



## RESEARCH ARTICLE OPEN ACCESS

# Indigenous Nigerian Vegetables Harbour Plant Species-Specific Seed Microbial Signatures for Bacteria and Fungi

Oluwatosin Ajibade<sup>1</sup> | Expedito Olimi<sup>2,3</sup> | Wisnu Adi Wicaksono<sup>2</sup> | Elijah Kolawole Oladipo<sup>1</sup> | Julius. K. Oloke<sup>1</sup> | Olubukola Monisola Oyawoye<sup>1</sup> | Tomislav Cernava<sup>2,3</sup> | Gabriele Berg<sup>2,4,5</sup>

<sup>1</sup>Department of Microbiology, Faculty of Science, Adeleke University, Osun State, Nigeria | <sup>2</sup>Institute of Environmental Biotechnology, Graz University of Technology, Graz, AT | <sup>3</sup>School of Biological Sciences, Faculty of Environmental and Life Sciences, University of Southampton, UK | <sup>4</sup>Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Potsdam, Germany | <sup>5</sup>Institute for Biochemistry and Biology, University of Potsdam, Potsdam, Germany

**Correspondence:** Expedito Olimi ([expedito.olimi@tugraz.at](mailto:expedito.olimi@tugraz.at)) | Gabriele Berg ([gabriele.berg@tugraz.at](mailto:gabriele.berg@tugraz.at))

**Received:** 14 July 2025 | **Revised:** 27 September 2025 | **Accepted:** 1 October 2025

**Funding:** The authors received funding from the Austrian Agency for International Cooperation in Education & Research (OeAD-GmbH) through the Ernst Mach-Stipendien, weltweit. This research was also funded in part by the Austrian Science Fund (FWF) under grant number 10.55776/J4753, awarded to Expedito Olimi, and by the EXCALIBUR project, with funding from the European Union's Horizon 2020 Research and Innovation Program under grant agreement No. 817946 (to Gabriele Berg). The authors also acknowledge the use of the IRIDIS 5 High Performance Computing Facility, and the associated support services at the University of Southampton, in the completion of this work.

**Keywords:** African indigenous vegetables | high-throughput sequencing | plant-microbe interactions | seed microbiome | sustainable agriculture

## ABSTRACT

**Introduction:** The seed microbiome is essential for plant health, development, and adaptation. Little is known regarding the constituents of seed microbiomes of Africa's indigenous vegetables, which enables the vertical transmission of a specific microbial heritage. Shedding light on this was the focus of the current study.

**Methodology:** Microbiome profiling using domain-specific amplicon sequencing was performed to examine the seed microbiome of five economically important vegetable crops in Nigeria. The selection included not yet studied *Amaranthus hybridus*, *Solanum macrocarpon*, *Corchorus olitorius*, *Celosia argentea*, and *Telfairia occidentalis*.

**Results:** The seeds of the five plant species harboured a diverse microbiome comprised of 2877, 765, and 56 amplicon sequence variants (ASVs), for bacterial, fungal, and archaeal communities, respectively. Plant species-specific seed microbial signatures were found for bacterial and fungal communities, which were highly abundant (i.e.,  $10^8$  to  $10^{11}$  gene copy numbers for bacterial and  $10^9$  to  $10^{11}$  for fungal community). In contrast, there was no separation for the less abundant archaea across the different plant species. The core seed microbiota was comprised by potentially plant-beneficial taxa including the archaeal genus *Methanobrevibacter* and bacterial genera like *Methylobacterium*, *Burkholderia-Caballeronia-Paraburkholderia*, *Bacillus*, *Paenibacillus*, *Stenotrophomonas*, *Pseudomonas*, *Serratia*, *Pantoea*, *Shingomonas*, and *Lactobacillus*; as well as fungal genera *Aspergillus* and *Fusarium*. Random forest classification revealed that the seed microbiome profiles can resolve plant species-specific differences.

**Conclusion:** Our study shows that indigenous African vegetables harbour diverse and abundant microbial communities which are potentially important for the sustenance of these plants, particularly in smallholder-dependent food production systems, like in Sub-Sahara Africa. The research also provides the basis for conservation of these unique plant genetic resources.

Expedito Olimi and Tomislav Cernava substantially contributed equally to this research.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Journal of Sustainable Agriculture and Environment* published by Global Initiative of Sustainable Agriculture and Environment and John Wiley & Sons Australia, Ltd.

## 1 | Introduction

The world's population continues to grow, necessitating the development of sustainable food production methods, particularly in Sub-Saharan Africa, where the population has been forecasted to increase (El Bilali et al. 2018; Kruk and Pate 2020). Agricultural systems are mainly managed by smallholder farmers in tropical countries, particularly in Sub-Saharan Africa, including Nigeria, and are characterized by low productivity due to global climate change leading to drought, pests, diseases, and floods (Ibeawuchi et al. 2015). Such consequences, however, could be partially mitigated by cultivating indigenous vegetables, which have a short life cycle and are less susceptible to irregular rainfall, making them reliable vegetable crops for food security (Taffner, Laggner, et al. 2020); as well as playing a role in addressing malnutrition in Africa.

These vegetables can be grown locally or gathered in the natural habitat (Abukutsa-Onyango 2010). Classical examples include African spinach (*Amaranthus hybridus* L.) (Akin-Idowu et al. 2017), African eggplant (*Solanum macrocarpon* L.) (Adeniyi et al. 2017), jute mallow (*Corchorus olitorius* L.), and Lagos spinach (*Celosia argentea* L.). They are consumed for their culinary and medicinal benefits, like their role as a source of protein, micronutrients, vitamins; as well as their antioxidant properties (Adeniyi et al. 2017). While the resilience of these vegetables can be attributed to genetic traits that are unique to these plants, the contribution of the plant microbiome in the enhancement of plant health cannot be underestimated (Badri and Vivanco 2009; Vandenkoornhuyse et al. 2015).

In Nigeria, malnutrition, undernutrition, and vitamin deficiencies are widespread due to poor diets associated with low economic status. These nutritional challenges can be mitigated by consuming indigenous vegetables (Icard-Vernière et al. 2015). Moreover, the cultivation and local consumption of vegetables offer an affordable, sustainable, and accessible means of improving the intake of essential nutrients (Ejoh et al. 2021). More recently it was found that edible plants are a potential source of pre/probiotics and could be essential in restoring the lost microbiome (Blaser 2017; Wicaksono et al. 2023). The sub-Saharan part of Africa is considered a hotspot for biodiversity which is resilient to climate change, and therefore serves as a potential source of novel probiotics which can enhance plant and human health. Maintaining their own batches of seed material by local farmers is a common practice that guarantees food security and sovereignty and genetic resources conservation. Despite the growing recognition of the nutritional and medicinal importance of indigenous vegetables in Nigeria, limited data exists on their microbial ecology—particularly in terms of the seed microbiome. Seed microbiomes are generally less explored than other plant compartments (Michl et al. 2023); however, for some plant species they remain entirely untapped. This knowledge gap is crucial, since understanding and harnessing plant- and seed microbiomes is essential for sustainable agriculture and crop improvement (Sulesky-Grieb et al. 2024; Garrido-Sanz and Keel 2025).

The plant microbiome: bacteria, archaea, fungi, viruses, microalgae, oomycetes, and protozoa—is critical for plant development, stress tolerance, and health (Pérez-Jaramillo et al. 2016;

Berg and Raaijmakers 2018; Sulesky-Grieb et al. 2024). Once thought to be primarily soil-derived, it is now recognized that seed endophytes are specifically acquired and maintained by plants. They play essential roles in early plant development and disease resistance (Shade et al. 2017; Berg and Raaijmakers 2018; Matsumoto et al. 2021; Sulesky-Grieb et al. 2024). Besides their role as a conduit for propagating plant genetic information, seeds are reservoirs of diverse and highly specialized microbial communities, which can highly vary between plant hosts (Simonin et al. 2022; Lobato et al. 2024). The occurrence of seed microbes is compartmentalized within seed tissues (endophytes) and on seed surfaces (epiphytes) (Barret et al. 2016; Nelson 2018; Abdelfattah et al. 2023); wherein endophytes constitute a major portion of the vertically transferred seed microbiome (Johnston-Monje et al. 2014; Abdelfattah et al. 2023), with a lasting impact across plant generations (Sulesky-Grieb et al. 2024). However, due to their cultivation-recalcitrant nature, many seed endophytes elude traditional microbiological methods, and comprehensive community-level analyses have immensely benefited from advances in next-generation sequencing technologies (Wassermann et al. 2019a; Lobato et al. 2025). While the seed microbiomes of many agriculturally important crops, including plants in natural and pristine environments have been examined (Johnston-Monje and Raizada 2011; Johnston-Monje et al. 2016; Adam et al. 2018; Wassermann et al. 2019b, 2022; Lobato et al. 2024), and while some studies already examined the microbiome of African green vegetables (Taffner, Laggner, et al. 2020), research on the seed microbiomes of most African crops remains scarce.

In this study, we examined the microbial community structure and diversity in the seeds of five indigenous vegetables. Specifically, we aimed to establish how the seed microbial communities (i.e., bacteria, fungi, and archaea) vary across plant species, and what proportion of taxa are shared and ubiquitous across them. Further, we examined if specific seed microbiota constituents can delimit the studied plants. We hypothesized that the seed microbiomes of the studied plants harbour distinct microbial communities, with key taxa potentially contributing to plant health and resilience. The overarching goal was to provide foundational insights into the microbial communities associated with the seeds of indigenous Nigerian vegetables, using an amplicon sequencing approach. These insights could be instrumental in improving crop varieties that can establish beneficial microbial interactions, thereby enhancing productivity in resource poor- and climate- constrained smallholder farming systems in Africa.

## 2 | Materials and Methods

### 2.1 | Experimental Design and Sample Materials

Seed samples ( $n = 6$ ) of each of the five indigenous African vegetables including: (i) African spinach (*Amaranthus hybridus*), (ii) African eggplant (*Solanum macrocarpon*), (iii) Jute mallow (*Corchorus olitorius* L.), (iv) Lagos spinach (*Celosia argentea*), and (v) Fluted pumpkin (*Telfairia occidentalis*) were used for experimentation. The samples were obtained from Osun State Agricultural Matters retail seed shop (Ede, Osun State) and Ede market seed shop (for seeds of Fluted pumpkin).

The six biological samples were composed of seed portions that were obtained from seed retailers and the market shop.

## 2.2 | Sample Processing

The seed microbiome samples (endophytes and epiphytes) were processed using in-house protocols at the Institute of Environmental Biotechnology, Graz University of Technology. Briefly, seeds were gently agitated on a shaker for 30 min with 0.85% NaCl buffer to facilitate the release of seed-associated microbiota, followed by crushing with a mortar and pestle using liquid nitrogen, and the addition of 4 mL of 0.85% NaCl. After thorough homogenization, the resulting microbial solution was centrifuged at 16,000 rpm for 20 min to concentrate the microbial communities into the resulting microbial pellets which were stored at  $-70^{\circ}\text{C}$  until DNA was extracted.

## 2.3 | Microbial DNA Extraction and Amplicon Library Construction

The microbial pellets obtained from the seeds were aseptically introduced into bead tubes from the FastDNA SPIN Kit for soil (MP Biomedicals, Solon, United States), and microbial DNA extracted according to the manufacturer's procedure. Briefly, samples were mechanically disrupted using the FastPrep Instrument (MP Biomedicals, Illkirch, France) for 30 s at  $6.0\text{ ms}^{-1}$ , then transferred into kit-supplied tubes for downstream implementation of microbial DNA extraction procedures. The quality of the DNA was examined using the Nanodrop 2000 (Thermo Fisher Scientific, Wilmington, DE, United States) and stored at  $-20^{\circ}\text{C}$  for domain-specific library preparation.

The seed bacterial microbiome was profiled by high-throughput amplicon sequencing targeting the V4 region of the 16S rRNA gene, using the primer pair (515 F/806 R) (Caporaso et al. 2012), carrying sample-specific tags. The PCR was performed in combination with peptide nucleic acid (PNA) blockers (pPNA: 5'-GGCTCAACCCTGGACAG-3'; mPNA: 5'-GGCAAGTGTTC TTCGA-3') to reduce the co-amplification of plasmid DNA (Lundberg et al. 2013). The reaction mixture for the PCR (30  $\mu\text{L}$ ) contained 1 $\times$  Taq-&GO (MP Biomedicals, Solon, OH, USA), 0.2  $\mu\text{M}$  of each primer, 0.75  $\mu\text{M}$  each of pPNA and mPNA (PNA Bio, Thousand Oaks, CA, USA), and 2  $\mu\text{L}$  of template DNA ( $96^{\circ}\text{C}$ , 5 min; 30 cycles of  $96^{\circ}\text{C}$ , 1 min;  $78^{\circ}\text{C}$ , 5 s for PNA annealing;  $54^{\circ}\text{C}$ , 1 min;  $74^{\circ}\text{C}$ , 1 min; and elongation at  $74^{\circ}\text{C}$ , 10 min).

The fungal community was analysed by amplifying ITS fragments using the primer pair ITS1f (5'-CTTGGTCATTTA GAGGAAGTAA-3') and ITS2r (5'-GCTGCGTTCCTCATCGA TGC-3') (White et al. 1990). The first PCR step was performed using ITS primers with attached linkers (to attach barcodes). The first PCR consisted of an initial denaturation s ( $95^{\circ}\text{C}$ , 5 min), followed by 30 cycles of denaturation ( $95^{\circ}\text{C}$ , 30 s), primer annealing ( $58^{\circ}\text{C}$ , 35 s), and elongation ( $72^{\circ}\text{C}$ , 40 s). The reaction mixture was composed of 2  $\mu\text{L}$  (5 $\times$  Taq & Go), 1.2  $\mu\text{L}$  (25 mM  $\text{MgCl}_2$ ), 0.1  $\mu\text{L}$  of each 10  $\mu\text{M}$  (ITS1/ITS2 primers), and 5.6  $\mu\text{L}$  PCR grade water, and 1  $\mu\text{L}$  template each in 10  $\mu\text{L}$  reaction.

The second PCR involved 15 cycles of denaturation ( $95^{\circ}\text{C}$ , 30 s), primer annealing ( $53^{\circ}\text{C}$ , 30 s), and elongation ( $72^{\circ}\text{C}$ , 30 s). It was performed using 2  $\mu\text{L}$  of the first PCR product in 30  $\mu\text{L}$  reaction mixture, composed of 6  $\mu\text{L}$  (5 $\times$  Taq & GO), 1.2  $\mu\text{L}$  each of 10  $\mu\text{M}$  (F/R barcode primers), and 19.6  $\mu\text{L}$  PCR grade water.

For archaeal community, a nested PCR using the archaeal-specific primer pair 344 f/915r and Phusion polymerase (for the first-step) was performed. This involved an initial denaturation at  $95^{\circ}\text{C}$  for 2 min; 10 cycles of  $96^{\circ}\text{C}$  for 30 s,  $60^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 1 min. This was followed by 15 cycles of denaturation at  $94^{\circ}\text{C}$  for 25 s,  $60^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 1 min; and a final elongation at  $72^{\circ}\text{C}$  for 10 min. The second round of PCR was performed using modified primers (349 f/519r) (Klindworth et al. 2013), each with a 10-bp primer-pad (forward: TATGG-TAATT; reverse: AGTCAGCCAG) and linker sequences (forward: GT; reverse: GG), following the Earth Microbiome Project (Walters et al. 2015). Thermal cycler conditions were: initial denaturation at  $95^{\circ}\text{C}$  for 5 min; 25 cycles of  $95^{\circ}\text{C}$  for 40 s,  $66^{\circ}\text{C}$  for 2 min, and  $72^{\circ}\text{C}$  for 1 min; and a final elongation at  $72^{\circ}\text{C}$  for 10 min. In the third PCR, sample-specific barcodes were added using cycler conditions:  $95^{\circ}\text{C}$  for 2 min; 10 cycles of  $95^{\circ}\text{C}$  for 30 s,  $56^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 30 s; final elongation at  $72^{\circ}\text{C}$  for 10 min. Triplicate reactions were performed per sample. PCR products were verified by electrophoresis, purified (Wizard SV Gel/PCR purification kit, Promega), pooled equimolarly, and sequenced on an Illumina MiSeq platform (2 $\times$  300 bp sequencing by Eurofins Genomics, Germany).

## 2.4 | Quantitative Real-Time PCR

Microbial abundance across the five vegetable species was quantified using quantitative PCR with primer pairs as follows: For bacterial quantification, the Unibac-II-515f/806r primer pair (10  $\mu\text{M}$ ; (Köberl et al. 2011)); for the archaea, 344 aF and 517uR (5  $\mu\text{M}$ ; Klindworth et al. 2013); and for fungal, ITS1f/ITS2r (White et al. 1990).

Quantitative PCR was performed in 10  $\mu\text{L}$  reactions containing 0.5  $\mu\text{L}$  each of forward and reverse primers, 5  $\mu\text{L}$  of KAPA SYBR Green master mix, 3  $\mu\text{L}$  PCR-grade water, and 1 of 1:10 diluted template. Amplification was carried out on a Rotor-Gene 6000 real-time analyser (Corbett Research, Sydney, Australia) using domain-specific thermal profiles: bacterial— $95^{\circ}\text{C}$  for 5 min; 40 cycles of  $95^{\circ}\text{C}$  for 20 s,  $54^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 30 s; with melt curve analysis ( $72^{\circ}\text{C}$ – $96^{\circ}\text{C}$ ); archaeal— $95^{\circ}\text{C}$  for 5 min; 40 cycles of  $95^{\circ}\text{C}$  for 15 s,  $60^{\circ}\text{C}$  for 30 s,  $72^{\circ}\text{C}$  for 30 s; with melt curve analysis; fungal— $95^{\circ}\text{C}$  for 5 min; 40 cycles of  $95^{\circ}\text{C}$  for 30 s,  $58^{\circ}\text{C}$  for 35 s,  $72^{\circ}\text{C}$  for 40 s; with melt curve analysis. Triplicate reactions were run per sample, including domain-specific standards for estimating the microbial quantities, as well as a negative control. As a standard, 16S rRNA gene standards from *Bacillus* sp. and *Haloferax denitrificans* were implemented for bacterial and archaea community, respectively; as well as *Penicillium* sp. for fungal ITS regions (Taffner, Bergna, et al. 2020a; Olimi et al. 2022). Standard regression curves from serial dilutions were used to calibrate qPCR assays. Gene copy numbers were interpolated and normalised to yield microbial abundance per gram of sample.

## 2.5 | Bioinformatic Data Processing

Raw amplicon sequencing dataset were processed using QIIME2 (v2023.9.1) environment (<https://qiime2.org>; (Bolyen et al. 2019)). Briefly, paired-end reads were demultiplexed and primers removed using cutadapt (Martin 2011). Demultiplexed sequence reads were quality checked using fastQC, followed by MultiQC-based summarization (Ewels et al. 2016). The datasets were then quality filtered, trimmed, denoised, merged, and chimeras removed using DADA2 v1.26.0 (Callahan et al. 2016), resulting into amplicon sequence variants (ASVs) table. The DADA2 pipeline for ASV-level resolution was run within the QIIME2 platform. For the bacterial and archaeal communities, representative sequences were assigned taxonomically by aligning against the SILVA132 reference database (Quast et al. 2013; Yilmaz et al. 2014), using VSEARCH algorithm (Rognes et al. 2016), while fungal taxonomy was assigned using the UNITE database v7.2 (release:2017-10-10) (Abarenkov et al. 2024); all implemented using the `q2-feature-classifier` command (Bokulich et al. 2018). An optimized, balanced configuration was applied, using a confidence threshold of 0.7 ( $c=0.7$ ) to ensure reliable classification as earlier described (Bokulich et al. 2018).

## 2.6 | Statistical Analysis and Modelling

Downstream analyses were performed using R v4.2.2 (R Core Team 2020) in RStudio v1.1.42, using the packages `phyloseq`, `vegan`, `tidyverse`, `ggplot2`; and other dependencies are specified in the publicly available script (<https://github.com/kaboyo/African-seed-microbiome>). The feature table, taxonomic information, and experimental metadata were imported into RStudio using `phyloseq` (McMurdie and Holmes 2013) and `VEGAN` v2.5.7 (Oksanen et al. 2020) packages. Filtering to exclude non-target taxa, like unassigned sequences, chloroplasts, and mitochondria was performed to retain the bacterial-, archaeal-, and fungal-associated sequence reads. Further filtering using a threshold prevalence of 0.25% of the total samples was performed to minimize community inflation caused by low abundant ASVs, as earlier described (Reitmeier et al. 2021). Sample reads were normalized by rarefying (using `rarefy_even_depth` function) to the minimum sampling depths of 10000, 5500, and 500 reads per sample for bacteria, fungal, and archaeal communities, respectively (Figure 1A,E,I), for the subsequent alpha diversity analyses; followed by `css` transformation of the non-rarefied `PHYLOSEQ`-formatted microbiome datasets for beta diversity analyses.

Alpha diversity analysis involved computing various indices, such as richness (observed ASVs) and Shannon index. Non-parametric (ranked-based) Kruskal-Wallis tests were performed to assess the statistical differences in alpha diversity indices across plant species, followed by Dunn's post hoc tests (i.e., `function:dunnTest`(diversity indices, variable, method = "bonferroni"), coupled with Bonferroni's correction of p-values ( $p < 0.05$ ). The letters representing statistical significances were obtained using the functions in the `rcompanion` package. Spearman correlation was performed to relate the bacterial and fungal community across the plant species.

Microbial (bacterial, fungal, and archaeal) community structure was examined by computing Bray-Curtis dissimilarity matrices using

`css`-transformed microbiome datasets, followed by Principal Coordinates Analyses (PCoA) for community structure visualization. The Bray-Curtis dissimilarity matrices were further subjected to permutational multivariate analysis (PERMANOVA, 999 permutations), using `model adonis2(Bray-Curtis_dist. ~ plant species, data=metadata, permutations=999)`, that examines the effects of plant species on the seed bacterial composition. To assess microbiome community compositional spread, beta dispersion test was performed on Bray-Curtis distances, using the function `vegan::betadisper(dist, phyloseq::sample_data(pseq)$plant species)`. Taxonomic community composition across plant species was represented using stacked bar plots, representing the top 500 amplicon sequence variants (ASVs), using the `ggplot2` package. The major taxa contributing towards community separation were obtained by linear discriminant analysis effect size (LEfSe) (Segata et al. 2011).

The core microbiome entailed determining the consistent members of the seed microbiome across samples, and was computed for 25% to 100% of the samples (as prevalence), and the minimum detection ranging between 0.00001% to 10%. The core microbiome results were represented using abundance-prevalence curves for precise representation of the core taxa, including the representation of unique and shared microbiota using UpSet plots. Briefly, abundance-prevalence curves were generated using the `microbiome` package for manipulating `phyloseq` objects, and calculating community relative abundance and prevalence using functions in `phyloseq` package. The `upset_pq(pyseq)` function was applied `phyloseq`-formatted datasets representing the top100 taxa for UpSet plot visualisation.

RandomForest (RF) classification (i.e., `set.seed: 57938`), `ntrees = 500`, and `cores = 3`), with unrestricted growth was implemented on microbiome datasets and predictive microbial features across plant species presented at genus level, using `RandomForest` package (v.4.6.14) in R with default parameters (Breiman 2001; Liu et al. 2012). The classifier model was fit on the microbiome data set with plant species as a variable, to identify the features that are important delimiters of plant species. Top 20 features, ordered according to mean decrease in accuracy values (so-called MeanDecreaseGini) were presented. The RF classification was applied using filtered datasets with ASVs minimum average relative abundance 0.05.

The seed microbial community functional potential was examined using a taxonomy-based functional prediction approach which was implemented using `FAPROTAX` v1.2.10 for prokaryotic functional predictions (Louca et al. 2016; Stilianos et al. 2017), as well as `FUNGuild` for fungal community functional potential predictions (Nguyen et al. 2016; Pölme et al. 2020). The functional counts data were normalised using the total read depth followed by heatmap visualisation.

## 3 | Results

### 3.1 | Bacterial and Fungal Community Diversity Varied Across Plant Species

After quality-filtering and removing chimeric sequences, followed by the subsequent removal of reads that were either

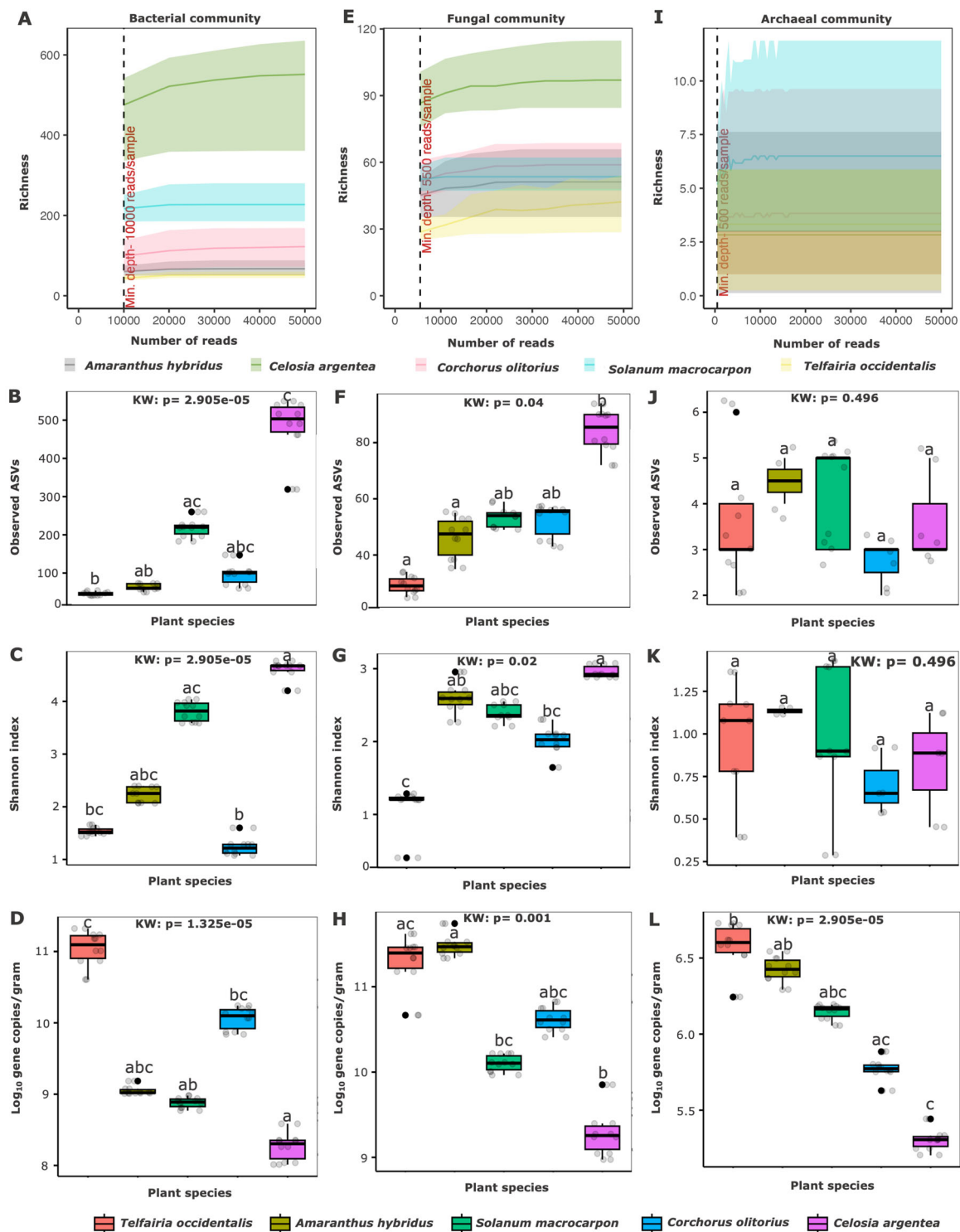


**FIGURE 1** | The seed microbial diversity and abundance across plant species of African vegetables. Panel A shows bacterial richness as a function of sequencing depth, seed bacterial richness (B), bacterial Shannon index (C), and bacterial abundance ( $\log_{10}$  gene copies) across plant species (D). Panels (E-H) and (I-L) represent the same for the fungal and archaeal community, respectively. Global significance differences across plant species, determined using Kruskal-Wallis (KW) statistical tests are shown on top of boxplots. Boxplots represent sample distribution (i.e., median and interquartile ranges) for each of the plant species. The transparent jitters represent samples per group category. The colours on the rarefaction curves represent the different plant species.

chloroplast and mitochondrial associated, the amplicon sequencing data set for the bacterial community yielded 2877 bacterial Amplicon sequence variants (ASVs). The bacterial ASVs were associated to 1,018,244 bacterial sequence reads, with an average of 33,942 reads per sample. For fungal community, the

quality-filtered data set contained 765 fungal ASVs that were obtained from 918,570 sequence reads, with the mean of 30,619 reads per sample. The Archaeal community was comprised by 56 ASVs which were generated from 51,646 reads with mean read composition of 5224 reads per sample (Figure 1A,E,I).

The seed bacterial and fungal community diversity varied across plant species with the observed richness ranging between 20 and 610 bacterial ASVs, 20 to 120 fungal ASVs, and 2 to 8 archaeal ASVs (Figure 1B,F,J). The Shannon indices ( $H'$ ) ranged from 1 to 5 and 1 to 3 for bacterial and fungal communities, respectively (Figure 1C,G). The archaeal diversity was generally similar across plant species, with richness ranging between 2 and 8 ASVs, and Shannon diversity estimated between 0.25 and 2 (Figure 1J,K). Significant differences in bacterial alpha diversity were found across plant species (Kruskal–Wallis, KW: richness;  $p = 2.9 \times 10^{-5}$ , Shannon index;  $p = 2.9 \times 10^{-05}$ ); wherein *Celosia argentea* and *Solanum macrocarpon* were significantly higher in bacterial richness as compared to *Amaranthus hybridus* and *Telfairia occidentalis* (Figure 1B); as well as a significantly higher Shannon index when compared to *Corchorus olitorius* and *Telfairia occidentalis*. Further, significant differences in alpha diversity (KW: richness;  $p = 0.04$ , Shannon index;  $p = 0.02$ ) were shown for the fungal community. For the fungal community, significant differences in richness were evidenced between *Celosia argentea* with *Amaranthus hybridus* and *Telfairia occidentalis*; as well as significant variation in Shannon diversity between *Corchorus olitorius* and *Telfairia occidentalis* (Figure 1F,G); while no significant differences across plant species were seen for the archaeal community (Figure 1J,K).

Microbiome correlation was performed to deduce microbiota relationship between the fungal and bacterial community across the different plant species. The seed microbiome correlations were species-specific, wherein negative correlation in richness between bacterial and fungal community were observed for *Amaranthus hybridus*, and *Solanum macrocarpon*, and positive correlations for the other plant species. For Shannon diversity, positive correlations were observed for all plant species, but *Corchorus olitorius* (Supporting Information S1: Figure S1A).

Seed microbial abundance significantly varied across plant species. Gene copy numbers of bacterial, archaeal, and fungal marker genes per gram of seed were ranged between  $10^8$  to  $10^{11}$  for bacterial community,  $10^9$  to  $10^{11}$  (fungal community), and  $10^4$  to  $10^7$  (archaeal community), respectively (Figure 1D,H,L). Overall, there was significant variations in bacterial (KW:  $p = 1.3 \times 10^{-05}$ ), fungal (KW:  $p = 0.001$ ), and archaeal abundance (KW:  $p = 2.9 \times 10^{-05}$ ) across plant species. The bacterial abundance was significantly high in *Telfairia occidentalis* as compared to the rest of the plant species, except in *Corchorus olitorius*. Significant difference in bacterial abundance was also observed between *Corchorus olitorius* and *Celosia argentea* (Figure 1D). The fungal abundance was significantly high for *Amaranthus hybridus* and *Telfairia occidentalis* as compared to *Celosia argentea*, but not significantly different from *Solanum macrocarpon* and *Corchorus olitorius* (Figure 1H). For archaeal community abundance, *Telfairia occidentalis* showed significantly high abundance in comparison to *Solanum macrocarpon*, *Corchorus olitorius* and *Celosia argentea*; but not significantly different from *Amaranthus hybridus* (Figure 1L).

The seed microbial community structure was distinct across plant species. Principal Coordinates Analysis (PcoA) demonstrated that microbial assemblages were primarily differentiated

along plant species; in which plant species strongly influenced the bacterial (Permanova:  $R^2 = 0.703$ ,  $p = 0.001$ ; dispersion: Kruskal–Wallis,  $p = 0.07$ ), and fungal community structures (Permanova:  $R^2 = 0.625$ ,  $p = 0.001$ ; dispersion: Kruskal–Wallis,  $p = 0.01$ ). Beta dispersion analyses showed that the seed bacterial community across plant species only differ in composition; while for fungal community, variations were evident both in community composition and spread (Figure 2C,D). In contrast, the archaeal community showed unclear plant species-dependent clustering; albeit significant differentiation in community composition ( $R^2 = 0.234$ ;  $p = 0.012$ ), no significant differences in community spread (Dispersion: Kruskal–Wallis,  $p = 0.111$ ; Figure 2E). These differences were further reflected in microbiota taxonomic composition, where the bacterial taxonomic composition revealed distinct species-specific patterns across the five African plant species.

The bacterial community was composed of 620 bacterial genera distributed across 53 classes and 27 phyla. The predominant phyla were Proteobacteria (79.9%), Firmicutes (11%), Actinobacteria (4.9%), and Bacteroidetes (3.5%). The dominant classes were Gammaproteobacteria (74.9%), Bacilli (10.6%), Actinobacteria (4.9%), Alphaproteobacteria (5%), Bacteroidia (3.5%), and Clostridia (0.3%), respectively (Figure 2A). Plants species including *Corchorus olitorius* and *Telfairia occidentalis* showed a high presence of Gammaproteobacteria (~80%–85%), while *Celosia argentea* exhibited balanced composition of Actinobacteria, Alphaproteobacteria, Bacilli, Bacteroidia, and Gammaproteobacteria. Major bacterial families were Enterobacteriaceae (52.6%), Burkholderiaceae (10.3%), Pseudomonadaceae (6.9%), Bacillaceae (6.7%), Xanthomonadaceae (3.4%), as well as Paenibacillaceae, Rhizobiaceae, Sphingomonadaceae, Streptomycetaceae, and Weeksellaceae.

The fungal community was composed of 133 genera, 18 Classes, and 9 Phyla. The dominant phyla included Ascomycota and Basidiomycota. Classes Eurotiomycetes (63%), Dothideomycetes (19%), Sordariomycetes (14%), and Saccharomycetes (1.4%). The dominant fungal families belonged to Aspergillaceae (61.9%), Corynesporascaceae (12.6%), Nectriaceae (10.1%), Botryosphaeriaceae (4.1%), Trichomonaceae, Diaporthaceae, Pleosporaceae, and Mycosphaerellaceae. The archaeal community was comprised by 20 genera, 5 classes, all distributed among 4 phyla. Contrasting pattern in Eurotiomycetes composition was observed across plant species, with a high presence in *Celosia argentea* and *T. occidentalis* (~80%–90%), while *Amaranthus hybridus* generally showed higher fungal composition (Figure 2B). Dominant seed archaeal classes were Halobacteria, Nitrososphaeria, Bathyarchaeia, and Methanobacteria. The dominant families included Heloferaceae, Holomicrobiaceae, Halococcaceae, and Nitrospiraceae (Supporting Information S1: Figure S1B). The highest bacterial abundance was observed for *Telfairia occidentalis* and the least for *Celosia argentea*.

Assessment of domain-specific shared microbiota across plants species revealed that *Corchorus olitorius* and *Solanum macrocarpon* overlapped in microbiome composition; while *Telfairia occidentalis* was composed of distinct bacterial members when compared to other plant species (Figure 3A,B). The bacterial genera including *Pantoea*, *Bacillus*, *Rhizobium*, *Pseudomonas*, *Bordetella*, and *Enterobacter* were ubiquitous across the seeds for different plant species (Figure 3B).

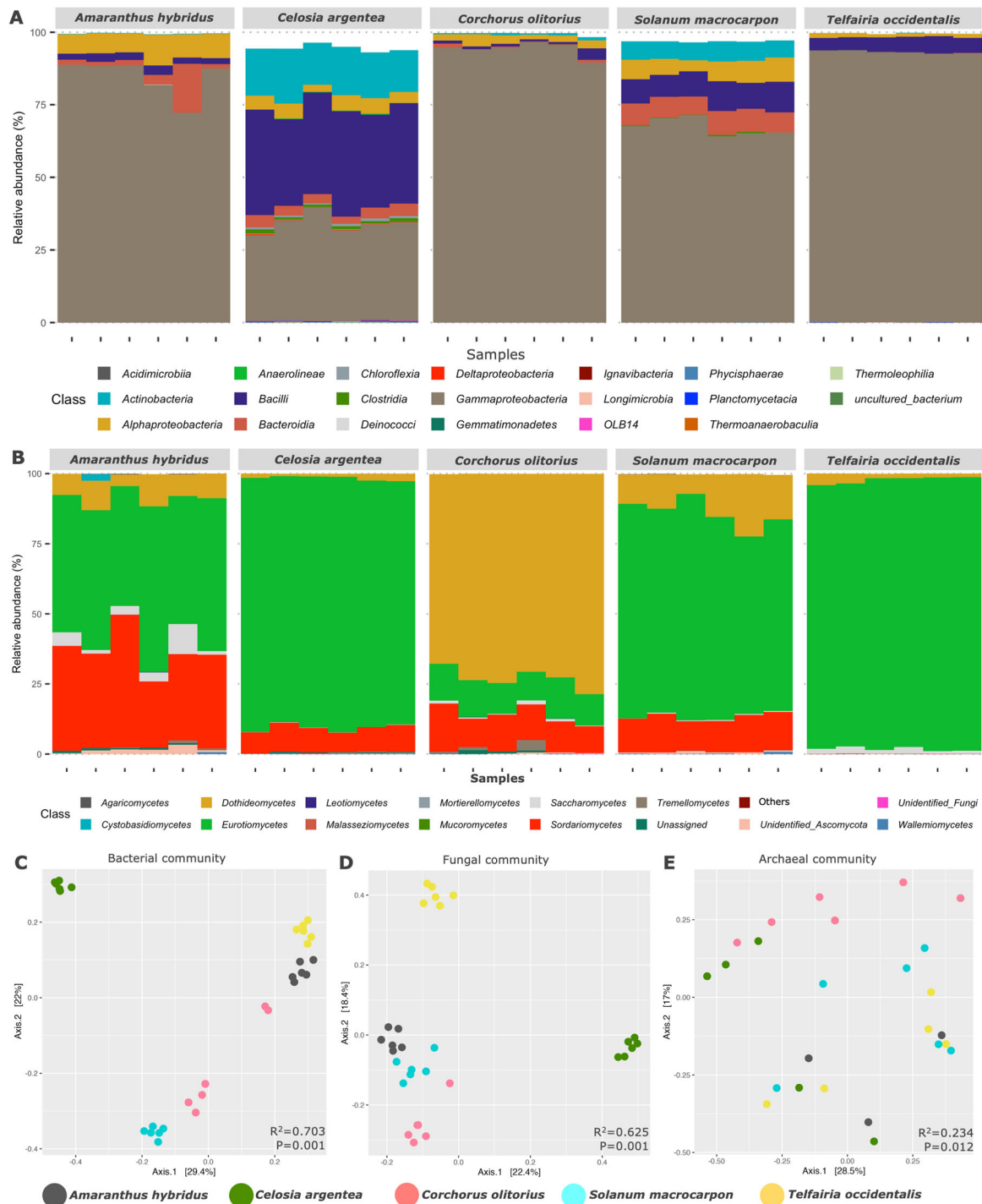


**FIGURE 2** | Seed microbial community structure varies between plant species. (A and B) Stacked barplot representation of bacterial and fungal community composition of top 500 ASVs across plant species. (C, D and E) are Principal Coordinates Analysis (PCoA) representations of the seed microbiome structure across plant species for bacterial, fungal, and archaeal communities, respectively.

Moreover, bacteria including *Methylobacterium*, *Enterococcus*, *Aquabacterium*, unidentified *Burkholderiaceae* lineage, *Aeromonas*, *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium* complex, *Paenibacillus*, and *Variovorax* also characterized the core seed microbiota (i.e., prevalence  $\approx 70\%$ – $90\%$ ) (Supporting Information S1: Figure S2A).

The fungal community was dominated by orders such as *Hypocreales*, *Capnodiales*, and *Eurotiales* which were observed across plant species (Figure 3C). Specifically, fungal genera

including *Aspergillus*, *Xeromyces*, and *Fusarium* were ubiquitous across the seed samples of the five plant species; while unidentified *Botryosphaeriaceae* were host-specific and only observed in *Telfairia occidentalis* (Figure 3D). A narrow core fungal microbiome mainly composed of *Aspergillus* ( $\approx 75\%$ ) and *Fusarium* ( $\approx 60\%$ ) was observed (Supporting Information S1: Figure S2B). The archaeal community mainly harboured taxa in order Halobacteriales; while members classified under Methanobacteriales and Nitrososphaerales were host-specific (Figure 3E–F). Moreover, some of the taxa that comprised the

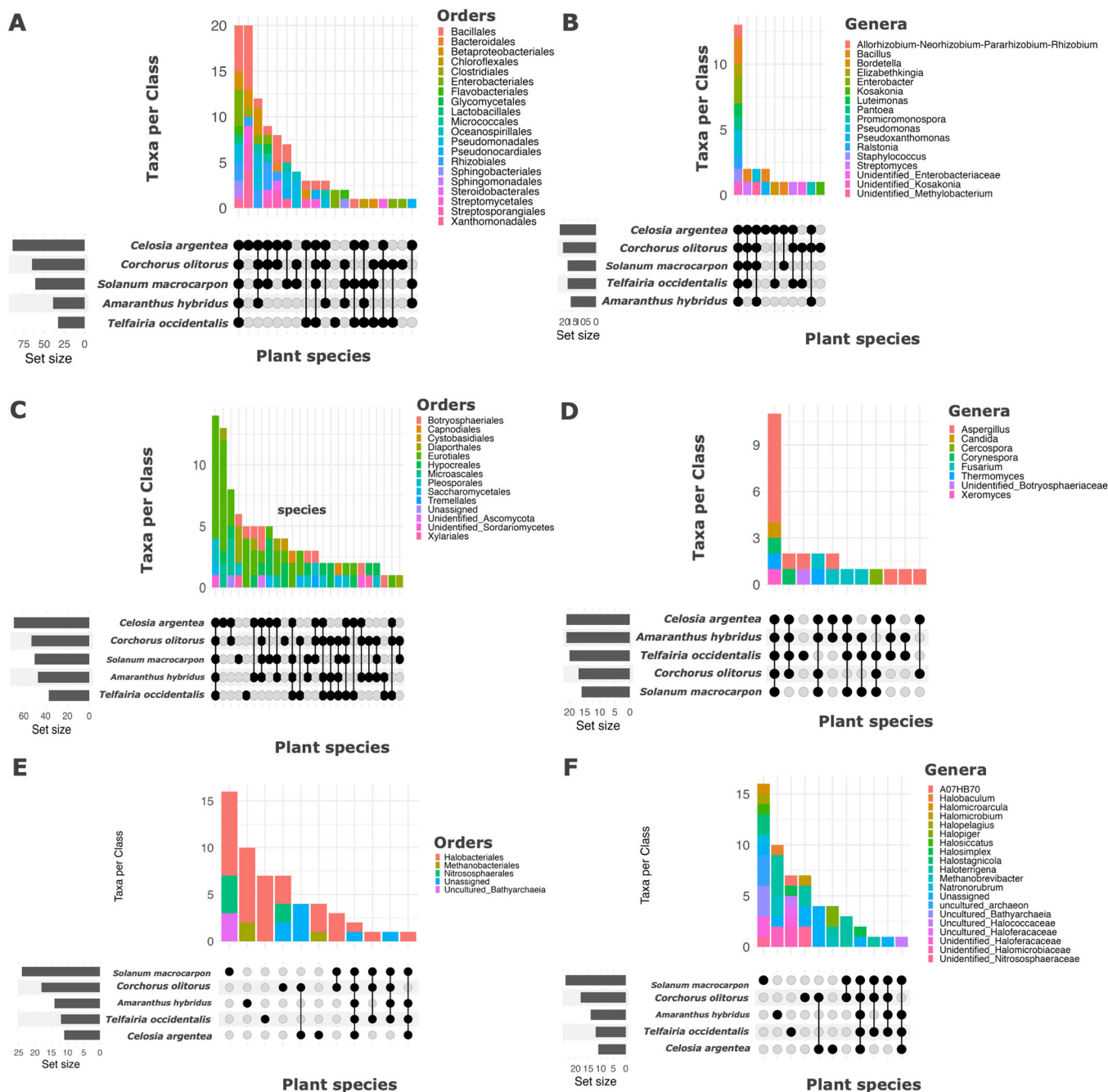


**FIGURE 3** | The shared seed microbiome across different plant species. (A and B) show UpSet plots representing the shared bacterial microbiota at order and genus levels, respectively; (C–D) and (E–F) show the same for the fungal and archaeal community, respectively.

core seed microbiome were also differentially enriched across plant species, both for bacterial and fungal communities, respectively (Supporting Information S1: Figure S2C and Figure 2D). No differentially enriched members of the archaeal community were observed.

### 3.2 | Identification of Seed Microbiome Biomarkers to Differentiate the Five Indigenous African Vegetables

To test if the seed microbiome can differentiate the different plant species, a pre-trained random forest classifier was

deployed to rank seed microbiota in the order of importance in delineating plant species. Specific plant-associated bacterial genera, like *Bacillus*, *Pseudomonas*, *Acinetobacter*, *Serratia*, *Acinetobacter*, and *Enterobacter* were important delimiters of plant species (Figure 4).

Meanwhile, the presence of potential animal-associated bacteria (including humans), like *Staphylococcus* in the seed microbiome data set were observed, representing potential influence owing to postharvest seed management. The fungal seed microbiota such as ASVs associated to the genera *Aspergillus*, *Eremothecium*, *Xeromyces*, *Xerochrysium*, and *Fusarium* were found to delimit the different plant species. For the archaeal

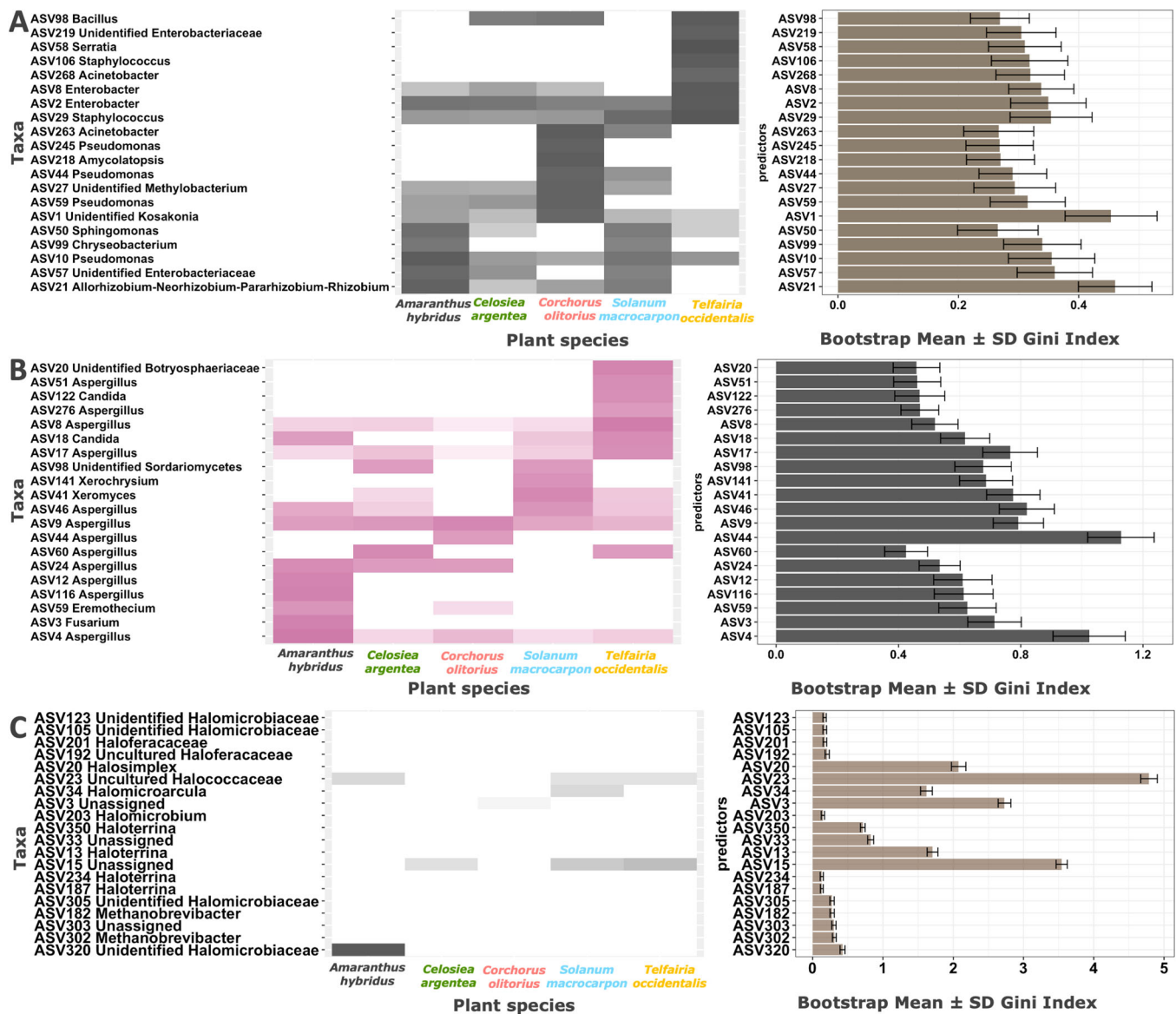


**FIGURE 4** | Seed microbiome profiles can be used as biomarkers to resolve variations between host species. (A–C) Heatmap-barplot representation of top 20 most important seed microbiome features (ASV and genus level) that could delimit the different plant species. (A, B, and C) show the bacterial, fungal, and archaeal communities, respectively. The important microbial features were selected based on MeanDecreaseGini (Gini index) values. DecreaseGini estimates the feature importance by quantifying how much the ASVs contribute to reducing “Gini impurity”, which is a measure of how mixed the classes (ASVs) are at a node.

community, one genus (uncultured Halococcaceae) was a strong delimiter of the different plant species.

### 3.3 | The Seed Microbiota Potentially Contributes Crucial Roles in Plant Health

To examine the potential functions that can be contributed by the seed microbiome, the taxonomy-based functional assignment tools FAPROTAX v1.2.10 (Louca et al. 2016; Stilianos et al. 2017) and FUNGuild (Nguyen et al. 2016; Pölme et al. 2020) were used for prokaryotic and fungal functional annotation.

The dominant bacterial functions were linked chemoheterotrophy which was observed among all plant species, with the

exception of *Telfairia occidentalis* (Figure 5A). Some of the bacteria that were linked to chemoheterotrophy include *Stenotrophomonas*, *Rhodanobacter*, *Pseudomonas*, *Acinetobacter*, *Serratia*, *Pantoea*, *Escherichia-Shigella*, *Enterobacter*, *Methylobacillus*, *Shingomonas*, *Bradyrhizobium*, *Ensifer*, *Mesorhizobium*, *Methylobacterium*, *Streptomyces*, *Paenibacillus*, and *Lactobacillus* (Supporting Information S2: Supplementary Data S1). The archaeal community functions were dominated by methanogenesis related roles of archaea, like methanogenesis by CO<sub>2</sub> reduction with H<sub>2</sub>, and hydrogenotrophic methanogenesis (Figure 5A). The archaeal functions were mainly linked to *Methanobrevibacter*, *Haloterrigena*, and *Halobaculum* (Supporting Information S2: Supplementary Data S2). The fungal community was dominated by a high presence of functional guild categories including endophytes, animal pathogens, plant saprotrophs, plant pathogens and wood saprotrophs; these

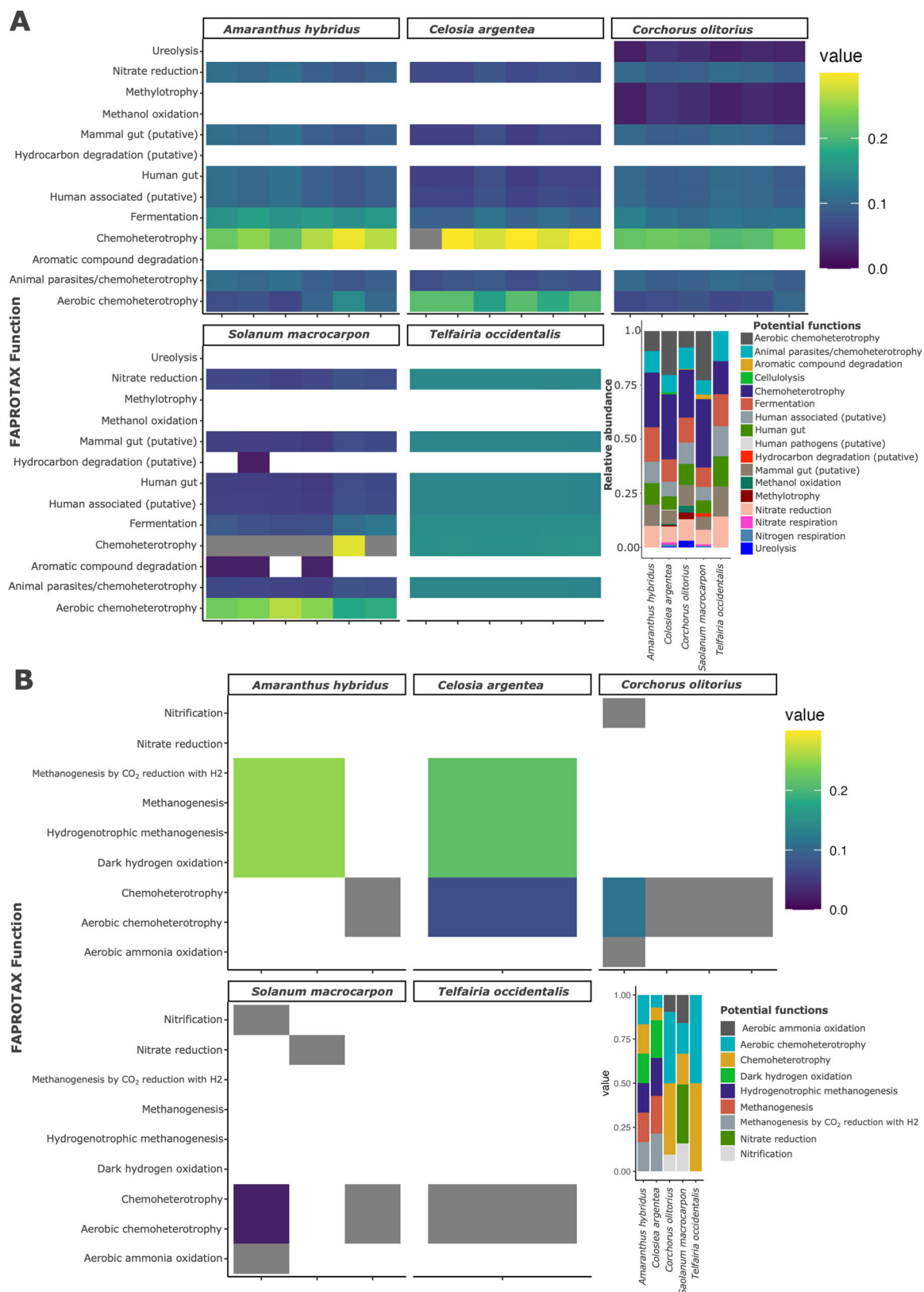


**FIGURE 5** | Taxonomy-based prediction of putative functions of the seed bacterial and archaeal communities across plant species, using the FAPROTAX database. (A and B) are heatmaps and stacked-bar compositional (inset) plots representing the proportions of bacterial and archaeal functions, respectively. The colour scale represents functional composition normalized to total read depths.

functions were particularly high in seed samples of *Celosia argentea*.

## 4 | Discussion

In the present study, the microbial community of five indigenous African vegetable seeds was explored to provide insight into their respective microbiomes. Their microbiomes likely include beneficial constituents on which the resilience of these vegetables is owed, and that have to be explored in more detail in the future. The seed microbiomes were plant-species specific, and were characterised by a high diversity and abundance of bacteria and fungi; albeit with a less structured archaeal seed community, which was also less diverse.

The host-specific variation of the seed microbiota composition in the present study suggests that host genetics may influence the seed bacterial and fungal community assembly, potentially through seed chemistry, physical structure, or plant immune responses that selectively favour or exclude certain microbiota. Put into context, the examined plant species not only varied in genetic make, but also in morphology and size. These findings aligns with previous studies that also assessed the influence of host genetics on seed microbiome composition of cultivated (Wolfgang et al. 2020; Wentzien et al. 2024; Lobato et al. 2024; Wassermann et al. 2022), and naturally occurring plant populations (Wassermann et al. 2019b).

The seed microbial (bacteria and fungi) diversity of the examined vegetables was high (i.e., up to 133 and 620 fungal and bacterial genera), reflecting that seed as the basic unit of a holobionts represent the essential fraction of the microbiota needed for the plants' early establishment. Specifically for the African indigenous vegetables, this microbiota may contribute crucial roles in enabling their resilience. More so, the seed environment is characterized by specific niches that support various bacteria for the continuity of complex plant-microbe interactions later after germination. Interestingly, the observed high bacterial and fungal seed diversity of African vegetables comparably ranks with some of the previously studied highly diverse plant species (Simonin et al. 2022). Additionally, there were major positive correlations between the fungal and bacterial seed diversity. Interestingly, plant species which were associated with high microbial diversity had low microbial abundance. Therefore, the seed microbial abundance might inversely relate to diversity.

Genotype-specificity of seed microbiome composition as observed in the present study was long observed by various authors in the past (Simonin et al. 2022; Lobato et al. 2024). The bacterial community across plant species included several prevalent bacterial genera, like *Pantoea*, *Methylobacterium*, *Sphingomonas*, *Paenibacillus*, *Variovorax*, *Bacillus*, *Rhizobium*, *Pseudomonas*, *Bordetella*, and *Enterobacter*. These bacteria are ubiquitous members of the seed microbiome as previously shown (Wassermann et al. 2019b; Matsumoto et al. 2021; Simonin et al. 2022; Abdelfattah et al. 2023; Sulesky-Grieb et al. 2024; Lobato et al. 2024). These bacteria are also valuable members of the plant microbiome which facilitate plant resilience against various stress constraints (Jiang et al. 2019;

Matsumoto et al. 2021; Wassermann et al. 2022, 2023). Analysis of the fungal community revealed the presence of ubiquitous genera such as *Aspergillus*, *Fusarium*, and *Penicillium* which are commonly found associated with seeds. Some of these genera can be pathogenic, but others like *Aspergillus* were recently shown to be beneficial for plant health (Fan et al. 2024). We also observed the presence of certain bacterial and fungal genera, like *Staphylococcus* and *Candida* which are majorly associated with animals, including humans. This observation could be a result of seed treatments or a specification of seeds grown in this region of the world, or likely attributed to postharvest seed management.

The archaeal community was lowest in alpha diversity, with values ranging between 25 and 56 ASVs (species). This low diversity might suggest that archaeal communities are not typical components of seed microbiomes, but unspecifically taken up during plant growth. This is in contrast to bacterial and fungal communities that were shown to be well structured in various crop plants (Taffner et al. 2020a; Michl et al. 2023). Additionally, the low richness of archaea may reflect their highly specialized metabolic requirements and restricted ecological niches in plant-associated environments such as seed. Therefore, specific environmental conditions can favour the thriving of specific archaeal taxa.

Our results cannot address the exact role of the seed microbiota; however, by using taxonomy-based functional predictions, we provide insights into potential roles of the seed microbiota of the five African vegetable plants. The potential functions were mainly categorised into guilds, like heterotrophy, methanogenesis, and endophytes, for bacterial, fungal, and archaeal communities, respectively. While seed bacterial functions have been fairly well studied (Torres-Cortés et al. 2018; Rodríguez et al. 2020; Sulesky-Grieb et al. 2024; Garrido-Sanz and Keel 2025), we are yet to understand the importance of seed-associated archaea. In this study, seed archaeal community functions were mainly dominated by methanogenesis, nitrification, and nitrate reduction; with *Methanobrevibacter*, *Halo-terrigena*, and *Halobaculum* as potential contributors to methanogenesis. Previous research suggest that archaea might be important in enabling microbial root colonization and the emission of volatiles for activation of plant systemic immunity and defence (Moissl-Eichinger et al. 2018; Taffner et al. 2018; Jung et al. 2020).

## 5 | Conclusion

We found that the seed microbiome varied across host genotype, especially for bacterial and fungal community, more than for archaeal community which was very unspecific. Given the central role of seeds in ensuring plant generational continuity and the transmission of seed microbiomes (Abdelfattah et al. 2023), understanding the factors that shape the seed microbiota is crucial in designing sustainable crop improvement programs that are environmentally safe (Simonin et al. 2022). Moreover, there is increasing evidence regarding the role of specific plant genes, so-called *M* genes, which are central for the recruitment and maintenance of plant-beneficial microbes and therefore could be key targets for sustainable crop breeding

frameworks (Zhan and Wang 2024; Cernava 2024). Overall, this study adds to the growing evidence that African indigenous vegetables harbour diverse microbiota, offering both bioresource opportunities for the management of various crop production stresses. However, further research is needed to elucidate the functional roles of these diverse microbial communities and their potential impacts on the growth, yield, and resilience of these important vegetable crops.

## Author Contributions

T.O., E.O., O.M.O., T.C., J.K.O., W.A.W., and G.B. conceptualized the study and designed the experiments. T.O., E.O., W.A.W., and T.C. performed data analysis. T.O., and E.O. wrote the manuscript and interpreted the data together with T.C., W.A.W., E.K.O., T.C., G.B.

## Acknowledgements

The authors received funding from the Austrian Agency for International Cooperation in Education & Research (OeAD-GmbH) through the Ernst Mach-Stipendien, weltweit. This study was also funded in part by the Austrian Science Fund (FWF) under grant number 10.55776/J4753, awarded to Expedito Olimi, and by the EXCALIBUR project, with funding from the European Union's Horizon 2020 Research and Innovation Program under grant agreement No. 817946 (to Gabriele Berg). The authors also acknowledge the use of the IRIDIS 5 High Performance Computing Facility, and the associated support services at the University of Southampton, in the completion of this study. Open Access funding provided by Technische Universität Graz/KEMÖ.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

All the raw sequencing data have been submitted to the National Centre for Biotechnology Information (NCBI) BioProject PRJNA133559 (<http://www.ncbi.nlm.nih.gov/bioproject/133559>). All the codes for generating the figures in the manuscript can be accessed at <https://github.com/kaboyo/African-seed-microbiome>.

## References

Abarenkov, K., R. H. Nilsson, K.-H. Larsson, et al. 2024. "The UNITE Database for Molecular Identification and Taxonomic Communication of Fungi and Other Eukaryotes: Sequences, Taxa and Classifications Reconsidered." *Nucleic Acids Research* 52: D791–D797.

Abdelfattah, A., A. J. M. Tack, C. Lobato, B. Wassermann, and G. Berg. 2023. "From Seed to Seed: The Role of Microbial Inheritance in the Assembly of the Plant Microbiome." *Trends in Microbiology* 31: 346–355.

Abukutsa-Onyango, M. 2010. *African Indigenous Vegetables in Kenya: Strategic Repositioning in the Horticultural Sector*. [https://www.researchgate.net/publication/235323508\\_African\\_indigenous\\_vegetables\\_in\\_Kenya\\_strategic\\_repositioning\\_in\\_the\\_horticultural\\_sector#fullTextFileContent](https://www.researchgate.net/publication/235323508_African_indigenous_vegetables_in_Kenya_strategic_repositioning_in_the_horticultural_sector#fullTextFileContent).

Adam, E., M. Bernhart, H. Müller, J. Winkler, and G. Berg. 2018. "The *Cucurbita pepo* Seed Microbiome: Genotype-Specific Composition and Implications for Breeding." *Plant and Soil* 422: 35–49.

Adeniyi, O., F. Olaifa, B. Emikpe, and S. Ogunbanwo. 2017. "Phytochemical Components and Antibacterial Activity of *Tamarindus indica* Linn. Extracts Against Some Pathogens." *Biotechnology journal international* 17: 1–9.

Badri, D. V., and J. M. Vivanco. 2009. "Regulation and Function of Root Exudates." *Plant, Cell & Environment* 32: 666–681.

Barret, M., J. F. Guimbaud, A. Darrasse, and M. A. Jacques. 2016. "Plant Microbiota Affects Seed Transmission of Phytopathogenic Microorganisms." *Molecular Plant Pathology* 17: 791–795.

Berg, G., and J. M. Raaijmakers. 2018. "Saving Seed Microbiomes." *ISME Journal* 12: 1167–1170.

El Bilali, H., C. Callenius, C. Strassner, and L. Probst. 2018. "Food and Nutrition Security and Sustainability Transitions in Food Systems." *Food and Energy Security*. <https://doi.org/10.1002/fes3.154>.

Blaser, M. J. 2017. "The Theory of Disappearing Microbiota and the Epidemics of Chronic Diseases." *Nature Reviews Immunology* 17: 461–463.

Bokulich, N. A., B. D. Kaehler, J. R. Rideout, et al. 2018. "Optimizing Taxonomic Classification of Marker-Gene Amplicon Sequences With QIIME 2's q2-Feature-Classifier Plugin." *Microbiome* 6: 90.

Bolyen, E., J. R. Rideout, M. R. Dillon, et al. 2019. "Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2." *Nature Biotechnology* 37: 852–857.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. "DADA2: High-Resolution Sample Inference From Illumina Amplicon Data." *Nature Methods* 13: 581–583.

Caporaso, J. G., C. L. Lauber, W. A. Walters, et al. 2012. "Ultra-High-Throughput Microbial Community Analysis on the Illumina HiSeq and Miseq Platforms." *ISME Journal* 6: 1621–1624.

Cernava, T. 2024. "Coming of Age for Microbiome Gene Breeding in Plants." *Nature Communications* 15: 6623.

E. akin-Idowu, P., O. A. Odunola, M. A. Gbadegesin, O. T. Ademoyegun, A. O. Aduloju, and Y. O. Olagunju. 2017. "Nutritional Evaluation of Five Species of Grain Amaranth—An Underutilized Crop." *International Journal of Sciences* 3: 18–27.

Ejoh, S. I., F. D. Wireko-Manu, D. Page, and C. M. Renard. 2021. "Traditional Green Leafy Vegetables as Underutilised Sources of Micronutrients in a Rural Farming Community in South-West Nigeria I: Estimation of Vitamin C, Carotenoids and Mineral Contents." *South African Journal of Clinical Nutrition* 34: 40–45.

Ewels, P., M. Magnusson, S. Lundin, and M. Käller. 2016. "MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report." *Bioinformatics* 32: 3047–3048.

Fan, X., H. Matsumoto, H. Xu, et al. 2024. "Aspergillus Cvjetkovicii Protects Against Phytopathogens Through Interspecies Chemical Signalling in the Phyllosphere." *Nature Microbiology* 9: 2862–2876.

Garrido-Sanz, D., and C. Keel. 2025. "Seed-Borne Bacteria Drive Wheat Rhizosphere Microbiome Assembly via Niche Partitioning and Facilitation." *Nature Microbiology* 10: 1130–1144.

Ibeawuchi, I., J. Obiefuna, and U. Iwuanyanwu. 2015. "Low External Input Agricultural Farming System for the Increase in Productivity of Resource Poor Farmers." *Journal of Biology, Agriculture and Healthcare* 5, no. 2. [www.iiste.org](http://www.iiste.org).

Icard-Vernière, C., F. Olive, C. Picq, and C. Mouquet-Rivier. 2015. "Contribution of Leafy Vegetable Sauces to Dietary Iron, Zinc, Vitamin A and Energy Requirements in Children and Their Mothers in Burkina Faso." *Plant Foods for Human Nutrition* 70: 63–70.

Jiang, L., J. C. Jeong, J.-S. Lee, et al. 2019. "Potential of *Pantoea dispersa* as an Effective Biocontrol Agent for Black Rot in Sweet Potato." *Scientific Reports* 9: 16354.

Johnston-Monje, D., D. S. Lundberg, G. Lazarovits, V. M. Reis, and M. N. Raizada. 2016. "Bacterial Populations in Juvenile Maize Rhizospheres Originate From Both Seed and Soil." *Plant and Soil* 405: 337–355.

Johnston-Monje, D., W. K. Mousa, G. Lazarovits, and M. N. Raizada. 2014. "Impact of Swapping Soils on the Endophytic Bacterial

- Communities of Pre-Domesticated, Ancient and Modern Maize.” *BMC Plant Biology* 14: 233.
- Johnston-Monje, D., and M. N. Raizada. 2011. “Conservation and Diversity of Seed Associated Endophytes in Zea Across Boundaries of Evolution, Ethnography and Ecology.” *PLoS One* 6: e20396.
- Jung, J., J.-S. Kim, J. Taffner, G. Berg, and C.-M. Ryu. 2020. “Archaea, Tiny Helpers of Land Plants.” *Computational and Structural Biotechnology Journal* 18: 2494–2500.
- Klindworth, A., E. Pruesse, T. Schweer, et al. 2013. “Evaluation of General 16S Ribosomal RNA Gene PCR Primers for Classical and Next-Generation Sequencing-Based Diversity Studies.” *Nucleic Acids Research* 41: e1.
- Köberl, M., H. Müller, E. M. Ramadan, and G. Berg. 2011. “Desert Farming Benefits From Microbial Potential in Arid Soils and Promotes Diversity and Plant Health.” *PLoS One* 6: e24452.
- Kruk, M. E., and M. Pate. 2020. “The Lancet Global Health Commission on High Quality Health Systems 1 Year On: Progress on a Global Imperative.” *Lancet Global Health* 8: e30–e32.
- Liu, Y., Y. Wang, and J. Zhang. 2012. “New Machine Learning Algorithm: Random Forest.” In *Information Computing and Applications*, edited by B. Liu, M. Ma, and J. Chang, 246–252. Springer.
- Lobato, C., A. Abdelfattah, G. Berg, and T. Cernava. 2025. “Defining the Cultured and Uncultured Bacterial Fractions in Cannabis Seeds.” *Environmental Microbiome* 20, no. 1: 68. <https://doi.org/10.1186/s40793-025-00731-4>.
- Lobato, C., J. M. de Freitas, D. Habich, I. Kögl, G. Berg, and T. Cernava. 2024. “Wild Again: Recovery of a Beneficial Cannabis Seed Endophyte From Low Domestication Genotypes.” *Microbiome* 12: 239.
- Louca, S., L. W. Parfrey, and M. Doebeli. 2016. “Decoupling Function and Taxonomy in the Global Ocean Microbiome.” *Science* 353: 1272–1277.
- Lundberg, D. S., S. Yourstone, P. Mieczkowski, C. D. Jones, and J. L. Dangl. 2013. “Practical Innovations for High-Throughput Ampli-con Sequencing.” *Nature Methods* 10: 999–1002.
- Martin, M. 2011. “Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads.” *EMBnet.Journal* 17: 10.
- Matsumoto, H., X. Fan, Y. Wang, et al. 2021. “Bacterial Seed Endophyte Shapes Disease Resistance in Rice.” *Nature Plants* 7: 60–72.
- McMurdie, P. J., and S. Holmes. 2013. “Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data.” *PLoS One* 8: e61217.
- Michl, K., G. Berg, and T. Cernava. 2023. “The Microbiome of Cereal Plants: The Current State of Knowledge and the Potential for Future Applications.” *Environmental Microbiome* 18: 28.
- Moissl-Eichinger, C., M. Pausan, J. Taffner, G. Berg, C. Bang, and R. A. Schmitz. 2018. “Archaea Are Interactive Components of Complex Microbiomes.” *Trends in Microbiology* 26: 70–85.
- Nelson, E. B. 2018. “The Seed Microbiome: Origins, Interactions, and Impacts.” *Plant and Soil* 422: 7–34.
- Nguyen, N. H., Z. Song, S. T. Bates, et al. 2016. “FUNGuild: An Open Annotation Tool for Parsing Fungal Community Datasets by Ecological Guild.” *Fungal Ecology* 20: 241–248.
- Oksanen, J., F. G. Blanchet, M. Friendly, et al. 2020. *Package ‘vegan’ Title Community Ecology Package Version 2.5-7*.
- Olimi, E., P. Kusstatscher, W. A. Wicaksono, A. Abdelfattah, T. Cernava, and G. Berg. 2022. “Insights Into the Microbiome Assembly During Different Growth Stages and Storage of Strawberry Plants.” *Environmental Microbiome* 17: 21.
- Pérez-Jaramillo, J. E., R. Mendes, and J. M. Raaijmakers. 2016. “Impact of Plant Domestication on Rhizosphere Microbiome Assembly and Functions.” *Plant Molecular Biology* 90: 635–644.
- Pölme, S., K. Abarenkov, R. Henrik Nilsson, et al. 2020. “Fungaltraits: A User-Friendly Traits Database of Fungi and Fungus-Like Stramenopiles.” *Fungal Diversity* 105: 1–16.
- Quast, C., E. Pruesse, P. Yilmaz, et al. 2012. “The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools.” *Nucleic Acids Research* 41: D590–D596.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. URL <https://www.R-project.org/>.
- Reitmeier, S., T. C. A. Hitch, N. Treichel, et al. 2021. “Handling of Spurious Sequences Affects the Outcome of High-Throughput 16S rRNA Gene Amplicon Profiling.” *ISME Communications* 1: 1–12.
- Rodríguez, C. E., L. Antonielli, B. Mitter, F. Trognitz, and A. Sessitsch. 2020. “Heritability and Functional Importance of the *Setaria viridis* Bacterial Seed Microbiome.” *Phytobiomes Journal* 4: 40–52.
- Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahé. 2016. “VSEARCH: A Versatile Open Source Tool for Metagenomics.” *PeerJ* 4: e2584.
- Segata, N., J. Izard, L. Waldron, et al. 2011. “Metagenomic Biomarker Discovery and Explanation.” *Genome Biology* 12: R60.
- Shade, A., M. A. Jacques, and M. Barret. 2017. “Ecological Patterns of Seed Microbiome Diversity, Transmission, and Assembly.” *Current Opinion in Microbiology* 37: 15–22.
- Simonin, M., M. Briand, G. Chesneau, et al. 2022. “Seed Microbiota Revealed by a Large-Scale Meta-Analysis Including 50 Plant Species.” *New Phytologist* 234: 1448–1463.
- Stilianos, L., S. M. Jacques, F. Pap, et al. 2017. “High Taxonomic Variability Despite Stable Functional Structure Across Microbial Communities.” *Nature Ecology & Evolution* 1: 0015. <https://www.nature.com/articles/s41559-016-0015>.
- Sulesky-Grieb, A., M. Simonin, A. F. Bintarti, B. Marolleau, M. Barret, and A. Shade. 2024. “Stable, Multigenerational Transmission of the Bean Seed Microbiome Despite Abiotic Stress.” *mSystems* 9: e00951-24.
- Taffner, J., A. Bergna, T. Cernava, and G. Berg. 2020a. “Tomato-Associated Archaea Show a Cultivar-Specific Rhizosphere Effect but an Unspecific Transmission by Seeds.” *Phytobiomes Journal* 4: 133–141.
- Taffner, J., A. Erlacher, A. Bragina, C. Berg, C. Moissl-Eichinger, and G. Berg. 2018. “What Is the Role of Archaea in Plants? New Insights From the Vegetation of Alpine Bogs.” *mSphere* 3: e00122-18.
- Taffner, J., O. Laggner, A. Wolfgang, D. Coyne, and G. Berg. 2020. “Exploring the Microbiota of East African Indigenous Leafy Greens for Plant Growth, Health, and Resilience.” *Frontiers in Microbiology* 11: 585690.
- Torres-Cortés, G., S. Bonneau, O. Bouchez, et al. 2018. “Functional Microbial Features Driving Community Assembly During Seed Germination and Emergence.” *Frontiers in Plant Science* 9. <https://doi.org/10.3389/fpls.2018.00902>.
- Vandenkoornhuyse, P., A. Quaiser, M. Duhamel, A. Le Van, and A. Dufresne. 2015. “The Importance of the Microbiome of the Plant Holobiont.” *New Phytologist* 206: 1196–1206.
- Walters, W., E. R. Hyde, D. Berg-Lyons, et al. 2016. “Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys.” *mSystems* 1. <https://doi.org/10.1128/msystems.00009-15>.
- Wassermann, B., A. Abdelfattah, W. A. Wicaksono, et al. 2022. “The *Brassica napus* Seed Microbiota Is Cultivar-Specific and Transmitted via Paternal Breeding Lines.” *Microbial Biotechnology* 15: 2379–2390.
- Wassermann, B., E. Adam, T. Cernava, and G. Berg. 2019a. “Understanding the Indigenous Seed Microbiota to Design Bacterial Seed Treatments.” In *Seed Endophytes*, edited by S. Verma, J. White, Jr. Springer. [https://doi.org/10.1007/978-3-030-10504-4\\_5](https://doi.org/10.1007/978-3-030-10504-4_5).

- Wassermann, B., T. Cernava, S. Goertz, et al. 2023. “Low Nitrogen Fertilization Enriches Nitrogen-Fixing Bacteria in the Brassica Seed Microbiome of Subsequent Generations.” *Journal of Sustainable Agriculture and Environment* 2: 87–98.
- Wassermann, B., T. Cernava, H. Müller, C. Berg, and G. Berg. 2019b. “Seeds of Native Alpine Plants Host Unique Microbial Communities Embedded in Cross-Kingdom Networks.” *Microbiome* 7: 108.
- Wentzien, N. M., A. J. Fernández-González, A. Valverde-Corredor, et al. 2024. “Pitting the Olive Seed Microbiome.” *Environmental Microbiome* 19: 17.
- White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. “Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics.” In *PCR Protocols*, 315–322. Elsevier.
- Wicaksono, W. A., T. Cernava, B. Wassermann, et al. 2023. “The Edible Plant Microbiome: Evidence for the Occurrence of Fruit and Vegetable Bacteria in the Human Gut.” *Gut Microbes* 15: 2258565.
- Wolfgang, A., C. Zachow, H. Müller, et al. 2020. “Understanding the Impact of Cultivar, Seed Origin, and Substrate on Bacterial Diversity of the Sugar Beet Rhizosphere and Suppression of Soil-Borne Pathogens.” *Frontiers in Plant Science* 11: 1.
- Yilmaz, P., L. W. Parfrey, P. Yarza, et al. 2014. “The SILVA and ‘All-Species Living Tree Project (LTP)’ Taxonomic Frameworks.” *Nucleic Acids Research* 42: D643–D648.
- Zhan, C., and M. Wang. 2024. “Disease Resistance Through M Genes.” *Nature Plants* 10: 352–353.

### Supporting Information

Additional supporting information can be found online in the Supporting Information section.

**Figure S1:** Correlation in diversity between the fungal and bacterial community across the different plant species. **Figure S2:** Abundance-occupancy curves showing the bacterial. **Figure S3:** Taxonomy-based prediction of the seed fungal community functions using FUNGuild database, respectively. **Data S1:** Bacterial taxonomy composition under the different functional categories. **Data S2:** Archaeal taxonomy composition under the different functional categories.