Assessing and Monitoring
the Impacts of Genetically modified plants on Agro-ecosystems

Workpackage No. 4: WP leader: Prof. Dr. Christoph Tebbe

Thünen Institute of Biodiversity, Federal Research Institute for Rural Areas, Forestry and Fisheries,

Braunschweig, Germany

Task 4.1 Soil metagenomics: Baselines for bacterial diversity and functional genes

Task leader: Prof. Dr. Christoph Tebbe, Thünen Institute of Biodiversity

Deliverable 4.1 – Report on the structural diversity of soil bacteria

Astrid Näther, Anja Dohrmann, and Christoph C. Tebbe*

Thünen Institute of Biodiversity, Federal Research Institute for Rural Areas, Forestry and Fisheries,

Braunschweig, Germany

*, corresponding author

May 2016

Dissemination level

PU Public X

Summary

Bacteria are the most abundant form of life on Earth. In soil, they represent a highly diverse group of organisms and they contribute with their metabolic versatility to biogeochemical processes and the sustainable agricultural use. They occur in relatively high densities in the immediate vicinity of plant roots, i.e., the rhizosphere, where they utilize carbon and energy supplied by the plants and, in turn, support plant growth. Within the "Soil fertility" work package (WP4) of the AMIGA project, an important objective was to utilize the new potentials of high-throughput DNA sequencing technologies for characterizing the diversity of rhizosphere bacterial communities of maize and potato, as cultivated in pairs of GM and non-GM at different field sites across Europe. The structural diversity of the communities were characterized from DNA, directly extracted from root material with adhering soil particles collected at the AMIGA field sites, on basis of the 16S ribosomal RNA (rRNA) genes amplified by PCR with phylogenetically highly conserved primers. The chosen primers potentially amplified genes of both prokaryotic domains, i.e. *Bacteria* and *Archaea*. To gain insight into the functional diversity, PCR-amplified *nirS* and *nirK* genes were analyzed. These genes encode for structurally different but functionally equivalent nitrite reductases. The rhizosphere samples came from the AMIGA maize and potato fields located in Spain, Slovakia, Denmark, Sweden, The Netherlands and Ireland, with each represented by two or three annual replications (2012 - 2014). Due to vandalism, sampling of the GM cultivar in The Netherlands was only possible in 2014. All samples of the AMIGA project collected for DNA sequencing came from crops at their flowering stage. Overall, the microbial communities as obtained by directly extracted DNA from rhizospheres of 200 maize and 152 potato samples were analyzed. A total of 10.6 Million quality filtered 16S rRNA amplicon was retrieved from maize and 5.1 Million from potato, respectively. For *nirK* 2.5 Million were obtained from maize and 1.3 Million from potato, while the number of quality sequences for *nirS* was much lower, due to technical problems with the sequencing (6,000 from maize, 4,127 from potato). At the level of operational taxonomic units (OTUs), which is an estimate for species, the

library coverage for maize and potato rhizospheres was 98 %, indicating that most of the prokaryotic community was included. A total of 13,081 prokaryotic OTUs were detected from maize and 9,693 from potato. The 10 most dominant OTUs of maize represented 27 % of the total community, for potato the corresponding value was 17 %. A core of 87 OTUs were consistently detected in maize rhizospheres, independent of the site, year or cultivar, for potato the core was only 28 OTUs, despite the fact that less field sites were analyzed. For *nirK*, a total of 1,942 OTUs was detected with maize, and 1,712 OTUs with potato. A single maize sample contained on average 12,275 OTUs, and for potato 9,151 OTUs, respectively. The ten most dominant OTUs represented 39 % (maize) and 19 % of all sequences. Considering both, 16S rRNA and *nirK* gene amplicons, no differences in the overall bacterial (prokaryotic) community structure were detected by the respective genetic modifications of maize and potato, respectively. In contrast, a clear distinction was possible for the particular sites, years of cultivation and cultivars. This deliverable informs about the preliminary taxonomic affiliation of the prokaryotic DNA sequences (16S rRNA, *nirK*, and *nirS* genes) retrieved in the AMIGA project.

1. Introduction

Prokaryotes include the taxonomic domains *Bacteria* and *Archaea*. In soil, bacterial cell numbers typically outnumber those of archaea by two or more orders of magnitude (see for example (1)). Both soil bacteria and archaea are highly diverse, metabolically versatile and capable of performing different biogeochemical transformations of inorganic and organic compounds. The rhizosphere, i.e. the soil compartment which is affected by plant roots, bacteria are also more abundant than archaea, and, together with the eukaryotic fungi (see also AMIGA Deliverable 4.2), they utilize carbon and energy sources released by the plant root cells. In turn for this plant investment, the enriched microbial communities provide the potential to support plant growth by facilitating the access to soil nutrients (nitrogen, phosphate, and iron), buffering of stress (as caused by drought or salinity),

defending against soil-borne plant pathogens (2). Thereby, the rhizosphere microbial communities directly affect plant growth and these interactions have a high agricultural potential.

The bacterial diversity in rhizospheres is characteristic for each plant species, but it is also influenced by the soil properties and environmental conditions under which the plants actually grew (3). In fact, the diversity of plant bacterial communities even changes with the growth stage (age) of the plant, indicating bacterial sensitivity to the quality of substrates provided by the roots (4). Therefore, it is not surprising that during the last two decades several studies on non-target effects of genetically modified (GM) plants have looked at potential differences between a GM and conventionally bred comparator cultivars (see for example (5, 6)). It was assumed that the release of recombinant proteins, e.g. the Cry1Ab protein, or even unintended effects on plant metabolism would be picked up by comparing the structural diversity of the bacterial (or fungal) communities.

Typically, most GM risk assessment studies focusing on microbial communities denied microbiological methods based on cultivation approaches, since the majority of soil microorganisms do not grow on common nutrient laboratory media. The most applied method for analyzing bacterial communities has been genetic fingerprinting, commonly based on profiling the diversity of PCR-amplified 16S ribosomal RNA (rRNA) gene fragments, which are then visualized by various electrophoretic detection methods (7). These methods usually did not reveal differences between GM and comparators, but they were limited to analyzing the dominant 20 to 50 community members ("phylotypes"). In order to identify the contributors to such communities, cloning and sequence of the PCR-amplicon DNA-sequences was required.

In the "soil fertility" workpackage of the AMIGA project, we analyzed the diversity of bacteria (and fungi, see D4.2) with the powerful techniques of high-throughput DNA sequencing, thereby increasing the sensitivity compared to the profiling techniques by several orders of magnitude, and, in addition, providing an assessment of the phylogenetic diversity contributing to these communities. This deliverable D4.1 reports about the diversity of bacteria found with maize and potatoes collected

at the AMIGA sites. This Alpha-diversity, i.e., the richness of “species” and their abundance in different samples, is complemented in this report with an initial view of Beta-diversity, which considers differences between communities as affected by field sites, cultivars (genetic background), spraying of fungicides (for potato only), and the genetic modifications.

The basis of the analyses was the DNA directly extracted from roots with adhering soil particles collected from the different AMIGA field sites. These were for maize located in Spain, Slovakia, Denmark, and Sweden, and for potato in The Netherlands and Ireland. Each site was sampled with two or three annual replications during the years 2012 to 2014. Due to vandalism, sampling of GM in The Netherlands for 2012 however was incomplete. All samples of the AMIGA project collected for DNA sequencing came from crops at their flowering stage in order to minimize variability caused by plant age.

The rhizosphere DNA was subjected to *in vitro* amplification by PCR, targeting the 16S rRNA genes for the assessment of the bacterial (prokaryotic) diversity and the *nirS*/*nirK* genes to assess the functional diversity, i.e. for the nitrite reductase, an enzyme involved in denitrification, the conversation of nitrate to nitrogen gas (N_2). It should be noted that prokaryotic species are not naturally defined by their biological characteristics, but by an arbitrary definition. While the 16S rRNA gene is the best marker gene to characterize the phylogenetic diversity, it cannot identify “species”. As an approximate for species, microbiology in general, and this report in particular, uses “OTUs”, i.e., operational taxonomic units, and assigns sequences which share a predefined threshold of sequence identity, e.g. for rRNA gens 97 %.

Both *nir*-genes encode for a structurally different but functionally equivalent version of the enzyme nitrite reductase, and they occur scattered within similar or different phylogenetic groups, indicating that the gene cannot directly be linked to phylogeny of the host organisms (8). The PCR amplicons did not cover the full gene lengths but significant fractions of the 16S rRNA or *nirS*/*nirK* sequence with a targeted size of approximately 251 base pairs, bp, for the 16S rRNA amplicons, 436

bp for *nirK*, and 382 bp for *nirS*, respectively. Paired-end multiplex sequencing was conducted using the Illumina MiSeq technology (9).

2. Materials and Methods

2.1 Location and design of the AMIGA field sites

Field sites for maize cultivation were located in Spain, Slovakia, Denmark and Sweden; field sites for potato were in Ireland and in The Netherlands. More information about the sites can be found in other AMIGA documents. It should be noted that in the subsequent years, the order of treatments and number cultivars were not altered. However, for the accurate positions of the potato plots at the Irish and Dutch sites were rotated within the respective sites. This was also done for maize at the site in Slovakia. For the other maize field sites, i.e., in Sweden, Denmark and Spain, the plots were in an identical position, with the exception of Sweden in 2012 and 2013, where plots 2, 4, 17, and 19 were not identical in regard to growing GM or conventional.

Maize event MON810 confers the capacity to produce the insecticidal Cry1Ab protein of *Bacillus thuringiensis* and provides resistance against the European corn borer *Ostrinia nubilalis* and other lepidopteran pests. It was cultivated as the genetically modified (GM) plant and compared to near isogenic non-modified cultivars. For Spain, the GM of event MON810 was the cultivar DKC6451YG (YG stands for “Yield Guard”, a brand name of Monsanto), and its near isogenic cultivar DKC6450, whereas in Slovakia, Sweden and Denmark, the MON810 event was DKC3872YG (Yield Guard), and the near isogenic variety was DKC3871. All field trials consisted of a randomized block design including 20 plots with maize and 10 replicates for each variety. The plot size was 10 x 10 m with an interspace of 5 m. The field site was bordered by a 5-m wide strip of conventional maize.

The GM potato cultivated at the sites in The Netherlands and Ireland was the cis-genic event A15-031 (Desirée + Vnt1), which confers resistance to the fungus *Phytophtora infestans*, the cause of

late blight. The cis-genic potato line was generated *via* insertion of the *Rpi-vnt1-1* gene into the genome of *Solanum tuberosum* cv. Desirée. The *Rpi-vnt1-1* gene was originally derived from the wild potato species *Solanum venturii* with its native promoter and terminator sequences. Thus, the gene was only expressed at exposure to *P. infestans*.

With the exception of 2012, the potato field trials in Ireland and in The Netherlands included three potato genotypes: the GM A15-031 (Desire + Vnt1), its conventional susceptible comparator (Desirée) and Sarpo Mira, a conventional cultivar. In 2012, a preliminary field evaluation was completed in Ireland on smaller plots (1m x 1m) and the conventional cultivar included in this evaluation was the variety King Edward. For the 2013 and 2014 potato study, rhizospheres from two treatments were analyzed: No spraying (**NS**) against late blight, and, conventional protection (**CP**) by weekly applications of preventive fungicides for late blight control. The randomized block design relevant to this study included 42 plots with potato, seven replicates for each of the three varieties and the two late blight control strategies. The 2012, the preliminary study in Ireland only contained a non-spray (NS) treatment and four replicates, thus a total of 12 plots. For 2013 and 2014, the plot size was 3 x 3 m in Ireland, and 6 x 6 m in The Netherlands, both with interspaces of 6 m. Field trials were sampled in the years 2012, 2013 and 2014, respectively. However due to vandalism, the site in The Netherlands could only fully be sampled in 2012.

2.2 Sampling strategy

The rhizosphere samples were collected during the flowering period of the respective crops to minimize the well-documented effect that plant age can modify the composition of the microbial communities in the rhizospheres (10, 11). Plants were carefully dug out of each plot and transferred immediately to the laboratory. Loosely adhering soil was removed by shaking. Microbial cells, spores and mycelia adhering to the roots of each plant were detached by suspending the fresh root material (8 g) in 30 ml of sterile saline for 30 min at 4 °C in an orbital shaker (Model 3040, GFL, Burgwedel,

Germany) at 10 rpm. The microbial cells were collected by centrifugation at 4,100 x *g* for 30 min at 4 °C and the cell pellets were stored at -80°C.

2.3 DNA extraction and purification

Total DNA was extracted from the frozen cell pellets using the FastDNA SPIN kit for soil (MP Biomedicals, Illkirch, France). The extraction included two bead beating steps (45 s at 6.5 m s^{-1}) on a FastPrep-24 (MP Biomedicals, Eschwege, Germany) and an additional washing of the binding matrix with 1 ml 5.5 M guanidine thiocyanate (Carl Roth, Karlsruhe, Germany). Extracted DNA was split into aliquots, one for bacterial sequencing and one for sequencing fungal genes (Deliverable 4.2).

2.4 Illumina library generation

The V4 region of the 16S rRNA gene (251 bp) was amplified using primers S-D-Arch-0519-a-S-15 (5'-**CAGCMGCCGCGGTAA**-3') and S-D-Bact-0785-a-A-21 (5'-**GACTACHVGGGTATCTAATCC**-3') covering about 85% and 83% of Bacteria and Archaea, respectively (12). Partial *nirK* (436 bp) and *nirS* (382 bp) genes were amplified using primers F1aCu (**ATCATGGTSCTGCCGCG**) / *nirK*1040R (**GCCTCGATCAGRTTRTGGTT**) and *nirS*Cd3a-F (**AAACGYSAAGGARACSGG**) / *nirS*R3cd-R (**GASTTCGGRTGSGTCTTSAYGAA**), respectively (13-15). The custom primers contained adapters specific to the Illumina flow cell, a unique 8-bp-index to allow for multiplexing, a 10-bp sequence to adjust the sequencing primer melting temperature to about 65 °C, a 2-bp link anti-complementary to known sequences and regions complementary to the conserved portions of the respective gene. The use of up to 12 forward and 30 reverse primers allowed for multiplexing of up to 360 samples by dual-indexing (9) for a single sequencing run. Two PCR amplifications were carried out with the FastStart High Fidelity PCR System (Roche Diagnostics, Mannheim, Germany) for each sample, using 50 µl reaction mixtures. Each reaction mixture contained 1 µl template DNA, 0.4 µM of each primer,

200 µM of each dNTP, 5% dimethyl sulfoxide and 2.5 U FastStart High Fidelity Enzyme Blend in a 1X reaction buffer containing 1.8 mM $MgCl_2$. The PCR conditions involved an initial denaturation step at 95°C for 2 min followed by 35 cycles of 95°C for 30 s, 50°C for 30 s, and 72°C for 1 min and ended with an extension step at 72°C for 5 min. Products from the two independent replicate amplifications were united and purified from agarose gels following the respective protocol of the HiYield PCR Clean-up & Gel-Extraction kit (SLG, Südlaborbedarf GmbH, Gauting, Germany) and quantified with the Quant-iT PicoGreen dsDNA assay (Invitrogen, Darmstadt, Germany). Equimolar amounts of the individual PCR products were pooled and sent to the company StarSEQ (Mainz, Germany) for 250 or 300-bp paired-end multiplex sequencing on a MiSeq. For 16S rRNA genes, two runs were completed; one with 140 maize samples of 2012 and 2013 (250-bp paired-end run) and another one with 60 maize samples of 2014 and 12, 56 and 84 potato samples of 2012, 2013 and 2014, respectively (300-bp paired-end run). For *nirK* and *nirS*, one 300-bp paired-end run with all 352 samples was accomplished, respectively. The three sequencing primers for each run were composed of (a) the forward pad, link and primer for read1; (b) the reverse pad, link and primer for read2; and (c) the reverse complement of the reverse pad, link and primer for the index read.

2.5 Sequence filtering and analysis

Reads were merged with VSEARCH (vsearch v1.9.5_linux_x86_64, 31.4GB RAM, 24 cores) (github.com/torognes/vsearch) asking for a minimum length of the overlap of 50 nucleotides (nt) and a minimum length of the merged read of 200 nt. Sequences with total expected errors E > 1 were discarded with the fastq_filter command of VSEARCH. Furthermore, the screen.seqs command of mothur (mothur v.1.31.2) (16) was used to remove sequences with any ambiguous base or more than six homopolymers and to select sequences in the size range of 251 – 252 nt. Thus, good quality sequences were retained. VSEARCH was applied to remove sequences that appeared only once (singletons) or chimeras which were identified by de novo chimera detection using the UCHIME

algorithm (17). Sequences for rRNA gene amplicons were clustered in OTUs (operational taxonomic units) with VSEARCH at a threshold of 97% sequence identity and for each cluster a consensus sequence was generated. For *nirK* and *nirS* amplicons, species-level threshold distances of 17 % and 18 % were applied (18). Sequences that represented ribosomal RNA genes were extracted with Metaxa2 (19). These curated sequences were taxonomically classified with mothur using the RDP reference database trainset14_032015 (20). The dataset of curated rRNA sequences was used as reference database to map all good quality sequences against it at a threshold of 97% identity applying USEARCH (usearch v8.1.1831_i86linux32) (21) with the usearch_global command. Good quality sequences were mapped to the seed. For phylogenetic comparisons of the bacterial communities the curated sequences were aligned with mothur and FastTree version 2.1.7 SSE3 (22) was used to calculate a phylogenetic tree. Weighted unifrac distances of the phylogenetic tree and NMDS were calculated with mothur.

3. Results

3.1 Sampling effort (overview)

For maize, a total of 200 individual samples were collected and analyzed for bacterial diversity with 16S rRNA gene, *nirS* and *nirK* amplicons, respectively. These originated from four AMIGA sites. The sites in Slovakia and Sweden were analyzed with three annual replicates, while the sites in Denmark and Spain were analyzed with two. Except for Spain for all years and Sweden for the first year of sampling (2012) the isogenic cultivars and, thus, the genetic background for MON810 analyzed, were identical (**Table 1A.**). Samples were collected from 10 independent replicates of both, the GEM cultivar and the isogenic comparator.

Table 1 A. Samples collected for the prokaryotic community analyses of **maize** rhizospheres according to field sites and cultivars, for sequencing of 16S rRNA genes as well as *nirK/nirS* genes

Site	Sampling date	Cultivar (BT, ISO)	No. of replicates	Total
Spain	July 24, 2012	DKC6451YG, DKC6450	10, 10	20
	July 24, 2013	DKC6451YG, DKC6450	10[a], 10	20
Slovakia	July 10, 2012	DKC3872YG, DKC3871	10, 10	20
	July 30, 2013	DKC3872YG, DKC3871	10[b], 10[a]	20
	July 21, 2014	DKC3872YG, DKC3871	10, 10	20
Sweden	August 15, 2012	DKC4442YG, DKC440	10[c], 10[b]	20
	August 12, 2013	DKC3872YG, DKC3871	10[d], 10[d]	20
	August 12, 2014	DKC3872YG, DKC3871	10, 10	20
Denmark	August 26, 2013	DKC3872YG, DKC3871	10, 10	20
	August 13, 2014	DKC3872YG, DKC3871	10, 10	20
Total				**200**

for *nirS* only [a]5 samples, [b]7 samples, [c]8 samples, [d]9samples

B. Samples collected for the prokaryotic community analyses of **potato** rhizospheres based on sequencing of 16S rRNA genes and *nirK/nirS* genes. GM, genetically modified (cis-genic); ISO, isogenic cultivar; CON, conventional cultivar; NS, no fungicides; CP, conventional protection by chemical treatments (fungicides)

Sampling date	Treatment	Cultivar	Replicates	Total number
Ireland				
October 18, 2012				
	NS	GM, ISO, CON	4, 4/3**, 4	12
August 13, 2013				
	CP	GM, ISO, CON	7, 7, 7/6**	21
	NS	GM, ISO, CON	7, 7, 7	21
July 28, 2014				
	CP	GM, ISO, CON	7/6**, 7/6**, 7	21
	NS	GM, ISO, CON	7, 7, 7	21
The Netherlands				
August 7, 2013				
	CP	CON	7	
	NS	CON	7	14
July 15, 2014				
	CP	GM,ISO, CON	7/6**, 7, 7/4*,**	21/18*
	NS	GM, ISO, CON	7/6*/5**, 7/2**, 7/6*/5**	21/19*
Total				**152/147***

*, refers to samples utilized for *nirK*-amplicon sequencing, **, refers to samples from which *nirS* amplicon sequences were obtained

For potato, a sampling was conducted in Ireland in 2012 with only four replicates and no spray of fungicides. In 2013 and 2014, the sampling in Ireland was extended to 7 replicated plots and in addition to no spray there was also conventional practice included. The sampling in The Netherlands in 2013 only included 7 replicates of the isogenic cultivar, since the cis-genic was not available after site vandalism. In contrast, the sampling in 2014 occurred as described for Ireland in the years 2013 and 2014, respectively (**Table 1B.**). More details on the number of sequences analyzed for each sample can be found in the Appendix section (Tables S1 – S4).

3.2 Prokaryotic diversity found in maize rhizospheres

3.2.1 Structural diversity (16S rRNA genes)

The number of raw sequences obtained from the 200 samples (libraries) was 13.444.961, and after quality filtering, 10,593,739 sequences remained. A sample included on average 52,969 ±20,597 sequences, with the smallest library of 20,841 and the largest with 118,337 sequences, respectively. The total number of OTUs was 13,081; the smallest number in a library was 1,464, the largest 4,706 OTUs (average 2,933 ± 741). At the OTU-level the library coverage (Good's coverage) was 97.7 ± 0.8 %. To allow comparisons, the rarefied libraries (subsamples) contained 20,841 sequences. Thus, a total of 4,168,200 sequences represented the full dataset with 200 samples. After subsampling, Good's coverage declined to 95.7 ± 0.7 %. The number of OTUs was accordingly 12,439 of which 41 were actually assigned by phylogenetic analyses to the domain *Archaea* (representing 0.3 % of all OTUs) while the other 99.7 % fell into the domain *Bacteria*.

Phylogenetic analyses of the amplicon sequences revealed that the most dominant phylum was *Proteobacteria*, which included more 54 % of all sequences, followed by *Actinobacteria* with 25 %. All other phyla were represented by less than 5 % (10 most dominant shown in **Table 2**). Only eight phyla included more than 1 % of all sequences. Overall, 27 phyla, among them two archaeal,

were detected. The most abundant phylum from the domain Archaea was *Thaumarchaeota*, representing 2 % of all sequences. At the class level, *Gammaproteobacteria* were most abundant with 30.1 % followed by *Actinobacteria* (23.2 %), *Alphaproteobacteria* (13.7 %), and *Betaproteobacteria* (7.2 %). *Nitrosphaerales* were the most dominant archaea with 2 %, which thus accounted for almost all archaeal sequences.

Table 2 Diversity of the ten most dominant prokaryotic phyla detected in maize rhizospheres as characterized by 16S rRNA gene sequencing

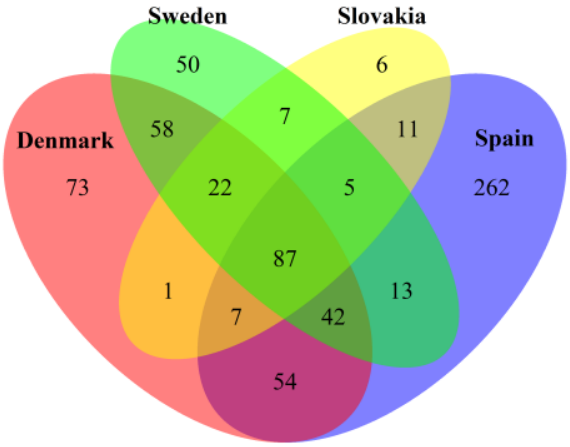
No.	Domain	Phylum	% of all sequences	Standard dev.
1	*Bacteria*	*Proteobacteria*	54.1	16.7
2	*Bacteria*	*Actinobacteria*	24.7	9.8
3	*Bacteria*	*Unclassified*	4.5	2.2
4	*Bacteria*	*Firmicutes*	4.5	1.9
5	*Bacteria*	*Bacteroidetes*	3.1	2.3
6	*Bacteria*	*Acidobacteria*	3.0	1.6
7	*Archaea*	*Thaumarchaeota*	2.4	1.6
8	*Bacteria*	*Verrucomicrobia*	1.9	1.0
9	*Bacteria*	*Planctomycetes*	0.9	0.5
10	*Bacteria*	*Gemmatimonadetes*	0.8	0.4

The ten most abundant OTUs represented 26.8 % of all sequences, the twenty most abundant 33.5 % and the fifty most abundant 44.5 %, respectively. Of the ten most abundant OTUs, seven were members of the *Gammaproteobacteria* and the tree others belonged to *Alphaproteobacteria*, *Betaproteobacteria*, and *Actinobacteria*. All these sequences were 100 % identical to sequences in the databases and were affiliated e.g. to *Pseudomonas*, *Enterobacter*, *Pantoea*, *Acinetobacter*, *Rahnella*, *Arthrobacter*, *Sphingobium*, or *Telluria*/*Massilia*, respectively.

Overall, 87 OTUs were always detected in the maize rhizosphere libraries, thus, these occurred independent of the field site, cultivar or year of cultivation. Considering the total of 12,439 OTUs detected, these represented only 0.7 %. However, considering their abundance (number of sequences within the total sequences of 4.2 Million sequences), the core contributed 37.3 + 12.4 % of the total community. Thus, approximately one third of the total community was shared

independent of the location and the other above mentioned factors. A Venn diagram (**Figure 1**) indicates the number of shared OTUs between sites, independent of the year of cultivation or GM or near-isogenic comparator.

Figure 1 Venn diagram of OTUs retrieved from maize rhizospheres, as detected from different field sites in the AMIGA project



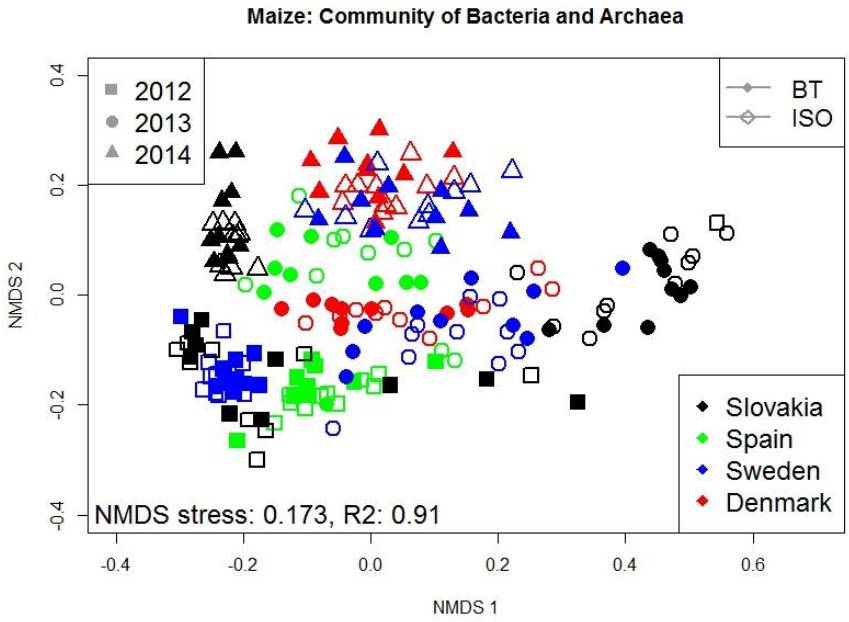


Figure 2 NMDS plots displaying the similarity/dissimilarity of the bacterial community structures of samples taken from maize rhizospheres cultivated at the four AMIGA field sites during 2012 to 2014.

Spain had the highest site-specific richness of site-specific OTUs (n=262), Slovakia the lowest (n=5!). The highest number of shared OTUs was, not surprisingly, between Denmark and Sweden (n=58) and, surprisingly, Denmark and Spain (n=54), while there between Sweden and Spain only 13 OTUs were shared (in addition to the overall core of 87 OTUs).

Multivariate-statistical analyses were applied to characterize the similarities of the overall prokaryotic community compositions from the respective samples (libraries). NMDS revealed that there was a clear clustering according to the field sites and a strong annual variation (Fig. 2). In contrast, there was no clear distinction between samples taken from the GM MON810 and their near-isogenic comparators. In some cases, like for Slovakia but also for Sweden, the annual variation appeared to be even stronger than the variation between sites from different geographical regions.

3.2.2 Functional diversity of nitrite reductases

3.2.2.1 Diversity of bacterial denitrifiers (*nirK*) - Maize

Out of 3.6 Million sequences obtained from *nirK* gene amplicons, a total of 2.5 Million quality filtered sequences were obtained. On average of the 200 individual samples, each contained 12,275 $\pm$ 4,550 ranging from 3,673 to 29,647 sequences for the individual samples. The dataset was subsampled to 3,000 sequences for further comparative analyses.
All sequences could be clustered into 1,942 OTUs at the species level, joining sequences of $\geq$ 83 % sequence identity. The average OTU number was 178 $\pm$ 42 (73 to 349) per rhizosphere sample. A coverage of 97.8 % indicated that not all sequences were captured.

The ten most dominant OTUs of the *nirK* library from all samples (rhizospheres) combined represented 39.6 % of all sequences obtained (Table 4). All sequences were most closely related to corresponding genes of *Rhizobiaceae* (*Alphaproteobacteria*), and affiliated to either *Mesorhizobium*, *Ensifer* (including the former genus *Sinorhizobium*), and *Bradyrhizobium*, respectively.

Table 3 DNA sequence similarity of the most dominant *nirK* OTUs amplified from maize rhizosphere DNA to known sequences in the Genbank database – preliminary taxonomic assignment

OTU No.	% of all sequences	Best BLAST hit	Sequence identity (%)
1	10.4	*Mesorhizobium* sp.	88
2	5.4	*Bradyrhizobium* sp.	91
3	5.1	*Mesorhizobium* sp.	95
4	4.3	*Ensifer adhaerens*	98
16	2.8	*Ensifer* sp.	96
11	2.6	*Mesorhizobium* sp.	87
5	2.3	*Bradyrhizobium* sp.	86
9	2.3	*Bradyrhizobium* sp.	91
57	2.3	*Mesorhizobium* sp.	100
7	2.1	*Ensifer* sp.	96

3.2.2.2 Diversity of bacterial denitrifiers (*nirS*)

Only 6,000 quality filtered sequences were obtained from *nirS* gene amplicons. This was less than 0.2 % of the sequences generated for *nirK* genes. The average number of sequences for each sample was 34, ranging from 3 to 163. Each rhizosphere sample contained an average of 12 OTU, ranging from 1 (excluding samples were no sequences were obtained) to a maximum of 38 (see also Appendix section). Applying BLAST searches, all dominant OTUs identified were most closely related to uncultured bacteria, the vast majority from soil, with relatively distant relationships to genera like *Azospirillum* (OTU1; *Alphaproteobacteria*), *Pseudomonas* (OTU2; *Gammaproteobacteria*), *Alicycliphilus* (OTU11; *Betaproteobacteria*), *Paracoccus* (OTU17; *Alphaproteobacteria*), or *Azoarcus* (OTU23; *Betaproteobacteria*), respectively. These genera are all not unusual in soil and known to have denitrifying capacity, as encoded by *nir* genes.

3.3 Prokaryotic diversity found in potato rhizospheres

3.3.1 Structural diversity (16S rRNA genes)

The number of raw sequences obtained from the 152 samples (libraries) was 7,249,532, and after quality filtering, 5,093,198 sequences remained. On average, a single sample was composed of 33,508 (± 9,338) sequence, with a minimum of 14,694 and a maximum of 61,198 sequences. At the OTU-level the library coverage (Good's coverage) was 97.6 ± 0.7 %. To allow comparisons, the rarefied libraries (subsamples) contained 14,694 sequences, which resulted in a total of 2,223,488 sequences remaining. After subsampling, Good's coverage declined to 95.7 ± 1.2 %. The number of OTUs was accordingly 9,323 of which 27 were actually assigned by phylogenetic analyses to Archaea (representing 0.3 % of all OTUs) while the other 99.7 % belonged to the bacterial domain. A single sample contained on average 2,042 OTUs (minimum 630; maximum 3491).

Table 4 Diversity of the ten most dominant prokaryotic phyla detected in potato rhizospheres as characterized by 16S rRNA gene sequencing

No.	Domain	Phylum	% of all sequences	Standard dev.
1	*Bacteria*	*Proteobacteria*	52.4	13.8
2	*Bacteria*	*Actinobacteria*	25.5	11.2
3	*Bacteria*	*Bacteroidetes*	7.1	4.1
4	*Bacteria*	*Firmicutes*	6.0	4.2
5	*Bacteria*	*unclassified*	3.1	1.8
6	*Bacteria*	*Acidobacteria*	2.3	1.7
7	*Bacteria*	*Verrucomicrobia*	1.0	0.7
8	*Bacteria*	*Planctomycetes*	0.8	0.5
9	*Bacteria*	*Gemmatimonadetes*	0.7	0.7
10	*Bacteria*	*Cyanobacteria/Chloropl.*	0.5	0.7

Phylogenetic analyses allowed the assignment of DNA sequences to 24 different bacterial and 2 different archaeal phyla. The most dominant phylum was *Proteobacteria*, which included more 52 % of all sequences, followed by *Actinobacteria* with 26 %, *Bacteriodetes* with 7 %, and *Firmicutes* with 6 %, respectively (**Table 4**). The most abundant phylum from the domain Archaea was

Thaumarchaeota, representing 0.4 % of all sequences. At the class level, *Actinobacteria* were most abundant with 23.9 % followed by *Alphaproteobacteria* (21.3 %), *Gammaproteobacteria* (19.8 %), *Betaproteobacteria* (9.4 %) and *Bacilli* (5.5 %), respectively. *Nitrosphaerales* were the most dominant archaea with 0.4 %, thus, accounting for the vast majority of archaeal sequences.

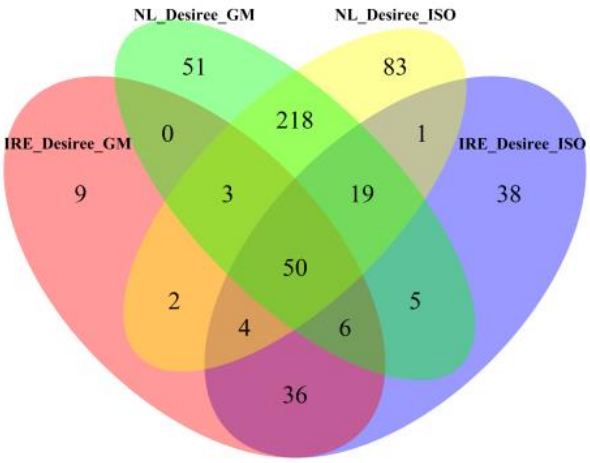
The ten most abundant OTUs accounted for 17.0 % of all sequences, the twenty most abundant 25.5 % and the fifty most abundant 38.4 %, respectively. Of the 10 most abundant OTUs, 7 were members of the *Proteobacteria*, among them four Gammaproteobacteria, two Alpha- two *Betaproteoteobacteria*, and two *Actinobacteria*, respectively. Except for one Bacillus sequence, which had 99 % sequence identity, all others were 100 % identical to sequences in the databases.

In contrast to maize, were five field sites were equally represented by samples (except some sites with two and some with three annual replications), the majority of samples for potato originated from Ireland, thus underrepresenting the other site located in The Netherlands. The identification of shared OTUs thus does not equally reflect the bacterial "core" as identified with maize. Nevertheless it gives an indication about the consistency of OTUs as they were found in the libraries for potato in this project, and thus they are reported here below.

The bacterial core, i.e., the OTUs which consistently were detected in potato rhizospheres independent of the cultivar, site or year of cultivation or application of fungicides, consisted of 28 OTUs, all within the bacteria domain. The five most dominant bacterial OTUs could be assigned to Arthrobacter (Actinobacteria; on average representing 3.1 % of all sequences), Sphingobium (Alphaproteobacteria; 2.0 %), Pseudomonas (Gammaproteobacteria; 2.0 %), Bacillus (Firmicutes, 1.7 %), and a member of the Oxalobacteraceae (Betaproteobacteria; 1.4 %), respectively.

A Venn diagram (**Figure 3**) indicates the number of shared OTUs at the Irish and Dutch sites between different cultivars independent of their year and cultivation.

Figure 3 Venn diagrams of OTUs retrieved from potato rhizospheres, as detected from the two field sites in the AMIGA project and according to the cultivar/ or genetic modification (cis-genesis). A. Differences according to cultivars/genetic modification as detected in Ireland. B. Differences between sites and cultivars



The highest numbers of OTUs (218) were shared between the GM and isogenic version of Desiree cultivated in The Netherlands. Surprisingly for the Irish site there were only 36 shared OTUs between both. Fifty OTUs were shared between all treatments and sites.

Non-metric multidimensional scaling (NMDS) was applied as a multivariate-statistical analysis to visualize the importance of field site, cultivar, genetic modification and chemical treatments of potatoes on the overall prokaryotic community structure in this AMIGA project (**Figure 4**). NMDS clear highlights the importance of the location of the field site, and, for Ireland, the impact of the different years of cultivation. In contrast, there were no pronounced differences visualized between the conventional variety, or the isogenic and cis-genic, respectively. Furthermore, the application of fungicides did not result in separation of the prokaryotic community structures within each site and sampling event. This NMDS-based interpretation was confirmed by ANOSIM, with highly significant results for separations according to site and year, but not the other factors.

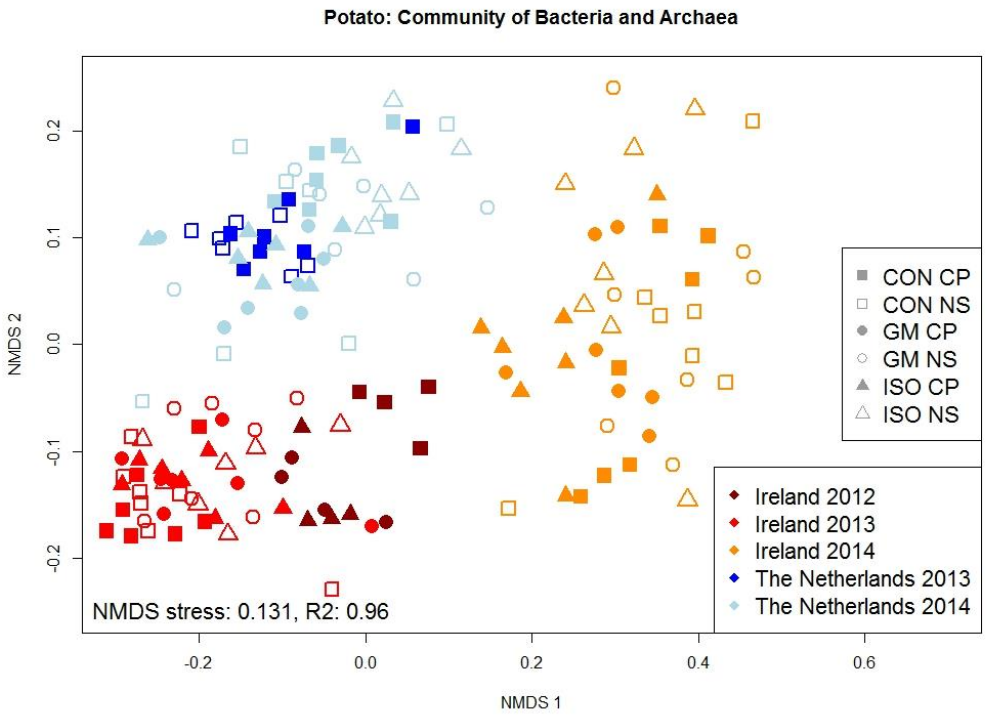


Figure 4 Non-metric multidimensional scaling (NMDS) for displaying the similarity/dissimilarity of the bacterial community structures of samples taken from potato rhizospheres cultivated in Ireland and The Netherlands

3.3.2 Functional diversity of nitrite reductases

3.3.2.1 Diversity of bacterial denitrifiers (*nirK*)

A total of 1.3 Million *nirK* sequences were obtained from the 147 samples included in this analysis. On average each sample contained 9,151 ± 3,186 sequences, with a minimum of 3,333 and a maximum of 16,244 sequences. The dataset was subsampled to 3,333 sequences for further comparative analyses. A total of 1,712 OTUs were retrieved from the dataset joining sequences of > 83 % sequence identity (corresponding to the species level). The average number of OTU was 180 ± 50 and ranged from 91 to 374. The average coverage was 98.0 ± 0.7 %. Thus, the sequences captured the major proportion, but not all sequences expected in these communities.

The ten most dominant OTUs of the *nirK* library from all samples (rhizospheres) combined showed that these represented 18.7 % of all sequences (**Table 5**). All sequences fell into the family *Rhizobiaceae* (*Alphaproteobacteria*) and were only represented by three genera, i.e., *Mesorhizobium*, *Ensifer* (including the former genus *Sinorhizobium*), and *Bradyrhizobium*, respectively.

Table 5 DNA sequence similarity of the most dominant *nirK* OTUs amplified from potato rhizosphere DNA to known sequences in the Genbank database – preliminary taxonomic assignment

OTU No.	% of all sequences	Best BLAST hit	Sequence identity (%)
19	4.8	*Mesorhizobium* sp.	92
1	2.4	*Bradyrhizobium* sp.	90
80	1.8	*Ensifer adhaerens*	98
14	1.7	*Ensifer adhaerens*	96
2	1.5	*Mesorhizobium* sp.	90
95	1.5	*Bradyrhizobium* sp.	86
3	1.4	*Mesorhizobium huakuii*	88
34	1.3	*Ensifer adhaerens*	100
11	1.2	*Bradyrhizobium* sp.	88
4	1.1	*Mesorhizobium* sp.	89

Sequencing of the *nirS* amplicons was almost a complete failure: Only 4,127 sequences were obtained for all samples together (see also Appendix). This was only 0.3 % of the sequences which were obtained with the *nirK* gene in the corresponding samples. Each rhizosphere sample contained an average of 32 sequences, ranging from 3 (excluding samples were no sequences were obtained) to a maximum of 291 (Table A7). The number of OTUs in the different samples ranged from 1 to 27. Only one of the 10 dominant OTUs showed the highest sequence identity to a cultivated isolate (*Pseudomonas* sp., *Gammaproteobacteria*, OTU1, 92 %). All other nine OTU were most closely related uncultured bacteria, the vast majority from soil, three of them with an indicated relationship

to *Cupriavidus* (OTU 5 and OTU10; *Betaproteobacteria*), or *Alicycliphilus* denitrificans (OTU7; *Betaproteobacteria*), respectively.

Short discussion & Conclusions

The vast majority of bacterial diversity was captured from the rhizospheres of maize and potato cultivated at the different AMIGA field sites. Library coverage values indicated that at the OTU level (approx. "species" level), above 97 % of the diversity was captured. A similar coverage was achieved for the *nirK* gene, but, due to technical problems, *nirS* gene sequences were less representative.

The maize dataset represents a solid view of bacterial diversity as it includes full data sets (MON810 compared to isogenic) with ten independent replicates, two to three annual repetitions and four different sites. In contrast the dataset for potato is less complete, one two sites were available and of those only one included a solid annual replication, because the other was destroyed by human "activists".

The datasets from the respective rhizospheres suggests that the diversity of bacterial community members is slightly higher in maize than potato, even though these comparisons have to consider that maize originated from five different locations, and potato only from two (**Table 6**). The ten most dominant OTUs were quantitatively more important for maize than potato, indicating a stronger influence of maize on the bacterial community structure. This was also indicated by the *nirK* gene, where the ten most dominant OTUs were twice abundant with maize than with potato.

Table 6 A. Summary of 16S rRNA data, comparison of maize and potato

	Maize	Potato
16S rRNA genes		
Total number of quality sequences	10.6 Million	5.1 Million
Total number of OTUs	13,081	9,693
Number of OTUs per rhizosphere	2,933 ± 741	2,042 ± 679
Ten most dominant OTUs represent …	26.8 %	17.0 %
nirK		
Total number of quality sequences	2.5 Million	1.3 Million
Total number of OTUs	1,942	1,712
Number of OTUs per rhizosphere	178± 42	180 ± 50
Ten most dominant OTUs represent …	39.6 %	18.7 %

Overall, the datasets elaborated for this deliverable demonstrate the influence of the plant species on the selection of bacterial communities in their rhizospheres. Independent of the site of cultivation a core of bacterial species (OTUs) is maintained. The abundance of the core community members can vary but they occur consistently. Each individual site contributes to the other part of the community which is variable. This variability is caused by a combination of factors, including physicochemical soil properties, climatic conditions, weather conditions during the year of cultivation, history of agricultural soil use and management practices. Consistent, site dependent differences between the GM varieties and their comparators could not be identified

Acknowledgements

We thank all persons involved in the AMIGA field trials for setting-up and maintaining the field experiments and supporting our work during sampling and sample preparations, namely Cristina Chueca (Spain), Tina d'Hertefeldt (Sweden), Ludovit Cagan (Slovakia), Gabor Lovey (Denmark), Bert Lotz (The Netherlands), Ewen Mullins (Ireland) and their respective teams.

Cited references

1. **Neumann D, Heuer A, Hemkemeyer M, Martens R, Tebbe CC.** 2013. Response of microbial communities to long-term fertilization depends on their microhabitat. FEMS Microbiology Ecology **86:**71-84.

2. **Mendes R, Garbeva P, Raaijmakers JM.** 2013. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. FEMS Microbiology Reviews **37:**634-663.

3. **Hartmann A, Schmid M, van Tuinen D, Berg G.** 2009. Plant-driven selection of microbes. Plant and Soil **321:**235-257.

4. **Marques JM, da Silva TF, Vollu RE, Blank AF, Ding GC, Seldin L, Smalla K.** 2014. Plant age and genotype affect the bacterial community composition in the tuber rhizosphere of field-grown sweet potato plants. FEMS Microbiology Ecology **88:**424-435.

5. **Schmalenberger A, Tebbe CC.** 2003. Genetic profiling of noncultivated bacteria from the rhizospheres of sugar beet (*Beta vulgaris*) reveal field and annual variability but no effect of a transgenic herbicide resistance. Canadian Journal of Microbiology **49:**1-8.

6. **Baumgarte S, Tebbe CC.** 2005. Field studies on the environmental fate of the Cry1Ab Bt-toxin produced by transgenic maize (MON810) and its effect on bacterial communities in the maize rhizosphere. Molecular Ecology **14:**2539-2551.

7. **Smalla K, Oros-Sichler M, Milling A, Heuer H, Baumgarte S, Becker R, Neuber G, Kropf S, Ulrich A, Tebbe CC.** 2007. Bacterial diversity of soils assessed by DGGE, T-RFLP and SSCP fingerprints of PCR-amplified 16S rRNA gene fragments: Do the different methods provide similar results? Journal of Microbiological Methods **69:**470-479.

8. **Heylen K, Gevers D, Vanparys B, Wittebolle L, Geets J, Boon N, De Vos P.** 2006. The incidence of *nirS* and *nirK* and their genetic heterogeneity in cultivated denitrifiers. Environmental Microbiology **8:**2012-2021.

9. **Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD.** 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and Environmental Microbiology **79:**5112-5120.

10. **Miethling R, Tebbe CC.** 2004. Resilience of a soil-established, genetically modified *Sinorhizobium meliloti* inoculant to soil management practices. Applied Soil Ecology **25:**161-167.

11. **Schmalenberger A, Tebbe CC.** 2002. Bacterial community composition in the rhizosphere of a transgenic, herbicide-resistant maize (*Zea mays*) and comparison to its non-transgenic cultivar Bosphore. FEMS Microbiology Ecology **40:**29-37.

12. **Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO.** 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Research **41**. doi: 10.1093/nar/gks808

13. **Throback IN, Enwall K, Jarvis A, Hallin S.** 2004. Reassessing PCR primers targeting *nirS*, *nirK* and *nosZ* genes for community surveys of denitrifying bacteria with DGGE. FEMS Microbiology Ecology **49:**401-417.

14. **Henry S, Bru D, Stres B, Hallet S, Philippot L.** 2006. Quantitative detection of the *nosZ* gene, encoding nitrous oxide reductase, and comparison of the abundances of 16S rRNA, *narG*, *nirK*, and *nosZ* genes in soils. Applied and Environmental Microbiology **72:**5181-5189.

15. **Kandeler E, Deiglmayr K, Tscherko D, Bru D, Philippot L.** 2006. Abundance of *narG*, *nirS*, *nirK*, and *nosZ* genes of denitrifying bacteria during primary successions of a glacier foreland. Applied and Environmental Microbiology **72:**5957-5962.

16. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF.** 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology **75:**7537-7541.

17. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R.** 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics **27:**2194-2200.

18. **Palmer K, Biasi C, Horn MA.** 2012. Contrasting denitrifier communities relate to contrasting N_2O emission patterns from acidic peat soils in arctic tundra. ISME Journal **6:**1058-1077.

19. **Bengtsson-Palme J, Hartmann M, Eriksson KM, Pal C, Thorell K, Larsson DGJ, Nilsson RH.** 2015. metaxa2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. Molecular Ecology Resources **15:**1403-1414.

20. **Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM.** 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Research **42:**D633-D642.

21. **Edgar RC.** 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics **26:**2460-2461.

22. **Price MN, Dehal PS, Arkin AP.** 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Molecular Biology and Evolution **26:**1641-1650.

Appendix to the AMIGA Deliverable 4.1

Table S1 **Number of PCR-amplified partial 16S rRNA gene sequences analyzed from rhizosphere DNA samples collected at the AMIGA project for maize**

Cultivar	Raw sequences
SPAIN 2012	
BT-1	59,113
BT-2	76,398
BT-3	78,617
BT-4	59,048
BT-5	69,308
BT-6	81,766
BT-7	61,673
BT-8	77,203
BT-9	63,101
BT-10	47,115
ISO-1	67,070
ISO-2	59,813
ISO-3	58,126
ISO4	64,911
ISO-5	108,573
ISO-6	100,679
ISO-7	106,811
ISO-8	143,702
ISO-9	121,041
ISO-10	116,758
SPAIN 2013	
BT-1	48,771
BT-2	50,787
BT-3	55,794
BT-4	36,090
BT-5	52,493
BT-6	53,328
BT-7	40,795
BT-8	77,514
BT-9	67,877
BT-10	65,687
ISO-1	72,978
ISO-2	64,448
ISO-3	71,254
ISO4	71,269
ISO-5	105,008
ISO-6	73,122
ISO-7	101,794
ISO-8	101,353
ISO-9	84,004
ISO-10	90,814

SLOVAKIA 2012	
BT-1	99,508
BT-2	87,572
BT-3	68,651
BT-4	95,911
BT-5	92,225
BT-6	88,359
BT-7	117,798
BT-8	104,193
BT-9	86,105
BT-10	66,503
ISO-1	97,723
ISO-2	52,247
ISO-3	92,917
ISO4	92,741
ISO-5	85,206
ISO-6	103,040
ISO-7	120,929
ISO-8	134,849
ISO-9	131,360
ISO-10	93,711
SLOVAKIA 2013	
BT-1	54,365
BT-2	62,333
BT-3	50,430
BT-4	53,452
BT-5	52,222
BT-6	53,726
BT-7	55,470
BT-8	68,274
BT-9	53,093
BT-10	63,218
ISO-1	54,401
ISO-2	45,053
ISO-3	46,281
ISO4	62,294
ISO-5	50,154
ISO-6	58,265
ISO-7	58,802
ISO-8	60,965
ISO-9	58,035
ISO-10	76,403

SLOVAKIA 2014	
BT-1	53,777
BT-2	59,932
BT-3	53,542
BT-4	49,452
BT-5	60,494
BT-6	43,624
BT-7	47,204
BT-8	57,759
BT-9	47,476
BT-10	60,932
ISO-1	47,885
ISO-2	54,015
ISO-3	60,561
ISO4	52,776
ISO-5	47,821
ISO-6	58,311
ISO-7	73,392
ISO-8	54,547
ISO-9	74,149
ISO-10	77,499
SWEDEN 2012	
BT-1	62,932
BT-2	51,323
BT-3	70,368
BT-4	41,044
BT-5	51,453
BT-6	46,675
BT-7	57,784
BT-8	33,882
BT-9	50,459
BT-10	44,231
ISO-1	46,182
ISO-2	47,037
ISO-3	55,748
ISO4	75,682
ISO-5	73,887
ISO-6	108,078
ISO-7	95,535
ISO-8	110,193
ISO-9	92,387
ISO-10	76,300

SWEDEN 2013	
BT-1	69,288
BT-2	74,393
BT-3	65,634
BT-4	77,497
BT-5	59,825
BT-6	73,855
BT-7	91,577
BT-8	70,762
BT-9	76,181
BT-10	96,095
ISO-1	69,559
ISO-2	100,893
ISO-3	77,856
ISO4	93,192
ISO-5	97,678
ISO-6	82,116
ISO-7	83,664
ISO-8	94,108
ISO-9	71,242
ISO-10	76,329

SWEDEN 2014	
BT-1	36,343
BT-2	35,937
BT-3	37,450
BT-4	36,023
BT-5	30,185
BT-6	35,569
BT-7	36,926
BT-8	37,835
BT-9	43,572
BT-10	33,214
ISO-1	40,810
ISO-2	42,580
ISO-3	40,038
ISO4	43,256
ISO-5	58,505
ISO-6	56,370
ISO-7	58,344
ISO-8	56,411
ISO-9	52,693
ISO-10	80,666

DENMARK 2013	
BT-1	91,857
BT-2	99,989
BT-3	90,176
BT-4	93,643
BT-5	83,981
BT-6	93,922
BT-7	107,080
BT-8	92,436
BT-9	90,313
BT-10	100,185
ISO-1	104,355
ISO-2	76,002
ISO-3	105,333
ISO4	101,611
ISO-5	103,818
ISO-6	81,014
ISO-7	83,528
ISO-8	78,043
ISO-9	76,986
ISO-10	73,711

DENMARK 2014	
BT-1	57,312
BT-2	42,205
BT-3	50,575
BT-4	52,084
BT-5	54,307
BT-6	61,667
BT-7	42,241
BT-8	37,789
BT-9	38,740
BT-10	42,494
ISO-1	43,336
ISO-2	42,785
ISO-3	52,640
ISO4	36,288
ISO-5	42,209
ISO-6	45,638
ISO-7	45,138
ISO-8	50,354
ISO-9	41,092
ISO-10	42,952

Table S2 **Number and DNA sequence quality of PCR-amplified partial *nirK* and *nirS* genes amplified from rhizosphere DNA of maize plants cultivated at the AMIGA field sites**

Cultivar	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirK SPAIN 2012						
BT-1	36,414	28.7	25,958	179	35.2	116
BT-2	42,604	33.3	28,410	190	36.8	120
BT-3	34,801	29.8	24,431	183	37.7	114
BT-4	29,419	27.3	21,373	177	37.9	110
BT-5	32,354	32.4	21,887	139	24.5	105
BT-6	21,740	35.7	13,974	173	34.7	113
BT-7	19,351	33.6	12,843	176	36.9	111
BT-8	21,368	31.7	14,592	174	34.5	114
BT-9	14,962	34.5	9,794	162	38.3	100
BT-10	18,245	32.3	12,352	172	33.7	114
ISO-1	18,989	31.7	12,977	164	34.8	107
ISO-2	16,105	29.3	11,383	151	33.8	100
ISO-3	19,404	33.3	12,949	189	40.7	112
ISO4	17,528	35.1	11,371	173	41.0	102
ISO-5	19,620	33.9	12,969	197	37.6	123
ISO-6	18,032	31.3	12,395	178	35.4	115
ISO-7	14,103	32.9	9,459	178	33.7	118
ISO-8	18,833	36.3	12,002	179	38.5	110
ISO-9	19,839	36.2	12,666	199	39.2	121
ISO-10	17,928	32.4	12,115	182	38.5	112
nirK SPAIN 2013						
BT-1	18,815	26.1	13,904	143	44.8	79
BT-2	15,366	32.3	10,405	187	38.5	115
BT-3	17,661	31.1	12,162	202	42.1	117
BT-4	14,038	33.1	9,398	119	42.0	69
BT-5	14,679	27.1	10,705	145	32.4	98
BT-6	15,027	24.3	11,369	158	40.5	94
BT-7	17,689	26.6	12,981	191	35.6	123
BT-8	16,248	21.9	12,696	155	36.8	98
BT-9	13,012	36.1	8,316	177	40.1	106
BT-10	17,360	46.2	9,339	147	36.1	94
ISO-1	13,526	35.1	8,780	194	39.2	118
ISO-2	18,040	26.2	13,312	155	41.3	91
ISO-3	16,779	31.1	11,568	168	41.1	99
ISO4	17,011	32.9	11,409	183	35.5	118
ISO-5	18,523	31.4	12,715	210	41.4	123
ISO-6	20,783	34.5	13,622	204	38.7	125
ISO-7	17,701	35.6	11,394	133	33.8	88
ISO-8	21,982	28.4	15,746	206	40.3	123
ISO-9	21,259	30.1	14,863	218	38.1	135
ISO-10	22,726	28.1	16,348	151	37.1	95

Cultivar	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirK SLOVAKIA 2012						
BT-1	37,241	20.4	29,647	111	24.3	84
BT-2	34,960	23.9	26,605	152	32.9	102
BT-3	37,900	26.9	27,707	159	35.2	103
BT-4	35,594	28.9	25,313	186	33.3	124
BT-5	25,569	24.5	19,292	150	34.0	99
BT-6	19,429	35.6	12,507	117	35.9	75
BT-7	16,614	33.9	10,983	170	38.8	104
BT-8	16,453	32.6	11,093	138	32.6	93
BT-9	15,445	35.2	10,010	153	30.1	107
BT-10	18,793	34.5	12,315	151	42.4	87
ISO-1	12,574	32.8	8,454	162	39.5	98
ISO-2	22,834	31.2	15,702	176	43.2	100
ISO-3	19,804	37.9	12,301	167	41.9	97
ISO4	17,328	25.9	12,841	154	37.0	97
ISO-5	19,367	30.5	13,469	178	43.3	101
ISO-6	15,603	34.9	10,162	147	30.6	102
ISO-7	17,585	35.3	11,384	180	41.7	105
ISO-8	20,324	32.2	13,781	175	36.0	112
ISO-9	20,687	33.2	13,823	182	37.9	113
ISO-10	21,881	37.7	13,632	161	37.9	100
nirK SPAIN 2013						
BT-1	8,718	22.3	6,772	165	34.5	108
BT-2	6,877	25.0	5,157	166	28.3	119
BT-3	7,096	27.7	5,132	169	32.0	115
BT-4	7,977	26.7	5,845	160	33.8	106
BT-5	7,708	18.8	6,261	134	23.9	102
BT-6	6,348	27.3	4,615	133	29.3	94
BT-7	5,493	28.1	3,948	166	28.3	119
BT-8	6,648	26.5	4,886	154	29.2	109
BT-9	8,057	27.0	5,882	177	34.5	116
BT-10	14,851	21.3	11,689	149	34.9	97
ISO-1	8,527	25.6	6,342	188	37.2	118
ISO-2	7,283	26.2	5,373	135	29.6	95
ISO-3	9,054	22.0	7,063	150	31.3	103
ISO4	18,417	23.8	14,032	126	34.1	83
ISO-5	16,945	26.8	12,405	145	33.1	97
ISO-6	6,202	25.3	4,635	139	23.7	106
ISO-7	22,540	24.0	17,136	74	36.5	47
ISO-8	13,632	27.9	9,834	162	37.0	102
ISO-9	11,039	31.9	7,516	72	40.3	43
ISO-10	17,640	28.8	12,551	80	27.5	58

nirK SPAIN 2014						
BT-1	7,181	38.7	4400	180	36.1	115
BT-2	8,256	41.2	4,852	203	38.4	125
BT-3	6,426	42.8	3,673	186	34.4	122
BT-4	6,950	39.8	4,182	142	31.7	97
BT-5	10,898	40.2	6,516	185	34.1	122
BT-6	13,451	43.7	7,569	170	39.4	103
BT-7	12,268	43.2	6,969	194	36.6	123
BT-8	11,508	43.0	6,558	156	32.7	105
BT-9	8,510	31.4	5,835	140	37.1	88
BT-10	11,785	30.1	8,236	211	37.9	131
ISO-1	15,258	40.5	9,076	328	36.6	208
ISO-2	15,851	40.3	9,464	324	42.0	188
ISO-3	15,824	42.8	9,047	305	40.0	183
ISO4	17,518	43.0	9,977	319	42.6	183
ISO-5	13,123	44.5	7,284	335	41.2	197
ISO-6	17,730	43.7	9,984	346	41.6	202
ISO-7	18,102	44.5	10,051	318	38.4	196
ISO-8	18,532	42.6	10,646	332	41.3	195
ISO-9	22,232	43.9	12,482	345	45.2	189
ISO-10	16,497	45.0	9,081	348	38.5	214

Cultivar	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirK SWEDEN 2012						
BT-1	30,095	32.3	20,385	183	38.3	113
BT-2	30,400	32.7	20,449	152	36.2	97
BT-3	27,198	26.0	20,121	169	29.0	120
BT-4	31,758	24.0	24,129	143	37.1	90
BT-5	23,556	30.0	16,496	171	31.6	117
BT-6	28,133	25.1	21,058	163	35.6	105
BT-7	12,734	29.9	8,924	187	35.3	121
BT-8	15,240	23.0	11,732	205	33.7	136
BT-9	12,672	25.1	9,493	195	34.9	127
BT-10	10,066	28.6	7,186	154	43.5	87
ISO-1	14,217	24.7	10,704	171	35.1	111
ISO-2	14,809	24.9	11,119	164	35.4	106
ISO-3	12,240	29.2	8,672	198	36.9	125
ISO4	12,555	24.5	9,475	193	33.7	128
ISO-5	11,398	29.5	8,034	155	27.7	112
ISO-6	13,824	21.8	10,816	183	30.6	127
ISO-7	14,474	28.4	10,363	172	32.0	117
ISO-8	12,982	20.5	10,326	193	36.3	123
ISO-9	7,492	23.8	5,711	162	36.4	103
ISO-10	8,032	23.2	6,169	206	38.3	127

nirK SWEDEN 2013						
BT-1	12,794	29.8	8,981	133	27.1	97
BT-2	16,923	28.7	12,067	150	29.3	106
BT-3	11,957	32.1	8,124	143	33.6	95
BT-4	23,154	29.1	16,421	164	31.7	112
BT-5	22,267	30.1	15,557	152	37.5	95
BT-6	22,069	29.4	15,582	173	37.0	109
BT-7	19,877	33.2	13,280	150	36.0	96
BT-8	23,764	31.5	16,284	155	34.2	102
BT-9	25,270	26.3	18,620	153	32.0	104
BT-10	24,412	23.9	18,588	170	29.4	120
ISO-1	18,207	24.7	13,710	185	36.2	118
ISO-2	15,922	26.4	11,725	178	34.8	116
ISO-3	14,092	29.3	9,963	164	32.9	110
ISO4	17,156	29.0	12,182	168	34.5	110
ISO-5	14,853	31.1	10,229	172	33.1	115
ISO-6	19,033	26.5	13,984	148	36.5	94
ISO-7	22,705	28.7	16,188	179	33.0	120
ISO-8	16,723	25.3	12,484	149	28.9	106
ISO-9	20,351	31.0	14,048	200	33.0	134
ISO-10	20,033	32.3	13,559	177	37.9	110
nirK SWEDEN 2014						
BT-1	13,041	25.7	9,688	161	32.3	109
BT-2	14,018	24.6	10,576	179	34.6	117
BT-3	13,775	26.6	10,113	170	42.4	98
BT-4	15,988	24.5	12,063	182	41.2	107
BT-5	14,380	32.6	9,687	141	38.3	87
BT-6	18,634	29.4	13,154	150	30.0	105
BT-7	15,646	40.9	9,244	192	32.8	129
BT-8	20,084	39.3	12,187	164	41.5	96
BT-9	19,277	40.2	11,528	174	41.4	102
BT-10	17,650	40.9	10,435	198	42.9	113
ISO-1	18,881	40.6	11,220	185	38.9	113
ISO-2	19,984	38.2	12,344	174	47.7	91
ISO-3	17,572	29.3	12,426	148	39.9	89
ISO4	21,253	32.7	14,298	197	35.0	128
ISO-5	20,256	33.4	13,490	167	29.9	117
ISO-6	21,676	30.7	15,029	162	39.5	98
ISO-7	20,640	47.2	10,890	181	30.4	126
ISO-8	19,253	43.3	10,908	185	37.3	116
ISO-9	15,506	35.0	10,079	156	39.1	95
ISO-10	19,218	38.7	11,774	175	38.9	107

Cultivar	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirK DENMARK 2013						
BT-1	16,198	23.5	12,389	153	27.5	111
BT-2	18,404	26.3	13,561	159	28.3	114
BT-3	21,925	25.5	16,329	175	34.9	114
BT-4	23,677	24.7	17,832	155	32.9	104
BT-5	18,270	33.0	12,250	157	33.8	104
BT-6	18,412	29.4	13,008	156	33.3	104
BT-7	14,369	29.6	10,121	152	34.2	100
BT-8	18,057	29.3	12,764	140	30.0	98
BT-9	18,655	31.5	12,778	157	29.9	110
BT-10	17,908	29.2	12,671	159	31.4	109
ISO-1	22,071	31.3	15,165	178	36.0	114
ISO-2	20,216	34.0	13,345	173	37.6	108
ISO-3	19,739	24.2	14,959	168	35.7	108
ISO4	19,358	26.3	14,267	175	33.1	117
ISO-5	24,025	26.3	17,709	178	27.5	129
ISO-6	22,514	28.7	16,045	175	32.6	118
ISO-7	21,536	34.5	14,100	177	32.8	119
ISO-8	22,738	33.4	15,145	174	32.8	117
ISO-9	19,845	30.8	13,727	166	37.3	104
ISO-10	24,682	31.5	16,912	170	38.8	104
nirK DENMARK 2014						
BT-1	20,073	36.3	12,795	181	38.1	112
BT-2	19,106	38.7	11,705	197	38.1	122
BT-3	21,208	38.6	13,012	184	43.5	104
BT-4	19,727	40.3	11,785	164	44.5	91
BT-5	16,561	28.7	11,800	172	37.8	107
BT-6	21,471	32.5	14,493	224	42.4	129
BT-7	21,152	35.6	13,622	198	40.4	118
BT-8	19,674	31.7	13,434	174	39.1	106
BT-9	18,342	38.0	11,375	176	37.5	110
BT-10	18,536	38.4	11,413	198	37.9	123
ISO-1	15,788	37.4	9,876	199	42.2	115
ISO-2	19,088	38.4	11,758	209	40.7	124
ISO-3	18,185	36.6	11,537	178	37.1	112
ISO4	16,214	39.3	9,843	211	40.3	126
ISO-5	19,108	36.8	12,070	206	41.3	121
ISO-6	19,182	38.2	11,863	187	44.4	104
ISO-7	18,350	30.1	12,822	177	36.7	112
ISO-8	19,760	32.4	13,365	159	32.7	107
ISO-9	19,429	34.0	12,815	199	41.7	116
ISO-10	19,440	30.2	13,560	187	35.8	120

Cultivar	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirS SPAIN 2012						
BT-1	1001	85,6	144	30	36,7	19
BT-2	1361	88,0	163	38	31,6	26
BT-3	868	88,1	103	25	40,0	15
BT-4	1081	90,2	106	34	44,1	19
BT-5	747	87,8	91	25	56,0	11
BT-6	526	89,5	55	24	37,5	15
BT-7	417	93,5	27	16	75,0	4
BT-8	471	90,0	47	20	65,0	7
BT-9	269	91,4	23	14	64,3	5
BT-10	384	90,9	35	18	61,1	7
ISO-1	320	90,0	32	8	25,0	6
ISO-2	438	87,4	55	22	68,2	7
ISO-3	206	90,8	19	11	81,8	2
ISO4	252	92,1	20	14	78,6	3
ISO-5	281	92,2	22	13	69,2	4
ISO-6	384	91,1	34	15	53,3	7
ISO-7	253	95,3	12	5	40,0	3
ISO-8	226	93,8	14	11	81,8	2
ISO-9	278	94,6	15	9	55,6	4
ISO-10	371	94,9	19	11	45,5	6
nirS SPAIN 2013						
BT-1	Missing					
BT-2	398	88,7	45	17	82,4	3
BT-3	Missing					
BT-4	956	91,7	79	31	45,2	17
BT-5	Missing					
BT-6	711	94,2	41	18	61,1	7
BT-7	Missing					
BT-8	173	93,1	12	8	50,0	4
BT-9	Missing					
BT-10	209	90,0	21	14	57,1	6
ISO-1	145	95,2	7	5	60,0	2
ISO-2	192	94,3	11	9	77,8	2
ISO-3	149	92,6	11	10	90,0	1
ISO4	249	93,6	16	12	75,0	3
ISO-5	178	91,6	15	9	66,7	3
ISO-6	124	94,4	7	7	100,0	0
ISO-7	515	92,0	41	23	60,9	9
ISO-8	273	93,8	17	11	45,5	6
ISO-9	562	91,5	48	20	50,0	10
ISO-10	388	92,8	28	12	66,7	4

Cultivar	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirS SLOVAKIA 2012						
BT-1	1604	90,5	152	18	38,9	11
BT-2	775	92,1	61	22	45,5	12
BT-3	1287	92,1	102	16	25,0	12
BT-4	1078	92,7	79	18	44,4	10
BT-5	1255	91,4	108	16	43,8	9
BT-6	421	91,7	35	11	45,5	6
BT-7	639	92,8	46	16	50,0	8
BT-8	251	91,6	21	8	37,5	5
BT-9	417	90,6	39	15	53,3	7
BT-10	381	92,9	27	14	50,0	7
ISO-1	580	91,6	49	17	58,8	7
ISO-2	489	91,2	43	15	60,0	6
ISO-3	471	92,1	37	10	40,0	6
ISO4	446	93,0	31	15	60,0	6
ISO-5	561	89,1	61	14	28,6	10
ISO-6	258	89,9	26	13	30,8	9
ISO-7	401	91,0	36	15	60,0	6
ISO-8	414	92,5	31	14	50,0	7
ISO-9	640	88,8	72	16	31,3	11
ISO-10	449	90,4	43	13	23,1	10
nirS SLOVAKIA 2013						
BT-1	621	91,0	56	13	53,8	6
BT-2	567	90,3	55	4	50,0	2
BT-3	577	91,7	48	7	28,6	5
BT-4	117	91,5	10	1	0,0	1
BT-5	186	89,8	19	3	66,7	1
BT-6	Missing					
BT-7	97	90,7	9	2	50,0	1
BT-8	Missing					
BT-9	177	88,1	21	3	66,7	1
BT-10	Missing					
ISO-1	148	93,2	10	9	88,9	1
ISO-2	Missing					
ISO-3	221	93,2	15	10	60,0	4
ISO4	Missing					
ISO-5	Missing					
ISO-6	167	93,4	11	8	75,0	2
ISO-7	Missing					
ISO-8	134	95,5	6	6	100,0	0
ISO-9	Missing					
ISO-10	774	93,4	51	22	54,5	10

nirS SLOVAKIA 2014						
BT-1	301	93,7	19	13	69,2	301
BT-2	340	95,0	17	8	50,0	340
BT-3	189	97,9	4	3	66,7	189
BT-4	204	95,6	9	7	71,4	204
BT-5	173	94,2	10	5	40,0	173
BT-6	266	94,4	15	7	85,7	266
BT-7	232	94,0	14	8	62,5	232
BT-8	228	94,7	12	5	40,0	228
BT-9	941	96,2	36	12	58,3	941
BT-10	496	95,6	22	13	76,9	496
ISO-1	977	96,0	39	19	57,9	977
ISO-2	713	95,2	34	17	76,5	713
ISO-3	240	94,6	13	8	75,0	240
ISO4	231	92,6	17	8	50,0	231
ISO-5	148	95,9	6	3	33,3	148
ISO-6	253	96,4	9	8	87,5	253
ISO-7	196	94,9	10	9	88,9	196
ISO-8	328	97,3	9	6	66,7	328
ISO-9	233	92,7	17	11	72,7	233
ISO-10	192	97,4	5	4	75,0	192

Cultivar	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirS SWEDEN 2012						
BT-1	833	90,9	76	24	45,8	13
BT-2	724	91,6	61	23	56,5	10
BT-3	881	92,4	67	17	41,2	10
BT-4	522	92,0	42	22	59,1	9
BT-5	1268	91,2	111	27	33,3	18
BT-6	775	93,0	54	18	38,9	11
BT-7	Missing					
BT-8	718	94,7	38	18	44,4	10
BT-9	Missing					
BT-10	271	92,6	20	11	63,6	4
ISO-1	Missing					
ISO-2	311	93,2	21	15	73,3	4
ISO-3	Missing					
ISO4	217	94,0	13	9	88,9	1
ISO-5	Missing					
ISO-6	276	94,6	15	10	70,0	3
ISO-7	190	94,2	11	10	90,0	1
ISO-8	2827	94,6	152	38	39,5	23
ISO-9	209	93,3	14	8	62,5	3
ISO-10	229	93,9	14	8	37,5	5

nirS SWEDEN 2013						
BT-1	Missing					
BT-2	164	92,7	12	9	77,8	2
BT-3	163	93,3	11	8	75,0	2
BT-4	157	91,7	13	8	62,5	3
BT-5	144	93,1	10	10	100,0	0
BT-6	233	94,8	12	9	66,7	3
BT-7	133	95,5	6	3	33,3	2
BT-8	171	95,3	8	7	85,7	1
BT-9	663	92,5	50	15	73,3	4
BT-10	393	88,8	44	12	33,3	8
ISO-1	Missing					
ISO-2	639	91,4	55	17	35,3	11
ISO-3	185	95,7	8	5	40,0	3
ISO4	264	92,4	20	14	64,3	5
ISO-5	137	93,4	9	4	75,0	1
ISO-6	265	94,3	15	8	62,5	3
ISO-7	208	89,4	22	9	44,4	5
ISO-8	388	91,2	34	13	53,8	6
ISO-9	217	92,2	17	5	40,0	3
ISO-10	242	92,6	18	11	72,7	3
nirS SWEDEN 2014						
BT-1	408	93,9	25	10	50,0	5
BT-2	321	89,1	35	5	40,0	3
BT-3	435	90,6	41	14	57,1	6
BT-4	483	89,2	52	15	66,7	5
BT-5	150	96,7	5	5	100,0	0
BT-6	176	93,2	12	10	80,0	2
BT-7	95	92,6	7	4	75,0	1
BT-8	165	90,3	16	9	88,9	1
BT-9	143	91,6	12	9	88,9	1
BT-10	151	89,4	16	7	71,4	2
ISO-1	198	89,9	20	2	50,0	1
ISO-2	154	92,9	11	6	66,7	2
ISO-3	659	90,1	65	14	42,9	8
ISO4	358	92,7	26	8	37,5	5
ISO-5	776	92,4	59	18	38,9	11
ISO-6	412	91,5	35	12	50,0	6
ISO-7	133	91,7	11	6	50,0	3
ISO-8	187	90,4	18	6	66,7	2
ISO-9	102	92,2	8	6	66,7	2
ISO-10	150	91,3	13	6	83,3	1

Cultivar	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirS DENMARK 2013						
BT-1	904	88,2	107	29	41,4	17
BT-2	500	87,6	62	20	35,0	13
BT-3	1058	90,2	104	25	44,0	14
BT-4	955	90,6	90	16	37,5	10
BT-5	338	87,3	43	18	44,4	10
BT-6	359	86,1	50	16	43,8	9
BT-7	187	91,4	16	9	55,6	4
BT-8	277	93,9	17	13	76,9	3
BT-9	263	90,1	26	12	50,0	6
BT-10	413	89,6	43	19	42,1	11
ISO-1	275	90,5	26	13	53,8	6
ISO-2	276	90,2	27	15	66,7	5
ISO-3	1063	90,4	102	27	44,4	15
ISO4	560	91,4	48	17	47,1	9
ISO-5	1041	91,8	85	19	31,6	13
ISO-6	613	92,2	48	15	40,0	9
ISO-7	300	90,3	29	17	70,6	5
ISO-8	321	90,7	30	17	64,7	6
ISO-9	214	89,7	22	13	38,5	8
ISO-10	274	90,5	26	13	46,2	7
nirS DENMARK 2014						
BT-1	94	95,7	4	4	100,0	0
BT-2	135	92,6	10	6	83,3	1
BT-3	118	92,4	9	5	80,0	1
BT-4	126	92,9	9	6	66,7	2
BT-5	508	93,9	31	18	66,7	6
BT-6	272	91,5	23	9	44,4	5
BT-7	686	93,4	45	19	52,6	9
BT-8	375	93,9	23	9	44,4	5
BT-9	153	95,4	7	5	60,0	2
BT-10	165	94,5	9	6	66,7	2
ISO-1	79	96,2	3	3	100,0	0
ISO-2	129	95,3	6	4	50,0	2
ISO-3	126	94,4	7	5	80,0	1
ISO4	171	93,6	11	6	66,7	2
ISO-5	77	88,3	9	5	80,0	1
ISO-6	123	95,1	6	2	0,0	2
ISO-7	407	93,4	27	13	61,5	5
ISO-8	254	93,3	17	6	66,7	2
ISO-9	487	93,0	34	14	42,9	8
ISO-10	360	93,6	23	13	69,2	4

Table S3 **Number of PCR-amplified partial 16S rRNA gene sequences analyzed from rhizosphere DNA samples collected in the AMIGA project for maize**

Cultivar/ Treatment	Raw sequences
IRELAND 2012	
CON1 – CP	65,329
CON2 – CP	68,985
CON3 – CP	70,076
CON4 – CP	63,714
ISO1 – CP	50,677
ISO2 – CP	48,510
ISO3 – CP	49,788
ISO4 – CP	54,363
GM1 – CP	52,198
GM2 – CP	59,753
GM3 – CP	42,124
GM4 – CP	39,078
IRELAND 2013	
GM_CP-1	60,114
GM_CP-2	54,049
GM_CP-3	63,736
GM_CP-4	71,550
GM_CP-5	77,836
GM_CP-6	56,500
GM_CP-7	67,813
ISO_CP-1	89,277
ISO_CP-2	65,021
ISO_CP-3	44,117
ISO_CP-4	57,724
ISO_CP-5	58,658
ISO_CP-6	68,017
ISO_CP-7	48,615
CON_CP-1	59,748
CON_CP-2	78,471
CON_CP-3	65,075
CON_CP-4	63,154
CON_CP-5	58,866
CON_CP-6	55,720
CON_CP-7	60,112
GM_NS-1	54,777
GM_NS-2	55,780
GM_NS-3	51,520
GM_NS-4	61,544
GM_NS-5	47,030
GM_NS-6	51,384
GM_NS-7	54,909
ISO_NS-1	55,569
ISO_NS-2	63,794
ISO_NS-3	47,374
ISO_NS-4	53,852
ISO_NS-5	56,788
ISO_NS-6	56,664
ISO_NS-7	52,132
CON_NS-1	43,262
CON_NS-2	50,783
CON_NS-3	48,046
CON_NS-4	47,382
CON_NS-5	55,625
CON_NS-6	52,406
CON_NS-7	54,889
IRELAND 2014	
GM_CP-1	33,426
GM_CP-2	30,817
GM_CP-3	29,438
GM_CP-4	29,168
GM_CP-5	38,720
GM_CP-6	26,859
GM_CP-7	36,179
ISO_CP-1	39,213
ISO_CP-2	35,371
ISO_CP-3	41,422
ISO_CP-4	32,876
ISO_CP-5	33,886
ISO_CP-6	36,313
ISO_CP-7	33,977
CON_CP-1	60,595
CON_CP-2	53,473
CON_CP-3	54,869
CON_CP-4	39,917
CON_CP-5	40,978
CON_CP-6	50,608
CON_CP-7	56,553
GM_NS-1	32,339
GM_NS-2	27,491
GM_NS-3	31,092
GM_NS-4	33,814
GM_NS-5	31,870
GM_NS-6	34,110
GM_NS-7	27,553
ISO_NS-1	61,468
ISO_NS-2	57,130
ISO_NS-3	58,748
ISO_NS-4	63,031
ISO_NS-5	69,037
ISO_NS-6	62,629
ISO_NS-7	40,754
CON_NS-1	36,976
CON_NS-2	39,517
CON_NS-3	21,401
CON_NS-4	39,181
CON_NS-5	37,030
CON_NS-6	31,429
CON_NS-7	31,348

Cultivar/ Treatment	Raw sequences
NETHERLANDS 2013	
CON_CP-1	60,049
CON_CP-2	56,700
CON_CP-3	57,245
CON_CP-4	59,285
CON_CP-5	57,702
CON_CP-6	51,745
CON_CP-7	57,756
CON_NS-1	43,077
CON_NS-2	56,128
CON_NS-3	57,004
CON_NS-4	62,763
CON_NS-5	67,968
CON_NS-6	47,691
CON_NS-7	41,179

NETHERLANDS 2014	
GM_CP-1	34,621
GM_CP-2	36,396
GM_CP-3	38,037
GM_CP-4	36,527
GM_CP-5	33,530
GM_CP-6	36,802
GM_CP-7	53,342
ISO_CP-1	53,276
ISO_CP-2	56,707
ISO_CP-3	48,698
ISO_CP-4	43,563
ISO_CP-5	46,696
ISO_CP-6	42,196
ISO_CP-7	23,614
CON_CP-1	24,995
CON_CP-2	25,039
CON_CP-3	38,054
CON_CP-4	30,107
CON_CP-5	19,705

CON_CP-6	28,764
CON_CP7	32,363
GM-NS-1	27,193
GM_NS-2	34,062
GM_NS-3	30,918
GM_NS-4	25,708
GM_NS-5	36,654
GM_NS-6	36,088
GM_NS-7	38,122
ISO_NS-1	56,917
ISO_NS-2	42,080
ISO_NS-3	44,865
ISO_NS-4	52,002
ISO_NS-5	45,283
ISO_NS-6	45,146
ISO_NS-7	45,255
CON_NS-1	47,253
CON_NS-2	40,202
CON_NS-3	40,003
CON_NS-4	47,326
CON_NS-5	45,967
CON_NS-6	44,695
CON_NS-7	55,940

Table S4 **Number and DNA sequence quality of PCR-amplified partial *nirK* and *nirS* genes amplified from rhizosphere DNA of maize plants cultivated at the AMIGA field sites**

Cultivar/ Treatment	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirK IRELAND 2012						
CON1 - CP	14,154	38.0	8,771	214	37.4	134
CON2 - CP	13,437	36.0	8,601	215	39.5	130
CON3 - CP	10,252	38.7	6,283	218	37.6	136
CON4 - CP	9,784	39.3	5,937	259	38.6	159
ISO1 – CP	10,740	37.9	6,673	230	40.0	138
ISO2 – CP	11,411	37.3	7,151	198	40.9	117
ISO3 – CP	12,454	38.2	7,694	236	38.1	146
ISO4 – CP	5,809	38.4	3,577	235	33.6	156
GM1 - CP	6,776	30.7	4,697	235	31.5	161
GM2 - CP	5,767	32.9	3,872	244	26.2	180
GM3 - CP	5,257	36.6	3,333	232	32.8	156
GM4 - CP	10,839	34.4	7,106	246	38.2	152
nirK IRELAND 2013						
GM_CP-1	22,496	30.7	15,590	177	32.8	119
GM_CP-2	18,810	30.9	13,001	184	27.7	133
GM_CP-3	14,368	32.5	9,700	211	33.6	140
GM_CP-4	24,498	34.8	15,977	187	32.6	126
GM_CP-5	23,426	30.7	16,244	169	25.4	126
GM_CP-6	20,452	30.4	14,235	187	31.6	128
GM_CP-7	22,911	36.1	14,646	202	36.6	128
ISO_CP-1	14,065	30.4	9,783	208	26.4	153
ISO_CP-2	10,234	27.0	7,474	160	32.5	108
ISO_CP-3	12,095	36.1	7,733	200	33.5	133
ISO_CP-4	8,438	36.2	5,387	194	29.4	137
ISO_CP-5	12,897	35.1	8,368	214	30.8	148
ISO_CP-6	20,241	34.2	13,309	198	35.9	127
ISO_CP-7	14,616	32.3	9,891	204	36.8	129
CON_CP-1	14,510	30.8	10,034	175	29.1	124
CON_CP-2	8,248	32.1	5,601	172	26.7	126
CON_CP-3	8,759	27.9	6,316	163	30.1	114
CON_CP-4	8,436	26.4	6,206	152	26.3	112
CON_CP-5	8,265	29.7	5,807	152	25.0	114
CON_CP-6	8,407	27.2	6,121	157	24.8	118
CON_CP-7	8,446	27.1	6,160	135	24.4	102
GM_NS-1	7,645	29.3	5,403	187	24.6	141
GM_NS-2	9,086	37.1	5,713	181	30.4	126
GM_NS-3	7,022	36.1	4,488	229	29.3	162
GM_NS-4	8,079	33.4	5,381	164	34.8	107
GM_NS-5	12,744	30.3	8,877	208	27.9	150
GM_NS-6	8,646	33.2	5,774	196	28.1	141
GM_NS-7	9,041	43.9	5,070	230	33.9	152
ISO_NS-1	9,265	28.2	6,654	215	32.1	146
ISO_NS-2	20,164	28.7	14,377	195	36.9	123

ISO_NS-3	19,919	30.1	13,928	229	33.6	152
ISO_NS-4	16,734	38.1	10,351	216	31.0	149
ISO_NS-5	17,570	33.2	11,742	206	38.3	127
ISO_NS-6	19,478	34.6	12,739	206	30.1	144
ISO_NS-7	14,580	34.0	9,622	218	34.9	142
CON_NS-1	15,906	29.3	11,246	179	33.0	120
CON_NS-2	15,252	30.6	10,592	190	30.5	132
CON_NS-3	14,342	32.8	9,631	189	32.8	127
CON_NS-4	15,548	32.6	10,487	169	30.2	118
CON_NS-5	14,501	37.5	9,067	178	40.4	106
CON_NS-6	14,427	30.3	10,054	173	30.1	121
CON_NS-7	11,032	26.5	8,111	173	33.5	115
nirK IRELAND 2014						,
GM_CP-1	19,349	46.7	10,310	137	42.3	79
GM_CP-2	19,801	49.6	9,985	157	37.6	98
GM_CP-3	15,772	38.9	9,644	184	38.0	114
GM_CP-4	21,251	28.5	15,194	150	48.0	78
GM_CP-5	22,252	40.0	13,345	143	47.6	75
GM_CP-6	16,354	37.9	10,151	174	35.6	112
GM_CP-7	18,769	36.8	11,861	227	48.0	118
ISO_CP-1	21,015	42.2	12,143	188	38.8	115
ISO_CP-2	18,535	32.6	12,496	201	38.3	124
ISO_CP-3	19,168	34.8	12,507	198	34.3	130
ISO_CP-4	19,405	35.2	12,577	188	42.0	109
ISO_CP-5	20,744	32.0	14,116	159	37.7	99
ISO_CP-6	19,122	38.7	11,730	157	39.5	95
ISO_CP-7	18,463	39.6	11,147	195	34.9	127
CON_CP-1	15,332	41.8	8,920	174	35.1	113
CON_CP-2	15,193	34.8	9,910	181	29.3	128
CON_CP-3	18,352	41.8	10,684	186	39.8	112
CON_CP-4	13,662	34.4	8,961	179	39.1	109
CON_CP-5	13,234	35.3	8,565	160	32.5	108
CON_CP-6	12,739	40.9	7,530	191	44.0	107
CON_CP-7	11,508	28.2	8,258	144	38.2	89
GM_NS-1	14,235	37.5	8,903	117	41.9	68
GM_NS-2	13,998	49.4	7,089	175	38.3	108
GM_NS-3	15,680	57.7	6,626	163	40.5	97
GM_NS-4	18,781	33.3	12,534	176	38.1	109
GM_NS-5	14,540	42.5	8,358	158	41.1	93
GM_NS-6	12,729	56.6	5,526	170	47.1	90
GM_NS-7	12,555	33.6	8,335	210	43.3	119
ISO_NS-1	14,244	42.6	8,183	216	36.1	138
ISO_NS-2	11,607	37.9	7,209	217	40.6	129
ISO_NS-3	14,206	36.8	8,979	223	40.8	132
ISO_NS-4	15,293	39.5	9,251	224	40.6	133
ISO_NS-5	17,411	36.3	11,097	130	40.0	78
ISO_NS-6	15,887	28.5	11,356	207	41.1	122
ISO_NS-7	16,411	33.5	10,912	193	41.5	113
CON_NS-1	16,326	33.7	10,828	204	37.3	128

CON_NS-2	13,473	27.5	9,770	179	34.1	118
CON_NS-3	15,009	28.1	10,790	165	40.6	98
CON_NS-4	12,184	31.5	8,347	150	33.3	100
CON_NS-5	13,381	27.7	9,675	156	34.0	103
CON_NS-6	15,185	31.5	10,409	166	31.9	113
CON_NS-7	14,617	33.7	9,695	174	37.4	109

Cultivar/ Treatment	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirK NETHERLANDS 2013						
CON_CP-1	13,007	26.3	9,592	105	26.7	77
CON_CP-2	10,202	26.7	7,474	129	36.4	82
CON_CP-3	11,495	35.7	7,396	104	38.5	64
CON_CP-4	11,131	29.6	7,840	123	36.6	78
CON_CP-5	12,007	34.4	7,871	136	37.5	85
CON_CP-6	13,273	32.6	8,948	127	35.4	82
CON_CP-7	22,877	39.9	13,758	160	34.4	105
CON_NS-1	9,593	38.3	5,918	167	38.9	102
CON_NS-2	13,127	36.4	8,350	138	31.2	95
CON_NS-3	12,108	28.2	8,699	108	37.0	68
CON_NS-4	13,145	32.7	8,845	141	34.0	93
CON_NS-5	14,757	29.7	10,380	136	42.6	78
CON_NS-6	14,993	30.2	10,467	137	40.1	82
CON_NS-7	16,639	36.1	10,631	178	42.1	103
nirK NETHERLANDS 2014						
GM_CP-1	16,248	31.7	11,098	129	37.2	81
GM_CP-2	16,610	28.0	11,967	112	40.2	67
GM_CP-3	5,962	26.1	4,406	110	40.0	66
GM_CP-4	7,508	46.7	4,003	368	41.8	214
GM_CP-5	6,156	30.0	4,311	226	42.5	130
GM_CP-6	6,399	42.5	3,677	374	40.6	222
GM_CP-7	5,599	36.2	3,570	121	34.7	79
ISO_CP-1	4,738	27.1	3,453	112	35.7	72
ISO_CP-2	4,826	29.1	3,423	100	37.0	63
ISO_CP-3	5,746	35.7	3,694	112	42.0	65
ISO_CP-4	5,205	27.6	3,771	121	40.5	72
ISO_CP-5	5,176	30.0	3,621	139	36.0	89
ISO_CP-6	5,149	30.6	3,575	204	34.8	133
ISO_CP-7	11,345	26.8	8,304	132	33.3	88
CON_CP-2	14,532	17.8	11,941	100	41.0	59
CON_CP-3	17,811	22.4	13,828	116	34.5	76
CON_CP-4	16,192	19.1	13,099	101	45.5	55
CON_CP-6	9,773	22.1	7,614	106	45.3	58
GM_NS-2	14,058	23.1	10,810	133	44.4	74
GM_NS-3	9,859	32.9	6,612	294	40.5	175
GM_NS-4	12,224	27.4	8,880	190	33.7	126
GM_NS-5	10,586	30.5	7,359	208	36.5	132
GM_NS-6	23,893	36.2	15,247	183	34.4	120

GM_NS-7	23,572	44.6	13,069	232	40.1	139
ISO_NS-1	20,264	21.4	15,929	100	35.0	65
ISO_NS-2	14,364	20.5	11,420	115	26.1	85
ISO_NS-3	24,733	37.1	15,567	206	35.9	132
ISO_NS-4	23,439	37.8	14,584	221	39.4	134
ISO_NS-5	11,291	23.1	8,684	91	42.9	52
ISO_NS-6	11,369	23.2	8,733	115	35.7	74
ISO_NS-7	13,290	24.1	10,091	138	45.7	75
CON_NS-1	10,000	30.3	6,972	125	31.2	86
CON_NS-2	19,301	41.0	11,387	311	43.7	175
CON_NS-3	16,026	25.5	11,946	215	30.7	149
CON_NS-4	13,304	37.9	8,266	293	40.3	175
CON_NS-5	15,612	28.5	11,155	139	27.3	101
CON_NS-7	13,838	38.8	8,464	320	42.5	184

Cultivar/ Treatment	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirS IRELAND 2012						
CON1 - CP	113	88.5	13	7	57.1	3
CON2 - CP	144	85.4	21	6	83.3	1
CON3 - CP	83	96.4	3	3	100.0	0
CON4 - CP	133	89.5	14	10	80.0	2
ISO1 – CP	Missing					
ISO2 – CP	204	87.7	25	14	50.0	7
ISO3 – CP	145	96.6	5	4	75.0	1
ISO4 – CP	140	88.6	16	10	70.0	3
GM1 – CP	659	89.1	72	24	58.3	10
GM2 – CP	335	90.7	31	12	66.7	4
GM3 – CP	738	90.2	72	19	21.1	15
GM4 – CP	562	90.4	54	19	36.8	12
nirS IRELAND 2013						
GM_CP-1	172	93.0	12	8	75.0	2
GM_CP-2	270	90.0	27	12	58.3	5
GM_CP-3	185	90.3	18	10	50.0	5
GM_CP-4	205	95.6	9	6	66.7	2
GM_CP-5	852	96.7	28	16	62.5	6
GM_CP-6	442	93.4	29	16	62.5	6
GM_CP-7	1032	95.3	48	18	38.9	11
ISO_CP-1	633	94.0	38	13	46.2	7
ISO_CP-2	329	97.0	10	6	66.7	2
ISO_CP-3	307	95.1	15	6	50.0	3
ISO_CP-4	113	96.5	4	4	100.0	0
ISO_CP-5	267	96.3	10	9	88.9	1
ISO_CP-6	199	95.0	10	7	71.4	2
ISO_CP-7	302	95.0	15	10	50.0	5
CON_CP-1	187	96.3	7	6	83.3	1
CON_CP-2	238	92.0	19	8	50.0	4
CON_CP-3	844	92.8	61	16	50.0	8

CON_CP-4	562	96.4	20	10	50.0	5
CON_CP-5	Missing					
CON_CP-6	758	95.6	33	15	60.0	6
CON_CP-7	354	95.8	15	11	63.6	4
GM_NS-1	340	94.1	20	12	75.0	3
GM_NS-2	97	96.9	3	3	100.0	0
GM_NS-3	239	96.2	9	8	87.5	1
GM_NS-4	188	95.2	9	5	60.0	2
GM_NS-5	267	96.3	10	7	57.1	3
GM_NS-6	221	95.9	9	7	71.4	2
GM_NS-7	273	95.2	13	9	55.6	4
ISO_NS-1	691	93.3	46	13	38.5	8
ISO_NS-2	266	92.9	19	12	66.7	4
ISO_NS-3	Missing					
ISO_NS-4	404	95.0	20	11	63.6	4
ISO_NS-5	117	91.5	10	8	75.0	2
ISO_NS-6	189	89.4	20	11	72.7	3
ISO_NS-7	98	93.9	6	3	33.3	2
CON_NS-1	161	91.3	14	8	50.0	4
CON_NS-2	237	92.0	19	11	54.5	5
CON_NS-3	362	93.6	23	11	45.5	6
CON_NS-4	220	93.6	14	9	66.7	3
CON_NS-5	286	91.3	25	11	45.5	6
CON_NS-6	902	95.5	41	19	63.2	7
CON_NS-7	419	94.7	22	11	54.5	5
nirS IRELAND 2014						
GM_CP-1	109	94.5	6	1	0.0	1
GM_CP-2	79	84.8	12	5	40.0	3
GM_CP-3	85	84.7	13	1	0.0	1
GM_CP-4	153	89.5	16	7	71.4	2
GM_CP-5	92	88.0	11	5	40.0	3
GM_CP-6	245	93.1	17	6	66.7	2
GM_CP-7	91	91.2	8	2	50.0	1
ISO_CP-1	107	93.5	7	4	50.0	2
ISO_CP-2	396	91.9	32	7	42.9	4
ISO_CP-3	312	90.7	29	4	25.0	3
ISO_CP-4	851	92.1	67	10	50.0	5
ISO_CP-5	347	94.5	19	6	66.7	2
ISO_CP-6	91	90.1	9	2	50.0	1
ISO_CP-7	180	90.0	18	7	71.4	2
CON_CP-1	94	88.3	11	1	0.0	1
CON_CP-2	162	93.2	11	2	0.0	2
CON_CP-3	121	91.7	10	4	50.0	2
CON_CP-4	124	87.1	16	4	50.0	2
CON_CP-5	115	90.4	11	2	50.0	1
CON_CP-6	104	91.3	9	1	0.0	1
CON_CP-7	485	82.1	87	7	71.4	2
GM_NS-1	287	80.8	55	10	50.0	5
GM_NS-2	1154	84.4	180	19	42.1	11

GM_NS-3	583	78.7	124	5	60.0	2
GM_NS-4	247	89.9	25	7	71.4	2
GM_NS-5	300	87.3	38	11	63.6	4
GM_NS-6	222	84.7	34	4	25.0	3
GM_NS-7	280	86.1	39	9	77.8	2
ISO_NS-1	227	77.1	52	10	40.0	6
ISO_NS-2	291	81.4	54	11	45.5	6
ISO_NS-3	295	83.1	50	10	70.0	3
ISO_NS-4	247	90.3	24	15	66.7	5
ISO_NS-5	1414	79.4	291	8	50.0	4
ISO_NS-6	570	83.7	93	10	60.0	4
ISO_NS-7	801	85.0	120	21	52.4	10
CON_NS-1	577	83.2	97	8	25.0	6
CON_NS-2	157	89.8	16	6	33.3	4
CON_NS-3	339	84.1	54	3	66.7	1
CON_NS-4	131	86.3	18	2	50.0	1
CON_NS-5	274	85.8	39	2	50.0	1
CON_NS-6	177	81.9	32	4	75.0	1
CON_NS-7	259	77.6	58	4	75.0	1

Cultivar/ Treatment	Raw sequences	Removed (%)	Quality sequences	OTU	Singletons (%)	OTU. no singletons
nirS NETHERLANDS 2013						
CON_CP-1	1343	96.3	50	7	14.3	6
CON_CP-2	450	93.6	29	6	33.3	4
CON_CP-3	240	94.2	14	6	50.0	3
CON_CP-4	391	94.6	21	7	57.1	3
CON_CP-5	109	95.4	5	3	33.3	2
CON_CP-6	187	94.1	11	5	40.0	3
CON_CP-7	186	94.1	11	4	50.0	2
CON_NS-1	312	95.2	15	4	50.0	2
CON_NS-2	170	95.3	8	3	0.0	3
CON_NS-3	260	95.8	11	6	50.0	3
CON_NS-4	931	93.1	64	10	40.0	6
CON_NS-5	584	92.8	42	8	37.5	5
CON_NS-6	953	91.9	77	8	25.0	6
CON_NS-7	799	92.0	64	10	20.0	8
nirS NETHERLANDS 2014						
GM_CP-1	198	93.4	13	8	75.0	2
GM_CP-2	Missing					
GM_CP-3	945	94.1	56	14	50.0	7
GM_CP-4	412	88.8	46	17	47.1	9
GM_CP-5	1074	88.5	124	27	40.7	16
GM_CP-6	702	91.7	58	21	42.9	12
GM_CP-7	199	91.0	18	6	16.7	5
ISO_CP-1	300	94.3	17	8	62.5	3

ISO_CP-2	183	90.7	17	7	42.9	4
ISO_CP-3	401	92.0	32	7	57.1	3
ISO_CP-4	228	89.5	24	11	54.5	5
ISO_CP-5	291	90.0	29	9	55.6	4
ISO_CP-6	242	93.0	17	5	40.0	3
ISO_CP-7	Missing					
CON_CP-2	Missing					
CON_CP-3	Missing					
CON_CP-4	Missing					
CON_CP-6	Missing					
GM_NS-2	198	93.9	12	4	75.0	1
GM_NS-3	372	95.2	18	12	58.3	5
GM_NS-4	Missing					
GM_NS-5	117	95.7	5	4	75.0	1
GM_NS-6	Missing					
GM_NS-7	377	90.5	36	14	50.0	7
ISO_NS-1	Missing					
ISO_NS-2	Missing					
ISO_NS-3	222	92.3	17	6	50.0	3
ISO_NS-4	230	86.5	31	7	28.6	5
ISO_NS-5	Missing					
ISO_NS-6	Missing					
ISO_NS-7	Missing					
CON_NS-1	95	89.5	10	6	83.3	1
CON_NS-2	185	93.0	13	8	75.0	2
CON_NS-3	87	88.5	10	4	50.0	2
CON_NS-4	503	90.9	46	22	54.5	10
CON_NS-5	Missing					
CON_NS-7	458	91.7	38	21	61.9	8