Impact of metagenomic DNA extraction procedures on the identifiable endophytic bacterial diversity in *Sorghum bicolor* (L. Moench)

Mapula Kgomotso Annah Maropola [a], Jean-Baptiste Ramond [a,b], Marla Trindade [a,*]

[a] *Institute for Microbial Biotechnology and Metagenomics (IMBM), Biotechnology Department, University of the Western Cape, Bellville 7535, Cape Town, South Africa*
[b] *Center for Microbial Ecology and Genomics (CMEG), Genomics Research Institute, University of Pretoria, South Africa*

A b s t r a c t

Culture-independent studies rely on the quantity and quality of the extracted environmental metagenomic DNA (mDNA). To fully access the plant tissue microbiome, the extracted plant mDNA should allow optimal PCR applications and the genetic content must be representative of the total microbial diversity. In this study, we evaluated the endophytic bacterial diversity retrieved using different mDNA extraction procedures. Metagenomic DNA from sorghum (*Sorghum bicolor* L. Moench) stem and root tissues were extracted using two classical DNA extraction protocols (CTAB- and SDS-based) and five commercial kits. The mDNA yields and quality as well as the reproducibility were compared. 16S rRNA gene terminal restriction fragment length polymorphism (t-RFLP) was used to assess the impact on endophytic bacterial community structures observed. Generally, the classical protocols obtained high mDNA yields from sorghum tissues; however, they were less reproducible than the commercial kits. Commercial kits retrieved higher quality mDNA, but with lower endophytic bacterial diversities compared to classical protocols. The SDS-based protocol enabled access to the highest sorghum endophytic diversities. Therefore, "SDS-extracted" sorghum root and stem microbiome diversities were analysed via 454 pyrosequencing, and this revealed that the two tissues harbour significantly different endophytic communities. Nevertheless, both communities are dominated by agriculturally important genera such as *Microbacterium*, *Agrobacterium*, *Sphingobacterium*, *Herbaspirillum*, *Erwinia*, *Pseudomonas* and *Stenotrophomonas*; which have previously been shown to play a role in plant growth promotion. This study shows that DNA extraction protocols introduce biases in culture-independent studies of environmental microbial communities by influencing the mDNA quality, which impacts the microbial diversity analyses and evaluation. Using the broad-spectrum SDS-based DNA extraction protocol allows the recovery of the most diverse endophytic communities associated with sorghum tissues and, as such, establishes a reliable basis for future study of endophytic communities.

1. Introduction

Endophytic bacteria play a crucial role in plant health and development, as they have either beneficial or detrimental effects. Beneficial endophytic bacteria – also known as plant growth promoting endophytic bacteria (PGPeB) – enhance the plant's growth and/or its ability to withstand stress (Schenk et al., 2012). They achieve this directly through production of plant-growth inducing phytohormones, such as indole acetic acid (IAA), gibberellins and auxins (Dodd et al., 2010), or indirectly via their varied nutrient-liberating metabolic activities such as nitrogen (N_2) fixation, phosphate (P) solubilisation and iron sequestration (Khan et al., 2009; Kraiser et al., 2011; Saha et al., 2012). PGPeB have also been implicated in enhancing the plant's resistance against pathogens through direct induction of the plant's defence system or production of antimicrobial compounds (Heydari and Pessarakli, 2010). Pathogenic and parasitic endophytic bacteria, on the other hand, cause plant disease and reduced fitness (Newton et al., 2010).

Studies of endophytic bacterial communities associated with grass species (*Poaceae* family) represent a well-established and continuously growing field (Seghers et al., 2004; Sun et al., 2008; Pereira et al., 2011; Sessitsch et al., 2012; Magnani et al., 2013); however, in-depth analyses in grasses such as sorghum are not available. Sorghum (and other grasses) is an important staple food in most developing countries (Taylor, 2004; Babalola and Glick, 2012) and, therefore, it is continuously being studied to increase yield of crop and its resistance to environmental and biological stress (Kapanigowda et al., 2013). The characterisation of the associated endophytic bacterial communities could make valuable contribution in this regard, and an understanding of their establishment and symbiotic roles within the plant has important implications in agriculture as it could lead to novel approaches in managing crop development and health (Schenk et al., 2012). Furthermore, endophytic

of agricultural applications (biofertilizers and biocontrol agents), production of industrial/medical bioproducts and/or for bioremediation processes (Ryan et al., 2007; Andrews et al., 2010).

One of the primary challenges faced in culture-independent studies of plant-associated microorganisms is the retrieval of good-quality metagenomic DNA (mDNA) that can allow reliable downstream analy-ses using PCR-based techniques (e.g. T-RFLP, DGGE, pyrosequencing)(Demeke and Jenkins, 2010). The mDNA extraction must (i) ensure lysis of all microbial cells, (ii) provide sufficient genomic material (Terrat et al., 2012), and (iii) efficiently remove plant-derived phyto-chemicals, which may result in PCR-inhibition (e.g. polysaccharides, polyphenolic compounds, and secondary metabolites) and enzymes (e.g. DNases, proteinases) (Wilson, 1997; Demeke and Jenkins, 2010). Lysis of all microbial cells is particularly critical since the genomic mate-rial retrieved will dictate the bacterial diversity observed (Sène et al., 2001; Terrat et al., 2012)This study evaluated the effects of seven DNA extraction protocols commonly used in endophytic bacterial community studies, to determine the sorghum root and stem community diversities. Two classical protocols (CTAB- or SDS-based) and five commercial kits were compared using 2 different weights of sorghum tissue. The 16S rRNA gene diversity of retrieved endophytic bacterial communities was assessed using terminal restriction fragment length polymorphism (t-RFLP), and the most diverse community determined (SDS-extracted) was further characterised by higher resolution 16S rRNA gene amplicon pyrosequencing.

2. Material and methods

2.1. Study site and sampling procedures

Sorghum (*Sorghum bicolor* L. (Moench)) plant samples were collect-ed from the Agricultural Research Council (ARC) academic farm situated in Potchefstroom (S26°44302′ E027°05584′, North West Province, South Africa) during the autumn season (April, 2012). At the time of sampling, sorghum plants were mature at 16 weeks old. This crop has been continuously planted on the field annually for four years. Water was primarily supplied via rainfall events. Prior to planting, the soil was fertilized with "3:2:1 (32) + ZN" at a rate of 150 kg N/ha. LAN 28 fertilizer was applied at a rate of 100 kg/ha when plants were at knee length, and the soil was also treated with the insecticide, Kombat® (Kombat, South Africa). The soil on this field was a mixture of clay, silt and fine sand.

Three mature and healthy plants were collected following a random sampling technique. Stems and roots were aseptically excised and placed in sterile bags. All samples were immediately placed on ice and transported to the Institute for Microbial Biotechnology and Metagenomics (IMBM, University of the Western Cape, South Africa), where they were stored at −80 °C prior to processing.

2.2. Plant tissue sterilisation and preparation for metagenomic DNA extraction

The plant organs were sterilised using a modified protocol designed by Mendes et al. (2007). The roots and stems were separately washed in autoclaved double-distilled water until all residual soil was removed from their surfaces. Plant organs were immersed in 500 mL 1× phos-phate buffer solution (PBS) for 1.5 h, shaking at 1 rcf. The samples were then sequentially washed by shaking in (i) 70% ethanol for 10 min, (ii) 2.5% sodium hypochlorite solution for 20 min and (iii) rinsed five times in autoclaved double-distilled water for 2 min. To con-firm sterility, 100 µL of the last wash water was plated on nutrient agar (NA) and R2A agar plates (Merck, Germany) and incubated at room temperature for three days. Sterility tests were conducted in triplicate. The plant tissues were stored, unshaken, in the last wash water at 4 °C during these 3 days.

Sterilisation was considered successful when no colonies were observed on either plate. If microbial growth was observed, the sterilisation process was repeated. Sterilisation was repeated at least once per sample in this study. Sterilised root or stem samples were pooled and aseptically ground to a fine powder in liquid nitrogen using autoclaved pestle and mortar. Ground tissue powder aliquots (100 mg) were then stored at −80 °C.

2.3. DNA extraction procedures

DNA was extracted from 0.1 g to 0.3 g ground root and stem tissues. Seven plant mDNA extraction protocols were tested; two classical pro-tocols (SDS- or CTAB-based) and five commercial kits (MoBio PowerPlant Pro® DNA Isolation Kit, Qiagen DNeasy[R] Plant Mini Kit, Fermentas GeneJet Plant Genomic DNA Purification Kit, MoBio PowerSoil™ DNA Purification Kit and MoBio UltraClean® Soil DNA Isolation Kit). These protocols have previously been used to study endophytic bacterial communities (Drabkova et al., 2002; Green et al., 1999; Krechel et al., 2002; West et al., 2010). All kit-based DNA extrac-tion protocols were performed according to the manufacturer's instruc-tions, with the exception that starting plant material quantities were always 0.1 g or 0.3 g, and the final elution was performed in 50 µL buffer for normalization purposes. All DNA extractions were carried out in triplicate.

The CTAB-based mDNA extraction was a modified version of a proto-col described by Murray and Thompson (1980). B r i e fly, 700 µL d o u b l e strength CTAB buffer (100 mM Tris–HCl; 1.2 M NaCl; 20 mM EDTA; 2%CTAB; 0.2% β-mercaptoethanol) was added to ground plant tissue. The mixture was vortexed for 20 s (maximum speed) and incubated at 65 °C for 1 h, followed by addition of 600 µL 24:1 (v/v) chloroform/isoamyl alcohol solution. The tubes were mixed ten times by inversion and centrifuged (13,000 rcf, 5 min). Equal volume ice-cold isopropanol and RNase A (50 µg·mL^{-1} final concentration) were added to the super-natant in a clean tube and mixed by inversion, followed by incubation at room temperature for 20 min and centrifuged (13,000 rcf, 5 min). The supernatant was discarded, and the pellets were air dried. DNA pellets were washed twice with 250 µL 70% ethanol, which was eluted follow-ing centrifugation (13,000 rcf, 5 min). The DNA pellets were air dried and then resuspended in 50 µL TE buffer (10 mM Tris–HCl; 1 mM EDTA) before storage at −20 °C.

The SDS-based method used in this study is a modified version of the protocol developed by Zhou et al. (1996). Five hundred microlitres of lysozyme buffer (25 mM Tris–HCl; 50 mM glucose; 10 mM EDTA; 25 mg·mL^{-1} lysozyme) and RNase A (50 µg·mL^{-1} final concentration) were added to ground plant tissue powders and vortexed for 20 s. The mixtures were incubated at 37 °C for 1 h, and then treated with Protein-ase K (1 mg·mL^{-1} final concentration) at 37 °C for 1 h. SDS was added to 1% final concentration, and mixed by inverting the tubes 10 times. Mixtures were incubated at 65 °C for 30 min. Tubes were centrifuged (14,000 rcf, 2 min) and the supernatants collected into new tubes. Equal volume phenol was added to each tube and mixed by inversion. Top aqueous phase containing DNA was collected after centrifugation (10,000 rcf, 1 min) and the bottom layer with organic phenol was discarded. The phenol extraction was repeated once. Equal volume 24:1 (v/v) chloroform/isoamyl alcohol solution was added to each tube and mixed by inversion. Top aqueous layer was collected and transferred to a new tube after centrifugation (10,000 rcf, 10 min). The tubes were placed on ice and equal volume ice-cold isopropanol was added, followed by incubation at 4 °C for 20 min. The tubes were centrifuged (10,000 rcf, 5 min) to recover mDNA and the isopropanol was discarded. DNA pellets were air dried under the laminar flow cabinet and then washed twice with 250 µL 70% ethanol, which was eluted after centrifugation (10,000 rcf, 5 min). The DNA pellets were air dried and then resuspended in 50 µL a ut oc la ved T Eb uf fer a nd st or ed at −20 °C.

2.4. PCR amplification and terminal restriction fragment length polymorphism (t-RFLP) analyses

Bacterial 16S rRNA gene amplification was conducted on mDNA with starting materials of 10 ng, 5 ng and 1 ng per reaction, using the universal primers E9F (5′-GAGTTTGATCCTGGCTCAG-3′) and U1510r (5′-GGTTACCTTGTTACGACTT-3′) (Marchesi et al., 1998; Turner et al., 1999). PCRs were carried out in 50 μL volumes containing 1× DreamTaq Buffer, 0.2 mM each dNTP, 0.5 M each primer, template DNA, 0.3 μL DreamTaq DNA polymerase (Fermentas, Lithuania) and deionised nuclease-free water, in a Labnet MultiGene™ Gradient PCR Thermal Cycler (Labnet International, Inc.). Cycling conditions were as follows: 4 min at 94 °C for initial denaturation; 30 cycles of 30 s denaturation at 94 °C, 30 s annealing at 52 °C and 105 s extension at 72 °C; and a final elongation step of 10 min at 72 °C. All PCRs were carried out in triplicate.

For t-RFLP analyses, the forward E9F primer (E9F) was 5′ end labelled with fluorescent dye fluorescein amidite (FAM). The labelled PCR amplicons were purified with the Illustra GFX™ PCR DNA and Gel Purification Kit (GE Healthcare, UK) according to manufacturer's instructions. Purified PCR products (200 ng) were digested with the restriction enzyme HaeIII (Fermentas, Lithuania) at 37 °C overnight. The lengths of fluorescently labelled terminally restricted fragments (t-RFs) were determined on an Applied Biosystems Genetic Analyzer sequencer (Applied Biosystems, Foster City, California, USA) at the Central Analytical Facility of the University of Stellenbosch, using the internal size standard marker, ROX1.1 (sizes in bp: 75, 100, 139, 150, 160, 200, 300, 340, 350, 400, 450, 490, 500, 583, 683, 782, 932, 991, 1121). T-RFLP patterns were analysed using the Peak Scanner™ Software V1.0 (Applied Biosystems). Valid peaks (between 35 and 1000 bp) were identified and analysed using the online T-REX software (http://trex.biohpc.org/) (Culman et al., 2009). T-RFs were characterised by peak height and aligned to create an operational taxonomic unit (OTU) data matrix. The term OTU refers to an individual t-RF, with recognition that one t-RF could comprise more than one ribotype (Blackwood et al., 2007).

2.5. Pyrosequencing

The bacterial 16S rRNA gene was amplified from 0.3 g sorghum root and stem mDNA extracted with the SDS-based protocol using multiplexed 8F (5′-CCATCTCATCCCTGCGTGTCTCCGAC-3′) and 518R (ATTACCGCGGCTGCTGG) primers (Turner et al., 1999; Muyzer et al., 1993). PCRs were carried out in 25 μL volumes containing 1× Phusion HF Buffer, 200 μM each dNTP, 0.5 μM each primer, 50 ng template DNA, 0.02 U/μL Phusion High Fidelity DNA Polymerase (Fermentas, Lithuania) and deionised nuclease-free water. PCR conditions were as follows: 2 min at 95 °C for initial denaturation; 25 cycles of 20 s denaturation at 98 °C, 15 s annealing at 75 °C and 90 s extension at 72 °C; 10 min extension at 72 °C. One composite mDNA sample was used per tissue. Five PCRs were prepared for each mDNA sample, and PCR products were pooled into single composite samples per tissue during the purification process and quantified. The amplicons were submitted to the Next Generation Sequencing Facility at the University of the Western Cape for pyrosequencing on the Roche 454 GS Junior System (Roche, Branford, Germany).

2.6. Data analysis

Statistical tests for DNA yield and quality as well as diversity indices were performed using the software Sigma-Plot, Version 11.0 (Systat Software, Inc.). Two-way analysis of variance (ANOVA) was used to compare yield and purity of DNA extracted using different DNA extraction protocols. Normality tests were performed on the data following the method of Kolmogorov–Smirnov, with Lillifor's correction (Justel et al., 1997). Data that did not pass the normality test was compared using the Holm–Sidak test that ranks the ordinal numbers and compares the median of the samples (Holm, 1979). Paired T-tests were conducted to compare differences in DNA yield at 0.1 g and 0.3 g starting plant material within individual tissues. Where normality criteria were not met, the Wilcoxon Signed Rank test was used (Wilcoxon, 1945).

The sample data matrix generated from t-RFLP profiles was analysed with the software Primer 6, version 6.1.11 (Primer E, Plymouth, UK). Diversity indices, i.e. species richness (S), Shannon index (H′) and the Simpson index $(1 - \lambda)$ (Clarke and Warwick, 2001), were calculated from untransformed data using the Diverse function. S is a direct count of operational taxonomic units (OTUs) observed. H′ measures the proportion of all OTUs in the whole community, and it is calculated as $H' = -\sum_i p_i \log(p_i)$, whereby p_i is the proportion of the total count arising from the ith OTU. $(1 - \lambda)$ measures community evenness (or equitability), which quantifies how evenly distributed OTUs are within a community. It is calculated as $1 - \lambda = 1 - [\sum_i N_i(N - 1)] / [N(N - 1)]$, where N_i is the number of OTUs that belong to species i.

The T-RFLP dataset was standardised by the presence–absence transformation, and used to calculate the Bray–Curtis similarity coefficients between samples (Bray and Curtis, 1957), which were used to create a resemblance matrix. The latter in fine led to the construction of 3D-Nonmetric Multi-Dimensional (nMDS) plots, which are ordinations of sample communities based on their relative similarities, i.e. the distance between two points reflects the degree of similarity between microbial community profiles (Clarke and Warwick, 2001; Shephard, 1962). Analysis of similarity (ANOSIM) was performed on the resemblance matrices to test for differences in structure, between a priori defined communities.

Pyrosequencing outputs from the 454 GS Junior software included a quality file (QUAL format) with information on the sequencing process and a metadata file (FASTA format) containing raw 16S rRNA sequences. The two files were processed using the CloVR-16S pipeline version 1.1., which comprises of a suite of phylogenetic tools (Angiuoli et al., 2011). Preprocessing, processing and analysis of data were done using modules in QIIME (Quantitative Insights into Microbial Ecology http://qiime.org), R (http://www.R-project.org/) and MOTHUR (Schloss et al., 2009).

Quality assessment and filtering of raw sequences were performed using Prinseq (Schmieder and Edwards, 2011). Multiplexed reads were split and assigned to samples based on their MID sequences using a Python script. Sequences were then trimmed and filtered to include only good-quality sequences of 200–470 bp. Sequences with high ambiguous base (N) occurrence and poly-A/T tails were removed. De novo chimera detection and OTU picking were performed with UCHIME and UCLUST, respectively (Edgar, 2010; Edgar et al., 2011). In this study, an OTU is defined as a cluster of sequences (Floyd et al., 2002), delimited at ≥97% sequence similarity. The representative sequences were aligned and used to create an OTU distance matrix with PyNAST (Caporaso et al., 2010). The distance matrix was converted to a phylogenetic tree with FastTree (Price et al., 2009). Taxonomy was assigned based on the Greengenes taxonomy and a Greengenes reference database (version 12.10) (McDonald et al., 2012), using the RDP Classifier (version 2.2) (Werner et al., 2012).

The OTU table was used to calculate the alpha diversity (within sample diversity). In this study, the Chao1 diversity index was calculated as a measure of true species diversity using the formula, $S_1 = S_{obs}(F_1^2/2F_2)$, where S_{obs} is the number of species observed, F1 is the number of singletons (occur once) and F2 is the number of doubletons (Gotelli and Colwell, 2011). The Simpson index $(1 - \lambda)$ was calculated to measure community evenness.

The taxonomic predictions for each OTU were used to create heatmaps and bar charts to reflect the distribution of phylotypes within each sample. UniFrac was then used to determine whether there was a significant difference in the structure of endophytic bacterial communities retrieved from the two tissues. Briefly, UniFrac measures the distance between each pair of environments as a fraction of the total

3

branch length in a phylogenetic tree, leading to sequences of one environment (Lozupone and Knight, 2005). To compare pairs of environments, the UniFrac value is calculated for each pair and a distance matrix is created (Lozupone and Knight, 2005; Lozupone et al., 2006, 2007). The two communities were then subjected to a statistical UniFrac test (F-test) to test for significance of difference at 95% confidence level, using the unweighted UniFrac metric for qualitative comparisons of β-diversities not affected by individual sequence abundance.

2.7. Characterisation of sorghum endophytic communities

The metabolic potential of sorghum endophytic communities was determined through literature survey. Properties of representative species of identified genera were recorded from published studies, in order to determine the agricultural and industrial potential of the sorghum endophytome. Sorghum endophytic communities analysed here were retrieved from unreplicated, but composite samples. Therefore, to exercise caution, only "dominant" phylotypes were characterised. In this case, "dominant" phylotypes constitute ≥0.1% of the total community, and these are most likely to be observed again in replicated studies (Charlson et al., 2012).

3. Results

3.1. Effects of DNA extraction protocols on mDNA yield and quality

Independently from starting plant material (0.1 g or 0.3 g) and plant tissue (stem or root), the classical CTAB- (ranging from 11.6 ± 2.8 to 19.1 ± 0.6 ng·mg^{-1}) and SDS-based (ranging from 48.0 ± 8.6 to 76.4 ± 4.3 ng·mg^{-1}) DNA extraction protocols retrieved significantly higher yields of sorghum mDNA than commercial kits (all ≤6.0 ± 0.3 ng·mg^{-1}) (ANOVA, $p < 0.001$) (Table 1), with also higher molecular weight (ranging between 10 kb and 20 kb). Appendices A1 and A2 show the visualisation of extracted mDNA by electrophoresis. The SDS-based protocol was particularly efficient when using 0.1 g starting plant material as it provided the highest mDNA yields for both stem and root tissue, with 48.0 (± 8.5) ng·mg^{-1} and 53.2 (± 17.7) ng·mg^{-1}, respectively. However, when compared to the other protocols, it was the least reproducible with respect to mDNA yields (Table 1). Contrastingly, the other protocols were highly reproducible, despite significantly lower yields. Increasing the starting plant material to 0.3 g resulted in a significant increase in mDNA yield and reproducibility only for SDS-extracted sorghum stem mDNA, to 76.4 (± 4.3) ng·mg^{-1}. The SDS-based protocol did not retrieve the most pure sorghum tissue mDNA (Table 1). In contrast, the CTAB-based protocol and Fermentas

GeneJet kit consistently retrieved pure mDNA, whilst the Qiagen DNeasy and MoBio PowerPlant kits generated inconsistent results (i.e., dependent on either the starting material size or the tissue extracted). Soil DNA extraction kits (MoBio PowerSoil and MoBio UltraClean) were only used on 0.1 g starting plant material as both failed to extract high-yield and pure mDNA (Table 1).

Successful 16S rRNA gene PCR amplification from kit-extracted mDNA was consistently observed (as shown in Appendices B1 and B2), indicating that minimal PCR-inhibiting compounds were co-extracted. Contrastingly, PCR inhibition was more frequently observed when CTAB- and SDS-extracted mDNA was used as template. However, there was no clear correlation between PCR inhibition and template mDNA concentration used.

3.2. Effects of mDNA extraction protocols on endophytic bacterial community diversity and structure

T-RFLP was used to compare the endophytic bacteria community diversities associated with sorghum tissues. Since the MoBio PowerPlant and MoBio PowerSoil kits retrieved significantly low mDNA yields, they were excluded from these analyses.

With 0.1 g starting material, the number of OTUs (S) retrieved with the classical protocols from sorghum stem and root [8 (± 1) to 10 (± 3)] was significantly higher than those retrieved with commercial kits [3(± 1) to 7(± 3)] (ANOVA, p-values < 0.05) (Fig. 1). With Shannon index values (H′) of less than 1.7 and Simpson index $(1 - \lambda)$ values ranging from 0.37 to 0.75, all sorghum endophytic communities retrieved had low diversity with low to moderate evenness.

Due to the consistent similarity in performance (mDNA yield and quality) observed between kit protocols (Table 1) in this study, only the Fermentas GeneJet kit was compared to classical protocols at 0.3 g starting plant material. An increase in starting plant material only resulted in a significant increase in community diversity with the SDS protocol (Fig. 1), with up to 31 (± 3.0) and 23 (± 3.0) OTUs retrieved from sorghum stem and root tissues, respectively. Contrastingly, CTAB-retrieved communities displayed significantly decreased richness and evenness when starting plant material was increased, whilst no changes in species richness or community evenness were observed when using the GeneJet kit (Fig. 1).

The 3D-Nonmetric Multi-Dimensional Scaling (nMDS) ordinations of the community t-RFLP profiles retrieved from the "0.1 g and 0.3 g communities" presented stress values of 0.11 and 0.09, respectively, indicating low levels of scatter and minimal prospects of misinterpretation (Fig. 2) (Clarke and Warwick, 2001). Using 0.1 g plant material, nMDS ordination revealed two significantly different clusters (ANOSIM,

Table 1
Yield (ng·mg^{-1}) and purity of sorghum mDNA retrieved with different DNA extraction protocols. Sorghum mDNA was extracted using classical protocols (SDS- or CTAB-based) and commercial kits (Fermentas GeneJet Plant Genomic DNA Purification Kit, Qiagen DNeasy^R Plant Mini Kit, MoBio PowerPlant Pro® DNA Isolation Kit, MoBio PowerSoil™ DNA Purification Kit and MoBio UltraClean® Soil DNA Isolation Kit). DNA is considered pure (shown in bold) when Nanodrop-estimated A$_{260}$/A$_{280}$ ratio is between 1.7 and 1.9.

Protocol	Sorghum stem				Sorghum root			
	0.1 g		0.3 g		0.1 g		0.3 g	
	Yield	Purity	Yield	Purity	Yield	Purity	Yield	Purity
SDS	48.0 ± 8.6^a	1.6 ± 0.0	76.4 ± 4.3^b	2.03 ± 0.1	53.2 ± 17.7^e	1.55 ± 0.1	60.0 ± 5.3^e	1.92 ± 0.1
CTAB	13.7 ± 1.1^c	$\mathbf{1.76 \pm 0.0}$	11.6 ± 2.8^c	$\mathbf{1.76 \pm 0.1}$	19.1 ± 0.6^f	$\mathbf{1.77 \pm 0.0}$	17.3 ± 1.2^f	$\mathbf{1.74 \pm 0.0}$
GeneJet	4.4 ± 0.6^d	$\mathbf{1.75 \pm 0.0}$	2.3 ± 0.2^d	$\mathbf{1.77 \pm 0.0}$	6.0 ± 0.3^g	$\mathbf{1.77 \pm 0.0}$	3.5 ± 0.4^g	$\mathbf{1.74 \pm 0.0}$
Dneasy	3.93 ± 0.2^d	1.62 ± 0.0	2.61 ± 0.21^d	$\mathbf{1.75 \pm 0.0}$	4.9 ± 0.2^g	$\mathbf{1.78 \pm 0.0}$	3.5 ± 0.1^g	1.67 ± 0.0
PowerPlant	4.1 ± 0.5^d	$\mathbf{1.70 \pm 0.0}$	3.7 ± 0.9^d	1.56 ± 0.0	5.9 ± 1.0^g	1.51 ± 0.0	4.8 ± 0.3^g	$\mathbf{1.77 \pm 0.0}$
PowerSoil	1.9 ± 0.1^d	1.60 ± 0.0	ND	ND	3.5 ± 0.2^g	1.62 ± 0.0	ND	ND
UltraClean	1.4 ± 0.0^d	1.57 ± 0.1	ND	ND	2.1 ± 0.6^g	1.62 ± 0.1	ND	ND

Values are given as mean ± standard deviation. Superscripts denote statistical comparisons in DNA yields (a–d for sorghum stem, and e–g for sorghum roots) (ANOVA, 95% significance level). DNA yield values with the same superscript letter are significantly similar. ND: not determined, as excluded from experiment.

Global R = 0.752, p-value = 0.001), clearly distinguishing sorghum root and stem communities, independently from the mDNA extraction protocol used (Fig. 2A). However, the structure of sorghum bacterial communities was also significantly affected by the DNA extraction protocol used (ANOSIM, Global R = 0.324; p-value = 0.001; Table 2). Communities retrieved with the SDS-based protocol from either tissue differed significantly from those retrieved with the other protocols (ANOSIM, $0.324 \leq R \leq 0.769$; $0.01 \leq$ p-values ≤ 0.05). The CTAB-protocol-retrieved communities were also significantly different to communities retrieved with commercial kits. When comparing commercial kits, a significant difference was only observed between communities retrieved with the Fermentas GeneJet kit and those with the MoBio UltraClean kit.

When starting plant material was increased to 0.3 g, sorghum root and stem communities remained significantly different (ANOSIM, Global R = 0.796; p-value = 0.002; Table 2), and community structures were even more significantly affected by the extraction protocol used (ANOSIM, Global R = 0.689; p-value = 0.001) (Fig. 2B, Table 2). SDS-retrieved communities were significantly different from CTAB- and GeneJet-retrieved communities, which were not significantly different (Table 2).

Interestingly, a generally strong clustering of the replicate communities from 3 different SDS-extracted mDNAs from either sorghum stem or root tissues is observed (Fig. 2B). This indicates that highly similar endophytic communities (structure and diversity) are reproducibly retrieved when using this protocol.

3.3. Diversity of sorghum endophytic communities by pyrosequencing

The most diverse sorghum associated bacterial communities (based on microbial community fingerprinting) were retrieved when sorghum mDNA was extracted with the SDS-protocol from 0.3 g starting plant material. Therefore, for a higher resolution analysis, bacterial 16S rRNA pyrosequencing of sorghum mDNA was performed using these mDNA extracts.

A total of 13086 and 5831 sequences were generated from sorghum stem and root samples, respectively. These were curated and clustered into 829 OTUs, of which 15.6% (129/829) was shared between the
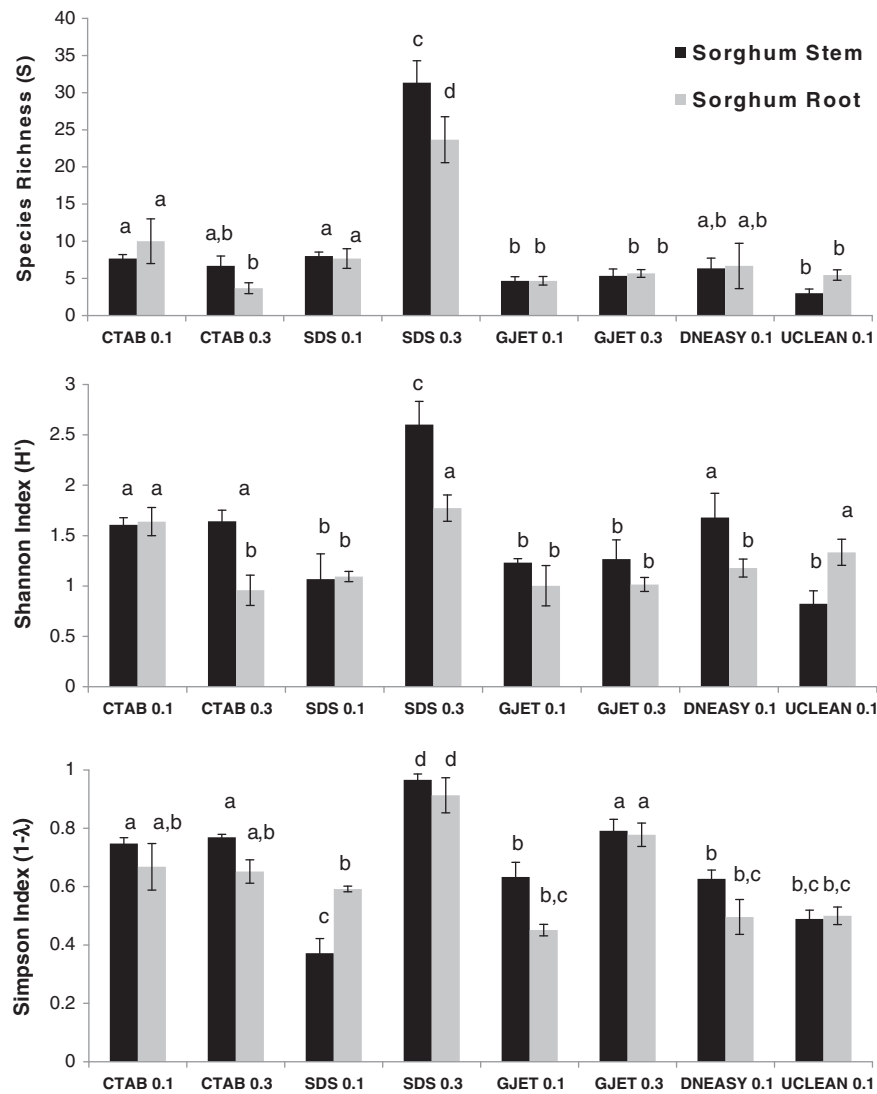

Fig. 1. Diversity of sorghum associated endophytic communities determined by t-RFLP. Diversity indices were calculated from t-RFLP profiles of stem and root communities retrieved with classical protocols and commercial kits from 0.1 g or 0.3 g plant tissue. Small letters (a-d) denote statistical (dis)similarities between endophytic communities (ANOVA, 95% confidence).
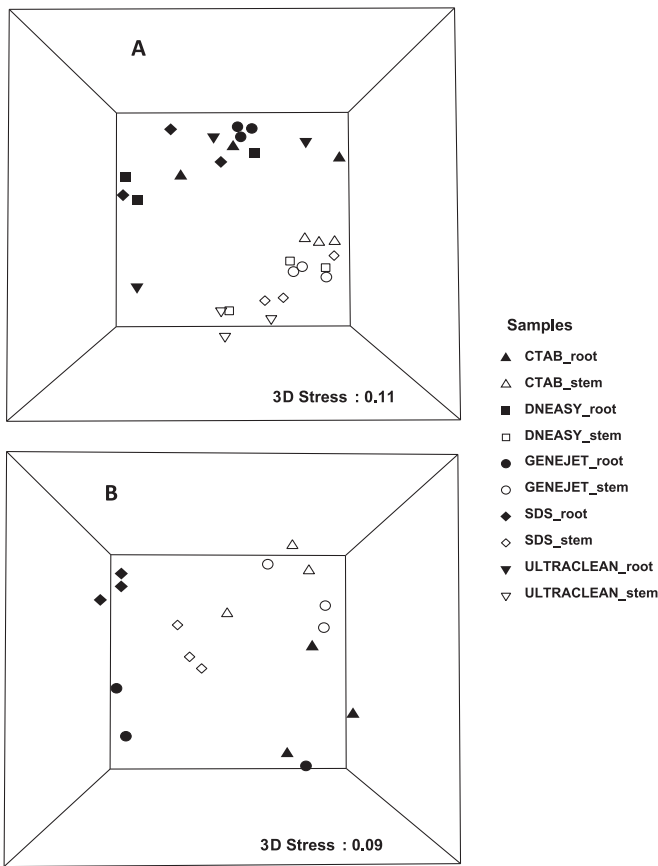
5

Fig. 2. Three-dimensional Non-Metric Multi-Dimensional Scaling (3D-nMDS) plot of Bray–Curtis similarity of bacterial community structures. Plots are based on the presence–absence transformed bacterial 16S rRNA gene t-RFLP profiles retrieved from 0.1 g (A) and 0.3 g (B) sorghum root and stem tissue using classical and commercial kit protocols.

Genera dominant only in the root tissue include *Leuconostoc* (0.1%), *Rhizobium* (1.0%) and *Sphingobium* (4.3%), whilst genera restricted to the stem tissue include *Rhodococcus* (1.5%), *Lactococcus* (6.7%) and *Swaminathania* (2.4%). Approximately 5% of the phylotypes found in each tissue and affiliated to the α-Proteobacteria class could not be classified at the genus level.

Table 3 highlights the metabolic potential of the dominant sorghum endophytes identified. Dominant sorghum-associated bacterial genera such as *Pseudomonas*, *Erwinia*, *Stenotrophomonas*, *Agrobacterium*, *Bacillus* and *Paenibacillus* have previously been associated with plant growth promoting attributes (e.g. nutrient acquisition, antimicrobial potential and phytohormone production) as well as pathogenicity (Table 3). There is little reported on the role of *Janthinobacteria* and *Methyloversatilis* previous endophytic analyses.

4. Discussion

DNA extraction constitutes one of the most critical components in any cultivation-independent study; particularly when determining microbial community diversity (Demeke and Jenkins, 2010). Accurate estimation of community profiles and diversity relies on the DNA's representativeness of the indigenous communities, and its usability in PCR (Demeke and Jenkins, 2010; Terrat et al., 2012). The current study established that different plant mDNA extraction procedures lead to highly variable mDNA quantity and quality, which is reflected in the endophytic bacterial community diversities observed.

Conventional SDS- or CTAB-based protocols retrieved high yields of sorghum mDNA compared to commercial kit protocols, as previously observed on fresh and historical plant samples (Drabkova et al., 2002; Niu et al., 2008; Sahu et al., 2012). Metagenomic DNA yield relies on the efficiency of the cell lysis step, which includes mechanical (e.g. grinding, dead-mill) and chemical (e.g. enzymatic lysis) processes for cell disruption (Moré et al., 1994). Therefore, it is reasonable to conclude that the combination of tissue grinding under liquid nitrogen and lysis buffers used in classical protocols was more efficient than the lytic procedures of commercial kits used in this study. It is important to note that plant DNA extracting kit manufacturers recommend the use of a bead-mill for increased mDNA yield, and its efficiency in this regard was previously demonstrated (Miller et al., 1999). However,

stem and root tissues (Appendix C). Chao1 values [645.76 (stem) and 502.82 (root)] demonstrated that the sorghum endophytic microbiome is diverse. Moreover, the evenness was approaching maximum equitability (Simpson index > 0.9) for both tissue communities. Rarefaction curves generated from retrieved sequence information are shown in Appendix D.

Pyrosequencing confirmed the t-RFLP results, which revealed that sorghum stem and root communities were significantly different (F-test, p-value < 0.05). The dominant phyla in both communities were Proteobacteria, Firmicutes and Actinobacteria (Fig. 3). Proteobacteria were the most dominant, with γ-Proteobacteria constituting approximately 53% of the total community in each tissue, α-Proteobacteria representing 28.1% and 23.6%, and β-Proteobacteria, 3.6% and 5.8% of stem and root communities, respectively. Firmicutes accounted for 7.2% and 3.3%, and Actinobacteria, 4.4% and 12.3%, for stem and root communities, respectively.

A total of 134 bacterial phylotypes in both tissues were identified (at genus level), and 50 (37.3%) of these were shared (Fig. 4). Of these, 20 were considered "dominant", i.e., constituting ≥ 0.1% in either tissue. These collectively represent approximately 90% of the sorghum endophytic community. Bacterial genera shown to be particularly dominant in both the root and stem tissues include *Pseudomonas* (46.5% and 10.1% in root and stem, respectively), *Agrobacterium* (8.7% and 14.2%), *Stenotrophomonas* (4.7% and 11.5%), *Erwinia* (0.1% and 30.2%), *Herbaspirillum* (3.4% and 2.6%), *Paenibacillus* (1.8% and 0.1%), *Microbacterium* (9.7% and 0.2%) and *Curtobacterium* (1.6% and 2.0%).

Table 2

Two-way crossed ANOSIM test results between plant tissue types and DNA extraction protocols. ANOSIM tests for differences between communities retrieved from stem and root sorghum tissues using classical and commercial kit DNA extraction protocols, based on Bray–Curtis similarity matrices from the presence–absence transformed bacterial t-RFLP profiles.

	0.1 g plant material		0.3 g plant material	
	R	P	R	P
Differences among plant tissue types				
Stem vs root	0.752	0.001[*]	0.796	0.002[*]
Differences among DNA extraction protocols				
Global Test	0.324	0.001[*]	0.689	0.001[*]
CTAB vs SDS	0.454	0.02[*]	0.963	0.01[*]
CTAB vs GeneJet	0.287	0.02[*]	0.306	0.09
CTAB vs DNeasy	0.454	0.02[*]	N/A	N/A
CTAB vs UltraClean	0.352	0.11	N/A	N/A
SDS vs GeneJet	0.769	0.01[*]	0.722	0.01[*]
SDS vs DNeasy	0.324	0.05[*]	N/A	N/A
SDS vs UltraClean	0.352	0.05[*]	N/A	N/A
GeneJet vs DNeasy	0.139	0.36	N/A	N/A
GeneJet vs UltraClean	0.287	0.03[*]	N/A	N/A
DNeasy vs UltraClean	−0.028	0.54	N/A	N/A

R: ANOSIM Statistic; p: probability level.
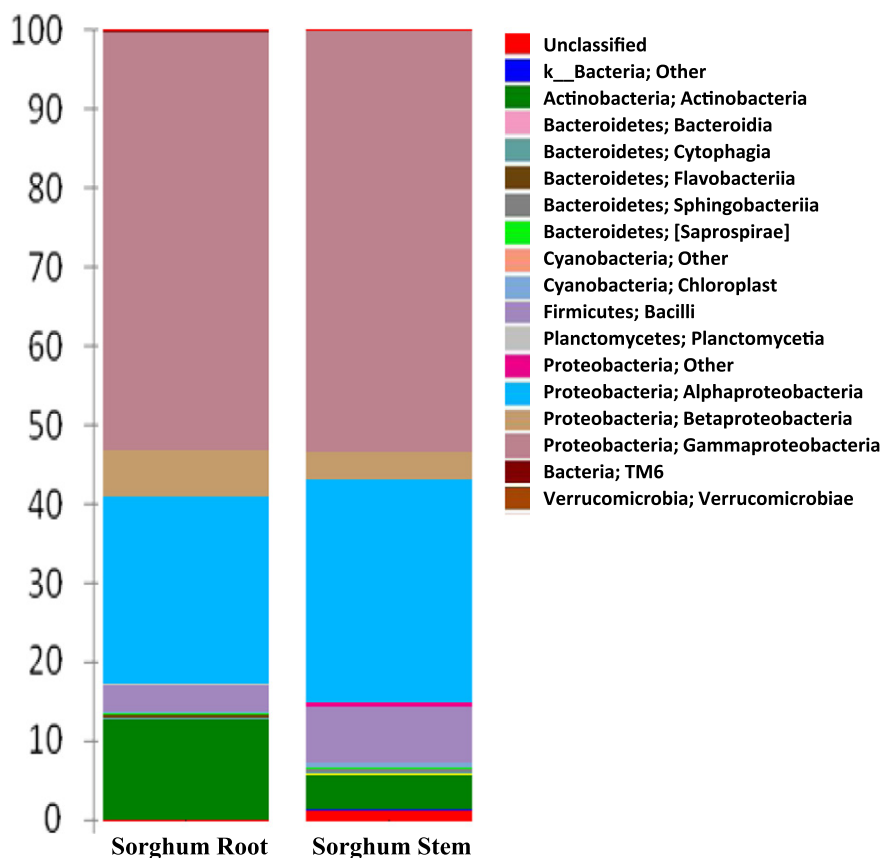[*] Statistically different (p < 0.05).

Fig. 3. Relative abundance of major bacterial lineages recovered from composite samples of sorghum root and stem tissues. Relative abundance is calculated as the percentage of sequences belonging to a particular lineage out of all 16S rRNA gene sequences recovered from each plant tissue.
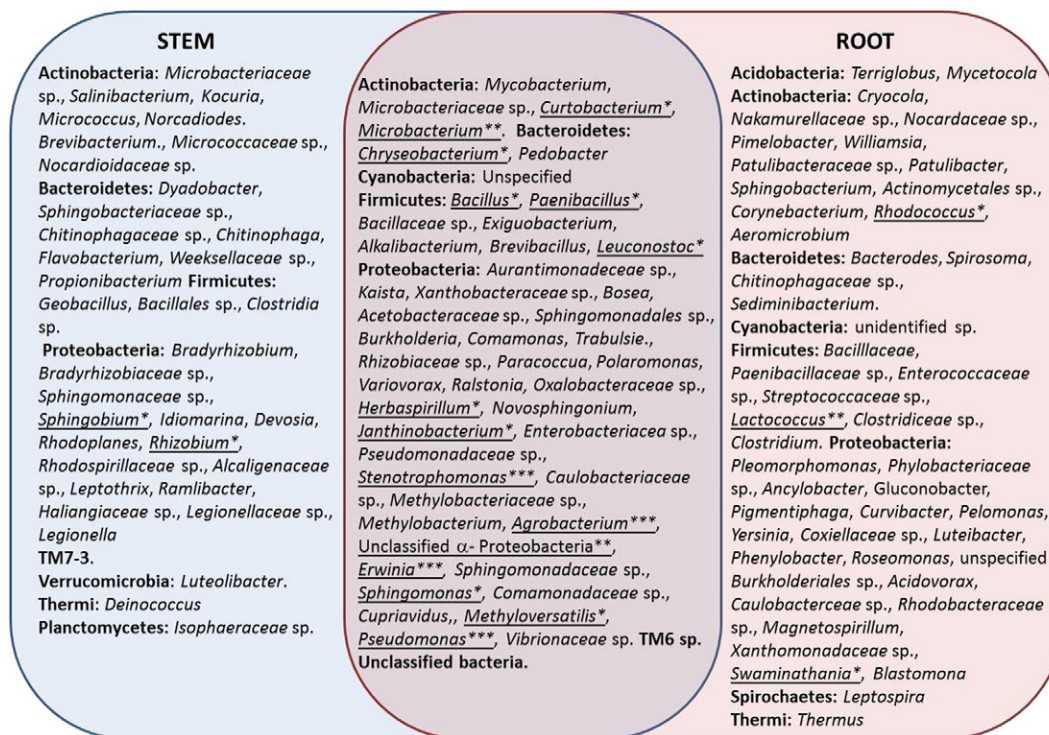

STEM

Actinobacteria: *Microbacteriaceae* sp., *Salinibacterium, Kocuria, Micrococcus, Norcadiodes. Brevibacterium., Micrococcaceae* sp., *Nocardioidaceae* sp.
Bacteroidetes: *Dyadobacter, Sphingobacteriaceae* sp., *Chitinophagaceae* sp., *Chitinophaga, Flavobacterium, Weeksellaceae* sp., *Propionibacterium* **Firmicutes:** *Geobacillus, Bacillales* sp., *Clostridia* sp.
Proteobacteria: *Bradyrhizobium, Bradyrhizobiaceae* sp., *Sphingomonaceae* sp., *Sphingobium*, Idiomarina, Devosia, Rhodoplanes, Rhizobium*, Rhodospirillaceae* sp., *Alcaligenaceae* sp., *Leptothrix, Ramlibacter, Haliangiaceae* sp., *Legionellaceae* sp., *Legionella*
TM7-3.
Verrucomicrobia: *Luteolibacter.*
Thermi: *Deinococcus*
Planctomycetes: *Isophaeraceae* sp.

Actinobacteria: *Mycobacterium, Microbacteriaceae* sp., *Curtobacterium*, Microbacterium**.* **Bacteroidetes:** *Chryseobacterium*, Pedobacter*
Cyanobacteria: Unspecified
Firmicutes: *Bacillus*, Paenibacillus*, Bacillaceae* sp., *Exiguobacterium, Alkalibacterium, Brevibacillus, Leuconostoc**
Proteobacteria: *Aurantimonadeceae* sp., *Kaista, Xanthobacteraceae* sp., *Bosea, Acetobacteraceae* sp., *Sphingomonadales* sp., *Burkholderia, Comamonas, Trabulsie., Rhizobiaceae* sp., *Paracoccua, Polaromonas, Variovorax, Ralstonia, Oxalobacteraceae* sp., *Herbaspirillum*, Novosphingonium, Janthinobacterium*, Enterobacteriacea* sp., *Pseudomonadaceae* sp., *Stenotrophomonas***, Caulobacteriaceae* sp., *Methylobacteriaceae* sp., *Methylobacterium, Agrobacterium***, Unclassified* α- *Proteobacteria**, Erwinia***, Sphingomonadaceae* sp., *Sphingomonas*, Comamonadaceae* sp., *Cupriavidus,, Methyloversatilis*, Pseudomonas***, Vibrionaceae* sp. **TM6 sp.**
Unclassified bacteria.

ROOT

Acidobacteria: *Terriglobus, Mycetocola*
Actinobacteria: *Cryocola, Nakamurellaceae* sp., *Nocardaceae* sp., *Pimelobacter, Williamsia, Patulibacteraceae* sp., *Patulibacter, Sphingobacterium, Actinomycetales* sp., *Corynebacterium, Rhodococcus*, Aeromicrobium*
Bacteroidetes: *Bacterodes, Spirosoma, Chitinophagaceae* sp., *Sediminibacterium.*
Cyanobacteria: unidentified sp.
Firmicutes: *Bacilllaceae, Paenibacillaceae* sp., *Enterococcaceae* sp., *Streptococcaceae* sp., *Lactococcus**, Clostridiceae* sp., *Clostridium.* **Proteobacteria:** *Pleomorphomonas, Phylobacteriaceae* sp., *Ancylobacter, Gluconobacter, Pigmentiphaga, Curvibacter, Pelomonas, Yersinia, Coxiellaceae* sp., *Luteibacter, Phenylobacter, Roseomonas,* unspecified *Burkholderiales* sp., *Acidovorax, Caulobacterceae* sp., *Rhodobacteraceae* sp., *Magnetospirillum, Xanthomonadaceae* sp., *Swaminathania*, Blastomona*
Spirochaetes: *Leptospira*
Thermi: *Thermus*

Fig. 4. Venn diagram representation of the composition of sorghum endophytic bacterial communities in stem and root tissues. A total of 134 phylotypes were identified from sorghum tissues through 16S rRNA gene amplicon pyrosequencing, with 50 phylotypes shared between the root and stem tissues. Twenty-one "dominant" phylotypes were observed (underlined, * indicates phylotypes accounting for over 0.1%, ** for over 5% and *** for over 10% of the total population in the root and/or stem tissue).

for the benefit of consistency in comparisons, this was not used in this study.

Despite lower DNA yields, commercial kits – notably, the Fermentas GeneJet kit – extracted mDNA of superior PCR-quality compared to classical protocols (Green et al., 1999; Drabkova et al., 2002). The observed low purity of mDNA retrieved with classical protocols and the low PCR-efficiency when it was used as template in PCR could result from a high concentration of co-extracted plant polyphenolics and polysaccharides which are known to bind to DNA, making it inaccessible to the polymerase enzyme (Horne et al., 2004; Varma et al., 2007; Demeke and Jenkins, 2010; Mornkham et al., 2012). Indeed, it was previously shown that chloroform and phenol in classical protocols were not consistent in efficiently removing potential PCR inhibitors. Moreover, reagents used in these processes such as CTAB, SDS, phenol, chloroform and ethanol, can also contaminate extracted mDNA and lead to PCR inhibition (Demeke and Jenkins, 2010).

T-RFLP was effective in evaluating the effects of plant mDNA extraction protocols. Classical protocols enabled access to more diverse endophytic bacterial communities than commercial kits. It is unclear from these findings, whether there is a relationship between plant mDNA yield and the diversity of endophytic bacteria retrieved. However, Scupham et al. (2007) and Salonen et al. (2010) have argued that DNA yield or purity was not always correlated to the diversity of communities retrieved, but that it was directly influenced by the lysis step of an extraction protocol. Robust lytic processes (e.g. SDS-treatment) are able to rupture a broader range of bacterial cells, including hard cell-walled and endospore-forming bacteria such as Firmicutes and Actinobacteria (Moré et al., 1994; Yuan et al., 2012).

The effects of plant material size on DNA recovery and endophytic community diversity varied with the extraction protocols used. With the SDS-protocol, mDNA yield and bacterial community diversity improved with an increase in plant tissue material, whilst the CTAB protocol and commercial kits had a limiting or no significant effect on mDNA yield or community diversity. To our knowledge, there is no published study that has assessed the effects of plant sample size on mDNA yield and bacterial community structure. However, previous studies on soil samples have indicated that correlations between sample size and DNA recovery or bacterial diversity are dependent on the type of soil

Table 3

Relative abundance (%) and potential metabolic properties of dominant bacterial phylotypes found in sorghum tissues. Functional characterisation is inferred from published literature, and the symbol (√) denotes that the specific genus has representative species/strains with respective metabolic characteristics, including: phytohormone production (H), N_2 fixation (N), P solubilisation (P), siderophore production (S), xenobiotic degradation (X), heavy metal tolerance (M), antimicrobial activity (A), human pathogen (C) and phytopathogen (D). The number of OTUs per phylotype is shown in brackets.

Phylum/class	Genus	H	N	P	S	X	M	A	C	D	Root	Stem	References
Actinobacteria	Curtobacterium	√	√	√	√		√			√	1.6 (4)	2.0 (7)	Aizawa et al. (2007), Rajkumar et al. (2009), Sowmya et al. (2013), Chen and Yin (2007)
	Microbacterium	√	√	√	√	√	√				9.7 (8)	0.2 (3)	Tsavkelova et al. (2006), Rajkumar et al. (2009), Lal et al. (2008)
	Rhodococcus	√	√	√	√	√	√	√	√	√	0.0	1.5 (4)	Tsavkelova et al. (2006), Ye et al. (2004), Ambrosini et al. (2012), Rajkumar et al. (2009), Cornelis et al. (2001), Al-Awadhi et al. (2009), Chen et al. (2006), Li et al. (2012)
Bacteroidetes	Chryseobacterium	√	√	√	√	√	√		√		0.3 (2)	0.0 (1)	Marques et al. (2010), Radianingtyas et al. (2003), Ambrosini et al. (2012), Luo et al. (2011a), Lucas Garcia et al. (2004), Singh et al. (2013), Berg et al. (2005)
Firmicutes	Bacillus	√	√	√	√	√	√	√	√		1.2 (4)	0.0 (2)	Tsavkelova et al. (2006), Rodriguez and Fraga (1999), Ye et al. (2004), Rajkumar et al. (2009), Cho et al. (2007), Şahin et al. (2004), Berg et al. (2005)
	Paenibacillus	√	√	√	√	√	√	√			1.8 (2)	0.1 (2)	Rajkumar et al. (2009), Cho et al. (2007), Daane et al. (2001), Lal and Tabacchioni (2009)
	Leuconostoc						√	√		√	0.1 (1)	0.0	Harding and Shaw (1990), Gillespie et al. (2002)
	Lactococcus					√	√				0.0	6.7 (7)	Barré et al. (2007), Rodriguez et al. (2005)
Alphaproteobacteria	Unclassified										5.3 (15)	5.2 (15)	
	Agrobacterium	√	√	√	√	√	√		√	√	8.7 (4)	14.2 (4)	Tsavkelova et al. (2006), Rodriguez and Fraga (1999), Ambrosini et al. (2012), Rajkumar et al. (2009), Ye et al. (2004), Mansfield et al. (2012), Xing et al. (2006), Amaya and Edwards (2003)
	Rhizobium	√	√	√	√	√	√	√			1.0 (1)	0.0	Tsavkelova et al. (2006), Franche et al. (2009), Rodriguez and Fraga (1999), Ambrosini et al. (2012), Wani et al. (2008), Karpouzas and Singh (2006), Hafeez et al. (2005)
	Swaminathania		√	√							0.0	2.4 (9)	Loganathan and Nair (2004)
	Sphingobium		√	√	√	√	√				4.3 (1)	0.0	Raina et al. (2008), Ambrosini et al. (2012), Pavić et al. (2011), Wang et al. (2013), Pereira and Castro (2014)
	Sphingomonas		√		√	√	√	√	√	√	0.3 (2)	2.7 (7)	Ye et al. (2004), Rajkumar et al. (2009), Franche et al. (2009), Berg et al. (2005), Buonaurio et al. (2002), Romanenko et al. (2007)
Betaproteobacteria	Herbaspirillum	√	√			√		√	√	√	3.4 (4)	2.6 (7)	Tsavkelova et al. (2006), Ambrosini et al. (2012), McGuinness and Dowling (2009), Berg et al. (2005), Bull et al. (2010), Ting et al. (2011)
	Janthinobacterium						√		√		0.3 (2)	0.1 (1)	Huang et al. (2012), Bull et al. (2010)
Gammaproteobacteria	Methyloversatilis										0.1 (3)	0.4 (4)	
	Erwinia	√	√	√	√	√				√	0.1 (5)	30.2 (15)	Tsavkelova et al. (2006), Rodriguez and Fraga (1999), Mansfield et al. (2012), Hiraishi (2003), Boddey et al. (2003), Tian et al. (2009)
	Pseudomonas	√	√	√	√	√	√	√	√	√	46.5 (34)	10.1 (29)	Tsavkelova et al. (2006), Miller et al. (2010), Radianingtyas et al. (2003), Ambrosini et al. (2012), Cho et al. (2007), Mansfield et al. (2012), Berg et al. (2005)
	Stenotrophomonas	√	√	√	√	√	√	√	√	√	4.7 (10)	11.5 (23)	Tsavkelova et al. (2006), Radianingtyas et al. (2003), Ambrosini et al. (2012), Rajkumar et al. (2009), Xiao et al. (2009), Berg et al. (2005), Wolf et al. (2002), Suckstorff and Berg (2003)
TOTAL											89.4	89.9	

used (Ranjard et al., 2003; Kang and Mills, 2006; Zhao and Xu, 2012; Leite et al., 2014). Therefore, we recommend investigating the sample size factor prior to any extensive environmental microbial community analysis. Furthermore, Ranjard et al. (2003) pointed out that whilst larger samples improve the reproducibility of fingerprinting analyses, smaller samples (between 0.1 g and 1 g) such as those used in the current study are able to reveal rare phylotypes that are otherwise obscured by dominant phylotypes when larger samples are used.

The use of the different protocols led to significantly different sorghum-associated bacterial communities. This indicates that the DNA extraction procedure introduces significant bias, thereby influencing the interpretation of the microbial community observed. PCR-based metagenomic approaches such as DGGE (Hardoim et al., 2012; Ramond et al., 2013), t-RFLP (Sessitsch et al., 2012; Ding et al., 2013) and next generation sequencing (Gottel et al., 2011; Lucero et al., 2011; Ìnceoğlu et al., 2011; Molina et al., 2012) are continuously used to elucidate the structure and ecological roles of plant-associated (and other) microbial communities. However, few, if any, of these studies share the same DNA extraction protocol and/or are preceded by an evaluation of DNA extraction procedures to select the most efficient. Therefore, discrepancies in community profiles retrieved by different procedures pose a challenge in the reproducibility of previous findings as well as comparisons between studies.

In the current study, the SDS-protocol was established as the most efficient protocol in retrieving diverse endophytic bacterial communities from sorghum tissues. This protocol was further used to extract sorghum root and stem mDNA to perform high-throughput pyrosequencing using 16S rRNA marker analysis. Sorghum-associated bacterial communities were shown to be highly diverse, and root and stem communities were significantly different (Castro-Carrera et al., 2014; Lagos et al., 2014).

To date, the diversity of the sorghum endophytome remains largely unexplored. In one study, bacterial isolates from sorghum tissues were affiliated to the phyla Firmicutes, Actinobacteria, α-Proteobacteria and β-Proteobacteria, with no further characterisation (Grönemeyer et al., 2012). Ramond et al. (2013) showed that the diversity of sorghum associated endophytic bacteria is lower than that of rhizospheric communities using t-RFLP and DGGE; however, the low resolution of the two fingerprinting techniques made it difficult to assign taxonomy for the recovered OTUs. Despite this, sorghum has been used widely as a host organism in the study of PGPeBs such as Gluconacetobacter species and Bacillus sp. SLS18, Agrobacterium sp. (Luna et al., 2010; Luo et al., 2011b; Wu et al., 2014) as well as pathogenic bacterial strains such as Herbaspirillum rubrisulbalbicans (James et al., 1997).

The most dominant OTU's in this study belong to 20 different bacterial phylotypes, 19 of which were identified at genus level (Table 3). These genera have previously been detected/isolated from tissues of other graminaceous and non-graminaceous plants including maize, sugarcane, rice, poplar, grapevine and sunflower (Pereira et al., 2011; Magnani et al., 2010; Sun et al., 2008; Ulrich et al., 2008; Compant et al., 2011; Ambrosini et al., 2012), thus indicating that these genera are probably ubiquitous in endophytic environments. This study also shows that the structures of sorghum stem and root communities are significantly different, and this could be explained by the differences in the selectivity of recruitment processes of microbes into each niche environment (Hardoim et al., 2008). The root tissue is the primary entry point for bacteria as it is in contact with the rhizosphere (Lugtenberg and Kamilova, 2009). All bacteria in the vicinity of the root can be recruited into the root, and as such, the root endophytome is diverse (Hardoim et al., 2008). Only the most competitive endophytes can survive within the plant tissue environment, and a selection of these will be able to move within the plant to other plant tissues. As such, bacterial communities in aerial parts of the plant are often uniquely composed, and significantly less diverse (Hardoim et al., 2008). However, in this study, sorghum stem communities were slightly more diverse than sorghum root communities. This further supports the hypothesis that only a small proportion of sorghum-associated endophytic bacteria are specifically recruited to different niches, whereas the bulk of the endophytic community is composed of opportunistic endophytes that are subjected to minimal selective pressure (Ramond et al., 2013).

Dominant phylotypes in sorghum included ecologically and biotechnologically significant bacterial genera. Genera including Pseudomonas, Erwinia, Stenotrophomonas, Paenibacillus, Bacillus and Herbaspirillum have previously been shown to play important and varied roles in plant life (Table 3). For example, representatives of Bacilli, Pseudomonads and Rhizobia species have been shown to produce plant growth promoting hormones (Matiru and Dakota, 2004; Sahu et al., 2012; Reis et al., 2011) and also include numerous diazotrophic bacteria that are able to mineralise atmospheric nitrogen and/or solubilise phosphate compounds, to the benefit of the plant (Reis et al., 2011; Kraiser et al., 2011). Genera with potential plant-growth promoting properties can be targeted and isolated for development of biofertilizer or biocontrol products. However, these bacterial genera also include well-known plant pathogens, which could affect crop production and food security, such as Pseudomonas syringae, Erwinia carotovora and Erwinia amylovora (Bender et al., 1999; Basset et al., 2000; Oh and Beer, 2005; Mohr et al., 2008), as well as human pathogens including Pseudomonas aeruginosa, Rhodococcus equi and Bacillus cereus (Govan and Deretic, 1996; Weinstock and Brown, 2002; Bottone, 2010).

5. Conclusion

The current study showed that DNA extraction protocols introduce significant biases in the endophytic community diversities. Conventional protocols provide high-yield mDNA; however, the quality of extracted mDNA is low. The opposite is true for commercial kit protocols; they retrieve low yield mDNA of good quality for application in PCR. It can be concluded that the cell lysis procedures of classical protocols were superior to those of commercial kits; whilst the latter, had more efficient mDNA purification processes in comparison. By this argument, future research should, therefore, seek a feasible compromise between cell lysis and DNA purification, towards the development of a broad-spectrum DNA extraction procedure that retrieves PCR-ready metagenomic DNA, which is representative of endophytic communities. Pyrosequencing analysis revealed that sorghum tissues harbour diverse bacterial phylotypes with the potential of being applied for the improvement of sorghum health and production. However, 16S rRNA analysis using pyrosequencing cannot adequately identify the diversity at the species level, which is the requirement in order to develop targeted species with plant growth promoting attributes. Therefore, future studies need to involve the identification and isolation of sorghum tissue specific endophytic species. This study provides the initial identification of bacteria recruited by sorghum tissues, and suggests that a strategy to develop sorghum specific applications is foreseeable.
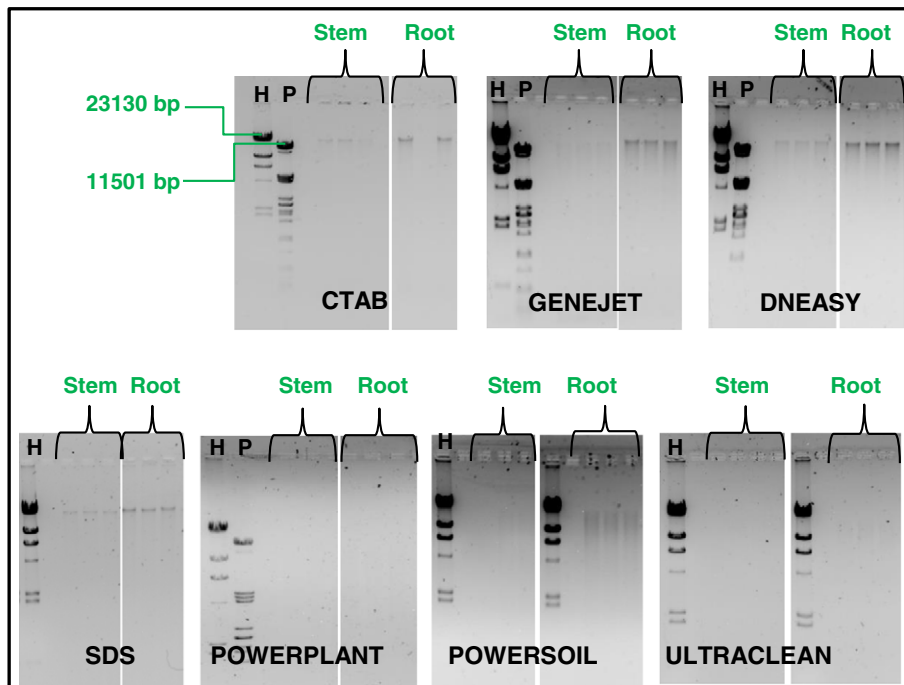
Conflict of interest

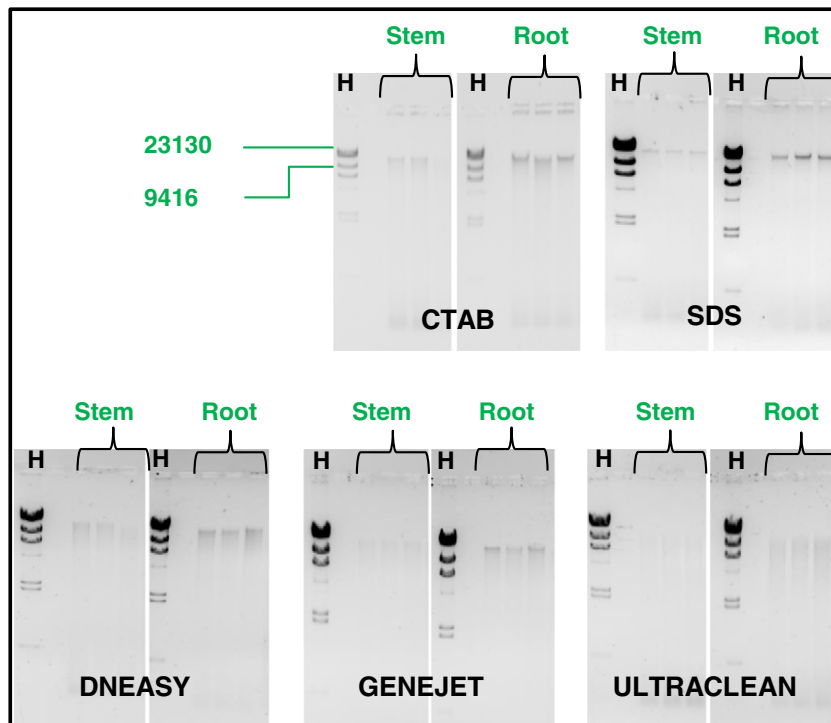The authors declare that there are no competing interests.

Appendix A1. Sorghum mDNA extracted from 0.1 g ground tissue. DNA was extracted from stem and root tissues using classical protocols (CTAB- or SDS-based) and commercial kits (Fermentas GeneJet Plant Genomic DNA Purification Kit, Qiagen DNeasy[R] Plant Mini Kit, MoBio PowerPlant Pro® DNA Isolation Kit, MoBio PowerSoil™ DNA Purification Kit and MoBio UltraClean® Soil DNA Isolation Kit), and 5 μL visualised by electrophoresis (80 V, 1 h) on 0.8% agarose gels. DNA molecular size was determined by comparison to molecular markers, lambda DNA digested with $Hind$III (H) and PstI (P) restriction enzymes. Extractions were performed in triplicate.

Appendix A2. Sorghum mDNA extracted from 0.3 g ground tissue. DNA was extracted from stem and root tissues using classical protocols (CTAB- or SDS-based) and commercial kits (Fermentas GeneJet Plant Genomic DNA Purification Kit, Qiagen DNeasy[R] Plant Mini Kit and MoBio PowerPlant Pro® DNA Isolation Kit), and 5 μL visualised by electrophoresis (80 V, 1 h) on 0.8% agarose gels. DNA molecular size was determined by comparison to molecular marker, lambda DNA digested with $Hind$III (H) restriction enzyme. Extractions were performed in triplicate.
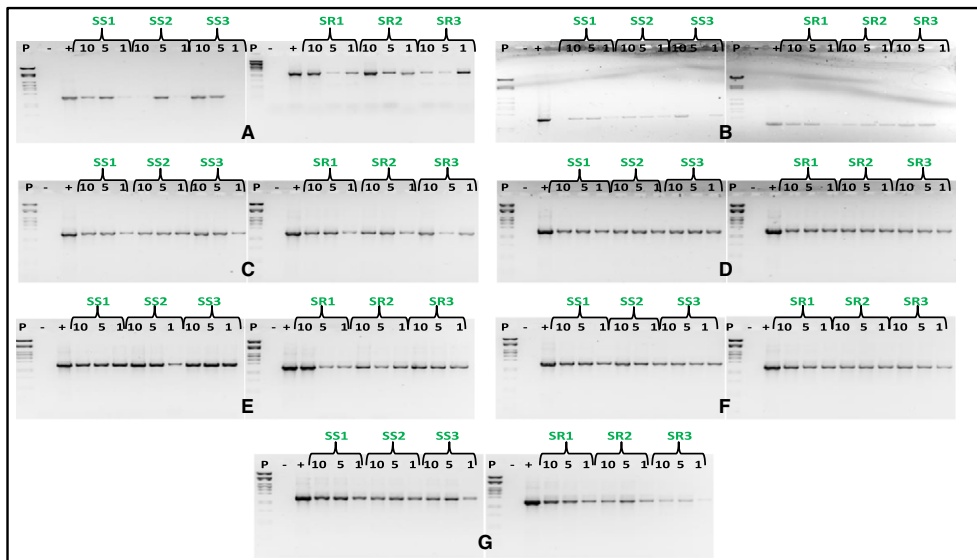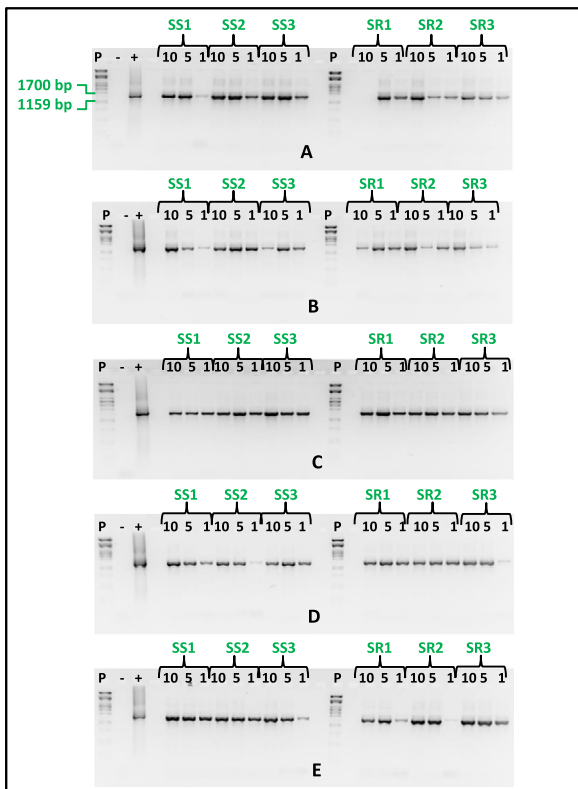
Appendix B1. Amplification of bacterial 16S rRNA gene from sorghum mDNA (from 0.1 g tissue). Bacterial 16S rRNA gene amplified from stem (SS1–SS3) and root (SR1–SR3) mDNA (triplicate samples) extracted using the CTAB-protocol (A), SDS-protocol (B) and commercial kits: Fermentas GeneJet Plant Genomic DNA Purification Kit (C), MoBio PowerPlant Pro® DNA Isolation Kit (D), Qiagen DNeasy^R Plant Mini Kit (E), MoBio PowerSoil™ DNA Purification Kit (F) and MoBio UltraClean® Soil DNA Isolation Kit (G). Template mDNA was tested at 10 ng, 5 ng and 1 ng per reaction. *Geobacillus* sp. genomic DNA (5 ng per reaction) and autoclaved double-distilled water were used as positive (+) and negative (−) controls, respectively. PCR products (5 μL) were visualised via electrophoresis (80 V, 30 min) on 1.5% agarose gels, and size determined by comparison to lambda *Pst*I marker.

Appendix B2. Amplification of bacterial 16S rRNA gene from sorghum mDNA (from 0.1 g tissue). Bacterial 16S rRNA amplified from stem (SS1–SS3) and root (SR1-SR3) mDNA (triplicate samples) extracted using the CTAB-protocol (A), SDS-protocol (B), Fermentas GeneJet Plant Genomic DNA Purification Kit (C), Qiagen DNeasy^R Plant Mini Kit (D) and MoBio UltraClean® Soil DNA Isolation Kit (E). PCR-products were visualised via electrophoresis (80 V, 30 min) on 1.5% agarose gels, and size determined by comparison to the lambda *Pst*I molecular marker.

Appendix C

Biom information of sequences retrieved from bacterial 16S rRNA pyrosequencing of sorghum mDNA processed with Qiime (http://qiime.org).

Statistic	Root	Stem
Number of samples	1	1
Number of sequences	5831.0	13086.0
Number of OTUs	331	498
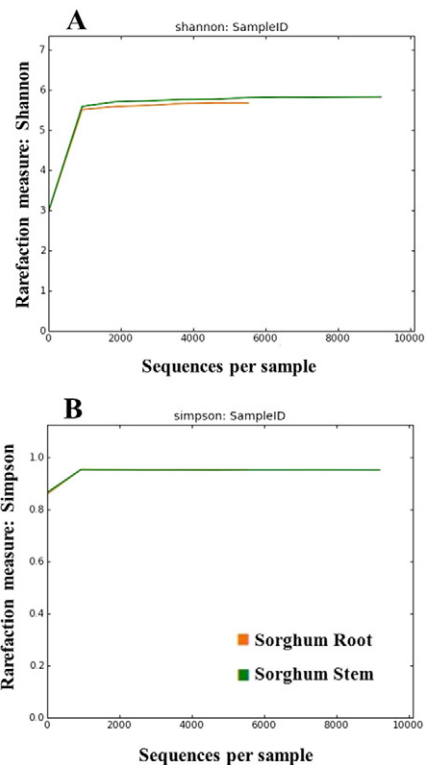
Appendix D. True diversity (A) and evenness (B) of endophytic bacterial communities recovered from sorghum root and stem tissues.

References

Aizawa, T., Ve, N.B., Kimoto, K., Iwabuchi, N., Sumida, H., Hasegawa, I., Sasaki, S., Tamura, T., Kudo, T., Suzuki, K., et al., 2007. *Curtobacterium ammoniigenes* sp. nov., an ammonia-producing bacterium isolated from plants inhabiting swamps in actual acid sulfate soil areas of Vietnam. Int. J. Syst. Evol. Microbiol. 57, 1447–1452.

Al-Awadhi, H., El-Nemr, I., Mahmoud, H., Sorkhoh, N.A., Radwan, S., 2009. Plant-associated bacteria as tools for the phytoremediation of oily nitrogen-poor soils. Int. J. Phytorem. 11, 11–27.

Amaya, R.A., Edwards, M.S., 2003. *Agrobacterium radiobacter* bacteremia in pediatric patients: a case report and review. Pedriatr. Infect. Dis. J. 22, 183–186.

Ambrosini, A., Beneduzi, A., Stefanski, T., Pinheiro, G., Vargas, L.K., Passaglia, L.M.P., 2012. Screening of plant growth promoting rhizobacteria isolated from sunflower (*Helianthus annuus* L.). Plant Soil 356, 245–264.

Andrews, M., Hodge, S., Raven, J.A., 2010. Positive plant microbial interactions. Ann. Appl. Biol. 157, 317–320.

Angiuoli, S.V., Matalka, M., Gussman, G., Galens, K., Vangala, M., Riley, D.R., Arze, C., White, J.R., White, O., Fricke, W.F., 2011. CloVR: a virtual machine for automated and portable sequence analysis from the desktop using cloud computing. BMC Bioinformatics 12, 356.

Babalola, O.O., Glick, B.R., 2012. Indigenous African agriculture and plant associated microbes: current practice and future transgenic prospects. Sci. Res. Essays 7, 2431–2439.

Barré, O., Mourlane, F., Solioz, M., 2007. Copper induction of lactate oxidase of *Lactococcus lactis* a novel metal stress response. J. Bacteriol. 189, 5947–5954.

Basset, A., Khush, R.S., Braun, A., Gardan, L, Boccard, F., Hoffmann, J.A., Lemaitre, B., 2000. The phytopathogenic bacteria *Erwinia carotovora* infects Drosophila and activates an immune response. Proc. Natl. Acad. Sci. U. S. A. 97, 3376–3381.

Bender, C.L., Alarcón-Chaidez, F., Gross, D., 1999. *Pseudomonas syringae* phytotoxins: mode of action, regulation, biosynthesis by peptide and polyketide synthetases. Microbiol. Mol. Biol. Rev. 63, 266–292.

Berg, G., Eberi, L., Hartmann, A., 2005. The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. Environ. Microbiol. 7, 1673–1685.

Blackwood, C., Hudleston, D., Zak, D., Buyer, J., 2007. Interpreting ecological diversity indices applied to terminal restriction fragment length polymorphism data: insights from simulated microbial communities. Appl. Environ. Microbiol. 73, 5276–5283.

Boddey, R.M., Urquiaga, S., Alves, B.J.R., Reis, V., 2003. Endophytic nitrogen fixation in sugarcane: present knowledge and future applications. Plant Soil 2003 (252), 139–149.

Bottone, E.J., 2010. *Bacillus cereus*, a volatile human pathogen. Clin. Microbiol. Rev. 23, 382–398.

Bray, J., Curtis, J., 1957. An ordination of the upland forest communities of Southern Wisconsin. Ecol. Monogr. 27, 325–349.

Bull, C.T., De Boer, S.H., Denny, T.P., Firrao, G., Fischer-Le Saux, M., Saddler, G.S., Scortichini, M., Stead, D.E., Takikawa, Y.L., 2010. Comprehensive list of names of plant pathogenic bacteria, 1980–2007. J. Plant Pathol. 92, 551–592.

Buonaurio, R., Stravato, V.M., Kosako, Y., Fujiwara, N., Naka, T., Kobayashi, K., Capelli, C., Yabuuchi, E., 2002. *Sphingomonas melonis* sp. nov., a novel pathogen that causes brown spots on yellow Spanish melon fruits. Int. J. Syst. Evol. Microbiol. 52, 2081–2087.

Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., Knight, R., 2010. PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics 26, 266–267.

Castro-Carrera, T., Toral, P.G., Frutos, P., McEwan, N.R., Hervás, G., Abecia, L., Pinloche, E., Girdwood, S.E., Belenguer, A., 2014. Rumen bacterial community evaluated by 454 pyrosequencing and terminal restriction fragment length polymorphism analyses in dairy sheep fed marine algae. J. Dairy Sci. 97, 1661–1669.

Charlson, E.S., Bittinger, K., Chen, J., Diamond, J.M., Li, H., Collman, R.G., Bushman, F.D., 2012. Assessing bacterial populations in the lung by replicate analysis of samples from the upper and lower respiratory tracts. PLoS ONE 7, e42786.

Chen, Y., Yin, Y., 2007. *Curtobacterium flaccumfaciens* pv. *beticola*, a new pathovar of pathogens in sugar beet. Plant Dis. 91, 677–684.

Chen, Y.P., Rekha, P.D., Arun, A.B., Shen, F.T., Lai, W.-A., Young, C.C., 2006. Phosphate solubilizing bacteria from subtropical soil and their tricalcium phosphate solubilizing abilities. Appl. Soil Ecol. 34, 33–41.

Cho, K.M., Hong, S.Y., Lee, S.M., Kim, Y.H., Kahng, G.G., Lim, Y.P., Kim, H., Yun, H.D., 2007. Endophytic bacterial communities in ginseng and their antifungal activity against pathogens. Microb. Ecol. 54, 341–351.

Clarke, K.R., Warwick, R.M., 2001. Change in Communities: An Approach to Statistical Analysis and Interpretation. 2nd edition. Primer-E Plymouth, Australia (Chapters 1-8).

Compant, S., Mitter, B., Colli-Mull, J.G., Gangl, H., Sessitsch, A., 2011. Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches colonization. Microb. Ecol. 62, 188–197.

Cornelis, K., Ritsema, T., Nijsse, J., Holsters, M., Goethals, K., Jaziri, M., 2001. The plant pathogen *Rhodococcus fascians* colonizes the exterior and interior of the aerial parts of plants. Mol. Plant Microbe Interact. 14, 599–608.

Culman, S.W., Bukowski, R., Gauch, H.G., Cadillo-Quiroz, H., Buckley, D.H., 2009. T-REX: software for the processing and analysis of T-RFLP data. BMC Bioinformatics 10, 177.

Daane, L.L., Harjono, I., Zylstra, G.J., Häggblom, M.M., 2001. Isolation and characterization of polycyclic aromatic hydrocarbon-degrading bacteria associated with the rhizosphere of salt marsh plants. Appl. Environ. Microbiol. 67, 2683–2691.

Demeke, T., Jenkins, G.R., 2010. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. Anal. Bioanal. Chem. 396, 1977–1990.

Ding, T., Palmer, M.W., Melcher, U., 2013. Community terminal restriction fragment length polymorphisms reveal insights into the diversity and dynamics of leaf endophytic bacteria. BMC Microbiol. 13, 1–11.

Dodd, I.C., Zinovkina, N.Y., Safronova, V.I., Belimov, A.A., 2010. Rhizobacterial mediation of plant hormone status. Ann. Appl. Biol. 157, 361–379.

Drabkova, L., Kirschner, J., Vlček, Č., 2002. Comparison of seven extraction and amplification protocols in historical herbarium specimens of Juncaceae. Plant Mol. Biol. Report. 20, 161–175.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27, 2194–2200.

Floyd, R., Abebe, E., Papert, A., Blaxter, M., 2002. Molecular barcodes for soil nematode identification. Mol. Ecol. 11, 839–850.

Franche, C., Lindström, K., Elmerich, C., 2009. Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. Plant Soil 329, 35–59.

Gillespie, R.S., Symons, J.M., McDonald, R.A., 2002. Peritonitis due to *Leuconostoc* species in a child receiving peritoneal dialysis. Pediatr. Nephrol. 17, 966–968.

Gotelli, N.J., Colwell, R.K., 2011. Estimating species richness. In: Magurran, A.E., McGill, B.J. (Eds.), Frontiers in Measuring Biodiversity. Oxford University Press, New York, pp. 39–54.

Gottel, N.R., Castro, H.F., Kerley, M., Yang, Z., Pelletier, D.A., Podar, M., Karpinets, T., Uberbacher, E., Tuskan, G.A., Vilgalys, R., et al., 2011. Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. Appl. Environ. Microbiol. 77, 5934–5944.

Govan, J.R., Deretic, V., 1996. Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. Microbiol. Mol. Biol. Rev. 60, 539–574.

Green, M.J., Thompson, D.A., MacKenzie, D.J., 1999. Easy and efficient DNA extraction protocol from woody plants for the detection of phytoplasmas by polymerase chain reaction. Plant Dis. 83, 484–485.

Grönemeyer, J.L., Burbano, C.S., Hurek, T., Reinhold-Hurek, B., 2012. Isolation and characterization of root-associated bacteria from agricultural crops in the Kavango region of Namibia. Plant Soil 356, 67–82.

Hafeez, F.Y., Naeem, F.I., Naeem, R., Zaidi, A.H., Malik, K.A., 2005. Symbiotic effectiveness and bacteriocin production by *Rhizobium leguminosarum* bv. *viciae* isolated from agriculture soils in Faisalabad. Environ. Exp. Bot. 54, 142–147.

Harding, C.D., Shaw, E.G., 1990. Antimicrobial activity of *Leuconostoc gelidum* against closely related species and *Listeria monocytogenes*. J. Appl. Bacteriol. 69, 648–654.

Hardoim, P.R., Van Overbeek, L.S., van Elsas, J.D., 2008. Properties of bacterial endophytes and their proposed role in plant growth. Trends Microbiol. 16, 463–471.

Hardoim, P.R., Hardoim, C.C.P., van Overbeek, L.S., van Elsas, J.D., 2012. Dynamics of seed-borne rice endophytes on early growth stages. PLoS ONE 7, e30438.

Heydari, A., Pessarakli, M., 2010. A review on biological control of fungal plant pathogens using microbial antagonists. J. Biol. Sci. 10, 273–290.

Hiraishi, A., 2003. Biodiversity of dioxin-degrading microorganisms and potential utilization in bioremediation. Microbes Environ. 18, 105–125.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6, 65–70.

Horne, E.C., Kumpatla, S.P., Patterson, K.A., Gupta, M., Thompson, S.A., 2004. Improved high-throughput sunflower and cotton genomic DNA extraction and PCR fidelity. Plant Mol. Biol. Report. 22, 83–84.

Huang, J.P., Mojib, N., Goli, R.K., Watkins, S., Waites, K.B., Rasik, R., Andersen, D.T., Bej, A.K., 2012. Antimicrobial activity of PVP from an Antarctic bacterium, *Janthinobacterium* sp. Ant5-2, on multi-drug and methicillin resistant *Staphylococcus aureus*. Nat. Prod. Bioprospect. 2, 104–110.

Ìnceoǧlu, Ö., Al-Soud, W.A., Salles, J.F., Alexander, V., Semenov, A.V., van Elsas, J.D., 2011. Comparative analysis of bacterial communities in a potato field as determined by pyrosequencing. Plos One 6, 1–11.

James, E.K., Olivares, F.L., Baldani, J.I., Dobereiner, J., 1997. *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue in leaves of *Sorghum bicolor* L. Moench. J. Exp. Bot. 48, 785–797.

Justel, A., Peña, D., Zamar, R., 1997. A multivariate Kolmogorov–Smirnov test of goodness of fit. Stat. Probab. Lett. 35, 251–259.

Kang, S., Mills, A.L., 2006. The effect of sample size in studies of soil microbial community structure. J. Microbiol. Methods 66, 242–250.

Kapanigowda, M.H., Perumal, R., Djanaguiraman, M., Aiken, R.M., Tesso, T., Prasad, P.V.V., Little, C., 2013. Genotyping variation in sorghum [*Sorghum bicolor* (L.) Moench] exotic germplast collections for drought and disease tolerance. SpringerPlus 2, 650.

Karpouzas, D., Singh, B.K., 2006. Microbial degradation of organophosphorus xenobiotics: metabolic pathways and molecular basis. Adv. Microb. Physiol. 51, 119–185.

Khan, A.A., Jilani, G., Akhtar, M.S., Naqvi, S.M.S., Rasheed, M., 2009. Phosphorus solubilizing bacteria: occurrence, mechanisms and their roles in crop production. J. Agric. Biol. Sci. 1, 48–58.

Kraiser, T., Gras, D.E., Gutiérrez, A.G., González, B., Gutiérrez, R.A., 2011. A holistic view of nitrogen acquisition in plants. J. Exp. Bot. 62, 1455–1466.

Krechel, A., Faupel, A., Hallmann, J., Ulrich, A., Berg, G., 2002. Potato-associated bacteria and their antagonistic potential towards plant-pathogenic fungi and the plant-parasitic nematode *Meloidogyne incognita*. Can. J. Microbiol. 48, 772–786.

Lagos, L.M., Navarrete, O.U., Maruyama, F., Crowley, D.E., Cid, F.P., Mora, M.L., Jorquera, M.A., 2014. Bacterial community structures in rhizosphere microsites of ryegrass (*Lolium perenne* var. Nui) as revealed by pyrosequencing. Biol. Fertil. Soils 50, 1253–1266.

Lal, S., Tabacchioni, S., 2009. Ecology and biotechnological potential of *Paenibacillus polymyxa*: a minireview. Indian J. Microbiol. 49, 2–9.

Lal, R., Dadhwal, M., Kumari, K., Sharma, P., Singh, A., Kumari, H., Jit, S., Gupta, S.K., Nigam, A., Lal, D., et al., 2008. *Pseudomonas* sp. to *Sphingobium indicum*: a journey of microbial degradation and bioremediation of hexachlorocyclohexane. Indian J. Microbiol. 48, 3–18.

Leite, D.C., Baliero, F.C., Pires, C.A., Madari, B.E., Rosado, A.S., Coutinho, H.L.C., Peixoto, R.S., 2014. Comparison of DNA extraction protocols from microbial communities from soil treated with biochar. Braz. J. Microbiol. 45, 175–183.

Li, J., Zhao, G., Huang, H., Qin, S., Zhu, W., Zhao, L., Xu, L., Zhang, S., Li, W., Strobel, G., 2012. Isolation and characterization of culturable endophytic actinobacteria associated with *Artemisia annua* L. A Van Leeuw J. Microbiol. 101, 515–527.

Loganathan, P., Nair, S., 2004. *Swaminathania salitolerans* gen. nov., sp. nov., a salt-tolerant, nitrogen-fixing and phosphate-solubilizing bacterium from wild rice (*Porteresia coarctata* Tateoka). Int. J. Syst. Evol. Microbiol. 54, 1185–1190.

Lozupone, C., Knight, R., 2005. Unifrac: a new phylogenetic method for comparing microbial communities. Appl. Environ. Microbiol. 71, 8228–8235.

Lozupone, C.A., Hamady, M., Knight, R., 2006. UniFrac — an online tool for comparing microbial community diversity in a phylogenetic context. BMC Bioinformatics 7, 371.

Lozupone, C.A., Hamady, M., Kelley, S.T., Knight, R., 2007. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial diversity. Appl. Environ. Microbiol. 73, 1576–1585.

Lucas Garcia, J.A., Probanza, A., Ramos, B., Barriuso, J., Gutierrez Mañero, F.J., 2004. Effects of inoculation with plant growth promoting rhizobacteria (PGPRs) and *Sinorhizobium fredii* on biological nitrogen fixation, nodulation and growth of *Glycine max* cv. Osumi. Plant Soil 267, 143–153.

Lucero, M.E., Unc, A., Cooke, P., Dowd, S., Sun, S., 2011. Endophytic microbiome diversity in micropropagated *Atriplex canescens* and *Atriplex torreyi* var *griffithsii*. Plos One 6, e16793.

Lugtenberg, B., Kamilova, F., 2009. Plant-growth-promoting rhizobacteria. Annu. Rev. Microbiol. 63, 541–556.

Luna, M.F., Galar, M.L., Aprea, J., Molinari, M.L., Boiardi, J.L., 2010. Colonization of sorghum and wheat by seed inoculation with *Gluconacetobacter diazotrophicus*. Biotechnol. Lett. 32, 1071–1076.

Luo, S., Chen, L., Chen, J., Xiao, X., Xu, T., Wan, Y., Rao, C., Liu, Y., Liu, C., Zeng, G., 2011a. Analysis and characterization of cultivable heavy metal-resistant bacterial endophytes isolated from Cd-hyperaccumulator *Solanum nigrum* L. and their potential use for phytoremediation. Chemosphere 85, 1130–1138.

Luo, S., Xu, T., Chen, L., Chen, J., Rao, C., Xiao, X., Wan, Y., Zeng, G., Long, F., Liu, C., Liu, Y., 2011b. Endophyte-assisted promotion of biomass production and metal-uptake of energy crop sweet sorghum by plant-growth-promoting endophyte *Bacillus* sp. SLS18. Appl. Microbiol. Biotechnol. 93, 1745–1753.

Magnani, G.S., Didonet, C.M., Cruz, L.M., Picheth, C.F., Pedrosa, F.O., Souza, E.M., 2010. Diversity of endophytic bacteria in Brazilian sugarcane. Genet. Mol. Res. 9, 250–258.

Magnani, G.S., Cruz, L.M., Weber, H., Bespalhok, J.C., Daros, E., Baura, V., Yates, M.G., Monteiro, R.A., Faoro, H., Pedrosa, F.O., Souza, E.M., 2013. Culture-independent analysis of endophytic bacterial communities associated with Brazilian sugarcane. Genet. Mol. Res. 12, 4549–4558.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A., et al., 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. Mol. Plant Pathol. 13, 614–629.

Marchesi, J.R., Sato, T., Weightman, A.J., Martin, T.A., Fry, J.C., Hiom, S.J., Wade, W.G., 1998. Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. Appl. Environ. Microbiol. 64, 795–799.

Marques, A.P.G.C., Pires, C., Moreira, H., Rangel, A.O.S.S., Castro, P.M.L., 2010. Assessment of the plant growth promotion abilities of six bacterial isolates using *Zea mays* as indicator plant. Soil Biol. Biochem. 42, 1229–1235.

Matiru, V.N., Dakota, F.D., 2004. Potential use of rhizobial bacteria as promoters of plant growth for increased yield in landraces of African cereal crops. Afr. J. Biotechnol. 3, 1–7.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., Hugenholtz, P., 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 6, 610–618.

McGuinness, M., Dowling, D., 2009. Plant-associated bacterial degradation of toxic organic compounds in soil. Int. J. Environ. Res. Public Health 6, 2226–2247.

Mendes, R., Pizzirani-Kleiner, A., Araujo, W., Raaijmarkers, J., 2007. Diversity of cultivated endophytic bacteria from sugarcane: genetic and biochemical characterization of *Burkholderia cepacia* complex isolates. Appl. Environ. Microbiol. 73, 7259–7267.

Miller, D.N., Bryant, J.E., Madsen, E.L., Ghiorse, W.C., 1999. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. Appl. Environ. Microbiol. 65, 4715–4724.

Miller, S.H., Browne, P., Prigent-Combaret, C., Combes-Meynet, E., Morrissey, J.P., O'Gara, F., 2010. Biochemical and genomic comparison of inorganic phosphate solubilisation in *Pseudomonas* species. Environ. Microbiol. Rep. 2, 403–411.

Mohr, T.J., Liu, H., Yan, S., Morris, C.E., Castillo, J.A., Jelenska, J., Vinatzer, B.A., 2008. Naturally occurring nonpathogenic isolates of the plant pathogen *Pseudomonas syringae* lack a type III secretion system and effector gene orthologues. J. Bacteriol. 190, 2858–2870.

Molina, L.G., da Fonseca, G.C., de Morais, G.L., de Oliveira, L.F.P., de Carvalho, J.B., Kulcheski, F.R., Margis, R., 2012. Metatranscriptomic analysis of small RNAs present in soybean deep sequencing libraries. Genet. Mol. Biol. 35, 292–303.

Moré, M.I., Herrick, J.B., Silva, M.C., Ghiorse, W.C., Madsen, E.L., 1994. Quantitative cell lysis of indigenous microorganisms and rapid extraction of microbial DNA from sediment. Appl. Environ. Microbiol. 60, 1572–1580.

Mornkham, T., Wangsomnuk, P.P., Wangsomnuk, P., Jogloy, S., Pattanothai, A., Fu, Y.B., 2012. Comparison of five DNA extraction methods for molecular analysis of Jerusalem artichoke (*Helianthus tuberosus*). Genet. Mol. Res. 11, 572–581.

Murray, M., Thompson, W., 1980. Rapid isolation of high molecular weight DNA plant DNA. Nucleic Acids Res. 8, 4321–4325.

Muyzer, G., De Waal, E.C., Uitierlinden, A.G., 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. 59, 695–700.

Newton, A.C., Fitt, B.D.L., Atkins, S.D., Walters, D.R., Daniell, T.J., 2010. Pathogenesis, parasitism and mutualism in the trophic space of microbe–plant interactions. Trends Microbiol. 18, 365–373.

Niu, C., Kebede, H., Auld, D.L., Woodward, J.E., Burow, G., Wright, R.J., 2008. A safe inexpensive method to isolate high quality plant and fungal DNA in an open laboratory environment. Afr. J. Biotechnol. 7, 2818–2822.

Oh, C., Beer, S.V., 2005. Molecular genetics of *Erwinia amylovora* in the development of fire blight. FEMS Microbiol. Lett. 253, 185–192.

Pavić, A., Stanković, S., Marjanović, Ž., 2011. Biochemical characterization of a sphingomonad isolate from the ascorarp of white truffle (*Tuber magnatum* Pico). Arch. Biol. Sci. 63, 697–704.

Pereira, S.I.A., Castro, P.M.L., 2014. Diversity and characterization of culturable bacterial endophytes from *Zea mays* and their potential as plant growth-promoting agents in metal-degraded soils. Environ. Sci. Pollut. Res. 21, 14110–14123.

Pereira, P., Ibáñez, F., Rosenblueth, M., Etcheverry, M., Martínez-Romero, E., 2011. Analysis of the bacterial diversity of the bacterial diversity associated with the roots of maize (*Zea mays* L.) through culture-dependent and culture-independent methods. ISRN Ecol. 2011 10 (Article ID 938546).

Price, M.N., Dehal, P.S., Arkin, A.P., 2009. FastTree: computing large minimum-evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. 26, 1641–1650.

Radianingtyas, H., Robinson, G.K., Bull, A.T., 2003. Characterization of a soil-derived bacterial consortium degrading 4-chloroaniline. Microbiology 149, 3279–3287.

Raina, V., Suar, M., Singh, A., Prakash, O., Dadhwal, M., Gupta, S.K., Dogra, C., Lawlor, K., Lal, S., van der Meer, J.R., 2008. Enhanced biodegradation of hexachlorohexane (HCH) in contaminated soils via inoculation with *Sphingobium indicum* B90A. Biodegradation 19, 27–40.

Rajkumar, M., Ae, N., Freitas, H., 2009. Endophytic bacteria and their potential to enhance heavy metal phytoextraction. Chemosphere 77, 156–160.

Ramond, J., Tshabuse, F., Bodpa, C.W., Cowan, D.A., Tuffin, M.I., 2013. Evidence of variability in the structure and recruitment of rhizospheric and endophytic bacterial communities associated with arable sweet sorghum (*Sorghum bicolor* (L) Moench). Plant Soil 372, 265–278.

Ranjard, L., Lejon, D.P.H., Mougel, C., Schehrer, L., Merdinoglu, D., Chaussod, R., 2003. Sampling strategy in molecular microbial ecology: influence of soil sample size on DNA fingerprinting analysis of fungal and bacterial communities. Environ. Microbiol. 5, 1111–1120.

Reis, V.M., Vanderleyden, J., Spaepen, S., 2011. N2-fixing endophytes of grasses and cereals. In: Polacco, J.C., Todd, C.D. (Eds.), Ecological Aspects of Nitrogen Metabolism in Plants. John Wiley and Sons Inc., pp. 231–253.

Rodriguez, H., Fraga, R., 1999. Phosphate solubilizing bacteria and their role in plant growth promotion. Biotechnol. Adv. 17, 319–339.

Rodriguez, E., Calzada, J., Arques, J.L., Rodriguez, J.M., Nunez, M., Medina, M., 2005. Antimicrobial activity of pediocin-producing *Lactococcus lactis* on *Listeria monocytogenes*, *Staphylococcus aureus* and *Escherichia coli* O157:H7 in cheese. Int. Dairy J. 15, 51–57.

Romanenko, L.A., Uchino, M., Frolova, G.M., Tanaka, N., Kalinovskaya, N.I., Latyshev, N., Mikhailov, V.V., 2007. *Sphingomonas molluscorum* sp. nov., a novel marine isolate with antimicrobial activity. Int. J. Syst. Evol. Microbiol. 57, 353–363.

Ryan, R.P., Germaine, K., Franks, A., Ryan, D.J., Dowling, D.N., 2007. Bacterial endophytes: recent developments and applications. FEMS Microbiol. Lett. 278, 1–9.

Saha, R., Saha, N., Donofrio, R.S., Bestervelt, L.L., 2012. Microbial siderophores: a mini review. J. Basic Microbiol. 53, 303–317.

Şahin, F., Cakmakci, R., Kantar, F., 2004. Sugar beet and barley yields in relation to inoculation with N₂-fixing and phosphate solubilizing bacteria. Plant Soil 265, 123–129.

Sahu, S.K., Thanaraj, M., Kathiresan, K., 2012. DNA extraction protocol for plants with high levels of secondary metabolites and polysaccharides without using liquid nitrogen and phenol. ISRN Mol. Biol. 12 (Article ID 205049).

Salonen, A., Nikkilä, J., Jalanka-Tuovinen, J., Immonen, O., Rajilic-Stojanovic, M., Kekkonen, R.A., Palva, A., de Vos, W., 2010. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J. Microbiol. Methods 81, 127–134.

Schenk, P.M., Carvalhais, L.C., Kazan, K., 2012. Unraveling plant–microbe interactions: can multi-species transcriptomics help? Trends Biotechnol. 30, 177–184.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., et al., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.

Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27, 863–864.

Scupham, A.J., Jones, J.A., Wesley, I.V., 2007. Comparison of DNA extraction methods for analysis of turkey cecal microbiota. J. Appl. Microbiol. 102, 401–409.

Seghers, D., Wittebolle, L., Top, E.M., Verstraete, W., Siciliano, S.D., 2004. Impact of agricultural practices on the *Zea mays* L. endophytic community. Appl. Environ. Microbiol. 70, 1475–1482.

Sène, M., Gallet, C., Doré, T., 2001. Phenolic compounds in a Sahelian sorghum (*Sorghum bicolor*) genotype (CE₁₄₅₋₆₆) and associated soils. J. Chem. Ecol. 27, 81–92.

Sessitsch, A., Hardoim, P., Döring, J., Weilharter, A., Krause, A., Woyke, T., Mitter, B., Hauberg-Lotte, L., Friedrich, F., Rahalkar, M., et al., 2012. Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. Mol. Plant Microbe Interact. 25, 28–36.

Shephard, R.N., 1962. The analysis of proximities: multidimensional scaling with an unknown distance function. Psychometrika 27, 125–140.

Singh, A.V., Chandra, R., Goel, R., 2013. Phosphate solubilization by *Chryseobacterium* sp. and their combined effect with N and P fertilizers on plant growth promotion. Arch. Agron. Soil Sci. 59, 641–651.

Sowmya, S., Rekha, P.D., Arun, A.B., 2013. Isolation, identification and characterization of Curtobacterium sp. YU-SS-C-67 for phosphate solubilization and uranium tolerance. Int. Res. J. Biol. Sci. 2, 102–106.

Suckstorff, I., Berg, G., 2003. Evidence for dose-dependent effects on plant growth by *Stenotrophomonas* strains from different origins. J. Appl. Microbiol. 95, 656–663.

Sun, L., Qiu, F., Zhang, X., Dai, X., Dong, X., Song, W., 2008. Endophytic bacterial diversity in rice (*Oryza sativa* L.) roots estimated by 16S rDNA sequence analysis. Microb. Ecol. 55, 415–424.

Taylor, J.R.N., 2004. Overview: importance of sorghum in Africa. Department of Food Science. University of Pretoria, Pretoria, South Africa, p. 21.

Terrat, S., Christen, R., Dequiedt, S., Leliévre, M., Nowak, V., Regnier, T., Bachar, D., Plassart, P., Wincker, P., Jolivet, C., et al., 2012. Molecular biomass and metataxogenomic assessment of soil microbial communities as influenced by soil DNA extraction procedure. Microb. Biotechnol. 5, 135–141.

Tian, F., Ding, Y., Zhu, H., Yao, L., Du, B., 2009. Genetic diversity of siderophore-producing bacteria of tobacco rhizosphere. Braz. J. Microbiol. 40, 276–284.

Ting, A.S.Y., Mah, S.W., Tee, C.S., 2011. Detection of potential volatile inhibitory compounds by endobacteria with biocontrol properties towards *Fusarium oxysporum* f. sp. *cubense* race 4. World J. Microbiol. Biotechnol. 27, 229–235.

Tsavkelova, E.A., Klimova, S.Y., Cherdynsteva, T.A., Netrusov, A.I., 2006. Microbial producers of plant growth stimulators and their practical use: a review. Appl. Biochem. Microbiol. 42, 117–126.

Turner, S., Pryer, K.M., Miao, V.P.W., Palmer, J.D., 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. J. Eukaryot. Microbiol. 46, 327–338.

Ulrich, K., Ulrich, A., Ewald, D., 2008. Diversity of endophytic bacterial communities in poplar grown under field conditions. FEMS Microbiol. Ecol. 63, 169–180.

Varma, A., Padh, H., Shrivastava, N., 2007. Plant genomic DNA isolation: an art or a science. Biotechnol. J. 2, 386–392.

Wang, C., Wang, F., Hong, Q., Zhang, Y., Kengara, F.O., 2013. Isolation and characterization of a toxic metal-tolerant phenanthrene-degrader *Sphingobium* sp. in a two-liquid-phase partitioning bioreactor (TPPB). Environ. Earth Sci. 70, 1765–1773.

Wani, P.A., Khan, M.S., Zaidi, A., 2008. Effect of metal-tolerant plant growth-promoting Rhizobium on the performance of pea grown in metal-amended soil. Arch. Environ. Contam. Toxicol. 55, 33–42.

Weinstock, D.M., Brown, A.E., 2002. *Rhodococcus equi*: an emerging pathogen. Clin. Infect. Dis. 34, 1379–1385.

Werner, J.J., Koren, O., Hugenholtz, P., DeSantis, T.Z., Walters, W.A., Caporaso, J.G., Angenent, L.T., Knight, R., Ley, R.E., 2012. Impact of training sets on classification of high-throughput bacterial 16S rRNA gene surveys. ISME J. 6, 94–103.

West, E.R., Cother, E.J., Steel, C.C., Ash, G.J., 2010. The characterization and diversity of bacterial endophytes of grapevine. Can. J. Microbiol. 56, 209–216.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biomet. Bull. 1, 80–83.

Wilson, I.G., 1997. Inhibition and facilitation of nucleic acid amplification. Appl. Environ. Microbiol. 63, 3741–3751.

Wolf, A., Fritze, A., Hagemann, M., Berg, G., 2002. *Stenotrophomonas rhizophila* sp., a novel plant-associated bacterium with antifungal properties. Int. J. Syst. Evol. Microbiol. 52, 1937–1944.

Wu, E., Lenderts, B., Glassman, K., Berezowska-Kaniewska, M., Christensen, H., Asmus, T., Zhen, S., Chu, U., Cho, M., Zhao, Z., 2014. Optimized *Agrobacterium*-mediated sorghum transformation protocol and molecular data of transgenic sorghum plants. In Vitro Cell. Dev. Biol. 50, 9–18.

Xiao, C., Chi, R., Huan, H., Zhang, W., 2009. Characterization of tricalcium phosphate solubilization by *Stenotrophomonas maltophilia* YC isolated from phosphate mines. J. Cent. South Univ. Technol. 16, 581–587.

Xing, Y., Yang, L., Huang, S., Li, Y., 2006. Identification of a new nitrogen-fixing endo-bacterium strain isolated from sugarcane stalk. Sugar Tech 8, 49–53.

Ye, J., Singh, A., Ward, O.P., 2004. Biodegradation of nitroaromatics and other nitrogen-containing xenobiotics. World J. Microbiol. Biotechnol. 20, 117–135.

Yuan, S., Cohen, D.B., Ravel, J., Abdo, Z., Forney, L.J., 2012. Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS ONE 7, e33865.

Zhao, F., Xu, K., 2012. Efficiency of DNA extraction methods on the evaluation of soil microeukaryotic diversity. Acta Ecol. Sin. 32, 209–214.

Zhou, J., Bruns, M.A., Tiedje, J.M., 1996. DNA recovery from soil of diverse composition. Appl. Environ. Microbiol. 62, 316–322.