Data in Brief 48 (2023) 109132

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Shotgun metagenomics dataset of *Striga hermonthica*-infested maize (*Zea mays* L.) rhizospheric soil microbiome

Olubukola Oluranti Babalola*, Olumayowa Mary Olowe, Ayansina Segun Ayangbenro

Food Security and Safety Focus Area, Faculty of Natural and Agricultural Sciences, North-West University, Private Mail Bag, Mmabatho, South Africa.

ARTICLE INFO

Article history:
Received 22 October 2022
Revised 29 March 2023
Accepted 30 March 2023
Available online 7 April 2023

Dataset link: Metagenomics sequence of Rhizosphere microbiome infected with Striga Harmontheca (Original data)
Dataset link: Metagenomics sequence of Rhizosphere microbiome infected with Striga Harmontheca (Original data)

Keywords:
Microbial community
Parasitic weed
Food security
Biotic stress
MG-RAST
Cereal

ABSTRACT

This dataset includes shotgun metagenomics sequencing of the rhizosphere microbiome of maize infested with *Striga hermonthica* from Mbuzini, South Africa, and Eruwa, Nigeria. The sequences were used for microbial taxonomic classification and functional categories in the infested maize rhizosphere. High throughput sequencing of the complete microbial community's DNA was performed using the Illumina NovaSeq 6000 technology. The average base pair count of the sequences were 5,353,206 bp with G+C content of 67%. The raw sequence data used for analysis is available in NCBI under the BioProject accession numbers PRJNA888840 and PRJNA889583. The taxonomic analysis was performed using Metagenomic Rapid Annotations using Subsystems Technology (MG-RAST). Bacteria had the highest taxonomic representation (98.8%), followed by eukaryotes (0.56%), and archaea (0.45%). This metagenome dataset provide valuable information on microbial communities associated with *Striga*-infested maize rhizosphere and their functionality. It can also

* Corresponding author.
E-mail address: olubukola.babalola@nwu.ac.za (O.O. Babalola).
Social media: @oloweolumayowa (O.O. Babalola), @mayolowe (O.M. Olowe), @AAyansina (A.S. Ayangbenro)

https://doi.org/10.1016/j.dib.2023.109132

be used for further studies on application of microbial resources for sustainable crop production in this region.

Specifications Table

Subject	Microbiology
Specific subject area	Microbial biotechnology
Type of data	Figures and Fastq files
How the data were acquired	Nine rhizosphere soil and 3 bulk soil samples were collected from each field (A total of 24 samples). DNA from soil samples were sequenced on Novaseq 6000 platform (2 × 150 paired end). Raw data were assembled and annotated using metagenomics rapid annotation online service (MG-RAST).
Data format	Raw data (Fastq.gz.file)
Description of data collection	Metagenomic DNA were extracted from rhizospheric soil samples collected from agricultural lands from Eruwa, Oyo state, Nigeria and Mbuzini, Mpumalanga Province, South Africa that were previously cultivated with maize and have the history of establishment of *Striga hermonthica*. Nucleospin soil genomic DNA purification kit was used for DNA isolation, sequenced on NovaSeq 6000 platform (Illumina) and metagenome sequence annotation through MG-RAST.
Data source location	Institution: North-West University, Mmabatho, North West Province, South Africa. Latitude and longitude for collected soil samples: Eruwa, Nigeria (7°28′2.034″N3°28′21.671″ E) Mbuzini, South Africa (25°55′30.90″S31°56′11.70″ E)
Data accessibility	Repository name: National Centre for Biotechnology Information Data identification numbers: PRJNA888840 and PRJNA889583 Direct URL to data: http://www.ncbi.nlm.nih.gov/bioproject/888840 [1] and http://www.ncbi.nlm.nih.gov/bioproject/889583 [2].

Value of the Data

- The dataset provides information about the composition and functional diversity of soil rhizosphere microbiome of *Striga Hermonthica*-infested maize.
- This dataset will be of interests to plant pathologist, environmental microbiologist, molecular biologist, and agriculturists.
- This gives insights into the potential of some microbes in alleviating the impacts of parasitic weeds.
- The dataset can be employed for further research into the genetic and molecular basis of host resistance and host-parasite association.
- Increased microbial diversity improves soil ecosystem functions and tolerance to seasonal and other disturbances.
- The dataset provides the prospect to potentially find new genes that could help end hunger and ensure food security

1. Objective

The recent development in next-generation DNA sequencing (NGS) technology, including metagenomics analysis has given the opportunities to enhance our understanding of the composition and function of soil microbial communities. Therefore, in this study, we aim to reveal the microbial diversity and composition within maize rhizosphere in *Striga hermonthica* infested field using a shotgun metagenomics approach in two locations with different seasons.

2. Data Description

This dataset is made up of raw data that was generated using shotgun sequencing of *Striga hermonthica*-infested maize rhizosphere from South Africa and Nigeria. All dataset obtained in fastq.gz file were deposited at the National Centre for Biotechnology Information (NCBI) SRA database (PRJNA888840 and PRJNA889583). Figs. 1 and 2 present information about the microbial community structure and functional categories of the soil rhizosphere microbial communities in *Striga hermonthica* -infested maize plants, respectively.

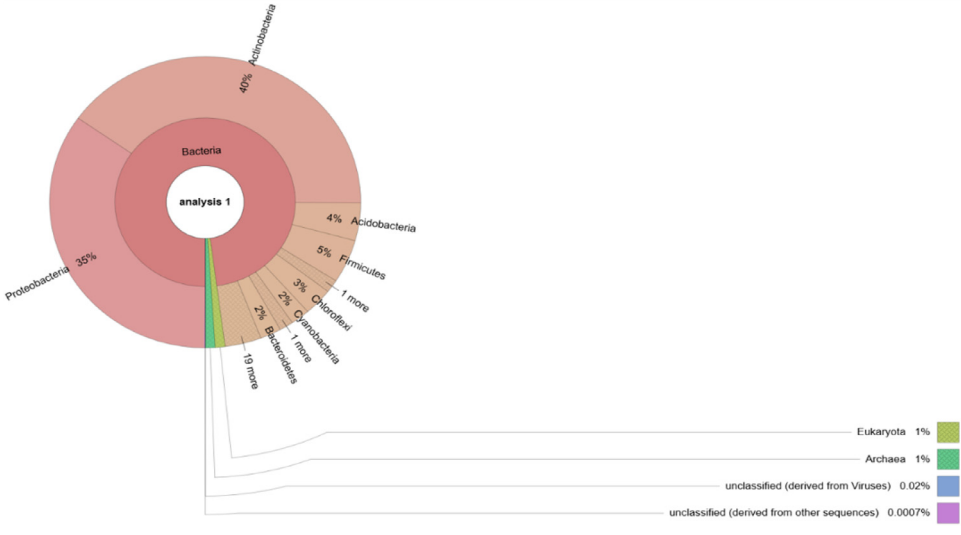

Fig. 1. Taxonomic structure of the microbiome in *Striga hermonthica*-infested maize rhizosphere.

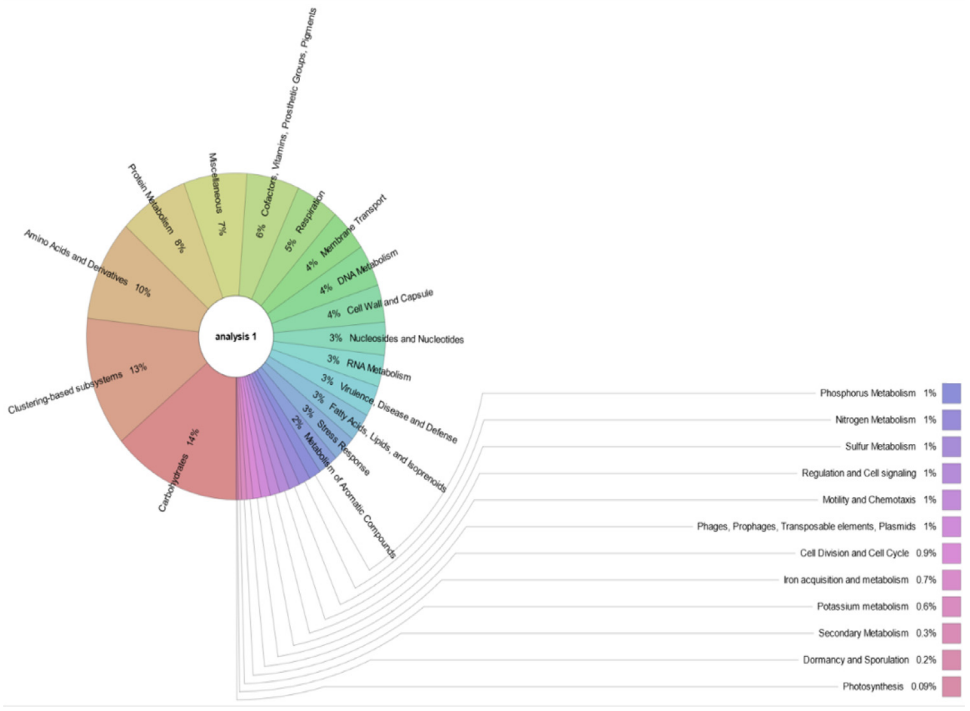

Fig. 2. Functional diversity based on subsystem in *Striga hermonthica*-infested maize rhizosphere soil metagenome.

3. Experimental Design, Materials and Methods

Rhizosphere soil samples were taken from the rhizosphere of Striga hermonthica-infested maize in Mbuzini, Mpumalanga Province, South Africa (25°55′30.90″ S 31°56′11.70″ E) and Eruwa, Oyo State, Nigeria (7°28′2.034″ N 3°28′21.671″ E). DNA was extracted using the Nucleospin soil genomic DNA purification kit according to the kit's protocol. Shotgun sequencing was performed using the Illumina sequencing platform at NovogeneAIT Genomics Singapore Pte Ltd. Libraries with 20–50 ng of DNA were prepared using a Nextera DNA flex library kit. Following that, adapter sequences were added after the samples had been simultaneously fragmented. Using the Agilent 2100-Bioanalyzer, the average library size was calculated, and the concentration of the libraries was quantified using the Qubit® dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA, USA). The libraries were pooled and diluted to 0.6 nM before being sequenced for 300 cycles on NovaSeq 6000 platform. For assessing the raw metagenome sequences, MG-RAST, an online metagenomic rapid annotation server, was employed (www.mg-rast.org) [3] Following quality control (QC), the sequences were annotated using BLAT algorithm [4] against the M5NR database [5], which includes non-redundant integration of many datasets, using the BLAT (the BLAST-like alignment tool) algorithm.

Ethics Statements

The current work follow the ethical requirements for publication in Data in Brief. It does not involve human subjects, animal experiments, or any data collected from social media platforms.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Metagenomics sequence of Rhizosphere microbiome infected with Striga Harmontheca (Original data) (NCBI).
Metagenomics sequence of Rhizosphere microbiome infected with Striga Harmontheca (Original data) (NCBI).

CRediT Author Statement

Olubukola Oluranti Babalola: Conceptualization, Data curation, Funding acquisition, Resources, Supervision, Writing – review & editing; **Olumayowa Mary Olowe:** Investigation, Methodology, Data curation, Writing – original draft, Visualization; **Ayansina Segun Ayangbenro:** Software, Validation, Writing – review & editing.

Acknowledgments

OOB appreciates National Research Foundation, South Africa, for the grant (UID132595) that have supported the project.

References

[1] O. Olowe, Metagenomic sequence of rhizosphere microbiome infected with *Striga hermontheca*, NCBI Bioproject, v1, PRJNA888840 (https://www.ncbi.nlm.nih.gov/bioproject/888840).
[2] O. Olowe, Metagenomic sequence of rhizosphere microbiome infected with *Striga hermontheca*, NCBI Bioproject, v1, PRJNA889583 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA889583).
[3] F. Meyer, D. Paarmann, M. D'Souza, R. Olson, E.M. Glass, M. Kubal, T. Paczian, A. Rodriguez, R. Stevens, A. Wilke, J. Wilkening, The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes, *BMC* Bioinform. 9 (2008) 386, doi:10.1186/1471-2105-9-386.
[4] W.J. Kent, BLAT—The BLAST-like alignment tool, Genome Res. 12 (2002) 656–664, doi:10.1101/gr.229202.
[5] A. Wilke, T. Harrison, J. Wilkening, D. Field, E.M. Glass, N. Kyrpides, K. Mavrommatis, F. Meyer, The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools, BMC Bioinform. 13 (2012) 141, doi:10.1186/1471-2105-13-141.