Wheat is the major crop in India and like other crops also subjected to influence by microbial communities of the rhizospheric region which are extremely diverse and undoubtedly play a central role in the nutrient cycle, plant productivity and growth promotion. In order to know how changes in the rhizospheric microbial community can make an impact on overall crop function, wheat rhizospheric soil samples from Ghazipur (25.913824 N 83.529715 E) regions of Eastern Uttar Pradesh (Eastern Indogangatic Plain), were collected and analyzed. Full length 16S rRNA gene amplification sequencing was performed to reveal the bacterial community in wheat rhizosphere. A total of 51,909 read were analyzed, out of that only 44,125 reads were classified and 7784 were unclassified using oxford nanopore sequencing and EPI2ME data analysis platform. MinION oxford nanopore sequencing uncovered that dominant phyla were Proteobacteria (68%), followed by firmicutes (13%), bacteroidetes (3%), actinobacteria (3%) and acidobacteria (3%). The data is available at the NCBI - Sequence Read Archive (SRA) with accession number: SRX5275271.

© 2019 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Data

The data in this dataset described the predominant bacterial communities present in the wheat rhizosphere. The wheat rhizospheric metagenome sequence was obtained using MinION oxford nanopore sequencing platform. Total 51,909 read were analyzed, out of that only 44,125 reads are classified and 7784 were unclassified (Table 1). Around 221 eukaryotes reads were also sequenced and they were cleaned out from the data. Fig. 1 describes the distribution of various bacterial communities in the wheat rhizosphere.

2. Experimental design, materials and methods

2.1. Sample collection

A total of ten wheat rhizospheric soil sample were randomly collected from the wheat field of the Ghazipur (25.913824 N 83.529715 E) region of eastern Uttar Pradesh, India. The collected soil samples were stored in sterile containers and processed for DNA isolation and sequencing.

| Specification table                   |
|---------------------------------------|
| Subject area                          | Biology                      |
| More specific subject area            | Metagenomics                 |
| Type of data                          | Metagenome sequence          |
| How data was acquired                 | MinION - Oxford Nanopore Sequencing platform |
| Data format                           | Raw data                     |
| Parameters for data collection        | Isolation of metagenome from wheat rhizospheric soil and full 16s rRNA gene sequencing was performed by using MinION oxford nanopore sequencing platform |
| Description of data collection       | 1 Collection of wheat rhizospheric soil. |
|                                      | 2 Isolation of DNA from soil.  |
|                                      | 3 Amplification of 16s rRNA gene |
|                                      | 4 Sequencing and data analysis |
| Data source location                 | Ghazipur (25.913824 N 83.529715 E)Uttar Pradesh, India |
| Data accessibility                   | The data is available at the NCBI - Sequence Read Archive (SRA) with accession number: SRX5275271 https://www.ncbi.nlm.nih.gov/sra/SRX5275271 |

Value of the data

- The data provide baseline information about the wheat rhizospheric bacterial community of Eastern Indo Gangetic Plains.
- The researchers working on wheat microbiome and its functional analysis shall be benefited with the data.
- The data provides information about the natural distribution of different bacterial phyla in the wheat rhizospheric community. This will be used as a baseline for further data related to spatial and temporal shift in the community under different wheat growing conditions and abiotic stresses.
- The data is also valuable in development of consortia based approach for soil health management.

Table 1
Summary statistics table.

| Reads            | Count |
|------------------|-------|
| Total analyzed reads | 51,909|
| Classified reads | 44,125|
| Unclassified reads | 7784  |
2.2. DNA extraction

Soil DNA was extracted by FastDNA spin kit (MP Biomedical). A total of 500 mg of sieved soil was lysed by MT buffer in lysing matrix. Soil suspension was centrifuged at 14000×g for 5 minutes and supernatant was treated with PPS to precipitate proteins. After centrifugation at 14000×g for 5 minutes binding matrix was used to bind the genomic DNA followed by washing of bound DNA with ethanol. Pure DNA was eluted out through filter in elution buffer containing Tris and EDTA. DNA was quantified using nanodrop and used for downstream processing.

2.3. Metagenome sequencing

To characterize and identify the microbial community associated with wheat rhizosphere, 16S rRNA gene in soil metagenome was amplified and metagenome sequencing was carried out by MinION oxford nanopore sequencer [1]. Library preparation for Nanopore sequencing was performed by using optimal fragmentation of high molecular weight genomic DNA followed by end repair/dA tailing of barcode adapters. A barcoding PCR allowed to barcode sequencing libraries, in which sheared DNA fragments repaired and dA tailed using dA tailing modules and then adapters possessing primer binding sites were ligated. Primers encoded with 5’ tag and barcode facilitates the attachments of adapters without ligases. The reaction samples then loaded on to flow cell of Nanopore device. Sequence raw reads obtained by sequencer were submitted in NCBI- Sequence Reads Archive (SRA).

2.4. Taxonomic analysis

Oxford Nanopore Technology (ONT’s) cloud based pipeline EPI2ME (WIMP rev. 3.2.3) workflow was used to retrieve the information 16S analysis up to genus level, along with detail insight to species and even sup species level by accessing basecalled file. Phylogenetic analysis information against the query
sequences was inferred by multiple sequence alignment. The workflow is designed in the way to BLAST [2] the basecalled sequences against 16S sequences of NCBI bacterial database with respective sequences. Here every single read is classified as per identity and percentage coverage. Total 51,909 read were analyzed, out of that only 44,125 reads are classified and 7784 were unclassified, high number of bacterial communities were identified followed by archea. Data revealed the Proteobacteria with 68% occurrence as dominant phylum, followed by the less dominant fircmuctes (13%), bacteroidetes (3%), actinobacteria (3%) and acidobacteria (3%). Further, on the basis of species level classification the data revealed *Escherichia coli* is most abundant species followed by *Candidatus solibacterusitatus* and *Achromobacter xylosoxidans*.

**Acknowledgement**

The present research is a part of Ph.D. work of Ruchi Srivastava. Authors are grateful to Director ICAR-NBAIM for providing necessary facilities for perusing research.

**Conflict of 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.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2019.105094.

**References**

[1] R.M. Leggette, M.D. Clark, A world of opportunities with nanopore sequencing, J. Exp. Bot. 68 (20) (2017) 5419–5429.
[2] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.