Evidence of variability in the structure and recruitment of rhizospheric and endophytic bacterial communities associated with arable sweet sorghum (Sorghum bicolor (L) Moench)

Jean-Baptiste Ramond[a], Freedom Tshabuse[a], Cyprien William Bopda[a], Don A. Cowan[a,b], Marla I. Tuffin[a,*]

[a] Institute for Microbial Biotechnology and Metagenomics (IMBM), Department of Biotechnology, University of the Western Cape, Bellville 7535, Cape Town, South Africa.

[b] Center for Microbial Ecology and Genomics (CEMEG), Department of Genetics, University of Pretoria, Pretoria 0002, South Africa

*: **Corresponding author**: Prof. Marla I. Tuffin (marlatuffin@gmail.com)

Keywords :

Core community, Endosphere, Molecular fingerprinting (T-RFLP/DGGE), Plant Growth Promoting Bacteria (PGPB), Rhizospheric environment, Sorghum

Abstract (167 words)

Sorghum is the second most cultivated crop in Africa and is a staple food source of many African communities. Exploiting the associated plant growth promoting bacteria (PGPB) has potential as an agricultural biotechnology strategy to enhance sorghum growth, yield and nutritional properties. Here we use Terminal-Restriction Fragment Length Polymorphism (T-RFLP) and Denaturing Gradient Gel Electrophoresis (DGGE) to evaluate the factors that potentially shape rhizospheric and endophytic bacterial communities associated with sorghum farmed in South Africa. Microbial diversity was typically higher in the rhizosphere and rhizoplane compared to the endophytic zones (root, shoot and stem). Geographical location was one of the main drivers in describing microbial community assemblages found in rhizospheric and endophytic sorghum-linked niches. NO_3-N, total nitrogen and pH were clearly identified as the main abiotic factors shaping sorghum-associated soil communities. Our results also suggest that specific bacterial taxa with potential N-fixing capacities (*Acetobacter* sp., *Azospirillum* sp., *Pantoea* sp., *Bacillus* sp. and cyanobacteria) are consistently detected in sorghum-created rhizospheric and endophytic environments, irrespective of environmental factor effects.

Introduction

Sorghum (*Sorghum bicolor* L.) is the world's fifth most cultivated cereal crop after wheat, rice, maize and barley, with a global production of 60 million tons (Dicko et al. 2006), and is the second most cultivated cereal grain in Africa after maize (Taylor 2003). It is a critically important food crop, as it is estimated that in Africa and Asia over 300 million people rely on this crop as an essential source of energy (Dicko et al. 2006).

The use of (engineered) bacteria to enhance plant productivity has been widely considered (e.g. Berg, 2009; Compant et al. 2005; Hafeez et al., 2006; Schenk et al. 2012), particularly the use of plant-associated bacteria as such species have evolved a structured and intimate relationship with the plant host (Berg, 2009; Compant et al. 2005; Schenk et al. 2012). Such bacteria are typically present in the "plant-created" rhizospheric and endophytic environments.

The endosphere is the micro-environment localized inside plant organs and tissues (Saito et al. 2007). The rhizospheric environment is the soil surrounding and influenced by plant roots (Morgan et al. 2005), and is created by the release of nutrient- and carbon-rich root exudates, making it a "hot-spot" for microbial growth (Morgan et al. 2005; Hartmann et al. 2009). The rhizosphere is subdivided into the (ecto)rhizosphere, corresponding to the soil influenced by the roots and its related exudates, and the rhizoplane, the soil in intimate contact with the plant root surface.

Plant-associated bacteria have been characterized as having either neutral, detrimental or beneficial effects on plants (for reviews, see Berg 2009; Schenk et al. 2012). Deleterious plant-associated bacteria inhibit plant-growth, produce phytotoxins, and/or compete for nutrients (e.g. certain *Pseudomonas* species) (Sturz and Christie 2003). Beneficial plant-associated bacteria, typically known as plant-growth promoting bacteria (PGPB), promote

plant growth and development by acting as biocontrol agents and/or biofertilizers (Berg, 2009). The biocontrol activities inhibit growth and colonization of phytopathogenic organisms, generally through the production of siderophores or antibiotics (Berg, 2009). Biofertilizers promote plant growth essentially *via* the production of phytohormones such as indole-3-acetic acid, gibberellins and cytokinin-like substances (Lugtenberg et al. 1991; Panchal and Ingle 2011) and/or by improving the plants' nutritional status, for example by contributing to the nitrogen fixation process (Bai et al. 2002). Endophytic microorganisms can also establish a mutualistic relationship with their host (Rajkumar et al. 2009), and are either obligate (e.g. some species of *Herbaspirillum* and *Burkholderia*) or facultative (e.g. some strains of *Azospirillum* ; Baldani et al. 1997; Rajkumar et al. 2009). Plant growth-promoting bacteria are therefore attractive vehicles for enhancing plant productivity since they are considered as natural, cheap and eco-friendly fertilizers (Schenk et al. 2012).

Based on culture-dependant studies, various sorghum-associated bacteria have been isolated and identified (Budi et al. 1999; Pedersen et al. 1978; Zinniel et al. 2002), some of which have exhibited PGP activities such as nitrogen-fixation (members of the Enterobacteriaceae related to *Klebsiella pneumoniae*, *Enterobacter cloacae* or *Erwinia herbicola*; Pedersen et al. 1978) and the capacity for biocontrol (*Paenibacillus* sp. strain B2; Budi et al. 1999). Here we use modern molecular fingerprinting methods to expand the known range of sorghum-associated microbial taxa, to correlate the presence of these assemblages with abiotic factors and to determine the core microbial community recruited by sorghum plants. Members of such a community, particularly those consistently (or obligately) associated with a crop plant species, would be valid targets for future biofertilizer development. **Materials and Methods**

Study site and sampling procedures

Mature and healthy sorghum plants (at approx. 100 m spacing) were harvested in farms from three different South African provinces: Free State, North West, and Limpopo (Figure 1, Table 1). For each plant, samples of root, shoot and stem tissues were aseptically excised and stored in sterile plastic bags. Rhizospheric soils were collected by unearthing individual plants and dislodging soil particles associated with the root structures. Soil particles remaining attached to the sorghum root (rhizoplanic) were collected with the root tissue samples. Open soil (control) samples, from a similar depth but not impacted by plant root systems, were collected at each site. All samples were kept on ice and transported to the University of the Western Cape (UWC, South Africa) where they were stored at -80°C prior to processing.

Soil characterization

The pH, total carbon (C), total nitrogen (N), ammonium (NH_4-N) and nitrate (NO_3-N) contents of soil samples were analyzed by Bemlab (Pty) Ltd (Strand, Western Cape, South Africa) (Table 1).

Plant tissue sterilization

Plant tissues (roots, shoots, stems) were surface sterilized using a modification of the protocol described by Mendes *et al*. (2007). Each tissue sample was washed five times with sterile distilled water to remove attached soil particles, and placed in 400 mL of 1X PBS buffer and incubated with shaking at room temperature for 2 hours. Samples were then sequentially washed by shaking in (i) a 70 % ethanol solution for 10 min, (ii) a 2 % (v/v) sodium perchlorate solution for 10 min, (iii) a 70 % ethanol solution for 5 min and (iv) rinsed three times with autoclaved distilled water for 1 min. To evaluate the efficiency of the sterilization

procedure, 100 µL volumes of the final dH_2O rinse were plated on TSA and R2A agar (Merck, Germany), supplemented with the fungicide actidione (100 $mg.mL^{-1}$) and incubated at 28 °C for 4 days. Where no colony growth was observed, the sterilization procedure was considered to be sufficient. Where colonies were observed, the complete sterilization process was repeated. Once sterilized, the tissue samples were stored at 4 °C for subsequent molecular analysis.

Soil and plant organ metagenomic DNA extraction

Total metagenomic DNA was extracted from 0.5 g soil samples with the Powersoil DNA isolation kit according to the manufacturer's instructions (MoBio laboratories, USA). DNA extractions from sorghum tissue samples were performed using a modified version of the method described by Murray & Thompson (1980). Plant tissues were ground to powder form in liquid nitrogen using sterilized mortars and pestles. A pre-heated solution of 700 µL of 2 % CTAB and 1 µL of β-mercaptoethanol was added to each tissue powder sample, vortexed at maximum speed for 20 s and incubated at 65 °C for 60 min. A chloroform/isoamyl alcohol (24:1 v/v) solution (600 µL) was added, mixed by inversion for 5 min and centrifuged (12000 rpm, 5 min). The supernatant was collected and an equal amount of ice-cold isopropanol was added with RNase A (10 $mg.mL^{-1}$ final concentration). The mixtures were incubated at room temperature for 20 min and centrifuged at 12000 rpm for 5 min. The DNA pellets were washed twice with 250 µL of 70 % ethanol and centrifuged at 12000 rpm for 5 min prior to drying in a laminar flow cabinet. The DNA was resuspended in 50 µL of 1X TE buffer and stored at 4 °C. Metagenomic DNA concentrations were measured with a NanoDrop spectrophotometer (NanoDrop Technologies, Montchanin, DE, USA).

PCR amplification, purification and restriction digestion

All polymerase chain reactions (PCRs) were carried out in a Bio-Rad Thermocycler (T100™ Thermal Cycler). Bacterial 16S rRNA encoding genes were amplified using the universal primers E9F (5′-GAGTTTGATCCTGGCTCAG-3′) and U1510R (5′-GGTTACCTTGTTACGACTT-3′) (Marchesi et al. 1998; Reysenbach and Pace, 1995). PCR was carried out in 50 µl reaction volumes. Each reaction contained 1X PCR buffer, 0.2 U Dream*Taq*™ polymerase (Fermentas, USA), 200 µM of each dNTP, 0.5 µM of each primer, 0.1% BSA and between 10 to 20 ng of metagenomic DNA. PCR amplification was carried out as follows: 4 min at 94°C for denaturation; 30 cycles of 30 s at 94°C, 30 s annealing at 52°C and 105 s at 72°C; and a final elongation step of 10 min at 72°C.

To perform T-RFLP, the E9F primer was 5'-end FAM-labelled and the PCR products were purified using the GFX™ PCR DNA and gel band purification kit as directed by the supplier (GE Healthcare, UK). Purified PCR products (200 ng) were digested with the restriction enzyme *Hae*III at 37°C overnight.

To perform DGGE, a nested-PCR was performed with the same 50 µL reaction mixture described above, using the primer set 341f-GC (5'-CCTACGGGAGGCAGCAG-3') / 534r (5'-ATTACCGCGGCTGCTG-3') (Muyzer et al., 1993), 1 µL of the amplicon obtained with the primer set E9F/U1510R as template DNA and as follows: 94 °C for 4 min; 20 cycles – 94 °C for 45 s; 65 °C for 45 s; 72 °C for 60 s; additional 20 cycles – 94 °C for 30 s; 55 °C for 30 s; 72 °C for 60 s; and a final elongation step at 72 °C for 10 min. A 40mer GC clamp was added to the 5' ends of the forward primers 341f-GC: GC clamp – CGCCCGCCGCGCGCGGCGGGCGGGGCGGGGGCACGGGGGG.

T-RFLP analysis

Bacterial community structures were assessed by T-RFLP fingerprinting using the 16S rRNA gene as a marker. The precise lengths of the T-RFs were determined by capillary electrophoresis using the Applied Biosystems DNA Sequencer 3130 (Applied Biosystems, Foster City, California, USA) and according to the molecular weight standard Rox1.1 (with an acceptable error of ±1 bp). T-RFLP patterns and quality were analyzed using the freeware PeakScanner™ (version 1.0) (Applied Biosystems, https://products.appliedbiosystems.com). Peak height was used to characterize each unique T-RF, and valid T-RF peaks (between 35 and 1000 bp) from triplicate T-RFLP profiles were identified, compiled and aligned to produce large data matrices using the online software T-REX (http://trex.biohpc.org/ ; Culman et al. 2009). T-RFs with intensities lower than 0.5%, which may have originated from background interference, were excluded from the matrices. The term OTU (Operational Taxonomic Unit) is used to refer to individual terminal restriction fragments (T-RF) in T-RFLP patterns, with recognition that each OTU may comprise more than one distinct bacterial ribotype (Nocker et al., 2007). The web-based tool MiCA (Microbial Community Analysis; Shyu et al., 2007), with the "RDP (R10,U27) 700,829 Good Quality (>1200) Bacterial" database, was used for the *in silico* affiliation of T-RFs. A ± 3 bp size margin was implemented to take into account potential differences between real and predicted T-RFs (Sercu et al., 2011).

DGGE analysis

PCR amplicons obtained with the nested primer sets (341f-GC/534r) were analyzed by DGGE as described previously (Rodriguez-Caballero et al., 2012). Selected DGGE bands were excised using sterile surgical blades and eluted in 50 µL of filter-sterilized water at 4 °C overnight. One microlitre of the supernatant was then analyzed again by PCR and DGGE to eliminate any residual contamination by 'parasite' bands. The remaining PCR products (~25

µL) were purified using the GFX^{TM} PCR DNA and gel band purification kit as directed by the supplier (GE Healthcare, UK). The purified PCR products from DGGE bands were sequenced with a Hitachi 3730xl DNA Analyzer (Applied Biosystems).

Statistical analysis

The community structures obtained by T-RFLP were analyzed by ordination using non-metric multidimensional scaling (nMDS) of Bray-Curtis similarity matrices of square-root transformed data with the software Primer 6 (Primer-E Ltd, UK). An analysis of similarity (ANOSIM), performed on the resemblance matrix, was used to test for differences in bacterial community structure between predefined groups (Clarke, 1993). BEST (Biota Environment STepwise matching; Clarke and Gorley 2006) analysis was performed to determine correlations between the soil bacterial T-RFLP profiles and the abiotic variables presented in Table 1 (Carson et al., 2007). BEST determines the rank correlation between the underlying similarity matrices for microbial community data and environmental variables using the Spearman coefficient (ρ). As ρ increases, the correlation between the microbial community data and environmental variables increases from no correlation (0) to complete correlation (1).

Results and Discussion

Using molecular tools (T-RFLP and DGGE), the soil and endophytic ebacterial community structures associated with *Sorghum bicolor* L., cultivated in farms from three South African provinces (Limpopo, Free State and North West) (Figure 1, Table 1) were examined.

Factors shaping sorghum-associated microbial communities

The clustering of the samples in the 3D-MDS plot presented in Figure 2 clearly demonstrates that there are significant differences (ANOSIM, Global R = 0.437, P < 0.001) between the soil (rhizosphere, rhizoplane and open soil) and the endophytic (root, shoot and stem) Communities. Rhizosphere and rhizoplane communities presented a higher bacterial species richness (~ 60 OTUs) than the endophytic communities (ranging from 4 to 37 OTUs) (Figure 3), as previously observed for maize and *Populus deltoides* associated communities (Gottel et al., 2011; Roesch et al. 2008; Seghers et al. 2004). These results confirmed that rhizospheric environments are "microbial hot-spots" (Hartmann et al. 2009) while endophytic communities are relatively low in prokaryote diversity (Rosenblueth and Martinez-Romero, 2006; Sturz et al. 1997). However, it is suggested that endophytic diversities may often be underestimated in metagenomic studies as the co-extracted plant DNA represents the majority of the total extracted DNA.

A clear geographical-dependent difference in the soil communities was also observed (Figure 4A; ANOSIM, Global R = 0.5, P = 0.022), which was not soil environment-dependent (i.e. open soil, rhizosphere or rhizoplane; ANOSIM, Global R = -0.083, P = 0.649). In contrast, and as indicated by the high stress value (0.15; Figure 4B), the MDS representation explaining the (dis)similarities in the sorghum endophytic communities was ambiguous (Clarke, 1993). However, ANOSIM revealed that the province-specific endophytic

communities were also significantly different (Global R = 0.593, P = 0.01), despite the fact that all pairwise comparisons between sites were not significant (P > 0.05). Therefore, as previously shown in different environments (Martiny et al. 2006), correlations between sorghum-associated microbial community assemblages and geographical origin were observed for the rhizospheric and the endophytic samples, indicating that biogeography more than the sorghum plant itself (and thus its associated environments) has a role in determining its associated microbial communities. Finally, despite the fact that all pairwise comparisons between plant tissues were not significant (P > 0.05), the sorghum endophytic communities were found to be significantly tissue specific (Global R = 0.519, P = 0.004), as recorded for the specific N-fixing community of cultivated rice (Prakamhang et al., 2009).

Depending on the province, the open and rhizospheric soil characteristics varied, with notably lower pH values in Free State soils, elevated total C and NO_3-N concentrations in North West soils and low total C and total N in Limpopo soils (Table 1). Multiple rank correlations (BEST analysis) of the abiotic factors and soil community diversity (measured by T-RFLP) demonstrated that NO_3-N was the principal abiotic factor defining the different soil community structures ($\rho = 0.381$) and that the best combination of edaphic variables included pH, NO_3-N and % N ($\rho = 0.405$) (data not shown; Fierer and Jackson 2005). Since one of the major roles of PGPBs is the facilitation of plant N-uptake (Berg, 2009), it can be hypothesized that the soil N-status may influence the N-fixing capacities and/or the compositions of different plant-associated microbial communities. Indeed, differences in agricultural practices (as in this study, Table 1) have been shown to have an effect on the physical and the chemical compositions of the soil, as well as on soil and endophytic microbial community structures (Girvan et al. 2003; Seghers et al. 2004).

Identification of sorghum-associated core communities

An endophytic core microbial community was not detected using T-RFLP (Figure 3A), whereas a sorghum-associated rhizospheric core microbial community was observed, independent of any abiotic factor or bio-geographical considerations (Figure 3B). 39 OTUs were observed in the sorghum rhizosphere from the three provinces sampled, 12 of which (with respective sizes of 35, 71, 76, 77, 120, 123, 192, 211, 227, 280, 290 and 374 bp) were consistently detected in the rhizospheres of the plants sampled. Similarly, 24 OTUs were identified in the sorghum rhizoplane of samples from the three provinces, with 4 (with sizes of 71, 192, 195 and 211 bp) observed in all the rhizoplane samples. Three of these (71, 192 and 211 bp), which were repeatedly detected in the sorghum-rhizospheric and rhizoplanic niches, were subjected to *in silico* identification. Predictive phylogenetic affiliations matched the 71 bp OTU mainly to *Acetobacter* and *Azospirillum* species, which have previously been shown to be N-fixing PGPBs (Kevin Vessey, 2003; Franche et al., 2009). The latter is known to be associated with sorghum (Franche et al., 2009). The 192 and 211 bp OTUs could neither be identified nor related to a single genus and were principally matched to uncultured bacteria.

To identify bacterial taxa which were ubiquitously associated with sorghum, DGGE coupled with post-electrophoretic phylogenetic analysis was also used (Figure 5, Table 2; Niepceron et al., 2010). Since endophytic microbial communities are typically characterized by a low taxonomic diversity (Gottel et al., 2011; Figure 3A), and since DGGE is less sensitive than T-RFLP (Nocker et al., 2007), the PCR products from similar plant-tissue samples were pooled prior to load on DGGE gels (Figure 5). The composite endophytic microbial community fingerprints from each province for each plant-tissue type is presented in Figures 5A, 5B and 5C, while a rhizospheric and rhizoplanic fingerprint for each plant sampled is shown in Figures 5D and 5E.

Except in the stem samples, co-migrated DGGE-bands (indicated by arrows in Figure 5) were observed in all the sorghum-associated micro-environments studied. All major co-migrating bands were re-sequenced (Table 2). The co-migrated bands A1, A2 and A3 detected in sorghum root metagenomic DNA showed highsequence identities (99 % to 100 %) with numerous cyanobacterial species (Figure 5A, Table 2). A range of N-fixingendophytic cyanobacteria (e.g. *Anabaena* spp.) have previously been characterized (Franche et al 2009, Terakado-Tonooka et al., 2008). The sorghum shoot sample B1, B2 and B3 bands also showed 99 % to 100 % sequence identity to N-fixing *Pantoea* species (Figure 5B, Table 2), a bacterial genus already described as a sugarcane endophytic PGPB (Loiret et al., 2004). The co-migrating bands observed in the rhizospheric metagenomic DNA samples showed high sequence homology to different taxonomic groups (bands C1 to C6, Figure 5D, Table 2), a disadvantage associated with DGGE where different phylotypes may possess similar electrophoretic mobilities (Nocker et al., 2007). The rhizoplanic bands D1 to D6 possessed high sequence identities (98 % to 100%) with the most abundant rhizospheric genus, *Bacillus* sp. (particularly *B. megaterium*, Table 2), which is known to possess various PGP activities (N-fixation, metabolite and phytohormone production, improvement of root performances; Bai et al., 2002; Hafeez et al., 2006; Lugtenberg et al., 1991; Saharan & Nehra, 2011).

In this study, two distinct sorghum core-communities were detected using two different methods (T-RFLP and DGGE). As previously demonstrated (van Felten et al., 2010), these molecular tools are complementary for targeting environmental core-communities or core-community members. Also, the consistently detected sorghum-associated rhizospheric and endophytic (cyano)bacterial taxa are known to have PGP capacities, particularly their capacity for N-fixation. Sorghum has important N-uptake capacities, particularly in N-starved conditions, and the mechanisms involved in conferring this physiological trait is unknown (Hirel et al., 2007). The role of the PGPB symbiosis in providing bioavailable N to plants is

well documented (Kevin Vessey, 2003; Berg, 2009; Franche et al., 2009). The identification of a core sorghum-associated community displaying potential N-fixation capabilities could suggest a mechanism which contributes to sorghum's elevated N-uptake capacity, i.e. through the consistent recruitment of specific endophytic and rhizoplanic PGPBs with N-fixing capacities.

Conclusion

In this study, differences in microbial community structures in micro-environments (rhizoplane / rhizosphere / root / shoot / stem) associated with sorghum farmed in different South African provinces (Limpopo, North West and Free State) were observed, with biogeography, soil characteristics (pH, NO_3-N and total N) and plant tissues being determining factors in shaping sorghum-associated microbial communities. A sorghum core community, composed of potentially N-fixing bacterial taxa (*Acetobacter* sp., *Azospirillum* sp., *Pantoea* sp., *Bacillus* sp., and cyanobacteria), was also detected. The presence of (cyano)bacterial taxa reliably associated with sorghum in South Africa is significant as they could be directly used as bio-inoculants, and possibly engineered with enhanced PGP activities to introduce various crop improvements. For example, antibiotic production capacities targeting sorghum-pathogens such as *Fusarium* sp. could be valuable objectives. With a worldwide annual growth rate of 10% in the "microbial inoculant" market (Berg, 2009), the sorghum-specific microorganisms identified in this study have significant potential in agricultural biotechnology as crop improvement tools.

Acknowledgments

We thank the South African National Research Foundation (NRF) for funding the study and Dale Wilcox for generating the map of South Africa presented in Figure 1. We would also like to acknowledge Dr Shagi from the North West academic farm of the ARC (Agricultural Research Council), Mr Osterhaizen (Free State farmer) and the community farm at Jane Furst (Limpopo) for allowing us to sample sorghum.

References

Bai Y, Aoust F D, Smith D L and Driscoll B T 2002 Isolation of plant growth promoting *Bacillus* strains from soybean root nodules. Can. J. Microbiol. 48, 230-238.

Baldani J, Caruso L, Baldani V, Goi S and Dobereiner J 1997 Recent advances in BNF with non-legume plants. Soil Biol. Biochem. 29, 911-922.

Berg G 2009 Plant–microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. Appl. Microbiol. Biotechnol. 84, 11–18

Budi S W, Van Tuinen D, Martinotti G and Gianinazzi S 1999 Isolation from the *Sorghum bicolor* mycorrhizosphere of a bacterium compatible with arbuscular mycorrhiza development and antagonistic towards soilborne fungal pathogens. Appl. Environ. Microbiol. 65, 5148-5150.

Carson JK, Rooney D, Gleeson DB & Clipson N (2007) Altering the mineral composition of soil causes a shift in microbial community structure. FEMS Microbiol Ecol. 61, 414–423.

Clarke K 1993 Non-parametric multivariate analysis of changes in community structure. Aust. J. Ecol. 18, 117-143.

Clarke K R and Gorley R N 2006 PRIMER v6: User Manual. Plymouth Marine Laboratory: Plymouth, UK.

Compant S, Duffy B, Nowak J, Clemet C and Barka E A 2005 Use of plant-growth promoting bacteria for biocontrol of plant diseases: Principles, mechanisms of action, and future prospects. Appl. Environ. Microbiol. 71, 4951-4959.

Culman S W, Bukowski R, Gauch H G, Cadillo-Quiroz H and Buckley D H 2009 T-REX: Software for the Processing and Analysis of T-RFLP data. BMC Bioinformatics. 10, 171.

Dicko M, Gruppen H, Traore A, Voragen A and van Berkel W 2006 Review: Sorghum grain as human food in Africa: relevance of content of starch and amylase activities. Afr. J. Biotechnol. 5, 384-395.

Fierer N and Jackson R B 2005 The diversity and biogeography of soil bacterial communities. Proc. Natl. Acad. Sci. USA . 103, 626-631.

Franche C., Lindström K., Elmerich C. 2009. Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. Plant Soil. 321, 35-59.

Girvan M S, Bullimore J, Pretty J N, Osborn M A and Ball A S 2003 Soil type is the primary determinant of the composition of the total and active bacterial communities in arable soils. Appl. Environ. Microbiol. 69, 1800-1809.

Gottel NR, Castro HF, Kerley M, 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.

Hafeez FY, Yasmin S, Ariani D, Mehboob-ur-Rahman Zafar Y & Malik KA, 2006. Plant growth-promoting bacteria as biofertilizer. Agronom. Sustain. Dev. 26**,**143-150

Hartmann A, Schimd M, van Tuinen D and Berg G 2009 Plant-driven selection of microbes. Plant Soil. 321, 235-237.

Hirel B., Le Gouis J., Ney B., Gallais A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. J. Exper. Bot. 58, 2369-2387.

Kevin Vessey J. 2003. Plant growth promoting rhizobacteria as biofertilizers. Plant and Soil. 255, 571-586.

Loiret F, Ortega E, Kleiner D, Ortega-Rodes P, Rodes R & Dong Z (2004) A putative new endophytic nitrogen-fixing bacterium *Pantoea* sp. from sugarcane. J. Appl. Microbiol. **97,** 504-511.

Lugtenberg B, De Weger L and Bennett J 1991 Microbial stimulation of plant growth and protection from disease. Curr. Opin. Biotech. 2, 457-464.

Marchesi J.R., Sato T., Weightman A.J., Martin T.A., Fry J.C., Hiom S.J. and Wade W.G.1998. Design and evaluation of useful bacterial primers that amplify genes coding for bacterial 16S rRNA. Applied and Environmental Microbiology. 64, 795-799.

Martiny J B H, Bohannan B J M, Brown J H, Colwell R K, Fuhrman J A, Green J L, Horner-Devine C M, Kane M, Krumins J A, Kuske C R, Morin P J, Naeem S, Øvreås L, Reysenbach A-L, Smith V H and Staley J T 2006 Microbial biogeography: putting microorganisms on the map. Nature Rev. 4, 102-112.

Mendes R, Pizzirani-Kleiner A, Araujo W and 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.

Morgan J A W and White P J 2005 Biological costs and benefits to plant-microbe interactions in the rhizosphere. J. Exp. Bot. 56, 1729-1739.

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

Muyzer, G., de Waal, E.C. & Uitterlinden, 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.

Niepceron, M., Portet-Koltalo, F., Merlin, C., Motelay-Massei, A., Barray, S., Bodilis, J., 2010. Both *Cycloclasticus* spp. and *Pseudomonas* spp. as PAH-degrading bacteria in the Seine estuary (France). FEMS Microb. Ecol. 71, 137-147.

Nocker A. Burr M., Camper A.K. (2007) Genotypic microbial community profiling: a critical technical review. Microb. Ecol.. 54, 276-289.

Panchal H and Ingle S 2011 Isolation and characterization of endophytes from the root of the medicinal plant *Chlorophytum borivilianum* (Safed musli). J. Adv. Dev. Res. 2, 205-209.

Pedersen WL, Chakrabarty K, Klucas RV & Vidave AK (1978) Nitrogen fixation (acetylene reduction) associated with roots of winter wheat and sorghum in Nebraska. Appl. Environ. Microbiol. 35, 129-135.

Prakamhang J., Minamisawa K., Teamtaisong K., Boonkerd N., Teaumroong N. 2009. The communities of endophytic diazotrophic bacteria in cultivated rice (*Oryza sativa* L.). Applied Soil Ecology. 42, 141-149.

Rajkumar M, Ae N and Freitas H 2009 A review. Endophytic bacteria and their potential to enhance heavy metal phytoextraction. Chemosphere. 77, 153-160.

Reysenbach, A.-L. & Pace, N.R. (1995). In Archaea: A Laboratory Manual – Thermophiles. Robb, F.T. & Place, A.R (Eds) pp 101-107. Cold Spring Harbour Laboratory Press: New York, USA.

Roesch L F W, Camargo F A O, Bento F M and Triplett E W 2008 Biodiversity of diazotrophic bacteria within the soil, root and stem of field-grown maize. Plant Soil. 302, 91–104.

Rodriguez-Caballero A., Ramond J.-B., Welz P.J., Cowan D.A., Odlare M., Burton S.G. 2012. Treatment of high ethanol concentration wastewater by biological sand filters: Enhanced COD removal and bacterial community dynamics. J. of Environ. Manag.. 109, 54-60.

Rosenblueth M and Martinez-Romero E 2006 A review. Bacterial endophytes and their interactions with hosts. Mol. Plant Microbe. 19, 827-837.

Saharan BS & Nehra V 2011. Plant growth promoting rhizobacteria: A critical review. Life Sci. Med. Res. 21**,** 1-30.

Saito A, Ikeda S, Ezura H and Minamisawa K 2007 A minireview. Microbial community analysis of the phytosphere using culture-independent methodologies. Microbes Environ. 22, 93-105.

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

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

Sercu B, van de Werfhorst L.C., Murray J.L.S., Holden P.A. (2011) Cultivation-independent analysis of bacteria in IDEXX quanti-tray/2000 fecal indicator assays. Appl. Environ. Microbiol. 77: 627-633.

Shyu, C., Soule, T., Bent, S.J., Foster, J.A., Forney, L.J. (2007) MiCA: A Web-Based Tool for the Analysis of Microbial Communities Based on Terminal-Restriction Fragment Length Polymorphisms of 16S and 18S rRNA Genes. J. Microb. Ecol. **53**:562-570.

Sturz A V, Christie B R, Matheson B G and Nowak J 1997 Biodiversity of endophytic bacteria which colonize red clover nodules, roots, stems and foliage and their influence on host growth. Biol. Fertil. Soil. 25, 13-19.

Sturz A and Christie B 2003 Beneficial microbial allelopathies in the root zone: the management of soil quality and plant disease with rhizobacteria. Soil Tillage Res. 72, 107-123.

Taylor J R N 2003 Overview: Importance of sorghum in Africa. http://www.afripro.org.uk/papers/Paper01Taylor.pdf.

Terakado-Tonooka J, Ohwaki Y, Yamakawa H, Tanaka F, Yoneyama T & Fujihara S (2008) Expressed *nifH* genes of endophytic bacteria detected in field-grown sweet patatoes (*Ipoema batatas* L.). *Microbes Environ*. **23:** 89-93.

van Felten A., Meyer J.B., Défago G., Maurhofer M. (2011) Novel T-RFLP method to investigate six main groups of 2,4-diacetylphloroglucinol-producing pseudomonads in environmental samples. Journal of Microbiological Methods. 84: 379-387.

Zinniel D K, Lambrecht P, Beth Harris N, Feng Z, Kuczmarski D, Higley P, Ishimaru C A, Arunakumari A, Barletta R G and Vivaver A K 2002 Isolation and characterization of endophytic colonizing bacteria from agronomic crops and prairie plants. Appl. Environ. Microbiol. 68, 2198-2208

Legends to Figures.

Figure 1. Map of South Africa displaying the different sampling sites. A: Limpopo. Plant 1: S24°38.620'/E029°52.484'; Plant 2: S24°39.375'/E029°53.593' **B: North West province.** Plant 1: S26°43.741'/E027°04.870'; Plant 2: S26°44.063'/E027°04.721' **C: Free State.** Plant 1: S27°02.975'/ E027°31.405'; Plant 2: S27°03.665'/E027°31.780'.

Figure 2. 3D-Nonmetric Multi-Dimensional Scaling plot of Bray Curtis similarity plot of soil (▲) and endophytic (△) bacterial community profiles associated with South African sorghum (stress = 0.1). The soil bacterial communities originate from open, rhizospheric, and rhizoplane soils and the endophytic communities from shoot, roots and stems of healthy sorghum plants harvested in Limpopo, North West and Free State provinces.

Figure 3. Venn Diagrams showing the distribution of T-RFs present in the different sorghum-associated environments. A: Endophytic environment. The numbers in italic, bold, or underlined indicate the number of T-RFs observed in sorghum root, shoot or stem respectively. The number in bracket indicate the number of OTUs present in the tissues of all the sorghum plants respectively compared. **B: Rhizospheric environment.** The numbers in bold or underlined indicate the numbers of T-RFs observed in rhizospheric or rhizoplanic soils respectively. The number in bracket indicate the number of OTUs present in the specific rhizospheric environments of all the sorghum plants respectively compared.

Figure 4. 2D-Nonmetric Multi-Dimensional Scaling plot of Bray Curtis similarity of microbial communities structures determined by T-RFLP analysis of 16S rRNA genes associated with sorghum in three South African provinces. Triangles represent bacterial communities from Free State, squares from North West and circles from Limpopo. Numbers

(1, 2) indicate sorghum plant. **A: Soil microbial communities (stress = 0.1).** Black: Open Soil / Grey: Rhizosphere / White: Rhizoplane. **B: Endophytic microbial communities (stress = 0.15).** Black: Root / Grey: Shoot / White: Stem.

Figure 5. DGGE profiles of the sorghum-associated microbial communities IN South Africa. A: Root endophytic communities. B: Shoot endophytic communities. C: Stem endophytic communities. D: Rhizosphere communities. E: Rhizoplane communities. Arrows and their associate reference indicate the co-migrated DGGE-bands sequenced and presented in Table 2. In the rhizospheric environments, numbers (1, 2) indicate the sorghum plant. NW: North West province / L: Limpopo / FS: Free State.

Figure 1.

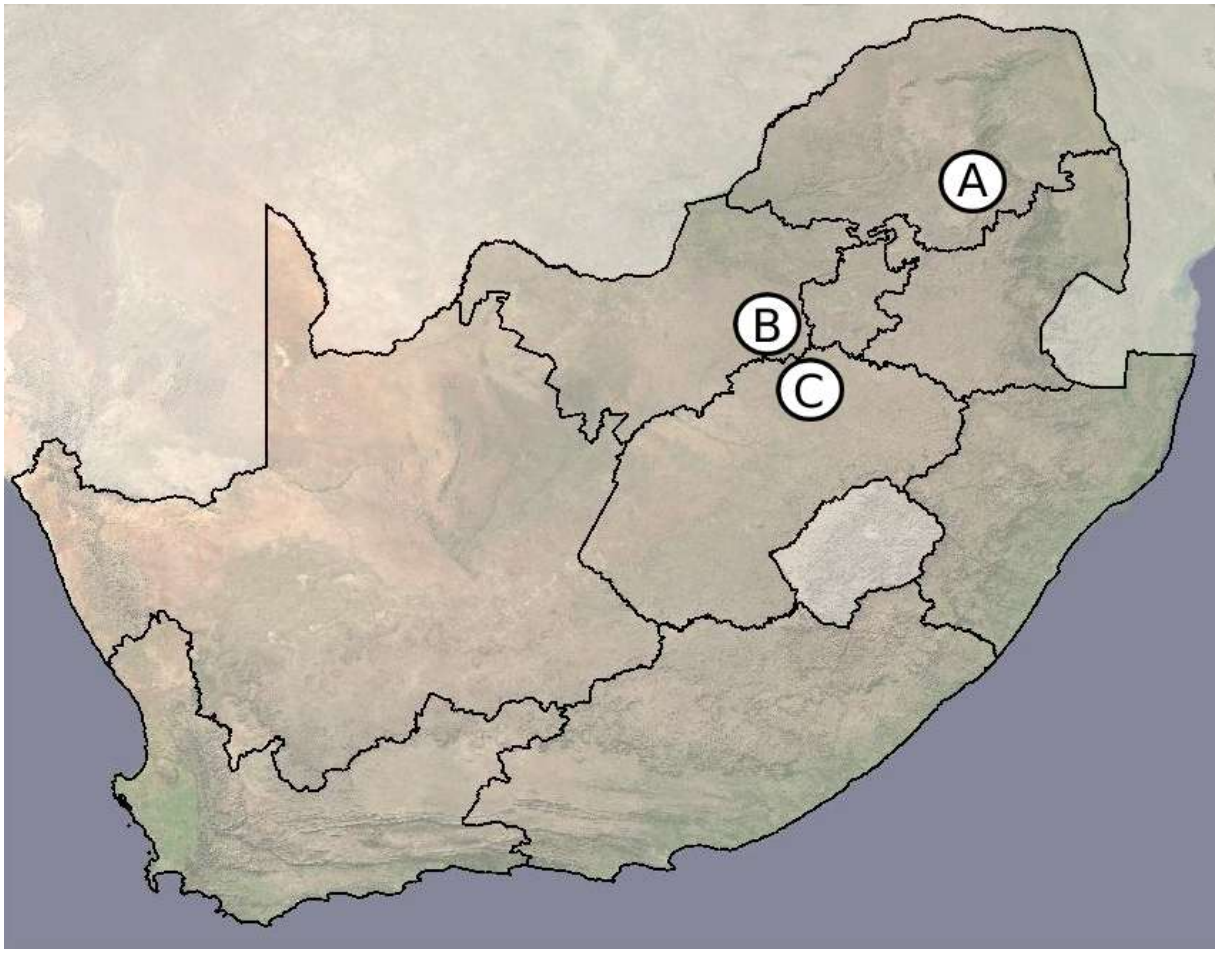


Figure 2.

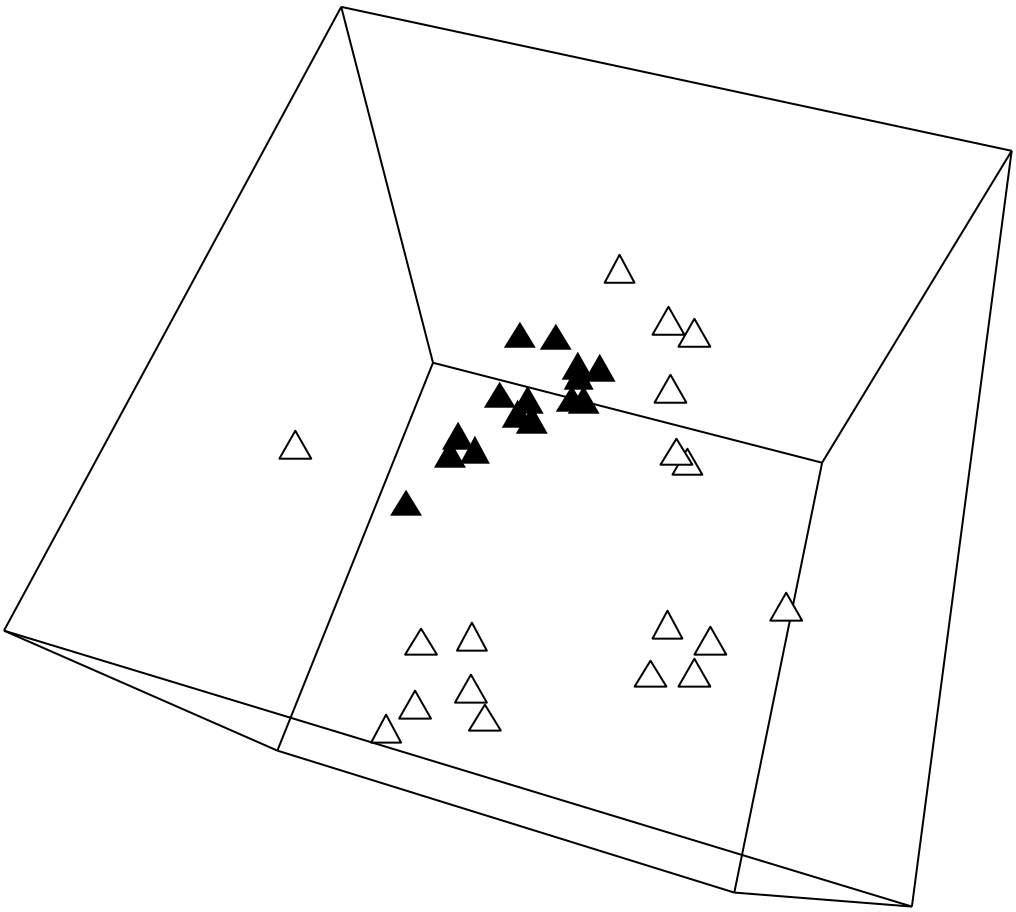
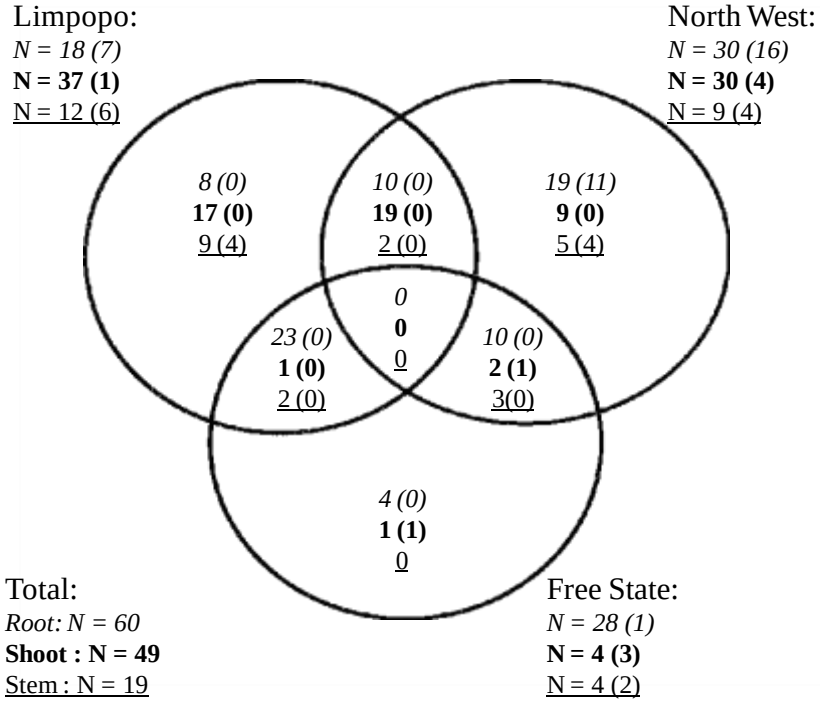


Figure 3.

A:


Limpopo:
N = 18 (7)
N = 37 (1)
N = 12 (6)
North West:
N = 30 (16)
N = 30 (4)
N = 9 (4)
8 (0)
17 (0)
9 (4)
10 (0)
19 (0)
2 (0)
19 (11)
9 (0)
5 (4)
0
0
0
23 (0)
1 (0)
2 (0)
10 (0)
2 (1)
3(0)
4 (0)
1 (1)
0
Total:
Root: N = 60
Shoot : N = 49
Stem : N = 19
Free State:
N = 28 (1)
N = 4 (3)
N = 4 (2)

B:

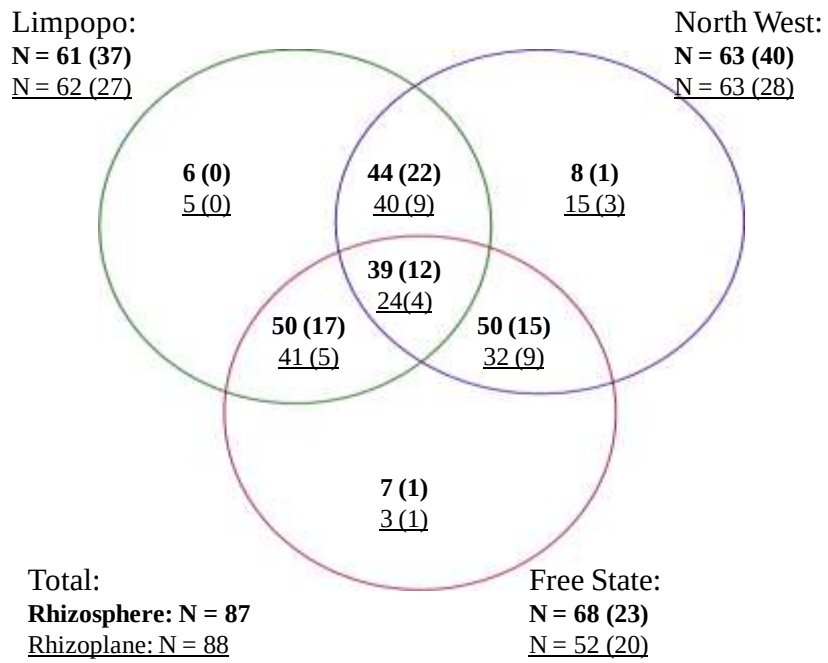
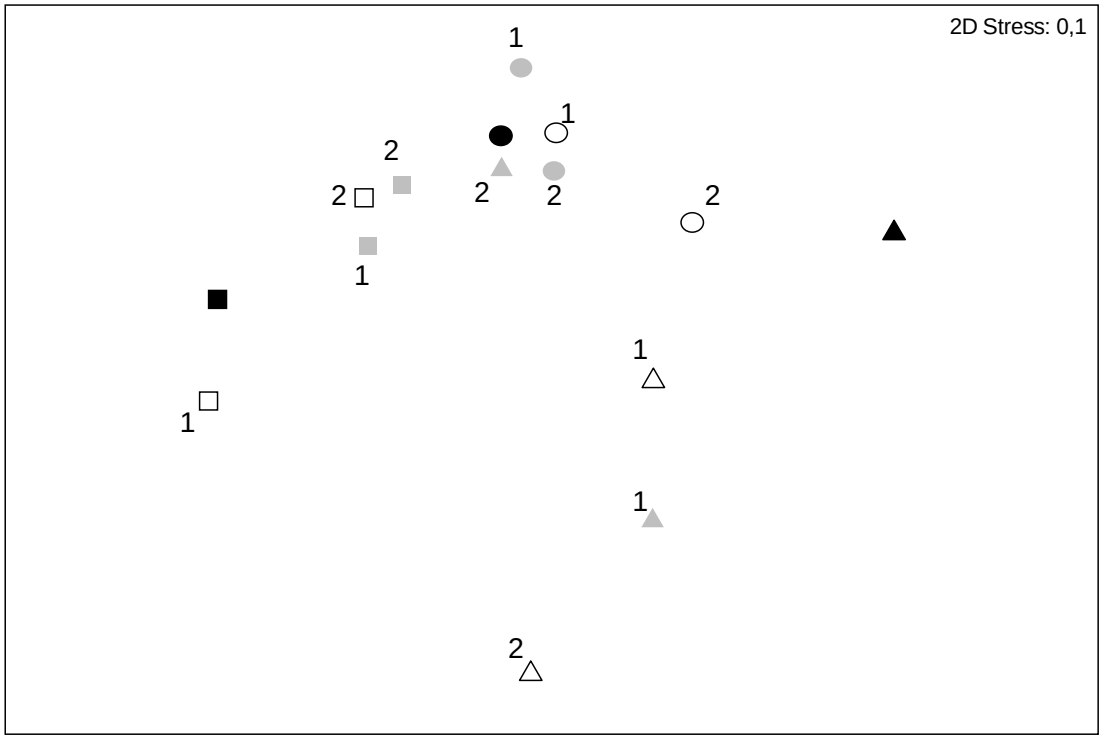

Limpopo:
N = 61 (37)
N = 62 (27)
North West:
N = 63 (40)
N = 63 (28)
6 (0)
5 (0)
44 (22)
40 (9)
8 (1)
15 (3)
39 (12)
24(4)
50 (17)
41 (5)
50 (15)
32 (9)
7 (1)
3 (1)
Total:
Rhizosphere: N = 87
Rhizoplane: N = 88
Free State:
N = 68 (23)
N = 52 (20)

Figure 4.

A:



B:

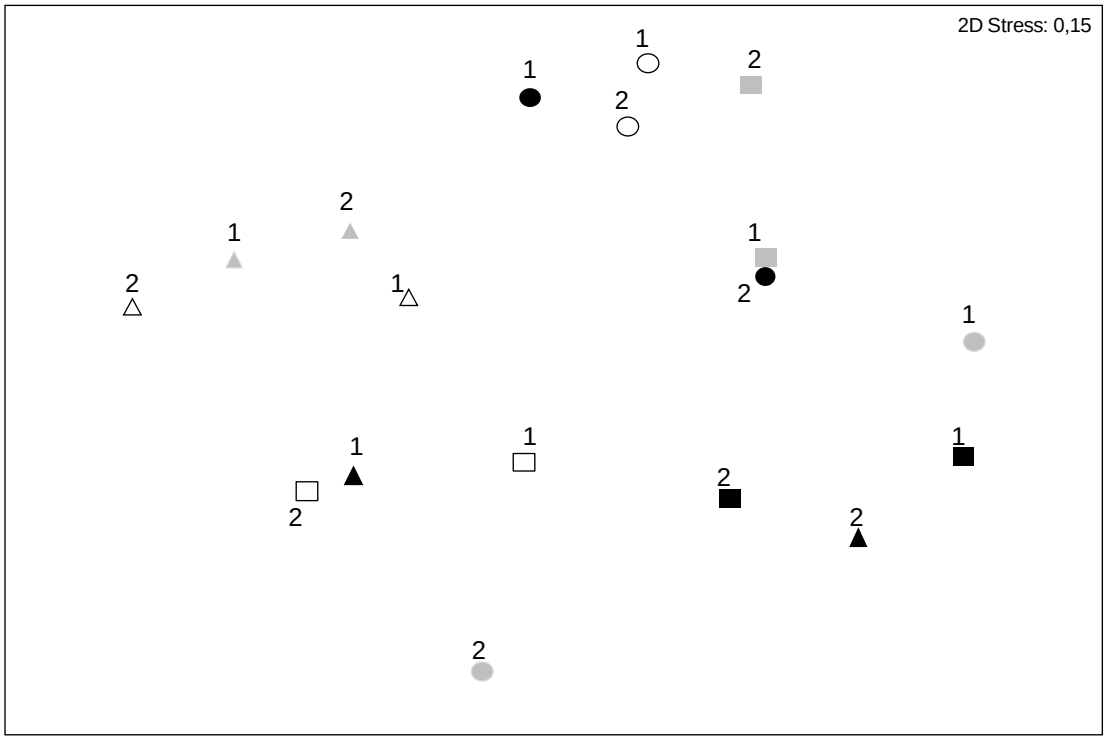


Figure 5

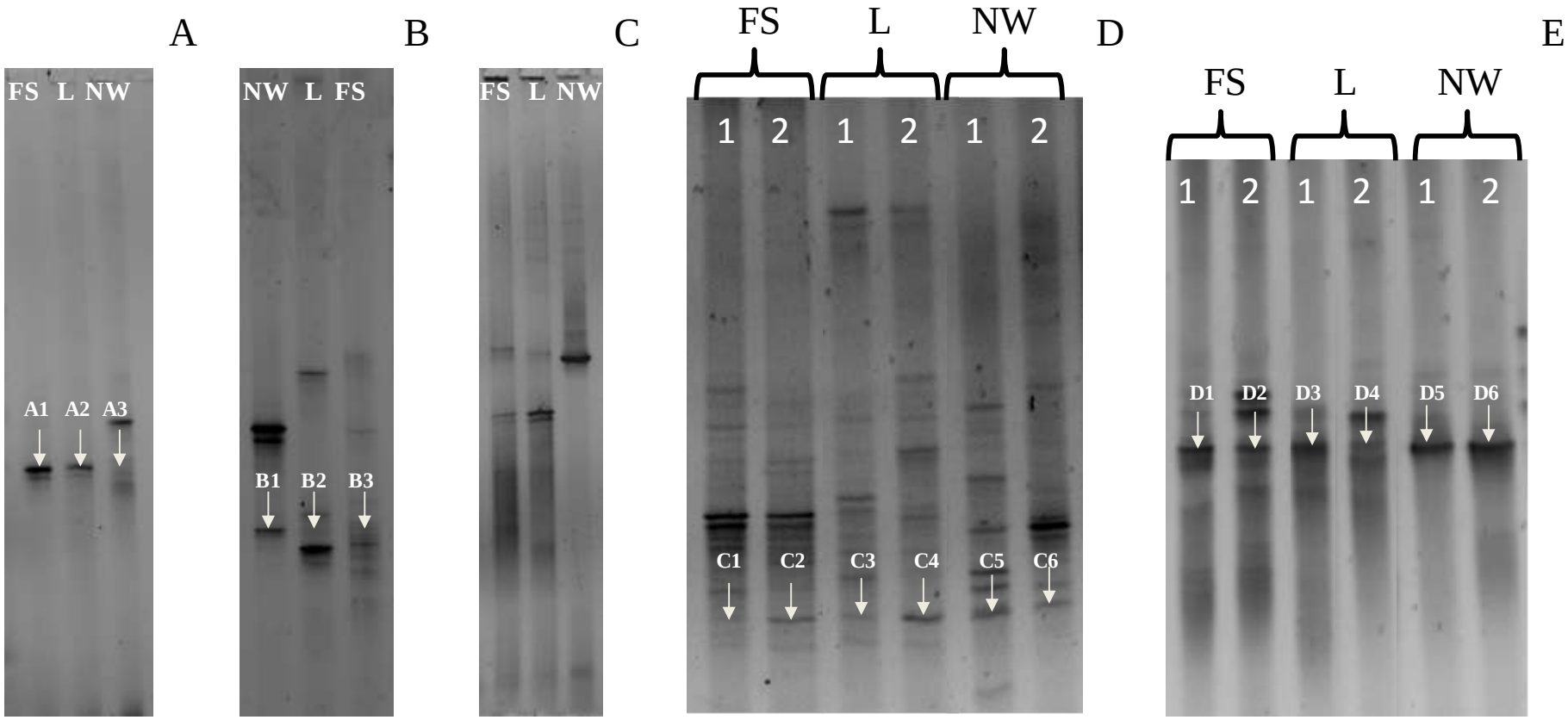

A
FS L NW
A1 A2 A3
B
NW L FS
B1 B2 B3
C
FS L NW
D
FS L NW
1 2 1 2 1 2
C1 C2 C3 C4 C5 C6
E
FS L NW
1 2 1 2 1 2
D1 D2 D3 D4 D5 D6

Table 1. Open and rhizospheric soil characteristics.

Province	Farm type	Agricultural practices	Soil type	pH	NH_4-N ($mg.kg^{-1}$)	NO_3-N ($mg.kg^{-1}$)	Total C (%)	Total N (%)
Free State	Commercial	Modern (N-fertilization)	Open soil	4.7	7.88	1.44	0.58	0.11
			Rhizosphere	4.2	8.68	0.52	0.40	0.11
Limpopo	Small household	Traditional (cow manure)	Open soil	5.4	9.6	3.72	0.19	0.09
			Rhizosphere	6.3	8.36	4.72	0.36	0.10
North West	Academic	Modern (N-fertilization)	Open soil	6.2	8.44	11.88	0.96	0.14
			Rhizosphere	6.0	9.60	5.80	0.94	0.14

C: Carbon. N: Nitrogen. NH_4-N: Ammonium. NO_3-N: Nitrate.

Table 2. Sequence similarities of excised DGGE-bands shown in Figure 5.

Ecological Niche	Province of origin	DGGE band	Most closely related sequence [Accession number][a]	% of Identity (number of bases)[b]	Origin	Taxonomic Group
Sorghum Root	Free State	A1	Various cyanobacterial 16S rRNA gene, including *Synechococcus* sp. clone R4CP3R1F09 [HQ018568.1]	100 % (162)	Sugarcane rhizosphere (Brazil)	Cyanobacteria
	Limpopo	A2		99 % (166)		
	North West	A3		100 % (170)		
Sorghum Shoot	Free State	B1	Various *Pantoea* strains, including *Pantoea dispersa* strain BH10 [JQ765428.1]	100 % (188)	Jasmine petal	γ-Proteobacteria
	Limpopo	B2	Various Enterobacteriacea, including Pantoea sp. CRPV0611B [FJ593752.1]	100% (182)	*Acromyrmex echinatior*	γ-Proteobacteria
	Free State	B3	*Pantoea ananatis* strain JB1/KB-10511[JQ513929.1]	99 % (188)	Rain water (Indonesia)	γ-Proteobacteria
Sorghum Rhizosphere	Free State	C1	*Bacillus pumilus* strain SL32 [JQ361041.1] *	98% (122)	Soil	Firmicutes
		C2	Uncultured *Escherichia* sp. [EF674507.1] *	89% (104)	Poultry farm	γ-Proteobacteria
	Limpopo	C3	*Escherichia fergusonii* ATCC 35469 [NR_027549.1] **	100% (139)	Human feces	γ-Proteobacteria
		C4	Uncultured Actinobacterium [EU300221.1] and Firmicutes [EF651750.1] clones *	100% (117)	Soil	Actinobacteria / Firmicutes
	North West	C5	Uncultured Actinobacterium [EU300221.1] and Firmicutes [EF651750.1] clones *	100% (117)	Soil	Actinobacteria / Firmicutes
		C6	*Escherichia coli* strain sch70 [JX294890.1] *	100% (139)	*Arion lusitanicus*	γ-Proteobacteria

Sorghum Rhizoplane	Free State	D1	Various *Bacillus* strains including *B. megaterium* strains UW2 [JX133188.1], UT3 [JX133180.1], AIMST 1.Hb.20 [HQ670443.1] and AIMST 3.24.2 [HQ694028.1] *	99% (129)	Leaf tissue / Rhizosphere	Firmicutes
		D2	Various *Bacillus* strains including *B. megaterium* strains UW2 [JX133188.1], UT3 [JX133180.1], AIMST 1.Hb.20 [HQ670443.1] and AIMST 3.24.2 [HQ694028.1] *	98% (121)	Leaf tissue / Rhizosphere	Firmicutes
	Limpopo	D3	Various *Bacillus* strains including *B. megaterium* strains UW2 [JX133188.1] and UT3 [JX133180.1] *	99% (110)	Rhizosphere	Firmicutes
		D4	Various *Bacillus* strains including *B. megaterium* strains UW2 [JX133188.1] and UT3 [JX133180.1] *	100% (121)	Rhizosphere	Firmicutes
	North West	D5	Various *Bacillus* strains including *B. megaterium* strains UW2 [JX133188.1] and UT3 [JX133180.1] *	100% (120)	Rhizosphere	Firmicutes
		D6	Various *Bacillus* strains including *B. megaterium* strains UW2 [JX133188.1] and UT3 [JX133180.1] *	100% (121)	Rhizosphere	Firmicutes

[a]: Data base blasted. *: Nucleotide collection (nr/nt) / **: 16S ribosomal RNA sequences (Bacteria and Archaea)

[b]: The number in parentheses correspond to the number of based used to calculate the levels of sequence identity