CABI Agriculture
and Bioscience

RESEARCH

Profiling bacterial communities of irrigation water and leafy green vegetables produced by small-scale farms and sold in informal settlements in South Africa

D. M. Kgoale[1,2], J. K. Gokul[1], S. Duvenage[2,3], E. M. Du Plessis[1,2] and L. Korsten[1,2]*

Abstract

Morogo is an African indigenous term used for leafy green vegetables harvested in the wild or cultivated in small-scale farms and consumed by the local populations of the region. Small-scale farmers have gained recognition as important suppliers of morogo to informal settlements. In commercial production systems, leafy green vegetables have increasingly been reported as associated with foodborne pathogens and disease outbreaks. Little is known of the presence of these organisms on leafy green vegetables in the informal unregulated food systems. This study aimed to profile bacterial communities in irrigation water (flooding and overhead irrigation water) and leafy green vegetables (*Brassica rapa L. chinensis* and *Brassica rapa* varieties of morogo)* to establish the natural bacterial flora at the water-fresh produce interface from five small-scale farms in two provinces in South Africa. Illumina MiSeq high-throughput sequencing showed that each farm exhibited a unique bacterial community composition, with an overall high relative abundance of Proteobacteria, Firmicutes and Actinobacteria, including prominent families such as *Burkholderiaceae* (48%), *Enterobacteriaceae* (34%), *Bacillales Family XII* (8%), *Rhodobacteraceae* (3%), *Micrococcaceae* (1.98%) and *Pseudomonadaceae* (1.79%). Specific *Enterobacteriaceae Serratia*, *Enterobacter*, *Salmonella*, *Shigella*, *Escherichia coli*, *Buchnera, Citrobacter, Klebsiella* and *Proteus* were identified, in addition to unique communities associated with plant or irrigation water source. These findings suggest that the edible plant microbiome can play an important role as transient contributor to the human gut and has the potential to affect overall health.

Keywords Morogo, Indigenous fresh produce, *Enterobacteriaceae*, Bacterial diversity, Irrigation water, Rape, Chinensis, One health, Food security, Food safety

*Correspondence:
L. Korsten
lise.korsten@up.ac.za
[1] Department of Plant and Soil Sciences, University of Pretoria, Hatfield, Pretoria, South Africa
[2] Department of Science and Innovation-National Research Foundation, Centre of Excellence in Food Security, Pretoria, South Africa
[3] Food and Markets Department, Natural Resources Institute, University of Greenwich, Chatham, UK

Introduction

Small-scale farming in South Africa accounts for 80% of fresh produce supplied to informal settlements (Bunce 2019). Despite this important food security role, small-scale farmers often struggle to produce optimal yields, mainly due to lack of access to adequate water resources and infrastructure, lack of land and funding, as well as technical support and know-how to consistently supply safe, quality food (Hlophe-Ginindza and Mpandeli 2020). The transmission of foodborne pathogens in animal-based food systems such as meat, poultry, eggs and

raw milk is well established (Abebe et al. 2020; Heredia and García 2018). However, foodborne pathogens outbreaks have increasingly been associated with fresh and minimally processed leafy green vegetables (Carstens et al. 2019, Herman et al. 2015). Little is known of the prevalence of foodborne pathogens in the informal food production systems. What is known is that these food systems carry an additional burden of being reliant on compromised water resources (Bhagwat 2019) without resources to manage and improve the quality of the water and having poor hygiene and sanitation systems operating in on an unregulated environment (Daniel 2022).

The diverse microbiomes of vegetables may serve as reservoirs for opportunistic and emerging pathogens (Berg et al. 2014). Bacteria in these matrixes occur as both beneficial and pathogenic variants when associated with plant, human and animal health. Pathogenic bacteria such as *Campylobacter* spp., *Clostridium botulinum*, *Clostridium perfringens*, *Bacillus cereus*, *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli*, *Listeria monocytogenes*, *Salmonella* spp., *Shigella* spp., enterotoxigenic *Staphylococcus aureus*, *Vibrio cholera*, and *Yersinia enterocolitica* can be significantly detrimental to human health in this one health context. Prevention of source contamination is predominantly advocated through more effective municipal waste management systems, improved hygiene practices and adoption of good agricultural practices within the farm-to-fork continuum. In the process of fresh produce e.g., leafy green vegetable production, contamination with pathogenic microorganisms may result through direct contact with contaminated irrigation water (groundwater, surface water, and human wastewater), improper composted manure, infected soil, or contact with food handlers not having access to basic potable water, hygiene and sanitation systems (Alegbeleye et al. 2018). While agricultural scientists are mindful of potential pathogenic microorganisms found naturally on the plant surface, the awareness of endophytic pathogens is only recently been shown (Dong et al. 2019, Mulaosmanovic et al. 2021). While most endophytes are likely non-pathogenic to humans, several pathogenic bacteria can become internalized and colonise the leaves as temporary endophytes (Mengistu 2020). From the perspective of consumer safety, these temporary pathogenic endophytic inhabitants and other epiphytes present within the vegetable biome can potentially cause disease outbreaks, particularly where no washing or a heat treatment step is practised after harvest.

Indigenous leafy green vegetables, also known as "morogo", naturally harbour a diverse microbial consortium in the phyllosphere (Taffner et al. 2020). They are considered highly nutritious, are popular with local communities and are mostly traded in informal markets.

However, little is known about the water-indigenous leafy green-vegetable nexus in informal food systems in South Africa. The presence of potential human pathogens in the microbial community of indigenous leafy green vegetables is of interest in the context of a product that has a high likelihood of contamination during production and further on in the supply chain during transport, sales and preparation. As such, it is imperative to determine the microbial ecology of the local indigenous leafy green small-scale production systems and informal supply chain. Next-generation 16S rRNA gene sequencing provides a tool to unravel the bacterial diversity of irrigation water, and fresh produce in the informal food system, detect the presence of potential foodborne pathogens amplicons and highlight the core microbiome that is prevalent in the one health context.

Materials and methods
Sample collection and pre-processing
A total of 36 samples of overhead or flood irrigation water was collected from multiple irrigation points across the plot ensuring water was in the proximity of both crops sampled. *Brassica rapa L. chinensis*, now referred to as chinensis (n = 25), and *Brassica rapa*, now referred to as rape (n = 25), plants were randomly collected from a single plot per cultivar, once off using systematic random sampling. Several leaves free from soil, from different plants were pooled to represent a single sample. Five small-scale farms in the Brits (North West Province, farming site A, B, C and E) and Delmas (Mpumalanga Province, farming site D) regions, South Africa (Additional file 1: Table S1) were sampled at. Water samples (1 L) were filtered through a 0.20 μm nitrocellulose filter membrane (Sartorius Stedim Biotech, Gottingen, Germany) and aseptically stored at 4 °C until DNA extraction. Chinensis and rape samples (25 g) were placed into stomacher bags containing 225 mL of Peptone Buffered Water (PBW) and subsequently homogenized for five min at maximum speed (level 4) using a Homogenizer Laboratory Blender (Thermo Fischer Scientific, Johannesburg, South Africa). The homogenates were transferred to 50 mL sterile falcon tubes and centrifuged at maximum speed (4000 rpm) for 10 min. The resulting pellets were stored at −20 °C until DNA extraction.

DNA extractions
Total community DNA from water samples were extracted from the filter membranes using the DNeasy® PowerWater® extraction kit (Qiagen, Johannesburg) following the manufacturer's guidelines. For fresh produce samples, DNA was extracted from fresh produce homogenate pellets using the DNeasy® PowerSoil® extraction kit (Qiagen) according to the manufacturer's

Kgoale *et al. CABI Agriculture and Bioscience* (2023) 4:36

Page 3 of 10

guidelines. All DNA samples were quantified and verified for purity using Qubit 2.0 fluorometer with both double stranded broad range and high sensitivity range, according to manufacturer's guidelines (Life Technologies, Johannesburg) and Nanodrop ND-2000 UV–VIS Spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA).

16S rRNA amplicon sequencing and processing

Library preparation and paired-end sequencing (2×300 bp) of the 16S rRNA V4-V5 variable gene region was performed by Molecular Research DNA (MR DNA, Shallowater, TX, USA) on the Illumina MiSeq platform using primers 515F (5'-GTGCCAGCMGCCGCG GTAA-3') and 909R (5'-CCCCGYCAATTCMTTTRAG T-3') (Wang and Qian 2009). Generated sequences were returned and pre-processed by demultiplexing and barcode and primer removal using the MR DNA in-house pipeline, and MR DNA freeware for binning and conversion from fastq to fasta and quality files. Sequence data are available at NCBI-SRA under submission numbers SUB12270895 and SUB12272756 for BioProject number PRJNA900001.

The bioinformatics open-source software Quantitative Insights Into Microbial Ecology 2 (QIIME2) version 2020.2 was utilised for raw data processing (Bolyen et al. 2019). Demultiplexing of sequences was performed according to the DADA2 workflow, using the q2-demux plugin (https://github.com/qiime2/q2-demux). Reads were trimmed (forward: 12 bp, end: 220 bp) prior to quality filtering and de-replication using the q2-dada2 plugin (Callahan et al. 2016). This process concurrently removes chimeric sequences and generates amplicon sequence variants (ASVs), using nucleotide quality scores. Taxonomic classification of ASVs was assigned using the Naïve Bayes classifier and the q2-feature-classifier plugin, which assigns taxonomies based on sequence k-mer frequencies (https://github.com/qiime2/q2-feature-class ifier), and the self-trained SILVA 138 16S rRNA gene database clustered at 99% similarity (Quast et al. 2013). Mitochondria, chloroplast and archaea sequences were removed. Where necessary, feature tables produced were normalised by rarefaction to a sampling depth of 24,230 reads for downstream diversity analysis.

Diversity measurements and statistical analyses

The alpha (α-) and beta (β-) diversity of bacterial communities was determined using the vegan function in R Statistical Software (R Core Team, 2021). The α-diversity indices measured using non-rarefied data included bacterial community richness, Pielou's evenness index (J), and Shannon diversity index (H). The bacterial community structure between samples (β-diversity) was visualised

using rarefied tables for non-metric multidimensional scaling (nMDS) ordination of the Bray–Curtis dissimilarity matrix (stress values: 0.083–0.165) and verified with permutational multivariate analysis of variance (PERMANOVA, 999 permutations) based on the factors source (overhead water or flooding water; chinensis or rape), area (Brits or Delmas) and farming site (A, B, C, D, E). Homogeneity of multivariate dispersions was determined with a permutational multivariate dispersions (PERMDISP) test using the Bray–Curtis similarity matrix (Anderson et al. 2006).

The Shapiro–Wilk normality test was used to determine the distribution of samples for further statistical testing. Where normal distribution was observed, a *T*-test or ANOVA was used followed by a post hoc Tukey test. Alternatively, the Kruskal–Wallis H sum test was used for samples not normally distributed, with further analysis using the Dunn test with a Bonferroni correction. Results with p values of ≤ 0.05 were considered statistically significant.

Results

Bacterial community diversity and composition

A total of 2,139,282 reads were recovered from the 16S V4-V5 region of the 68 samples after paired-end alignment, quality filtering and deletion of chimeric sequences. Four samples from farming site D (chinensis samples 3, 4, 5 and rape sample 3) were excluded from the analysis during filtering as these only matched to plant chloroplast and mitochondrial DNA. The 64 remaining samples were subsequently assigned to a total of 2249 amplicon sequence variants (ASVs). Taxonomic composition of bacteria from the 64 samples showed approximately 28 phyla (excluding unidentified and unassigned taxa) composed of 67 classes, 176 orders, 243 families, 298 families, 566 bacterial genera and 260 bacterial species.

The α-diversity analysis of irrigation water (Fig. 1) showed a significant difference in bacterial richness, Shannon diversity and Pielou's evenness within samples when analysed by farming site, with $p = 0.015$, $p = 0.003$ (ANOVA) and $p = 0.029$ (Kruskal Wallace), respectively. A post hoc TukeyHSD test further highlighted significant community diversity difference for bacterial richness between farming sites A–D ($p = 0.022$), A–E ($p = 0.014$), B–E ($p = 0.032$), C–D ($p = 0.016$) and C–E ($p = 0.011$), while the Shannon diversity showed significantly different community compositions between farming sites A–D ($p = 0.030$), A–E ($p = 0.005$) and B–E ($p = 0.015$). Irrigation water source was also determined to be a significant contributor to bacterial richness (ANOVA: $p = 0.003$) and H (ANOVA: $p = 0.018$), with the sampling area being significant only for bacterial richness (ANOVA: $p = 0.049$).
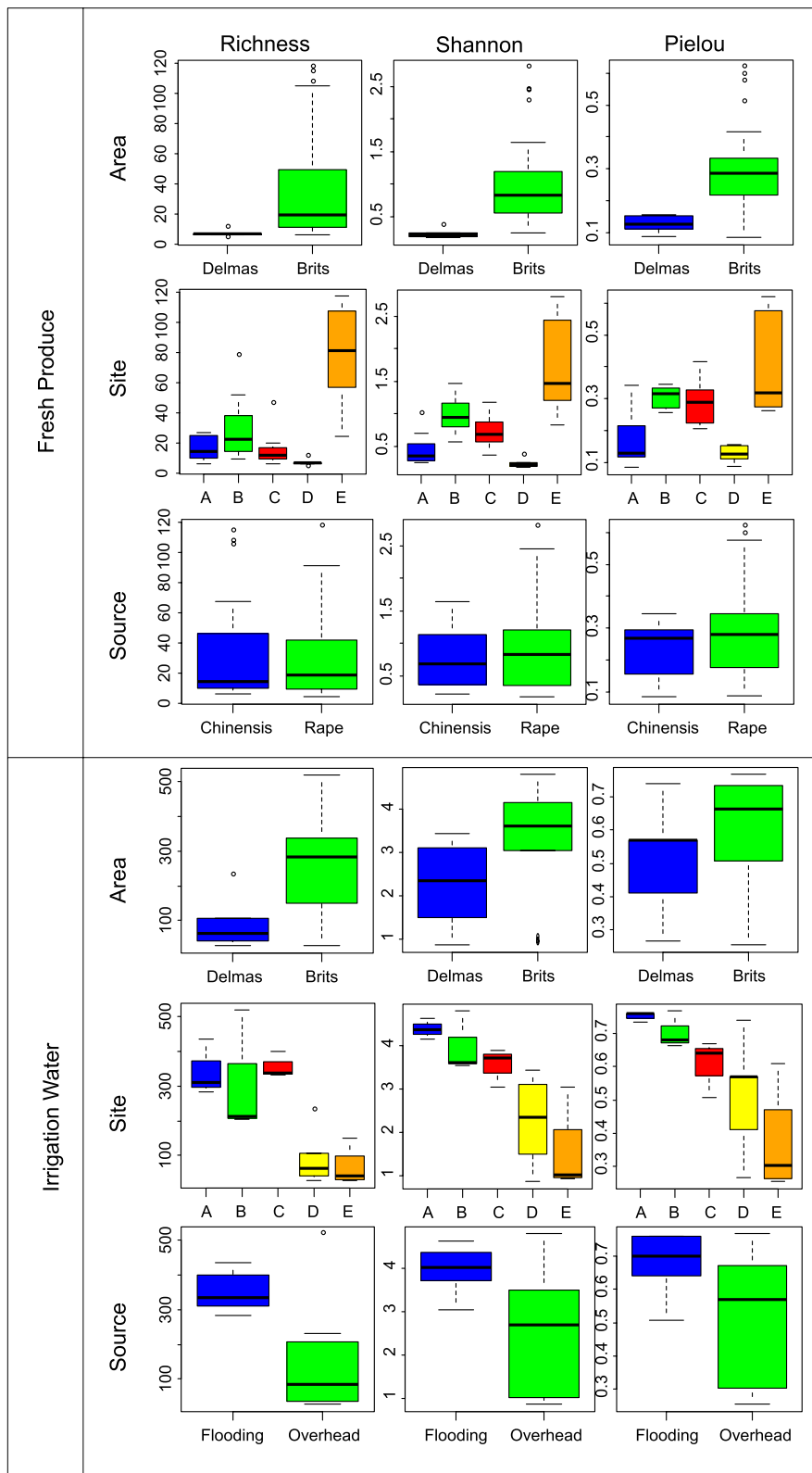
Fig. 1 Diversity indices of irrigation water and fresh produce based on the area (Brits, Delmas), farming site (**A**–**E**) and source (flooding, overhead irrigation, chinensis, rape)

Kgoale *et al. CABI Agriculture and Bioscience*   (2023) 4:36

Page 5 of 10

The bacterial community showed distinct clustering when evaluated against area (PERMANOVA: $R^2 = 0.145$, $p = 0.012$), farming site (PERMANOVA: $R^2 = 0.523$, $p = 0.001$) and irrigation water source (PERMANOVA: $R^2 = 0.171$ $p = 0.01$,) (Fig. 2).

Overall, the observed alpha bacterial diversity in chinensis and rape (morogo) showed a significant difference in community composition based on area and farming site for ASV richness (Kruskal–Wallis rank sum test: area, $p = 9.737 \times 10^{-4}$; farming site, $p = 1.161 \times 10^{-5}$), Shannon diversity (H, Kruskal–Wallis rank sum test: area, $p = 2.006 \times 10^{-4}$; farming site, $p = 6.736 \times 10^{-7}$) and Pielou's evenness (J, Kruskal–Wallis rank sum test: $p = 1.945 \times 10^{-3}$; farming site, $p = 2.882 \times 10^{-5}$) (Fig. 1).

The post hoc Dunn test showed that farming sites A–D (richness: $p = 0.017$), A–E (richness: $p = 0.005$), C–E (richness: $p = 0.0004$; H: $p = 0.009$), D–E (richness: $p = 0.000$; H: $p = 0.0003$) demonstrated the most significant differences between community composition for the measured beta diversity indices. No significant differences in community diversity were observed for chinensis and rape. Non-metric multidimensional scaling (nMDS) ordinations of the fresh produce bacterial

communities showed distinct and significant separation for the farming site (PERMANOVA: $R^2 = 0.276$, $p = 0.001$; PERMDISP: $F = 3.518$, $p = 0.016$) (Fig. 2), however, no evidence of bacterial clustering by area or source was evident, despite being statistically significant in their composition (PERMANOVA: area, $R^2 = 0.049$, $p = 0.002$; source, $R^2 = 0.045$, $p = 0.003$,). It should be noted that no statistical differences between the dispersions were present for area (PERMDISP: $F = 0.000$, $p = 0.997$) and source (PERMDISP: $F = 0.258$, $p = 0.589$).

Bacterial community composition of irrigation water and morogo

The predominant phyla observed in all samples (irrigation water, morogo) were Proteobacteria (69.68% relative abundance), Firmicutes (16.47% relative abundance), Actinobacteria (6.06% relative abundance), Bacteroidetes (2.95% relative abundance), Verrucomicrobia (0.75% relative abundance), Cyanobacteria (0.43% relative abundance), Chloroflexi (0.27% relative abundance), Planctomycetes (0.27% relative abundance) and Gemmatimonadetes (0.18% relative abundance) (Fig. 3). Elusimicrobia, Latescibacteria and Nitrospinae were exclusive to
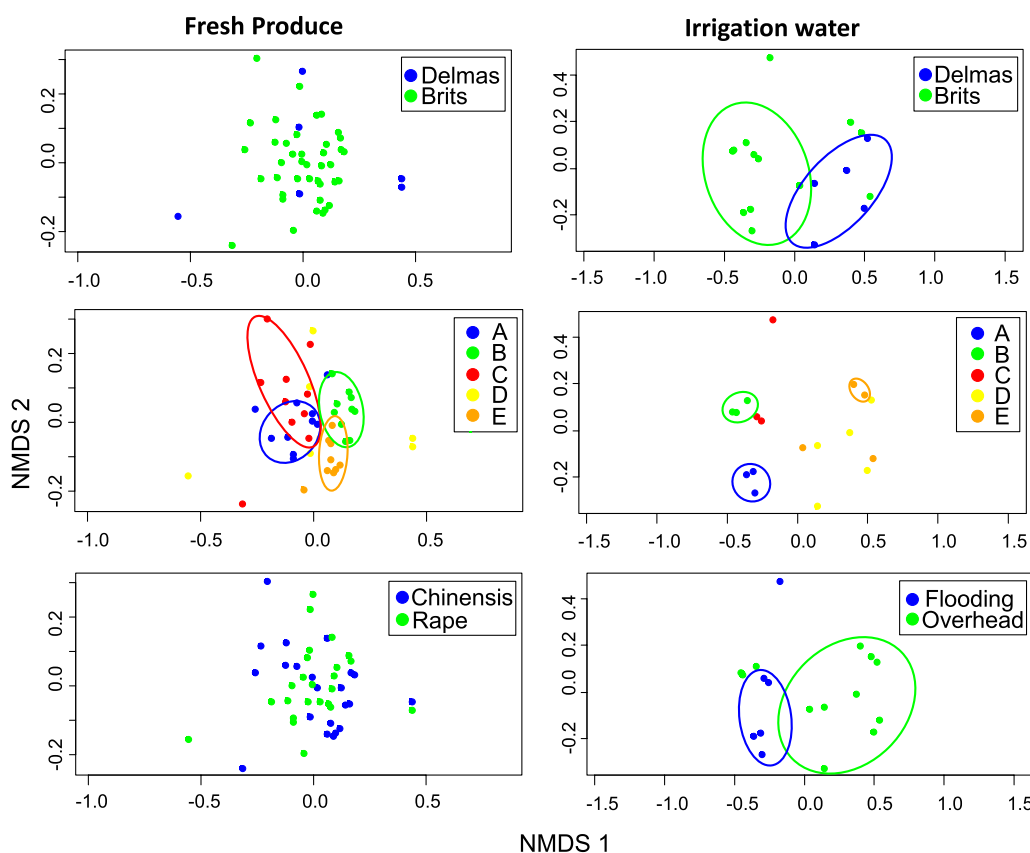
Fig. 2 Non-metric multidimensional scaling (nMDS) ordinations of bacterial communities from irrigation water and fresh produce samples assessed on area (Brits, Delmas), farming site (**A**–**E**) and source (flooding, overhead irrigation, chinensis, rape)
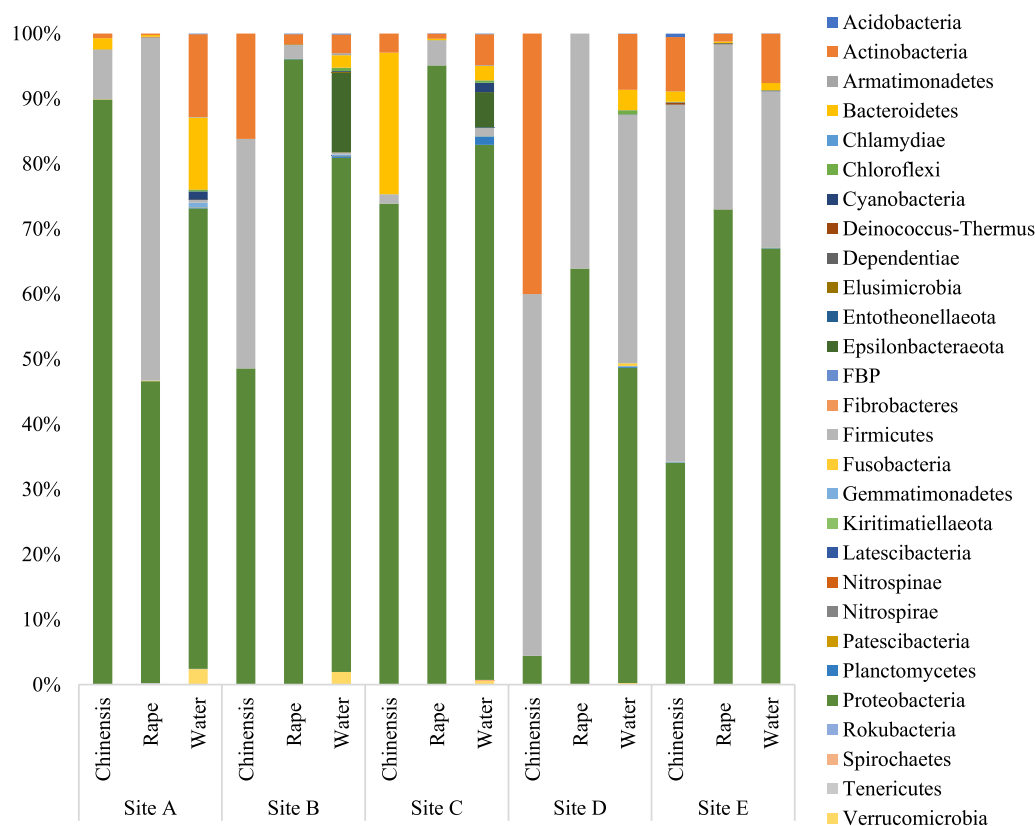
Fig. 3 Relative abundance of different bacterial phyla detected on chinensis, rape and irrigation water samples from informal morogo small-scale farms (farming site **A**–**E**) in South Africa

farming site B's irrigation water while Fibrobacteres, Bacteriodetes and Planctomycetes (FBP) and Spirochaetes were only exclusive to farming site C's irrigation water. Chlamydiae, Dependentiae, Elusimicrobia, Epsilonbacteraeota, FBP, Fibrobacteres, Kiritimatiellaeota, Latescibacteria, Nitrospinae, Patescibacteria and Spirochaetes were only prevalent in irrigation water at all farming sites (Fig. 3). Nitrospirae, Rokubacteria and Tenericutes were only prevalent in irrigation water and rape, while Armatimonadetes were only prevalent in irrigation water and chinensis (Fig. 3). Further analysis of relative abundances at family level in all samples revealed *Burkholderiaceae* (48.00% relative abundance), *Enterobacteriaceae* (34.00% relative abundance), *Bacillales Family XII* (8.00% relative abundance), *Rhodobacteraceae* (3.00% relative abundance), *Micrococcaceae* (1.98% relative abundance), *Pseudomonadaceae* (1.79% relative abundance), *Bacillaceae* (1.33% relative abundance) and *Moxaxellaceae* (0.43% relative abundance) as the most dominant bacterial families.

Core analysis was performed at 100% similarity across all irrigation water and morogo samples (Additional file 1: Table S2). A single *Micrococcaceae* ASV was a common taxon in chinensis samples (15/22) and flooding irrigation water (6/6), while *Exiguobacterium* was the only common taxon in rape samples (19/24). Flooding irrigation water exhibited a unique core community dominated by 23 taxa (Additional file 1: Table S2), with *Burkholderiaceae* being the highest abundant of those present, while overhead irrigation water showed a common seven taxa, with *Ralstonia* (10/12), *Staphylococcus* (9/12) and *Escherichia-Shigella* (10/12) as the highest contributing members.

Thirteen species within the *Enterobacteriaceae* family were identified in the overall community analysis: *Serratia* (0.25% relative abundance), *Raoultella* (0.06% relative abundance), *Enterobacter* (0.27% relative abundance), *Salmonella* (4.34% relative abundance), *Erwinia* (0.56% relative abundance), *Shigella* (0.06% relative abundance), *E. coli* (26.48% relative abundance), *Buchnera* (1.60% relative abundance), *Citrobacter* (23.00% relative abundance), *Klebsiella* (3.57% relative abundance), *Cedecea* (2.11% relative abundance), *Pantoea* (37.48% relative abundance) and *Proteus* (0.06% relative abundance). *Salmonella* and *E. coli* were most prevalent in irrigation water at all farming sites. Overall,

Raoultella, *Shigella* and *Proteus* were identified in irrigation water samples, while *Erwina*, *Buchnera*, *Citrobacter* and *Cedecea* were identified in fresh produce.

Upon closer inspection of the ASVs at each farming site, farming site A showed distinctive *Enterobacteriaceae* species in irrigation water (*Salmonella* and *E. coli*). This is in contrast to fresh produce, which showed a dominance of *Enterobacter*, *Erwinia*, *Buchnera* and *Pantoea* species. In farming site B, *Serratia*, *Salmonella*, *E. coli* and *Klebsiella* were identified in both irrigation water and fresh produce, while *Raoultella* was identified in irrigation water only and *Erwinia*, *Buchnera*, *Cedecea* and *Pantoea* were identified in fresh produce. Farming site C harboured distinctive *Enterobacteriaceae* species *Serratia*, *Salmonella*, *E. coli* and *Klebsiella* in irrigation water (Fig. 4). This contrasts with fresh produce, which showed high abundances of *Buchnera*, *Cedecea* and *Pantoea*. In farming site D, *E. coli* was identified in both irrigation water and fresh produce, while *Raoultella*, *Salmonella*, *Shigella*, *Klebsiella* and *Pantoea* were identified in irrigation water. In farming site E, *Serratia*, *Salmonella*, *Enterobacter*, *E. coli* and *Klebsiella* were observed in both fresh produce and irrigation water; *Shigella* and *Proteus* were present in irrigation water only, and *Erwina*, *Buchnera*, *Citrobacter*, *Cedecea* and *Pantoea* were identified in fresh produce only.
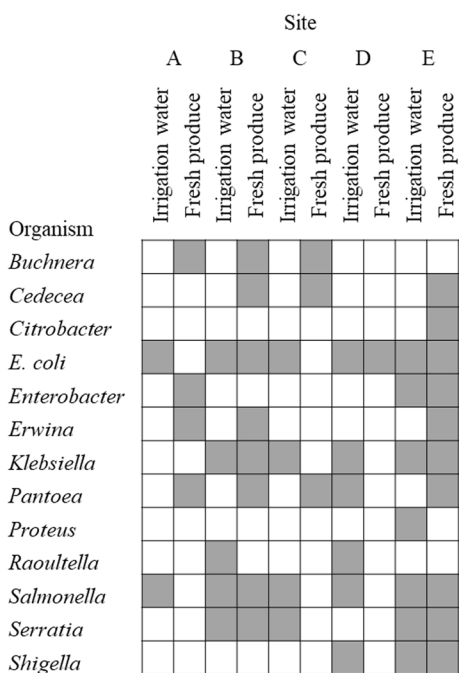
Fig. 4 *Enterobacteriaceae* genera were identified in irrigation water and leafy green vegetables in the five small-scale farms

Discussion

In this study, we used high-throughput Illumina sequencing to characterize the microbial community of irrigation water and fresh produce from regional small-scale farming systems in South Africa. We were further able to identify potentially pathogenic bacterial ASVs from small-scale farming systems and informal settings. Overall, we were able to identify a highly diverse bacterial community that includes copiotrophic phyla such as Proteobacteria, Firmicutes, Bacteriodetes, and Actinobacteria, which are normally detected on other indigenous African plant and crop species (Berg et al. 2016; Bulgarelli et al. 2013; Ho et al. 2017; Taffner et al. 2020). At the genus level, it was evident that both fresh produce and irrigation water harboured unique bacterial communities which may be driven by specific adaptations to distinct environmental factors in the farming area and farming activities at each farming site as well as plant species specificity.

A core community was identified for each source, demonstrating that indigenous leafy green vegetables such as chinensis and rape do not harbour a core microbiome, which is in complete contrast to the irrigation water sources. Flooding irrigation water appeared to possess a core microbiome that consisted of common environmental bacteria isolated from irrigation and soil. However, it is in close proximity to soil, which harbours a complex and dynamic bacterial community (Mäder et al. 2002). *Ralstonia* and *Escherichia-Shigella* are amongst the highest contributing core members in overhead irrigation water. Previously reported disease outbreak-associated bacteria in South Africa were matched to genera dominating the bacterial families identified in this study (Additional file 1: Table S3) and include *Clostridiaceae* and *Enterococcaceae* known to include genera that indicate human or animal faecal contamination.

In this pioneering study, we identified 28 different phyla, excluding those that were unidentified or unassigned in both fresh produce and irrigation water. Our results correspond with Gu et al. (2019) that shows Proteobacteria, Actinobacteria, Bacteroidetes and Firmicutes are the dominant phyla in irrigation water, regardless of the source (groundwater, wastewater and roof-harvested rainwater). In the present study, we additionally identified Firmicutes, which mostly dominates fresh produce (Leff et al. 2013), and which is a phylum prominent in irrigation water in small-scale farming systems. Leafy green vegetables such as spinach and lettuce are dominated by Proteobacteria, Actinobacteria, Firmicutes and Bacteroidetes (Berg et al. 2014; Leff et al. 2013). Morogo (rape and chinensis) show homologous bacterial communities differing in

abundance, complementary to previous studies on lettuce and spinach (Jackson et al. 2015; Leff et al. 2013).

The high diversity observed in all farming sites, production areas and plant sources show the potential for these microbial communities to contain bacteria which have previously been associated with foodborne illnesses. Nevertheless, analysis of the community structure at all sampling points in this study showed the presence of *Enterobacteriaceae*, known to be associated with vegetables (Berg et al. 2014) and some members are potential foodborne pathogens (Janda and Abbott 2021). In this study, faecal indicators and potential foodborne pathogens such as *Enterobacter*, *E. coli*, *Klebsiella*, *Pantoea*, *Salmonella*, *Serratia* and *Shigella* were also prevalent.

It is important to take note that flooding water had higher bacterial diversity than overhead irrigation water. The microbiota of crops are directly affected by irrigation application methods (Rodrigues et al. 2020, Banach and van der Fels-Klerx 2020), therefore, contaminated irrigation water could lead to morogo contamination. This is especially relevant when considering the large number of reports associated with foodborne illness outbreaks involving consumption of leafy green vegetables (Browne et al. 1962; Jaklevic 2021; Marshall et al. 2020). This highlights the need for extensions in surveillance of foodborne pathogens in the production chain of small-scale farming environments. Further to this, the World Health Organisation (WHO) lists genera and species of emergent *Enterobacteriaceae* as potential superbugs (WHO 2017) due to the presence of extended-spectrum beta-lactamases (ESBLs). These bacteria are of clinical importance due to its health care impact rendering antibiotics ineffective (Teklu et al. 2019). In culture-dependent studies performed in parallel with this study, ESBLs or AmpC-producing *Enterobacteriaceae* were isolated, including *E. coli*, *Serratia fonticola*, *Klebsiella pneumoniae*, *Enterobacter* spp., *Citrobacter* spp., *Rahnella aquatilis*, *Cedecea davisae* and *Proteus vulgaris* (Du Plessis et al. 2021). Furthermore, multi-virulent and multidrug resistant *S. enterica* strains were isolated from source and irrigation water and fresh produce (Du Plessis et al. 2021). These multidrug resistant bacteria are natural in the environment, however, their presence is indicative of human or animal faecal contamination (Delahoy et al. 2018). These microbes are a potential threat to consumers, especially if the leafy green vegetables are consumed unwashed, eaten raw or not cooked. Future studies should focus on the prevalence of foodborne pathogen surveillance systems in unregulated farming systems and informal systems.

Conclusion

In this study we have presented for the first time, a diverse and robust bacterial microbiome profile of irrigation water and indigenous leafy green vegetables produced in small-scale farming systems and sold in informal settings. The current study provides an important framework for further research into the microbial communities of small-scale farming systems and the informal supply chains and its link with public health. Mining this data permitted the detection of potential foodborne pathogen ASVs that arise from the microbial assemblages in irrigation water of small-scale farming systems. The presence of *Enterobacteriaceae* as well as other putative pathogens commonly associated with human illness in local farming systems demonstrates that irrigation water and morogo cannot be ruled out as potential sources of foodborne pathogens and antimicrobial resistance. From this, plant-associated bacteria play a crucial role in the host microbiome and human health, especially when the product is consumed raw. The maintenance and support of microbial diversity are of key importance to stabilise agroecosystems. In addition, an extension of molecular surveillance in small-scale informal supply chains will be an integral component to establish a "fit-for-purpose" food safety management system to reduce the likelihood of foodborne disease outbreaks.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43170-023-00176-0.

Additional file 1: Table S1. Samples analysed for bacterial community characterisation. **Table S2.** Taxonomic breakdown of core bacterial taxa present in flooding irrigation water. **Table S3.** Bacterial families that are associated with isolation and outbreaks in South Africa.

Kgoale *et al. CABI Agriculture and Bioscience*   (2023) 4:36

Page 9 of 10

References

Aaron J, van Zyl LJ, Dicks LMT. Isolation and characterization of lytic *Proteus* virus 309. Viruses. 2022;14(6):1309. https://doi.org/10.3390/v14061309.

Abebe E, Gugsa G, Ahmed M. Review on major food-borne zoonotic bacterial pathogens. J Trop Med. 2020. https://doi.org/10.1155/2020/4674235.

Alegbeleye OO, Singleton I, Sant'Ana AS. Sources and contamination routes of microbial pathogens to fresh produce during field cultivation: a review. Food Microbiol. 2018;73:177–208. https://doi.org/10.1016/j.fm.2018.01.003.

Anderson MJ, Ellingsen KE, McArdle BH. Multivariate dispersion as a measure of beta diversity. Ecol Lett. 2006;9(6):683–93. https://doi.org/10.1111/j.1461-0248.2006.00926.x.

Bamford C, Milligan P, Kaliski S. Dangers of *Clostridium perfringens* food poisoning in psychiatric patients. S Afr J Psychiatr. 2019. https://doi.org/10.4102/sajpsychiatry.v25i0.1339.

Banach JL, van der Fels-Klerx HJ. Microbiological reduction strategies of irrigation water for fresh produce. J Food Prot. 2020;83(6):1072–87. https://doi.org/10.4315/JFP-19-466.

Berg G, Erlacher A, Smalla K, Krause R. Vegetable microbiomes: is there a connection among opportunistic infections, human health and our 'gut feeling'? Microb Biotechnol. 2014;7(6):487–95. https://doi.org/10.1111/1751-7915.12159.

Berg G, Rybakova D, Grube M, Köberl M. The plant microbiome explored: Implications for experimental botany. J Exp Bot. 2016;67(4):995–1002. https://doi.org/10.1093/jxb/erv466.

Bhagwat VR. Safety of water used in food production. Food Safety Hum Health. 2019. https://doi.org/10.1016/B978-0-12-816333-7.00009-6.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852–7. https://doi.org/10.1038/s41587-019-0209-9.

Browne AS, Lynch G, Leonard AR, Stafford G. A clostridial or enterococcal food poisoning outbreak. Public Health Rep. 1962;77(6):533–6. https://doi.org/10.2307/4591538.

Bulgarelli D, Schlaeppi K, Spaepen S, van Themaat EVL, Schulze-Lefert P. Structure and functions of the bacterial microbiota of plants. Annu Rev Plant Biol. 2013;64:807–38. https://doi.org/10.1146/annurev-arplant-050312-120106.

Bunce B. Fresh produce production under irrigation by small-scale farmers in South Africa Reducing Maternal and Child Mortality through Strengthening Primary Health Care in South Africa Programme (RMCH) View project GTAC/CBPEP/EU project on Employment-intensive rural land reform in South Africa: policies, programmes and capacities' View project. 2019. https://doi.org/10.13140/RG.2.2.14449.92003/1

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581–3. https://doi.org/10.1038/NMETH.3869.

Carstens CK, Salazar JK, Darkoh C. Multistate outbreaks of foodborne illness in the United States associated with fresh produce from 2010 to 2017 Fronti Microbiol. Frontiers Media SA. 2019;10:2667. https://doi.org/10.3389/fmicb.2019.02667.

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

Daniel J. The importance of a food safety framework for the informal sector - Centre of Excellence. 2022. https://foodsecurity.ac.za/news/the-importance-of-a-food-safety-framework-for-the-informal-sector/. (Accessed 9 Dec 2022)

Delahoy MJ, Wodnik B, McAliley L, Penakalapati G, Swarthout J, Freeman MC, et al. Pathogens transmitted in animal feces in low- and middle-income countries. Int J Hyg Environ Health. 2018;221(4):661. https://doi.org/10.1016/j.ijheh.2018.03.005.

Dong CJ, Wang LL, Li Q, Shang QM. Bacterial communities in the rhizosphere, phyllosphere and endosphere of tomato plants. PLoS ONE. 2019. https://doi.org/10.1371/journal.pone.0223847.

Dramowski A, Aucamp M, Bekker A, Mehtar S. Infectious disease exposures and outbreaks at a South African neonatal unit with review of neonatal outbreak epidemiology in Africa. Int J Infect Dis. 2017;57:79–85. https://doi.org/10.1016/j.ijid.2017.01.026.

Du Plessis EM, Duvenage S, Korsten L, Sigge G. Measurement of water pollution determining the sources and changes of microbial contamination and impact on food safety from farming to retail level for fresh vegetables (WRC Report No. 2706/1/21). Water Research Commission. 2021. https://www.wrc.org.za/?mdocs-file=61886

Essel V, Tshabalala K, Ntshoe G, Mphaphuli E, Feller G, Shonhiwa AM, et al. A multisectoral investigation of a neonatal unit outbreak of *Klebsiella pneumoniae* bacteraemia at a regional hospital in Gauteng Province, South Africa. South Afr Med J. 2020;110(8):783–90. https://doi.org/10.7196/samj.2020.v110i8.14471.

Goszczynska T, Moloto VM, Venter SN, Coutinho TA. Isolation and identification of *Pantoea ananatis* from onion seed in South Africa. Seed SciTechnol. 2006;34(3):655–68. https://doi.org/10.15258/sst.2006.34.3.12.

Gu G, Yin HB, Ottesen A, Bolten S, Patel J, Rideout S, et al. Microbiomes in ground water and alternative irrigation water, and spinach microbiomes impacted by irrigation with different types of water. Phytobiomes J. 2019;3(2):137–47. https://doi.org/10.1094/PBIOMES-09-18-0037-R.

Heredia N, García S. Animals as sources of food-borne pathogens: a review animal nutrition. Keai Communicat Co. 2018;4(3):250–5. https://doi.org/10.1016/j.aninu.2018.04.006.

Herman KM, Hall AJ, Gould LH. Outbreaks attributed to fresh leafy vegetables, United States, 1973–2012. Epidemiol Infect. 2015;143(14):3011–21. https://doi.org/10.1017/S0950268815000047.

Hlophe-Ginindza NS, Mpandeli NS. The role of small-scale farmers in ensuring food security in Africa IntechOpen. Food Secur Afric. 2020. https://doi.org/10.5772/intechopen.91694.

Ho A, di Lonardo DP, Bodelier PLE. Revisiting life strategy concepts in environmental microbial ecology. FEMS Microbiol Ecol. 2017;93(3):6. https://doi.org/10.1093/femsec/fix006.

Ismail H, Govender NP, Singh-Moodley A, van Schalkwyk E, Shuping L, Moema I, et al. An outbreak of cutaneous abscesses caused by Panton-Valentine leukocidin-producing methicillin-susceptible *Staphylococcus aureus* among gold mine workers, South Africa November 2017 to March 2018. BMC Infect Dis. 2017. https://doi.org/10.1186/s12879-020-05352-5.

Jackson C, Stone B, Tyler H. Emerging perspectives on the natural microbiome of fresh produce vegetables. Agriculture. 2015;5(2):170–87. https://doi.org/10.3390/agriculture5020170.

Jaklevic MC. Trends identified in foodborne illnesses from leafy greens. JAMA. 2021;325(19):1932–1932. https://doi.org/10.1001/jama.2021.7165.

Janda JM, Abbott SL. The changing face of the family Enterobacteriaceae (Order: 'Enterobacterales'): New members, taxonomic issues, geographic expansion, and new diseases and disease syndromes. Clin Microbio Rev. 2021. https://doi.org/10.1128/jcm.21.1.46-76.1985.

Kgoale *et al. CABI Agriculture and Bioscience*   (2023) 4:36

Page 10 of 10

Leff JW, Fierer N, Hunter G, Bavinger J, Pearson M. Bacterial communities associated with the surfaces of fresh fruits and vegetables. Berg G, editor. PLoS One. 2013;8(3):59310. https://doi.org/10.1371/journal.pone.0059310.

Lowe M, Singh-Moodley A, Ismail H, Thomas T, Chibabhai V, Nana T, et al. Molecular characterisation of *Acinetobacter baumannii* isolates from bloodstream infections in a tertiary-level hospital in South Africa. Front Microbiol. 2022. https://doi.org/10.3389/fmicb.2022.863129.

Mäder P, Fliessbach A, Dubois D, Gunst L, Fried P, Niggli U. Soil fertility and biodiversity in organic farming. Science. 2002;296(5573):1694–7. https://doi.org/10.1126/science.1071148. (**PMID: 12040197**).

Marshall KE, Hexemer A, Seelman SL, Fatica MK, Blessington T, Hajmeer M, et al. Lessons learned from a decade of investigations of shiga toxin-producing *Escherichia coli* outbreaks linked to leafy greens, United States and Canada. Emerg Infect Dis. 2020;26(10):2319–28. https://doi.org/10.3201/eid2610.191418.

McCarthy KM, van Nierop W, Duse A, von Gottberg A, Kassel M, Perovic O, et al. Control of an outbreak of vancomycin-resistant *Enterococcus faecium* in an oncology ward in South Africa: effective use of limited resources. J Hosp Infect. 2000;44(4):294–300. https://doi.org/10.1053/jhin.1999.0696.

Mengistu AA. Endophytes: colonization, behaviour, and their role in defense mechanism. Int J Microbiol. 2020. https://doi.org/10.1155/2020/6927219.

Mulaosmanovic E, Lindblom TUT, Windstam ST, Bengtsson M, Rosberg AK, Mogren L, et al. Processing of leafy vegetables matters: damage and microbial community structure from field to bag. Food Control. 2021. https://doi.org/10.1016/j.foodcont.2021.107894.

Ntshoe GM, Thomas J, Cengimbo A, Muvhango N, Ekermans P, Fourie F. Suspected outbreak of shigellosis in Nelson Mandela Bay health district, Eastern Cape Province - South Africa, November 2012 to February 2013. Int J Infect Dis. 2014;21:243–4. https://doi.org/10.1016/j.ijid.2014.03.927.

Opperman CJ, Moodley C, Lennard K, Smith M, Ncayiyana J, Vulindlu M, et al. A citywide, clonal outbreak of *Pseudomonas aeruginosa*. Int J Infect Dis. 2022;117:74–86. https://doi.org/10.1016/j.ijid.2022.01.039.

Pillay DG, Karas AJ, Sturm AW. An outbreak of Shiga *Bacillus dysentery* in Kwa-Zulu/Natal. South Africa J Infect. 1997;34(2):107–11. https://doi.org/10.1016/S0163-4453(97)92376-8.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 2013;41(D1):590–6.

Ramalwa N, Page N, Smith A, Sekwadi P, Shonhiwa A, Ntshoe G, et al. has foodborne disease outbreak notification and investigation changed since the listeriosis outbreak in South Africa? A review of foodborne disease outbreaks reported to the national institute for communicable diseases, March 2018-August 2020. 2020; 18(2).

Ramsamy Y, Mlisana KP, Amoako DG, Allam M, Ismail A, Singh R, et al. Pathogenomic analysis of a novel extensively drug-resistant *Citrobacter freundii* isolate carrying a blandm-1 carbapenemase in South Africa. Pathogens. 2020. https://doi.org/10.3390/pathogens9020089.

Rodrigues C, da Silva ALBR, Dunn LL. Factors Impacting the Prevalence of Foodborne Pathogens in Agricultural Water Sources in the Southeastern United States. Water. 2020;12(1):51. https://doi.org/10.3390/w12010051R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Said M, van Hougenhouck-Tulleken W, Naidoo R, Mbelle N, Ismail F. Outbreak of *Ralstonia mannitolilytica* bacteraemia in patients undergoing haemodialysis at a tertiary hospital in Pretoria. South Africa Antimicrob Resist Infect Control. 2020. https://doi.org/10.1186/s13756-020-00778-7.

Setlhare G, Malebo N, Shale K, Lues R. Identification of airborne microbiota in selected areas in a health-care setting in South Africa. BMC Microbiol. 2014. https://doi.org/10.1186/1471-2180-14-100.

Sholeye AR. Bacterial diversity of South African soils. Master's dissertation. North West University. 2020.

Shonhiwa AM, Ntshoe G, Essel V, Thomas J, Mccarthy K, Lapen D. A review of foodborne diseases outbreaks reported to the outbreak response unit, national institute for communicable diseases, South Africa, 2013–2017 Using quantitative microbial risk assessment to explore waterborne pathogen infection risks associated with interventions of agricultural management practices. 2019; https://doi.org/10.1016/j.ijid.2018.11.186

Taffner J, Laggner O, Wolfgang A, Coyne D, Berg G. Exploring the Microbiota of East African indigenous leafy greens for plant growth, health, and resilience. Front Microbiol. 2020;11:2968. https://doi.org/10.3389/fmicb.2020.585690/bibtex.

Teklu DS, Negeri AA, Legese MH, Bedada TL, Woldemariam HK, Tullu KD. Extended-spectrum beta-lactamase production and multi-drug resistance among Enterobacteriaceae isolated in Addis Ababa, Ethiopia. Antimicrob Resist Infect Control. 2019;8(1):1–12. https://doi.org/10.1186/s13756-019-0488-4/figures/3.

van Nierop WH, Duse AG, Stewart RG, Bilgeri YR, Koornhof HJ. Molecular epidemiology of an outbreak of *Enterobacter cloacae* in the neonatal intensive care unit of a provincial hospital in Gauteng, South Africa. J Clin Microbiol. 1998;36(10):3085–7. https://doi.org/10.1128/jcm.36.10.3085-3087.1998.

Wang Y, Qian PY. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PLoS ONE. 2009;4(10):7401. https://doi.org/10.1371/journal.pone.0007401.

World Health Organisation (WHO). Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. 2017. https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.