JMBFS
Journal of Microbiology, Biotechnology and Food Sciences
International peer-reviewed scientific online journal

Published by Faculty of Biotechnology and Food Sciences

CHLOROPLAST EXCLUDING PRIMERS FOR METAGENOMIC ANALYSIS OF BACTERIA IN PLANT TISSUES

Renata Artimová[1], Lukáš Hleba[2], Soňa Javoreková[2] Jana Maková[2] Janka Medová[3] Juraj Medo[2]*

Address(es): Ing. Juraj Medo, PhD.
[1] University of Agriculture in Nitra, Research Centre AgroBioTech, Tr. A. Hlinku 2, 94976, Nitra, Slovakia.
[2] University of Agriculture in Nitra, Faculty of Biotechnology and Food Sciences, Institute of Biotechnology, Tr. A. Hlinku 2, 94976, Nitra, Slovakia.
[3] Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, 949 74 Nitra, Slovakia.

*Corresponding author: juraj.medo@uniag.sk

https://doi.org/10.55251/jmbfs.9650

ARTICLE INFO

Received 2. 6. 2022
Revised 17. 11. 2022
Accepted 30. 11. 2022
Published 1. 12. 2022

Regular article

OPEN ACCESS

ABSTRACT

Plant microbiomes are responsible for the growth, health, nutrition, and resistance to biotic and abiotic stress. Moreover, some plant microbiome members, especially the *Enterobacteriaceae* family, can act as human pathogens. In recent years, metagenomic analysis of amplified 16S rRNA gene sequenced on Illumina second generation sequencers became the most widely used method for bacterial microbiome studies. The 16S r RNA gene of the bacterium is phylogenetically similar to the 16S genes of chloroplasts and mitochondria since they have common prokaryotic ancestors. Thus, plant microbiome analysis is affected by unwanted non-target chloroplast and mitochondrial sequences. The solution to this contamination uses specially designed primer pairs that exclude chloroplasts and mitochondria and provide undistorted information. In this study, we analyzed chloroplast excluding primer 799R (reverse complement of 799F) which contains 4 mismatches of nucleotides against the chloroplast sequence modified for use with Illumina sequencers. We used primer 799R with universal bacterial forward primers 341F and 515F and compared the results obtained with 799R to the results obtained with the universal 806R primer. Microbiomes of 12 samples from 4 distinct parts of the tomato plants (*Solanum lycopersicon* L.) leaves, fruits, roots, and rhizosphere, were amplified using all combinations of these primers. The combination of universal forward primers 515F or 341F with the universal bacterial primer 806R amplified 60-85% of chloroplast sequences in samples of plant tissue. The 799R primer effectively reduced the number of chloroplast sequences to less than 1%. However, some sequences derived from mitochondria remained, with a higher proportion using the 341F primer (66-72%) than 515F (21-28%). The ability to describe microbial population diversity using 799R was similar to 806R, although there is a clear difference in the proportions of amplified microorganism groups using these primers. Analysis revealed the highest frequency of *Enterobacteriaceae* when primer pair 515F+799R was used. Based on the results, the combination of primers 515F+799R was the most appropriate, met all parameters, and can be recommended for routine analysis of plant microbiomes using second generation sequencers.

Keywords: Microbiome, Chloroplasts, PCR primer, 16S rRNA, Metagenomics, Solanum lycopersicum

INTRODUCTION

Plant-associated microbiome can be defined as a community of microorganisms inhabiting a particular niche on or within a plant. Plant microbiomes are important for plant growth, health, and productivity but they are also important from consumers' point of view (**Rastogi *et al.*, 2012**). Human pathogenic bacteria inhabiting the plants may pose a threat to consumers (**Ali *et al.*, 2017**). These bacteria can naturally inhabit the soil, or they are applied to the soil with organic fertilizers or untreated water (**Berg *et al.*, 2015**). Analysis of plant microbiomes is crucial for understanding how human pathogens can withstand in a plant environment, translocate, and internalize within plant tissues (**van Overbeek *et al.*, 2021**). Knowledge about plant microbiomes is still limited, due to technical issues associated with analysis. Microbiological techniques based on cultivation appear less effective because around 99% of microbial species in the soil/plant microbiome are not cultivable (**Rastogi & Sani, 2011**). Molecular techniques have changed the paradigm in the analysis of the microbial community and have become widely used in research (**Berg *et al.*, 2014**). High throughput parallel sequencing of PCR amplicons is the most widely used technique for microbiome analysis in recent years. Among next generation sequencers, Illumina platform is the most used for these purposes. Besides shotgun metagenomics, mainly metabarcoding techniques that use sequencing of DNA markers like 16S rRNA gene or ITS region are the most frequently used (**Gołębiewski & Tretyn, 2020**). This technique has a standardized workflow for the human or soil environment (**Thompson *et al.*, 2017**) but there are still some problems in plant associated studies. The sequencing of 16S PCR amplicons is used for bacterial community analysis because the 16S rRNA synthesis gene is present in all bacterial species and includes hypervariable regions that can be used for bacterial genera or species identification. Universal PCR primers with complementarity to conserved part of 16S rRNA were designed for amplification of variable regions resulting in identification of bacteria (**Caporaso *et al.*, 2011**; **Klindworth *et al.*, 2012**; **Parada *et al.*, 2016**). However, random mutations appear even in conserved areas, so none of these primers are absolutely universal. Thus, various primer pairs describe bacterial communities differently.

Main issue with 16S rRNA gene analysis in plant microbiomes is the similarity of bacterial 16S to the 16S rRNA of plant chloroplasts. Chloroplasts are organelles within plant cells responsible for photosynthesis. Endosymbiotic theory supposes that eukaryotic cells acquired photosynthetic bacteria which lost their ability to live outside host cells and specialized to photosynthesis (**Moreira *et al.*, 2000**). Similar process led to internalization of mitochondria (**Gray, 2012**). Chloroplast and mitochondria retain their circular DNA also with 16S rRNA gene which is very similar to the bacterial one. Both of them cannot be synthetized *de-novo* and must be inherited from parental cells and thus they are present in every plant cell including non-green cells. More than 200 chloroplasts and more than 10000 copies of chloroplast DNA can be found in a single plant cell, which is significantly more than the number of bacterial DNA associated with plant cells (**Morley & Nielsen, 2016**). As the universal primers for amplification of 16S rRNA gene successfully amplify plant DNA majority (70-99%) of acquired sequences may belong to chloroplast resulting to unreliable picture about bacterial community or acquisition of suitable results need high costs (**Medo *et al.*, 2018**; **Regalado *et al.*, 2020**; **Žiarovská *et al.*, 2020**). Various techniques were considered to solve the problem. Specific oligo blocker for elimination of chloroplast 16S rRNA amplification (**Mayer *et al.*, 2021**), peptide nucleic acid blocker (**Víquez-R *et al.*, 2020**), or suicide PCR (**Green & Minz, 2005**) were tried with various results. Size selection of amplified PCR fragments may help in lowering of chloroplast contamination (**Sakai *et al.*, 2004**). Change of universal primers to primers not amplifying chloroplasts may be a viable solution as it is easy to use and does not require additional steps in analysis.

Chelius and Triplett (2001) made pioneer work in the decrease of chloroplast contamination in 16S rRNA PCR from environmental samples when they found the region harboring 4 nucleotides with mismatch to chloroplast. This region contains 2 nucleotides on position 798-799 (numbering according to standard *Escherichia coli* 16S gene) and another 2 on positions 783 and 784. Primer containing these sequences named 799F in combination with different reverse

primers was successfully used for amplification of chloroplast free 16S rRNA gene regions **(Beckers *et al.*, 2016; Hanshew *et al.*, 2013; Tian & Zhang, 2017)**. However, this primer was only partially optimized for use with Illumina sequencer and direct comparison with other widely used universal primer combinations is missing.

The aims of this study are to design and modify primers capable of excluding chloroplast amplification from plant tissue during sequencing on Illumina Miseq platform. Then, evaluate the primer pairs according to their ability to describe diversity and composition of plant bacterial community with special attention to human pathogenic bacteria

MATERIAL AND METHODS

Preparation of samples

For evaluation of primers ability to describe plant bacterial microbiome, samples from tomato plants *Solanum lycopersicum L.* variety Cristal F1 were used. It is an indeterminate type of high yield, long shelf life hybrid with increased resistance to diseases. Plants were pre cultivated in pots with gardening substrate and then transplanted into soil. Soil on the experimental plot was clay-loam fluvisol with pH 6,5. Soil was fertilized using cow manure in dose 40 t/ha before plant transplantation. Samples were taken from 3 different plants (i. e. replications).

It is common to analyze various plant parts and surrounding environments, mainly soil associated, during experiments with plant microbiomes and it is important that primers should be usable in all these conditions. Thus, we used 4 types of sample types for primers evaluation. Rhizosphere samples were obtained from the soil attached to tomato roots. Plants were dug up, shaken from free soil, and transported to the laboratory. Then 10 g of roots with attached soil was washed in 20 ml of sterile saline solution in 50 ml tubes. Suspension was centrifuged and the pellet was used for DNA extraction. Then, all remaining soil was removed from roots using a brush, roots were 3 times washed by tap water, surface sterilized by 2% NaOCl for 2 minutes, and then 3 times washed by sterile distilled water. Surface sterilized roots were used for DNA extraction. Leaves and fruits were not surface sterilized. They were cut by a sterile scalpel to approx. 3 mm pieces before DNA extraction

Hundred mg of samples prepared as mentioned above were frozen in 2 ml microtubes and homogenized using 4 mm zirconium beads in Beadbug homogenizer (Benchmark scientifics, New Jersey, USA). DNA extraction was done using EZ 10 Column Plant Genomic DNA Purification Kit (Biobasics, Markham, Canada). Together 12 DNA samples were obtained.

Primers, PCR and sequencing

Primer 799R was designed as a simple reverse complement of forward primer 799F proposed by **Chelius and Triplett (2001)**. As it contains 2 chloroplast mismatched nucleotides on the 3` end chloroplast exclusion should be kept. Bioinformatics analysis using a probe test in SILVA database (testprobe tool) showed that 78,2% coverage for *Bacteria*, Primer 799R had 4 mismatches (positions 1,2,16,17) to all available tomato chloroplast in Genbank.

For selection of the best primer pair usable for plant microbiome assays we used 2 forward primers 341F and 515F and two reverse primers, 806R, and chloroplast excluding 799R (**table 1**). These combinations were used for amplification of V4 or V3-V4 variable regions of 16S rRNA of bacteria (**table 2**). Each primer combination was used for amplification of each DNA sample.

Primers were enhanced by a barcode which is used for recognition of samples within multiplexed reading of Illumina sequencers. Barcodes with different length 6-8 nt were used as the Illumina sequencers suffer from low quality of signal when large parts of parallel read sequences contain the same base in a certain position. To ensure that the barcode did not interfere with primer, 2 nt spacer mismatched to common bacterial sequences was used between primer and barcode.

Table 1 Primers used for amplification of plant microbiomes.

primer	Sequence	Author
341F	XXXXXXXX[1]-GG[2]-CCTACGGGNGGCWGCAG[3]	(**Klindworth *et al.*, 2012**)
515F	XXXXXXXX -GT-GTGCCAGCMGCCGCGGTAA	(**Caporaso *et al.*, 2011**)
806R	XXXXXXXX -CC-GGACTACHVGGGTWTCTAAT	(**Caporaso *et al.*, 2011**)
799R	XXXXXXXX -TA-CMGGGTATCTAATCCKGTT	This study

[1] barcode for identification of sample multiplex reading (different for each sample)
[2] spacer non-complementary to bacterial 16S rRNA gene
[3] Primer sequence complementary to bacterial 16S rRNA gene

Table 2 Primer pairs used in the study

Forward primer	Reverse primer	Approx. length of product	Chloroplast amplification
341F	806R	450 bp	Yes
341F	799R	450 bp	No
515F	806R	300 bp	Yes
515F	799R	300 bp	No

The same PCR mixture and reaction conditions were used for all primer combinations. 15 µl of PCR mixture contained 0,15 µl of Q5 HS HF DNA polymerase (New England Biolabs, Ipswich, USA), 3 µl of 5X Q5 reaction buffer, 0,3 µl of 2 mM dNTPs, 0.75 µl of 10µM forward and reverse primers and 1 µl od DNA. PCR ran on SureCycler 8800 (Invitrogen, Waltham, USA) with following conditions: initial denaturation 98°C for 30 s followed by 40 cycles consisting of denaturation 98°C for 3 s, annealing, 60°C for 20 s, and polymerization 72°C for 20 s.

Resulting PCR products were quality and length checked on agarose gel, purified by AMPURE XP (Beckman Coulter, Brea, USA) paramagnetic beads, quantified by Qubit fluorometric assay (Thermo, Waltham, USA) and mixed in equimolar ratio. Sequencing library was prepared by TruSeq LT PCR free kit (Illumina, San Diego, USA), where fragmentation and size selection were omitted as the PCR products were used as input DNA. Prepared library was quantified using NEBnext Library Quantification kit (New England Biolabs, Ipswich, USA), and sequencing was done using Illumina V3 2x300bp kit on an Illumina MiSeq system.

Data processing

Acquired sequences were analyzed in the SEED2 environment (**Vetrovský *et al.*, 2018**). Analysis was done separately for each primer combination because of different settings. Firstly, sequences were joined using join2fastqc. Fifty bp overlap and max 8% mismatch were set up for samples amplified with primer 515F (approx. 300 bp products), while only 40 bp overlap was set for samples with primer 341F. Unusually long or short sequences were removed. Only sequences with lengths in the range 400-500 bp for primer 341F and 250-350 bp for primer 515F were retained. Sequences with overall quality score lower than Q30 were removed from further analysis. Then individual samples were recognized according to the barcodes. Then, barcode and primer parts were removed from sequences. Sequences were checked for chimera using *de-novo* algorithm in Vsearch software (**Rognes *et al.*, 2016**) and chimeras were removed. Operational taxonomic units (OTUs) were generated by Vsearch at similarity level 97%. The most abundant sequence from each OTU was found and identified using the RDP classifier (**Wang *et al.*, 2007**).

Within each OTU table, chloroplast and mitochondrial sequences of the host (*Solanum* spp.) were identified on class level. Numbers of such sequences were used for analysis of primer`s ability to avoid host sequence amplification. Chloroplast and mitochondria sequences were discarded for analysis of alpha and beta diversity. Diversity was analyzed on genus level due to the fact that 341F primer combinations provide longer sequence reads and thus more OTUs. Sequences were rarified to the lowest count in the dataset for comparison of alpha diversity indices. Alpha diversity indices were computed in R (**R Core Team, 2019**), compared using ANOVA followed by Tukey test. Non-metric multidimensional scaling (NMDS) and permutational analysis of variance (PERMANOVA) were done in R using package vegan (**Oksanen *et al.*, 2013**). Comparison of equality of phyla detection was done by pairwise Wilcoxon signed-rank test in R.

RESULTS

Amount of bacterial and unwanted sequences

Totally 365871 sequences were successfully joined for all samples and 341236 (7109 per sample) of them was retained after quality filtering, length selection and chimera removal. Tested primer pairs showed very significant results regarding amplification of wanted (*Bacteria* and *Archaea*) and unwanted (mitochondria and chloroplast) sequences. However, their discrimination power is also affected by the sample type. Very small portions of chloroplast sequences and almost no mitochondria were detected in rhizosphere samples while *Archaea* were very sporadic in samples of leaves and fruits. On the other hand, chloroplast frequency became around 60-70 % in fruits and leaves samples when universal primer pairs were used (Tab 3).

Primers 515F+799R resulted in an increase of mitochondrial sequences frequency by 10% and 15 % in fruit and leaves samples, respectively. Enormous increase was found when 799R was used with 341F. Percentage of mitochondrial sequences increased from 18 to 67 % in fruit and from 23 to 77% in leaf samples.

Table 3 Relative frequencies of chloroplast, mitochondria, *Archaea* and *Bacteria* in samples of tomato rhizosphere, roots, leaves, and fruits amplified using universal and chloroplast excluding primer pairs.

Chloroplast	**341F+806R**		**515F+806R**		**341F+799R**		**515F+799R**		**Average**	
Fruits	64.6	Cb*	67.0	Cb	0.1	Aa	0.4	Aa	33.0	C
Leaves	72.3	Cc	61.4	Cb	0.0	Aa	0.0	Aa	33.5	C
Roots	24.5	Bb	21.7	Bb	0.4	Aa	0.0	Aa	11.7	B
Rhizosphere	1.5	Ac	0.6	Ab	0.0	Aa	0.0	Aa	0.5	A
Average	40.7	b	37.7	b	0.1	a	0.1	a		
Mitochondria	**341F+806R**		**515F+806R**		**341F+799R**		**515F+799R**		**Average**	
Fruits	16.0	Ba	21.2	Ba	66.4	Bb	21.9	Ba	31.4	B
Leaves	21.5	Ba	24.6	Ba	71.5	Bb	28.3	Ba	36.5	B
Roots	7.0	Aa	7.7	Aa	0.9	Aa	0.1	Aa	3.9	A
Rhizosphere	0.1	Aa	0.4	Ab	0.1	Aa	0.0	Aa	0.2	A
Average	11.2	a	13.5	a	34.7	b	12.6	a		
Archaea	**341F+806R**		**515F+806R**		**341F+799R**		**515F+799R**		**Average**	
Fruits	0.1	Aa	0.1	Aa	3.0	Aa	0.3	Aa	0.9	A
Leaves	0.2	AB	0.1	Aa	3.4	Aa	0.1	Aa	0.9	A
Roots	4.7	Ba	1.3	Aa	7.9	Aa	9.6	Aa	5.9	B
Rhizosphere	1.6	ABa	6.0	Bbc	3.4	Aab	7.9	Ac	4.7	AB
Average	1.6	a	1.9	a	4.4	a	4.5	a		
Bacteria	**341F+806R**		**515F+806R**		**341F+799R**		**515F+799R**		**Average**	
Fruits	19.3	Aa	11.8	Aa	30.5	Aa	77.4	Ab	34.8	A
Leaves	5.9	Aa	13.9	Aab	25.1	Ab	71.6	Ac	29.1	A
Roots	63.7	Ba	69.3	Ba	90.8	Ba	90.3	Aa	78.5	B
Rhizosphere	96.8	Cb	92.9	Bab	96.5	Bb	92.1	Aa	94.6	C
Average	46.4	a	47.0	a	60.7	b	82.9	c		

Averages accompanied with the same letter (uppercase for columns; lowercase for rows) are not statistically significantly different (ANOVA; Tukey test; α=0,05; n=3)

Alpha and beta diversity

Analysis of alpha diversity (i.e. Shannon's index) was done after removal of chloroplast and mitochondria sequences. It revealed differences between sample types (ANOVA; P=0.0116) but also between primer pairs (P=0.0073). The lowest average value was found after the use of 341F+806R primer pair (Tab 4). Insignificantly higher value was obtained using 341+799R. Both were significantly lower than 5.64 and 6.01 values for pairs 515F +806R and 515F+799R, respectively. Lower indices were related to the amount of chloroplast and mitochondria sequences. In samples where their frequencies were high, the low number of remaining *Bacteria* / *Archaea* sequences resulted in lower diversity indices.

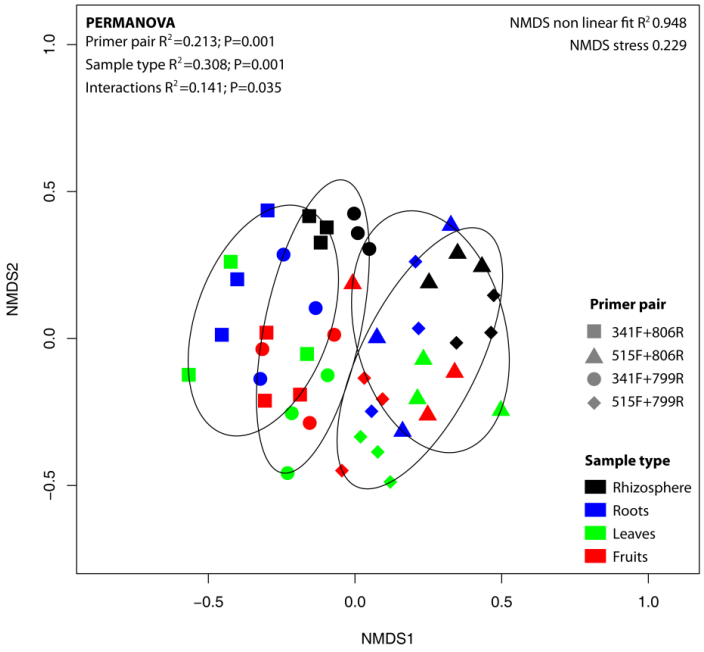

Figure 1 NMDS plot of microbiome similarity in four types of plant associated samples analyzed by sequencing of 16S rRNA gene amplified using universal and chloroplast excluding primer pairs.

Table 4 Shannon's indices of diversity for bacterial/archeal microbiome of tomato rhizosphere, roots, leaves, and fruits analyzed using universal and chloroplast excluding primer pairs.

	341F+806R	**515F+806R**	**341F+799R**	**515F+799R**	**Average**
Fruits	4.17Aa*	5.53Aab	5.7ABab	6.27Ab	5.59AB
Leaves	4.88Aa	5.10Aa	4.79Aa	5.72Aa	4.95A
Roots	5.22Aa	5.76Aa	5.90Ba	5.90Aa	5.70B
Rhizosphere	6.07Ab	6.18Ab	4.81ABa	6.12Ab	5.80B
Average	5.08a	5.64ab	5.30ab	6.01b	

*Averages accompanied with the same letter (uppercase for columns; lowercase for rows) are not statistically significantly different (ANOVA; Tukey test; α=0,05; n=3)

Community of bacteria amplified during PCR showed significant dependence on primer combination (PERMANOVA P<0,001; R2=0,21) despite the effect of sample type was stronger (P<0,001; R2=0,31). According to the NMDS analysis (Fig 1), all primer combinations were able to clearly distinguish communities among sample types in a very similar way, but effect of primer combination is visible.

Equality of microbial group detection

In detailed examination, we found that certain primers increase or decrease amplification of 16S rRNA gene from particular groups of microorganisms (Fig 2). Primers 341F+799R overestimated *Saccharibacteria* phyla in the rhizosphere while 341+806R underestimated *Proteobacteria* and overestimated *Firmicutes* in leaves. Using pairwise analysis (Tab 5), we analyzed which group of microorganisms is differently amplified between primers. Differences between universal primer pairs were low, only in 3 of 22 phylogenetic groups. Using chloroplast excluding primer 799R instead of universal 806R primer resulted in significant change of 5 and 6 bacterial phyla for primer 341F and 515F respectively.

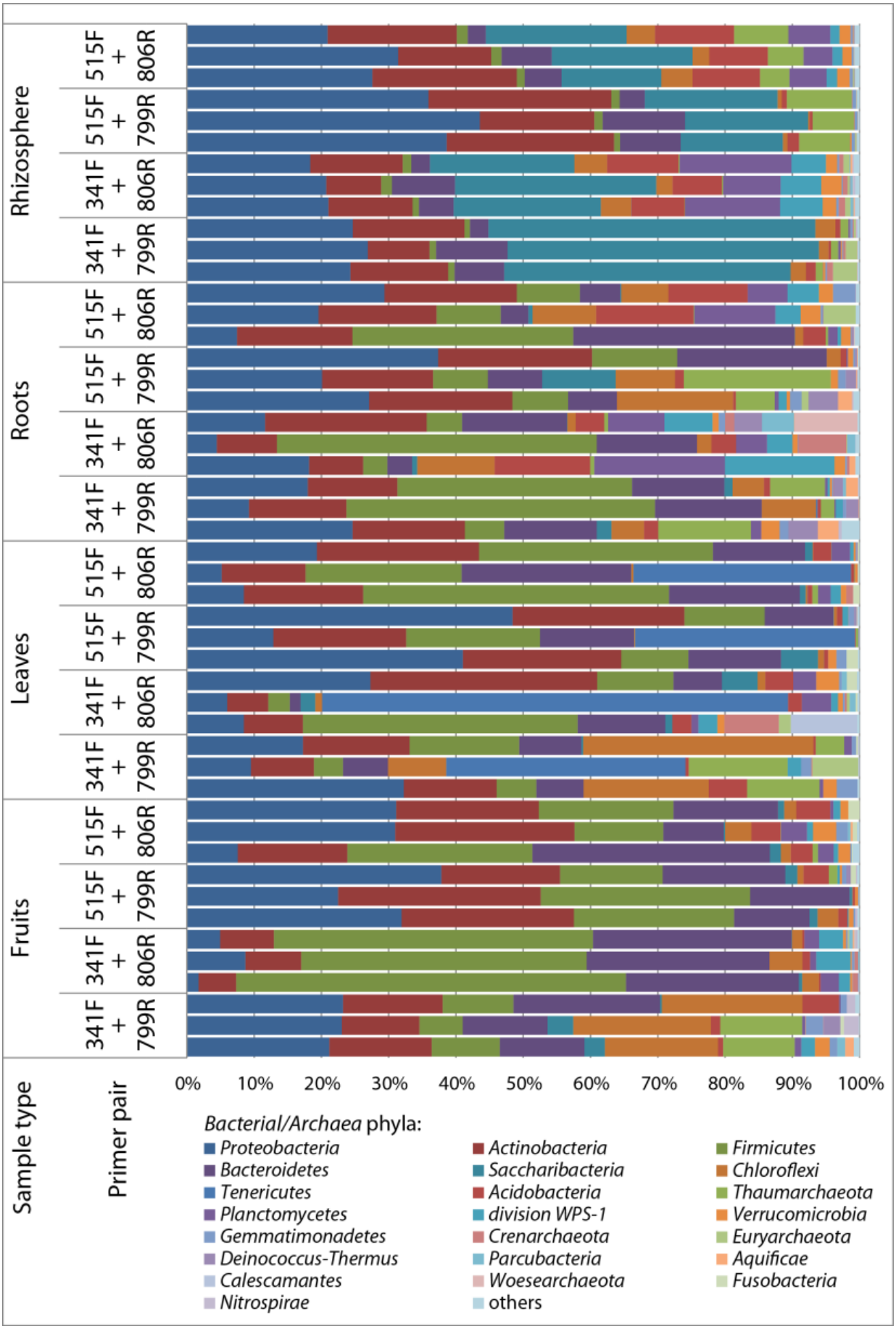

Figure 2 Frequency of *Bacteria*/*Archaea* phyla in tomato associated samples analyzed by sequencing of 16S rRNA gene amplified using universal and chloroplast excluding primer pairs.

Table 5 Significance of difference in frequency of phyla detection between primer pairs (P values; Paired student T-test n=12)

Phylum	Statistically tested combination of primer pairs					
	341F+806R vs. 515F+806R	**341F+806R vs. 341F+799R**	**341F+806R vs. 515F+799R**	**341F+799R vs. 515F+806R**	**341F+799R vs. 515F+799R**	**515F+806R vs. 515F+799R**
Acidobacteria	0.361	0.051	0.017	0.016	0.199	0.002
Actinobacteria	0.009	0.498	0.001	0.001	0.001	0.005
Aquificae	0.231	0.127	0.548	0.117	0.087	0.357
Bacteroidetes	0.674	0.412	0.722	0.314	0.557	0.470
candidate.division.WPS.1	0.019	0.006	0.002	0.048	0.404	0.015
Candidatus.Calescamantes	0.330	0.338	0.330	0.339	0.386	0.339
Candidatus.Saccharibacteria	0.093	0.091	0.617	0.063	0.142	0.376
Crenarchaeota	0.050	0.076	0.052	0.832	0.085	0.204
Deinococcus.Thermus	0.384	0.357	0.650	0.069	0.333	0.140
Euryarchaeota	0.889	0.305	0.189	0.461	0.170	0.447
Firmicutes	0.597	0.161	0.100	0.314	0.982	0.157
Fusobacteria	0.452	0.465	0.777	0.132	0.389	0.402
Gemmatimonadetes	0.221	0.030	0.075	0.277	0.205	0.808
Chloroflexi	0.956	0.023	0.846	0.023	0.029	0.933
Nitrospirae	0.116	0.141	0.549	0.090	0.200	0.014
Parcubacteria	0.085	0.125	0.07	0.378	0.213	0.339
Planctomycetes	0.113	0.004	0.002	0.006	0.003	0.002
Proteobacteria	0.051	0.008	0.001	0.708	0.001	0.003
Tenericutes	0.339	0.342	0.339	0.308	0.303	0.364
Thaumarchaeota	0.075	0.002	0.046	0.040	0.470	0.160
Verrucomicrobia	0.601	0.156	0.039	0.048	0.479	0.001
Woesearchaeota	0.332	0.330	0.330	0.339	0.218	0.339

***Enterobacteriaceae* resolution**

Differences were found in detection of possible human pathogenic bacteria from the *Enterobacteriaceae* family. Fourteen genera of *Enterobacteriaceae* were found among all samples across all primer combinations. Relative frequency and numbers of detected genera were dependent on sample and primer combination (Tab 6). The most prevalent genera were *Enterobacter*, *Escherichia*/*Shigella*, *Klebsiella*, *Citrobacter*, and *Raoultella*. They were found in all types of samples including leaves and fruits. Primer combination 341F+806R was not able to detect *Enterobacter* in all samples except the single rhizosphere. On the other hand, primer 341F allows deeper identification of bacteria as the sequences are longer, thus it can provide better resolution of genera within *Enterobacteriaceae* family.

Table 6 Relative frequencies and number of detected genera (in parentheses) of *Enterobacteriaceae* family in samples of tomato rhizosphere, roots, leaves, and fruits amplified using universal and chloroplast excluding primer pairs.

Sample type	341F+806R	515F+806R	341F+799R	515F+799R
Fruits	0.33 (5)	0.81 (5)	0.23 (5)	0.79 (3)
Leaves	0.13 (1)	0.34 (4)	0.09 (2)	0.74 (4)
Roots	0.25 (3)	0.22 (5)	0.22 (6)	0.58 (1)
Rhizosphere	0.2 (4)	0.12 (1)	0.13 (2)	0.32 (2)

DISCUSSION

The main purpose of this study was to evaluate the use of chloroplast exclusive primer 799R and compare it to common primer 806R. As it was predicted, use of the chloroplast excluding primer 799R resulted in decrease of chloroplast amplification and their frequency was lowered almost to zero values. Similar results were achieved by **Beckers *et al.* (2016)** who tested primers 799F and 783Rabc (mixture of 3 primers targeting the same positions like 799F) for analysis of bacterial community in rhizosphere roots, stem and leaves of hybrid poplars (*Populus tremula* x *P. album*) using pyrosequencing. They favored combinations 799F+1391R or 799F+1193R over combination 341F+783Rabc. However, primers 341F and 515F are now preferred for analysis of soil microbiome in standardized protocols like Earth microbiome project (**Thompson *et al.*, 2017**). There is also a limitation of length of the reading on Illumina platform to maximum approx. 450 bp which disqualifies primer 799F in combination with any reverse primer binding on position higher than 1200 (numbered based on *Escherichia coli* standard genome). **Thijs *et al.* (2017)** tested the combination of 341F with reverse primer 785R (**Klindworth *et al.*, 2012**) which is similar to 806R shifted by 2 bases and 96.1% of acquired sequences belonged to bacteria. High coverage of this combination should lead to the most informative results. However, we found strong affinity of these primers to chloroplast and also mitochondrial 16S rRNA gene amplification. Mitochondria amplification was apparent in samples processed using universal primers and it even increased after using chloroplast excluding 799R primer. Potential solution may be using primers which also decrease mitochondrial amplification. Primers 335F+769R developed bioinformatically by **Dorn-In *et al.* (2015)** were evaluated by **Nakano (2018)** for amplification of 16S rRNA on sample of potato salad and animal feed. Primers successfully excluded chloroplasts and mitochondria from amplification in comparison to 341F+785R. However, authors offered only limited insight to performance of these primers and their ability to amplify bacteria involved in soil-plant relationships are not clear.

Samples amplified with 341F primer shoved the lowest diversity index which was probably caused by lower number of sequences. It was caused by lower base quality on the end of reads that did not allow correct joining of forward and reverse reads. It is in contrast with other studies where samples amplified with 341F+785R repeatedly showed high diversity indices (**Fadeev *et al.*, 2021**; **Liu *et al.*, 2020**; **Thijs *et al.*, 2017**) and our results seem to be affected by lower read quality. However, more than 2000 sequences were still detected in these samples. Despite number of sequences acquired by next generation sequencer (sequencing depth) may affect results of microbiome research (**Sanchez-Cid *et al.*, 2022**) even less than 1000 sequences correctly described diversity and composition of microbial community (**Dully *et al.*, 2021**; **Shirazi *et al.*, 2021**).

There are not any primers capable to describe bacterial community exactly because there is not a position in 16 S rRNA gene where all members of community share the same sequence. In all studies comparing primers for microbial community description some differences were found (**Nakano, 2018**; **Parada *et al.*, 2016**; **Thijs *et al.*, 2017**). In spite of differences in community between primer pairs in our study, the main members of the community were detectable using all primer pairs and they reflect the actual bacterial community in samples. For example, one sample of leaves was apparently attacked by stolbur phytoplasma despite it was not symptomatic. All primer pairs were able to amplify it and 3 pairs confirmed phytoplasma in a similar extent, i.e. around 33%, while combination 341F+806R showed almost 70%. Generally, some loss of accuracy of description is acceptable trade-off for exclusion of chloroplast sequences which may account for more than 90% of sequences and make the analysis impossible.

We particularly targeted *Enterobacteriaceae* family of bacteria as it is considered to be the most common human pathogens able to survive in plant and surrounding soil (**Holden *et al.*, 2009**). They are considered an integral part of plant microbiomes and besides direct harmful effects to human health they also act as reservoirs of antibiotic resistance (**Cernava *et al.*, 2019**). Primer 515F showed ability to amplify higher portion of *Enterobacteriaceae* sequences thus its combination 515F+799R seems to be a good candidate for analysis of human pathogenic bacteria in plant microbiome. Although 515F forward primer provided more sequences of *Enterobacteriaceae*, 341F offered better genera resolution. However, PCR with primers 341F+799R requests the use of 2x300bp Illumina sequencing kit and results are strongly affected by the quality of sequence reads.

CONCLUSION

In present study we proved that primer 799R can be used in microbiome analysis using sequencing of 16 rRNA gene on Illumina platform. Despite small differences in community description ability of tested primer pairs they main purpose, i.e. exclusion of chloroplast amplification, was successful. Combination 515F+799R appears to be more universal with less amplification of mitochondria while combination 341F+ 799R may be used for longer read sequencing on Illumina platform.

Acknowledgments: This research was supported by project of Ministry of Education, Science, Research and Sport of the Slovak Republic, grant no. VEGA 1/0661/19 "Plant microbiome and safe food". Authors would thank Natália Žácka for her help in laboratory.

REFERENCES

Ali, M. A., Naveed, M., Mustafa, A., & Abbas, A. (2017). The good, the bad, and the ugly of rhizosphere microbiome *Probiotics and plant health* (pp. 253-290): Springer.https://doi.org/10.1007/978-981-10-3473-2_11

Beckers, B., Op De Beeck, M., Thijs, S., Truyens, S., Weyens, N., Boerjan, W., & Vangronsveld, J. (2016). Performance of 16s rDNA Primer Pairs in the Study of Rhizosphere and Endosphere Bacterial Microbiomes in Metabarcoding Studies. *Frontiers in Microbiology, 7*. https://doi.org/10.3389/fmicb.2016.00650

Berg, G., Erlacher, A., & Grube, M. (2015). The Edible Plant Microbiome: Importance and Health Issues. In B. Lugtenberg (Ed.), *Principles of Plant-Microbe Interactions: Microbes for Sustainable Agriculture* (pp. 419-426). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-08575-3_44

Berg, G., Grube, M., Schloter, M., & Smalla, K. (2014). Unraveling the plant microbiome: looking back and future perspectives. *Frontiers in Microbiology, 5*. https://doi.org/10.3389/fmicb.2014.00148

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., . . . Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the national academy of sciences, 108*(Supplement 1), 4516-4522. https://doi.org/10.1073/pnas.1000080107

Cernava, T., Erlacher, A., Soh, J., Sensen, C. W., Grube, M., & Berg, G. (2019). Enterobacteriaceae dominate the core microbiome and contribute to the resistome of arugula (Eruca sativa Mill.). *Microbiome, 7*(1), 13. https://doi.org/10.1186/s40168-019-0624-7

Dorn-In, S., Bassitta, R., Schwaiger, K., Bauer, J., & Hölzel, C. S. (2015). Specific amplification of bacterial DNA by optimized so-called universal bacterial primers in samples rich of plant DNA. *J Microbiol Methods, 113*, 50-56. https://doi.org/10.1016/j.mimet.2015.04.001

Dully, V., Wilding, T. A., Mühlhaus, T., & Stoeck, T. (2021). Identifying the minimum amplicon sequence depth to adequately predict classes in eDNA-based marine biomonitoring using supervised machine learning. *Computational and Structural Biotechnology Journal, 19*, 2256-2268. https://doi.org/10.1016/j.csbj.2021.04.005

Fadeev, E., Cardozo-Mino, M. G., Rapp, J. Z., Bienhold, C., Salter, I., Salman-Carvalho, V., . . . Boetius, A. (2021). Comparison of Two 16S rRNA Primers (V3–V4 and V4–V5) for Studies of Arctic Microbial Communities. *Frontiers in Microbiology, 12*. https://doi.org/10.3389/fmicb.2021.637526

Gołębiewski, M., & Tretyn, A. (2020). Generating amplicon reads for microbial community assessment with next-generation sequencing. *Journal of Applied Microbiology, 128*(2), 330-354. https://doi.org/10.1111/jam.14380

Gray, M. W. (2012). Mitochondrial evolution. *Cold Spring Harb Perspect Biol, 4*(9), a011403. https://doi.org/10.1101/cshperspect.a011403

Green, S. J., & Minz, D. (2005). Suicide polymerase endonuclease restriction, a novel technique for enhancing PCR amplification of minor DNA templates. *Appl Environ Microbiol, 71*(8), 4721-4727. https://doi.org/10.1128/aem.71.8.4721-4727.2005

Hanshew, A. S., Mason, C. J., Raffa, K. F., & Currie, C. R. (2013). Minimization of chloroplast contamination in 16S rRNA gene pyrosequencing of insect herbivore bacterial communities. *J Microbiol Methods, 95*(2), 149-155. https://doi.org/10.1016/j.mimet.2013.08.007

Holden, N., Pritchard, L., & Toth, I. (2009). Colonization outwith the colon: plants as an alternative environmental reservoir for human pathogenic enterobacteria. *FEMS Microbiology Reviews, 33*(4), 689-703. https://doi.org/10.1111/j.1574-6976.2008.00153.x

Chelius, M. K., & Triplett, E. W. (2001). The Diversity of Archaea and Bacteria in Association with the Roots of Zea mays L. *Microbial Ecology, 41*(3), 252-263. https://doi.org/10.1007/s002480000087

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2012). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research, 41*(1), e1-e1. https://doi.org/10.1093/nar/gks808

Liu, P. Y., Wu, W. K., Chen, C. C., Panyod, S., Sheen, L. Y., & Wu, M. S. (2020). Evaluation of compatibility of 16S rRNA V3V4 and V4 amplicon libraries for clinical microbiome profiling. BioRxiv. https://doi.org/10.1101/2020.08.18.256818

Mayer, T., Mari, A., Almario, J., Murillo-Roos, M., Syed M. Abdullah, H., Dombrowski, N., . . . Agler, M. T. (2021). Obtaining deeper insights into microbiome diversity using a simple method to block host and nontargets in amplicon sequencing. *Molecular Ecology Resources, 21*(6), 1952-1965. https://doi.org/10.1111/1755-0998.13408

Medo, J., Žiarovská, J., Medová, J., Javoreková, S., Kyseľ, M., & Hricová, A. (2018). Endophytic bacterial diversity decrease in amaranth mutant lines after radiation mutagenesis. *Cereal Chemistry, 95*(1), 109-116. https://doi.org/10.1002/cche.10006

Moreira, D., Le Guyader, H., & Philippe, H. (2000). The origin of red algae and the evolution of chloroplasts. *Nature, 405*(6782), 69-72. https://doi.org/10.1038/35011054

Morley, S. A., & Nielsen, B. L. (2016). Chloroplast DNA Copy Number Changes during Plant Development in Organelle DNA Polymerase Mutants. *Frontiers in Plant Science, 7*. https://doi.org/10.3389/fpls.2016.00057

Nakano, M. (2018). 16S rRNA Gene Primer Validation for Bacterial Diversity Analysis of Vegetable Products. *Journal of Food Protection, 81*(5), 848-859. https://doi.org/10.4315/0362-028x.jfp-17-346

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R., . . . Wagner, H. (2013). Package 'vegan'. *Community ecology package, version, 2*(9).

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology, 18*(5), 1403-1414. https://doi.org/10.1111/1462-2920.13023

R Core Team. (2019). *R: A language and environment for statistical computing*: R Foundation for Statistical Computing, Vienna, Austria.

Rastogi, G., & Sani, R. K. (2011). Molecular Techniques to Assess Microbial Community Structure, Function, and Dynamics in the Environment. In I. Ahmad, F. Ahmad & J. Pichtel (Eds.), *Microbes and Microbial Technology: Agricultural and Environmental Applications* (pp. 29-57). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-7931-5_2

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. https://doi.org/10.1038/ismej.2012.32

Regalado, J., Lundberg, D. S., Deusch, O., Kersten, S., Karasov, T., Poersch, K., . . . Weigel, D. (2020). Combining whole-genome shotgun sequencing and rRNA gene amplicon analyses to improve detection of microbe–microbe interaction networks in plant leaves. *The ISME Journal, 14*(8), 2116-2130. https://doi.org/10.1038/s41396-020-0665-8

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ, 4*, e2584. https://doi.org/10.7717/peerj.2584

Sakai, M., Matsuka, A., Komura, T., & Kanazawa, S. (2004). Application of a new PCR primer for terminal restriction fragment length polymorphism analysis of the bacterial communities in plant roots. *J Microbiol Methods, 59*(1), 81-89. https://doi.org/10.1016/j.mimet.2004.06.005

Sanchez-Cid, C., Tignat-Perrier, R., Franqueville, L., Delaurière, L., Schagat, T., & Vogel, T. M. (2022). Sequencing Depth Has a Stronger Effect than DNA Extraction on Soil Bacterial Richness Discovery. *Biomolecules, 12*(3), 364. https://doi.org/10.3390/biom12030364

Shirazi, S., Meyer, R. S., & Shapiro, B. (2021). Revisiting the effect of PCR replication and sequencing depth on biodiversity metrics in environmental DNA metabarcoding. *Ecology and evolution, 11*(22), 15766-15779. https://doi.org/10.1002/ece3.8239

Thijs, S., Op De Beeck, M., Beckers, B., Truyens, S., Stevens, V., Van Hamme, J. D., . . . Vangronsveld, J. (2017). Comparative Evaluation of Four Bacteria-Specific Primer Pairs for 16S rRNA Gene Surveys. *Frontiers in Microbiology, 8*. https://doi.org/10.3389/fmicb.2017.00494

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., . . . The Earth Microbiome Project, C. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature, 551*(7681), 457-463. https://doi.org/10.1038/nature24621

Tian, X.-Y., & Zhang, C.-S. (2017). Illumina-Based Analysis of Endophytic and Rhizosphere Bacterial Diversity of the Coastal Halophyte Messerschmidia sibirica. *Frontiers in Microbiology, 8*. https://doi.org/10.3389/fmicb.2017.02288

van Overbeek, L., Duhamel, M., Aanstoot, S., van der Plas, C. L., Nijhuis, E., Poleij, L., . . . Andreo-Jimenez, B. (2021). Transmission of Escherichia coli from Manure to Root Zones of Field-Grown Lettuce and Leek Plants. *Microorganisms, 9*(11), 2289. https://doi.org/10.3390/microorganisms9112289

Vetrovský, T., Baldrian, P., Morais, D., & Berger, B. (2018). SEED 2: a user-friendly platform for amplicon high-throughput sequencing data analyses. *Bioinformatics, 1*, 3. https://doi.org/10.1093/bioinformatics/bty071

Víquez-R, L., Fleischer, R., Wilhelm, K., Tschapka, M., & Sommer, S. (2020). Jumping the green wall: The use of PNA-DNA clamps to enhance microbiome sampling depth in wildlife microbiome research. *Ecology and evolution, 10*(20), 11779-11786. https://doi.org/10.1002/ece3.6814

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology, 73*(16), 5261-5267. https://doi.org/10.1128/aem.00062-07

Žiarovská, J., Medo, J., Kyseľ, M., Zamiešková, L., & Kačániová, M. (2020). Endophytic Bacterial Microbiome Diversity in Early Developmental Stage Plant Tissues of Wheat Varieties. *Plants, 9*(2), 266. https://doi.org/10.3390/plants9020266