | 1 | 4 | 3 | 1 |
| TM7 ^b | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Verrucomicrobia</i> | 15 | 13 | 3 | 19 | 15 | 0 | 8 | 5 | 1 |
| WS3 ^c | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 0 | 0 |
| WYO ^d | 4 | 2 | 0 | 1 | 1 | 0 | 5 | 3 | 0 |

^a Bacterial Rice Cluster 1 (Derakshani et al., 2001)

^b Torf, mittlere Schicht 7 (Rheims et al., 1996)

^c WurtSmith Aquifer 3 (Dojka et al., 1998)

^d Wyoming soil clone (Mummey and Stahl, 2003)

As expected by the length of the sequenced amplicons, a species affiliation was possible in 26.2% (36 OTUs), 7.7% (34 OTUs) and 16.1% (48 OTUs) of the sequenced of the three samples. These were mainly associated with the most abundant genera, and therefore no separate evaluation was carried out. Furthermore, no correlation was found between the abundance of rRNA gene sequences detected in this study and the mean number of rRNA genes present in the respective genus (Lee et al., 2008). As expected, the standard deviations between the samples and thus the intra-sample variation increased with the taxonomic level from phylum to genus. This could be attributed to the length of the sequenced amplicons. In this study a method published within the Earth Microbiome Project was applied to evaluate it for studying the plant microbiota, too. For further decreasing the intra-sample variation it could be considered to study longer amplicons, as they might facilitate a higher taxonomic resolution. Thus it was demonstrated that the established procedure consisting of bacterial biomass harvest, DNA isolation and next generation sequencing is indeed suitable for the biodiversity analysis of the epiphytic bacterial microbiota of lettuce.

CONCLUSION

In this study a complete procedure for the assessment of the bacterial leaf surface microbiota of lettuce was successfully established. Comparable results with low intra-sample variation are obtained with the established protocol by analysing three samples of the same lettuce head. In the current study, *γ-Proteobacteria* were most abundant, as already reported for other lettuce varieties and spinach. Contrary to these studies, the genus *Alkanindiges* comprised approximately a quarter of the abundant 16S rRNA gene sequences. While the implications of the variations in the bacterial biodiversity remain to be elucidated, this study has provided a sound protocol for the extraction, performance and analysis of 16S rRNA gene sequencing data of epiphytic bacterial DNA from delicate leaf structures, which greatly improves the understanding of the lettuce phyllosphere microbiota.

Acknowledgments: The authors thank Florian Fricke, University of Hohenheim, and Mitja N.P. Remus-Emsermann, Agroscope, for helpful discussions.

This research project was partially supported by the German Ministry of Economics and Technology (via AiF) and the FEI (Forschungskreis der Ernährungsindustrie e.V., Bonn), project number AiF 17122 N.

This cooperation was part of COST Action FA 1202 "A European Network for Mitigating Bacterial Colonisation and Persistence on Foods and Food Processing Environments".

REFERENCES

ARONESTY, E. 2011. ea-utils: Command-line tools for processing biological sequencing data. <http://code.google.com/p/ea-utils/>

CAPORASO, J.G., LAUBER, C.L., WALTERS, W.A., BERG-LYONS, D., HUNTLEY, J., FIERER, N., OWENS, S.M., BETLEY, J., FRASER, L., BAUER, M., GORMLEY, N., GILBERT, J.A., SMITH, G., KNIGHT, R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, 6(8), 1621-1624. <http://dx.doi.org/10.1038/ismej.2012.8>

CRUAUD, P., VIGNERON, A., LUCCHETTI-MIGANEH, C., CIRON, P.E., GODFROY, A., CAMBON-BONAVITA, M.-A. 2014. Influence of DNA extraction method, 16S rRNA targeted hypervariable regions, and sample origin on microbial diversity detected by 454 pyrosequencing in marine chemosynthetic ecosystems. *Applied and Environmental Microbiology*, 80(15), 4626-4639. <http://dx.doi.org/10.1128/aem.00592-14>

DELMOTTE, N., KNIEF, C., CHAFFRON, S., INNEREBNER, G., ROSCHITZKI, B., SCHLAPBACH, R., VON MERING, C., VORHOLT, J.A. 2009. Community proteomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 106(38), 16428-16433. <http://dx.doi.org/10.1073/pnas.0905240106>

DERAKSHANI, M., LUKOW, T., LIESACK, W. 2001. Novel bacterial lineages at the (sub)division level as detected by signature nucleotide-targeted recovery of the 16S rRNA genes from bulk soil and rice roots of flooded rice microcosms. *Applied and Environmental Microbiology*, 67(2), 623-631. <http://dx.doi.org/10.1128/aem.67.2.623-631.2001>

DESANTIS, T.Z., HUGENHOLTZ, P., LARSEN, N., ROJAS, M., BRODIE, E.L., KELLER, K., HUBER, T., DALEVI, D., HU, P., ANDERSEN, G.L. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, 72(7), 5069-5072. <http://dx.doi.org/10.1128/aem.03006-05>

DOJKA, M.A., HUGENHOLTZ, P., HAACK, S.K., PACE, N.R. 1998. Microbial diversity in a hydrocarbon- and chlorinated-solvent-contaminated aquifer undergoing intrinsic bioremediation. *Applied and Environmental Microbiology*, 64(10), 3869-3877.

EDGAR, R.C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19), 2460-2461. <http://dx.doi.org/10.1093/bioinformatics/btq461>

ENYA, J., SHINOHARA, H., YOSHIDA, S., NEGISHI, T.T.H., SUYAMA, K., TSUSHIMA, S. 2007. Culturable leaf-associated bacteria on tomato plants and their potential as biological control agents. *Microbial Ecology*, 53(4), 524-536. <http://dx.doi.org/10.1007/s00248-006-9085-1>

ERLACHER, A., CARDINALE, M., GROSCHE, R., GRUBE, M., BERG, G. 2014. The impact of the pathogen *Rhizoctonia solani* and its beneficial counterpart *Bacillus amyloliquefaciens* on the indigenous lettuce microbiome. *Frontiers in Microbiology*, 5, Article 175. <http://dx.doi.org/10.3389/fmicb.2014.00175>

FINKEL, O.M., BURCH, A.Y., ELAD, T., HUSE, S.M., LINDOW, S.E., POST, A.F., BELKIN, S. 2012. Distance-decay relationships partially determine diversity patterns of phyllosphere bacteria on tamarix trees across the Sonoran desert. *Applied and Environmental Microbiology*, 78(17), 7818-7818. <http://dx.doi.org/10.1128/aem.02505-12>

FINKEL, O.M., BURCH, A.Y., LINDOW, S.E., POST, A.F., BELKIN, S. 2011. Geographical location determines the population structure in phyllosphere microbial communities of a salt-excreting desert tree. *Applied and Environmental Microbiology*, 77(21), 7647-7655. <http://dx.doi.org/10.1128/aem.05565-11>

GREEN, P. 2005. *Methylobacterium*. In: BRENNER, D.J., KRIEG, N.R., STALEY, J.T. (Eds), *Bergey's Manual of Systematic Bacteriology*. Second edition, vol. 2. Springer; New York, Dordrecht, Heidelberg, London: 567-571. ISBN: 0-387-24145-0. http://dx.doi.org/10.1007/0-387-29298-5_136

HAUBEN, L., SWINGS, J. 2005. Genus XIII. *Erwinia*. In: BRENNER, D.J., KRIEG, N.R., STALEY, J.T. (Eds), *Bergey's Manual of Systematic Bacteriology*. Second edition, vol. 2. Springer; New York, Dordrecht, Heidelberg, London: 670-679. ISBN: 0-387-24145-0.

HUNTER, P.J., HAND, P., PINK, D., WHIPPS, J.M., BENDING, G.D. 2010. Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (*Lactuca species*) phyllosphere. *Applied and Environmental Microbiology*, 76(24), 8117-8125. <http://dx.doi.org/10.1128/aem.01321-10>

IZHAKI, I., FRIDMAN, S., GERCHMAN, Y., HALPERN, M. 2013. Variability of bacterial community composition on leaves between and within plant species. *Current Microbiology*, 66(3), 227-235. <http://dx.doi.org/10.1007/s00284-012-0261-x>

KADIVAR, H., STAPLETON, A.E. 2003. Ultraviolet radiation alters maize phyllosphere bacterial diversity. *Microbial Ecology*, 45(4), 353-361. <http://dx.doi.org/10.1007/s00248-002-1065-5>

KIM, H., NISHIYAMA, W., KUNITO, T., SENOO, K., KAWAHARA, K., MURAKAMI, K., OYAZU, H. 1998. High population of *Sphingomonas* species on plant surface. *Journal of Applied Microbiology*, 85(4), 731-736. <http://dx.doi.org/10.1111/j.1365-2672.1998.00586.x>

KNIEF, C., DELMOTTE, N., CHAFFRON, S., STARK, M., INNEREBNER, G., WASSMANN, R., VON MERING, C., VORHOLT, J.A. 2012. Metaproteomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *The ISME Journal*, 6(7), 1378-1390. <http://dx.doi.org/10.1038/ismej.2011.192>

KUNIN, V., ENGELSBREKTSON, A., OCHMANN, H., HUGENHOLTZ, P. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*, 12(1), 118-123. <http://dx.doi.org/10.1111/j.1462-2920.2009.02051.x>

LEE, Z.M.-P., BUSSEMA, C., SCHMIDT, T.M. 2008. rrnDB: documenting the number of rRNA and tRNA genes in bacteria and archaea. *Nucleic Acids Research*, 37(suppl. 1), D489-D493. <http://dx.doi.org/10.1093/nar/gkn689>

LINDOW, S.E., BRANDL, M.T. 2003. Microbiology of the phyllosphere. *Applied and Environmental Microbiology*, 69(4), 1875-1883. <http://dx.doi.org/10.1128/aem.69.4.1875-1883.2003>

LOPEZ-VELASCO, G., WELBAUM, G.E., BOYER, R.R., MANE, S.P., PONDER, M.A. 2011. Changes in spinach phylloepiphytic bacteria communities following minimal processing and refrigerated storage described using pyrosequencing of 16S rRNA amplicons. *Journal of Applied Microbiology*, 110(5), 1203-1214. <http://dx.doi.org/10.1111/j.1365-2672.2011.04969.x>

MUMMEY, D.L., STAHL, P.D. 2003. Candidate division BD: phylogeny, distribution and abundance in soil ecosystems. *Systematic and Applied Microbiology*, 26(2), 228-235. <http://dx.doi.org/10.1078/072320203322346074>

PRICE, M.N., DEHAL, P.S., ARKIN, A.P. 2010. FastTree 2 - approximately maximum-likelihood trees for large alignments. *Public Library of Science one*, 5(3), e9490. <http://dx.doi.org/10.1371/journal.pone.0009490>

RASTOGI, G., SBODIO, A., TECH, J.J., SUSLOW, T.V., COAKER, G.L., LEVEAU, J.H.J. 2012. Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *The ISME Journal*, 6(10), 1812-1822. <http://dx.doi.org/10.1038/ismej.2012.32>

REDFORD, A.J., BOWERS, R.M., KNIGHT, R., LINHART, Y., FIERER, N. 2010. The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environmental Microbiology*, 12(11), 2885-2893. <http://dx.doi.org/10.1111/j.1462-2920.2010.02258.x>

- REISBERG, E.E., HILDEBRANDT, U., RIEDERER, M., HENTSCHEL, U. 2012. Phyllosphere bacterial communities of trichome-bearing and trichomeless *Arabidopsis thaliana* leaves. *Antonie Van Leeuwenhoek*, 101(3), 551-560. <http://dx.doi.org/10.1007/s10482-011-9669-8>
- RHEIMS, H., RAINEY, F.A., STACKEBRANDT, E. 1996. A molecular approach to search for diversity among bacteria in the environment. *Journal of Industrial Microbiology*, 17(3-4), 159-169. <http://dx.doi.org/10.1007/bf01574689>
- THUNBERG, R.L., TRAN, T.T., BENNETT, R.W., MATTHEWS, R.N., BELAY, N. 2002. Microbial evaluation of selected fresh produce obtained at retail markets. *Journal of Food Protection*, 65(4), 677-682.
- VOKOU, D., VARELI, K., ZARALI, E., KARAMANOLI, K., CONSTANTINIDOU, H.I.A., MONOKROUSOS, N., HALLEY, J.M., SAINIS, I. 2012. Exploring biodiversity in the bacterial community of the mediterranean phyllosphere and its relationship with airborne bacteria. *Microbial Ecology*, 64(3), 714-724. <http://dx.doi.org/10.1007/s00248-012-0053-7>
- VORHOLT, J.A. 2012. Microbial life in the phyllosphere. *Nature Reviews Microbiology*, 10(12), 828-840. <http://dx.doi.org/10.1038/nrmicro2910>
- WALTERS, W.A., CAPORASO, J.G., LAUBER, C.L., BERG-LYONS, D., FIERER, N., KNIGHT, R. 2011. PrimerProspector: *de novo* design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics*, 27(8), 1159-1161. <http://dx.doi.org/10.1093/bioinformatics/btr087>
- WHITE, J.R., NAVLAKHA, S., NAGARAJAN, N., GHODSI, M.R., KINGSFORD, C., POP, M. 2010. Alignment and clustering of phylogenetic markers – implications for microbial diversity studies. *BioMed Central Bioinformatics*, 11, 152. <http://dx.doi.org/10.1186/1471-2105-11-152>
- WILLIAMS, T.R., MOYNE, A.-L., HARRIS, L.J., MARCO, M.L. 2013. Season, irrigation, leaf age, and *Escherichia coli* inoculation influence the bacterial diversity in the lettuce phyllosphere. *Public Library of Science one*, 8(7), e68642. <http://dx.doi.org/10.1371/journal.pone.0068642>
- WILSON, K. 1997. Preparation of genomic DNA from bacteria. *Current Protocols in Molecular Biology*, 2.4.1-2.4.5. <http://dx.doi.org/10.1002/0471142727.mb0204s56>
- WULFKUEHLER, S., KURFISS, L., KAMMERER, D.R., WEISS, A., SCHMIDT, H., CARLE, R. 2013. Impact of different washing procedures on quality of fresh-cut iceberg lettuce (*Lactuca sativa* var. *capitata* L.) and endive (*Cichorium endivia* L.). *European Food Research and Technology*, 236(2), 229-241. <http://dx.doi.org/10.1007/s00217-012-1878-5>
- WULFKUEHLER, S., STARK, S., DIETZ, J., SCHMIDT, H., WEISS, A., CARLE, R. 2014. Effect of water jet cutting and moderate heat treatment on quality of fresh-cut red oak leaf lettuce (*Lactuca sativa* L. var. *crispa*). *Food and Bioprocess Technology*, 7(12), 3478-3492. <http://dx.doi.org/10.1007/s11947-014-1360-4>