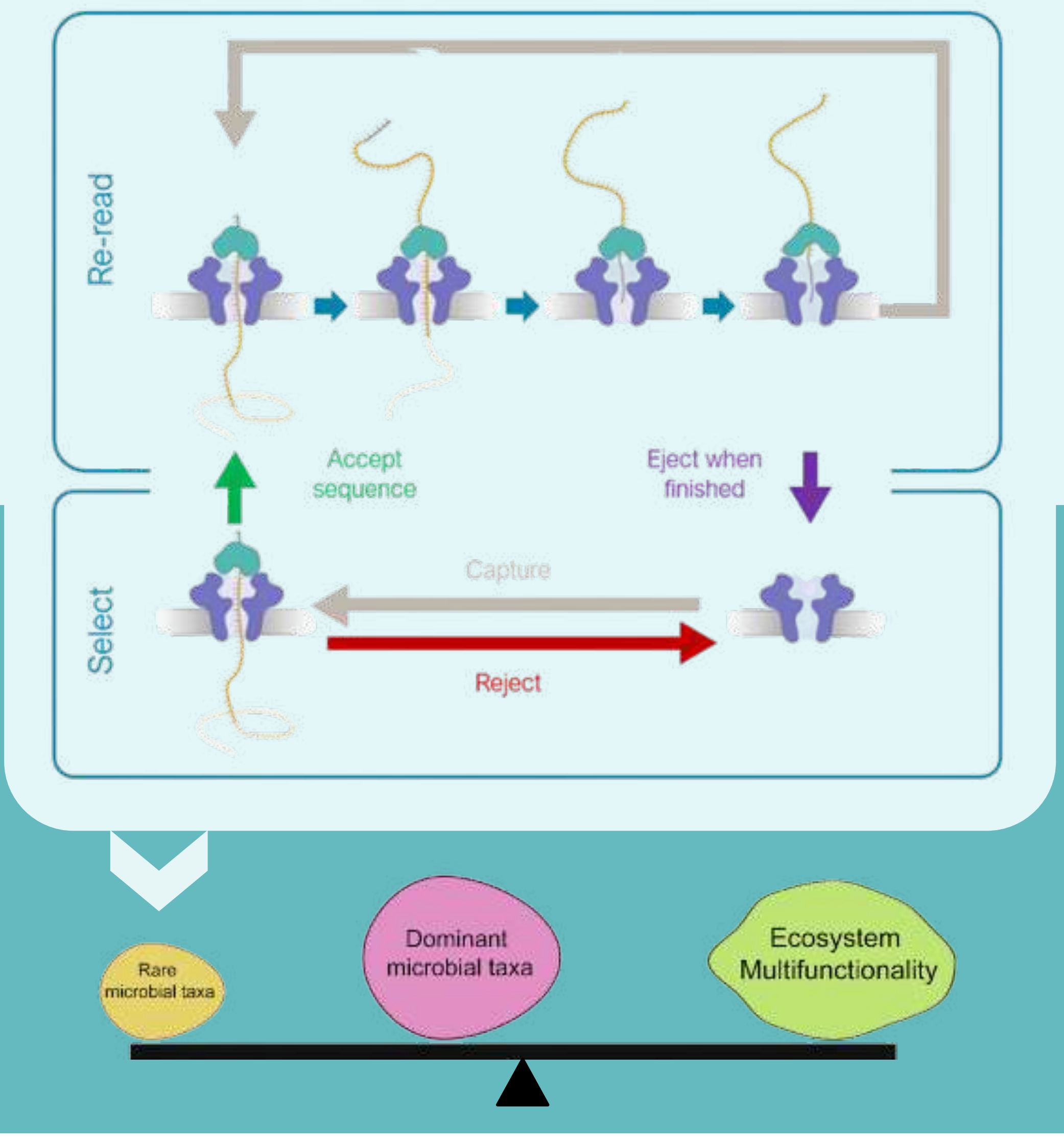
Genome enrichment of rare and unknown species from complicated microbiome by nanopore selective sequencing

AUTHORS

Yuhong Sun, Zhanwen Cheng, Xiang Li, Qing Yang, Bixi Zhao, Ziqi Wu, Yu Xia *

School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, 518055, China
*: Corresponding author



INTRODUCTION

In metagenomic sequencing, the low-abundance microorganisms (sometimes referred to as rare biosphere) are often missed or simply neglected due to low sequencing coverage. However, pieces of evidence show the low-abundance microbiota may play essential roles in community functionality. Benefiting from the selective sequencing enabled by nanopore sequencing, we rejected DNA of microorganisms with high abundance and enriched rare taxa when metagenomic sequencing of the community of thermophilic anaerobic digester (TAD). By structurally altering the sequencing community in this way, we obtained more low abundance and high-quality metagenome-assembled genomes (HQ MAGs).

METARUPORE PIPELINE

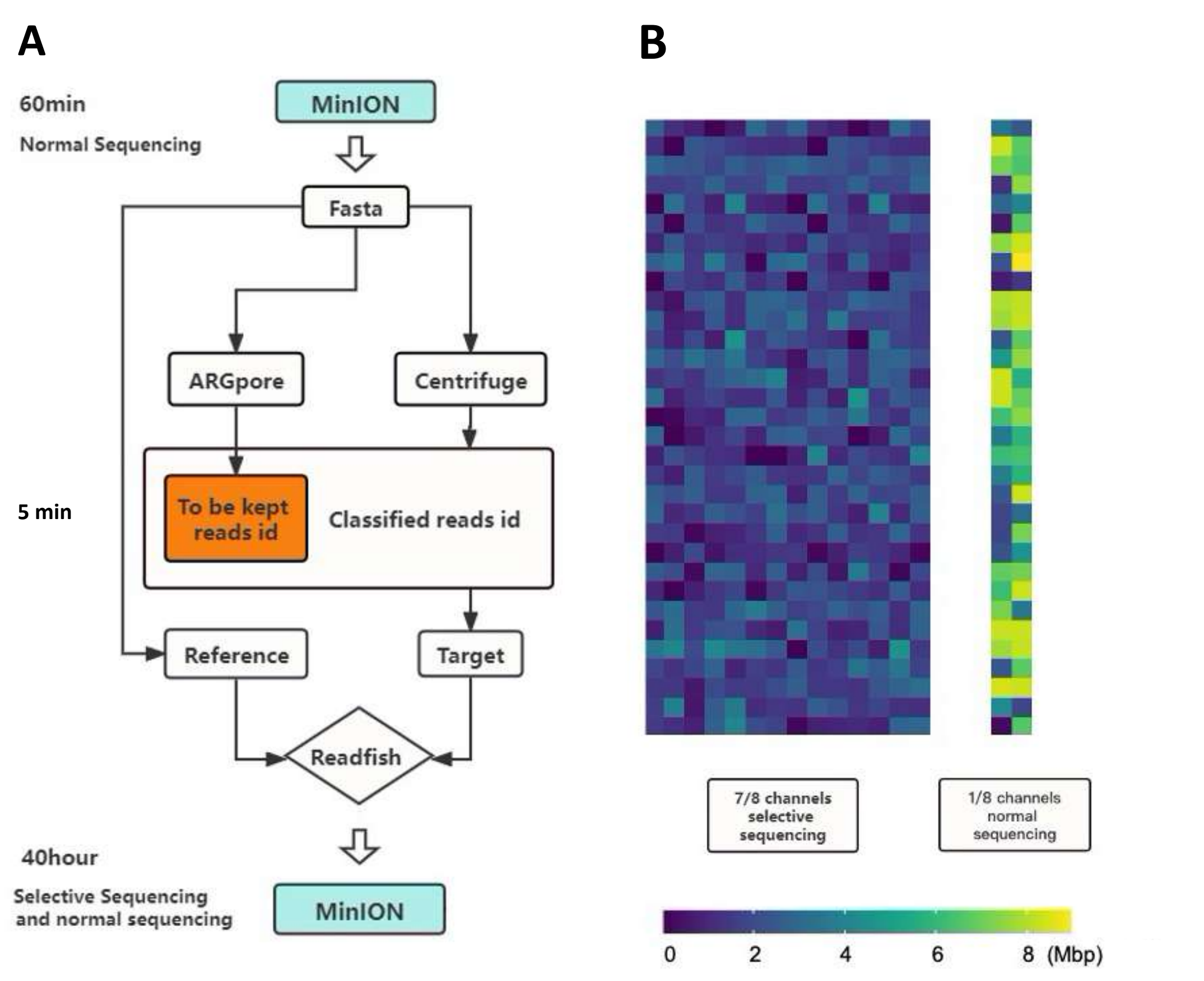


Figure 1. (A) The workflow of metaRUpore. (B) A MinION flow cell in metaRUpore is configured into two parts, 1/8th of the channels for normal sequencing and the remaining channels for selective sequencing.

The protocol consists of three consecutive steps:

- **A short-time normal sequencing** to obtain an overall picture of the community structure and the genomic profile of the dominant populations.
- **Bioinformatics analysis** to determine the reference and target dataset for optimized RU configuration.
- **A 40h selective sequencing** for enriching rare populations in the sample.

CONTRIBUTION TO TAD COMMUNITY

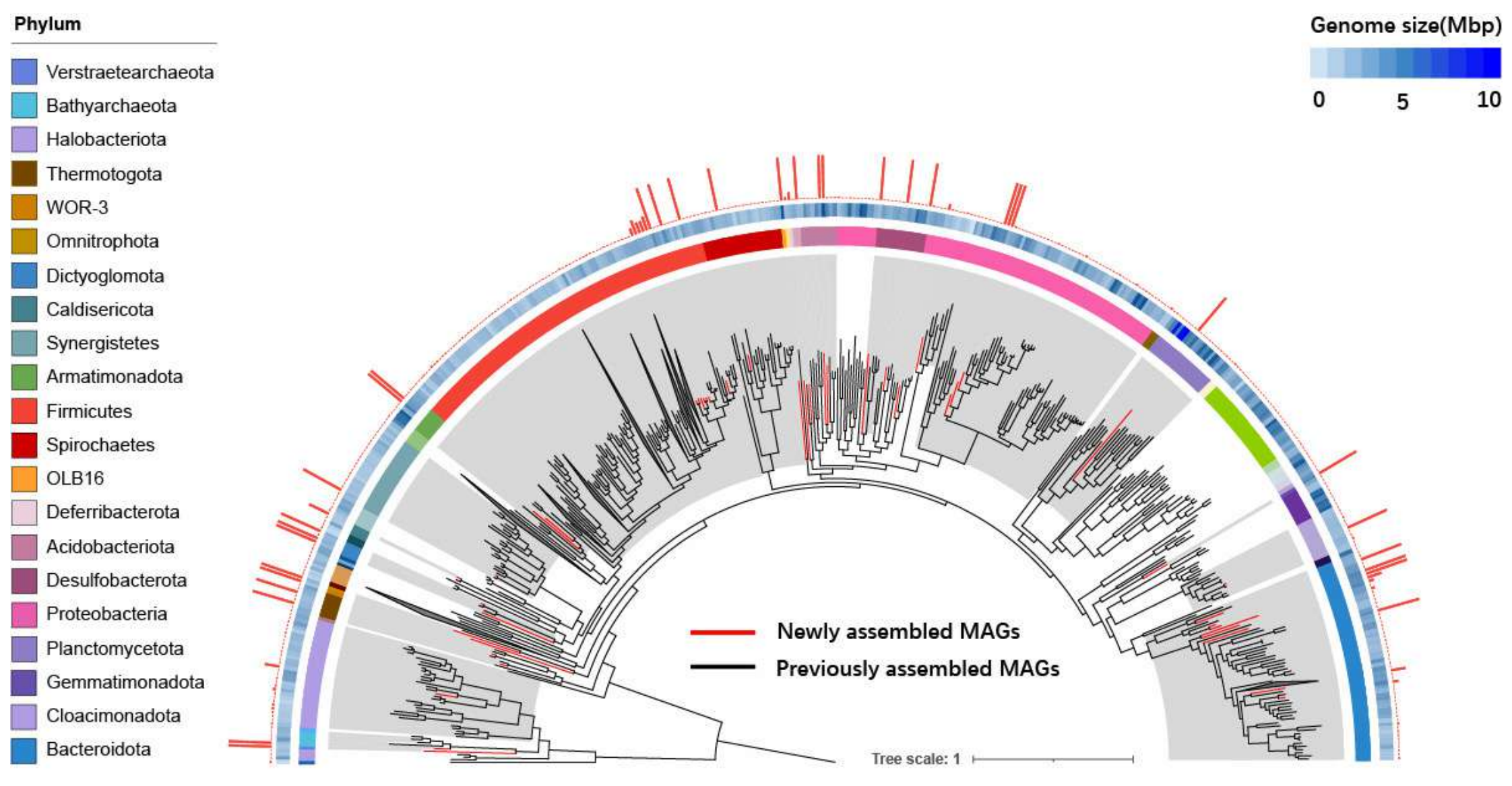
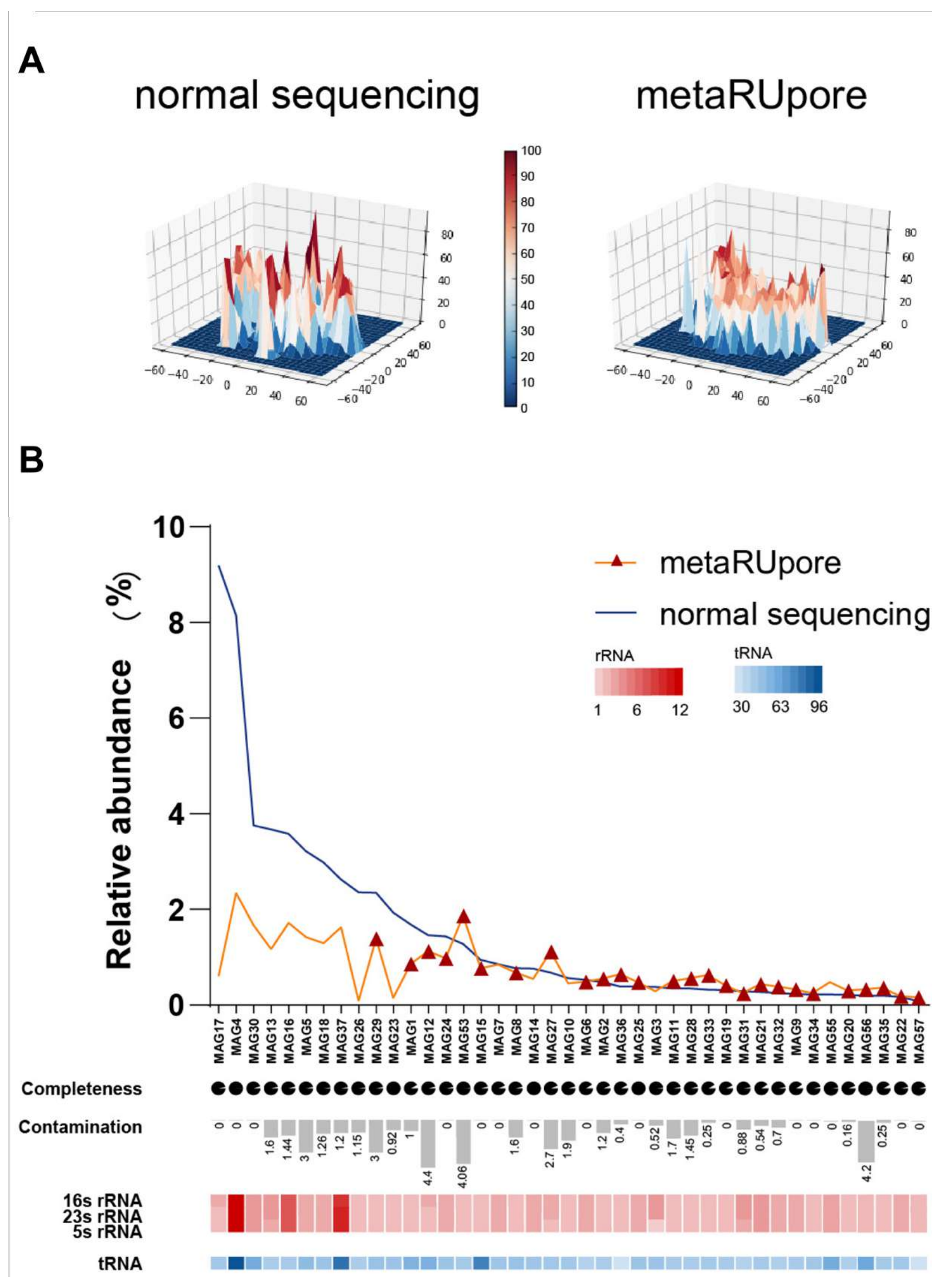


Figure 3. A phylogenetic tree was constructed from 41 HQ-MAGs derived by metaRUpore (red branches) and 1,108 HQ-MAGs collection derived from other AD systems (black branches). The bar plot represents the genome continuity, calculated as the reciprocal number of contigs. The grey-shaded areas indicate phyla with near-finished genomes obtained by metaRUpore, and the name of each phylum is in the legend on the left.

The 41 HQ MAGs derived by metaRUpore :

- **Introduce 5 new phyla** into the global HQ genome collection of the anaerobic digester (AD) microbiome.
- Represent **the first near-finished** MAGs for 21 phyla in the AD.
- Show **much better integrity and continuity** than those in the previous collection assembled with SRs in terms of N50, number of contigs as well as the integrity of rRNA operon.

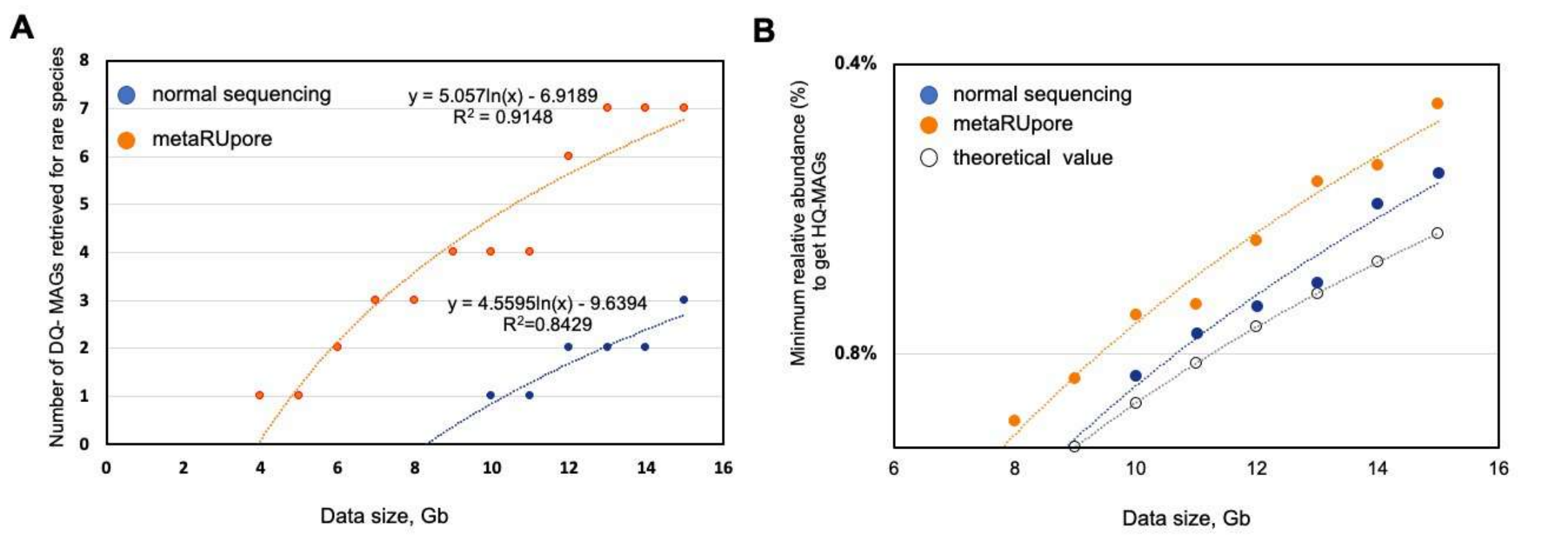
PERFORMANCE OF METARUPORE IN TAD COMMUNITY



- In the 3D density plot, several density peaks of the original TAD community were depleted in the metaRUpore datasets, indicating **DNA of the high abundance populations was effectively ejected** and the **community becomes homogeneous** with coverage of different populations become much more unified.
- **Evident coverage reduction** was observed in the dominant populations.
- 26 HQ MAGs that are additionally obtained by metaRUpore were mainly from the **rare populations** of the TAD community, this proves that metaRUpore **facilitates the recovery of more rare species genomes.**

Figure 2. (A) 3D density plots of t-SNE downscaling results for normal sequencing datasets and selective sequencing datasets by metaRUpore at four base frequencies, showing that metaRUpore renders the TAD community structure homogenous. (B) The distribution of 41 retrieved HQ MAGs in normal and RU sequencing dataset. The red triangles indicate MAGs that could only be assembled in the metaRUpore dataset.

SIMUNATION AND EVALUATION OF METARUPORE



- metaRUpore dataset can recover roughly **twice more MAGs of rare species** than normal sequencing.

Figure 4. (A) The number of draft-quality (DQ) MAGs retrieved for rare species by metaRUpore and normal nanopore sequencing. (B) Significant promotion in the minimum relative abundance to get HQ MAGs by metaRUpore as compared to normal nanopore sequencing.

CONCLUSION

In conclusion, we have demonstrated the effectiveness of metaRUpore which normalizes sequenced genome coverage and increases coverage of low-abundance species, facilitating the assembly of additional nf-MAGs of rare species within a complicated natural microbiota. It could be expected that by enhancing sequencing effort, HQ-MAGs could be obtained for populations with even lower prevalence using the metaRUpore protocol. Furthermore, metaRUpore protocol is robust and requires minimal modification to the experimental procedure of nanopore library construction and sequencing, making it easily applicable to metagenomic investigations of other environmental microbiomes to improve the time-to-answer in terms of sequencing costs and computational requirements.

REFERENCE：https://nanoporetech.com/