令和 6 年度 京都大学化学研究所 スーパーコンピュータシステム 利用報告書

SAR11 細菌の単一細胞ゲノム解析
Single-cell genomics analysis of SAR11 bacteria

京都大学 化学研究所 化学生命科学研究領域 Shaanaav Daniel

Throughout my 1 year at the Ogata Lab, I have used the Institute for Chenical Research Supercomputer system for various purposes. During the period between April and September 2024, I minimally utilized the system for basic programming and bioinformatics exercises, where I familiarized myself with the various databases and tools in the supercomputer, and also did work related to metagenomics and phylogenetic analysis. While most of the work done was admittedly quite elementary, this helped me to understand important processes related to the supercomputer, such as submitting/deleting tasks, writing BASH scripts, loading tools, and accessing databases.

From October 2024 onwards, my research began to be more focused on single-cell genomics analysis of bacteria, so I mainly operated tools such as BLAST+ to search for targeted conserved rRNA regions within SAGs. Other than that, I did numerous types of analysis on my data, such as gene prediction and annotation, clustering, amino/nucleic acid identity analysis, among others. Due to this, tools such as KOfamKOALA, Prokka, SuperPang, FASTANI, ezAAI and CD-HIT were utilised. I expect to utilize the supercomputer system more in the months to come as I will also begin working with long-read metagenomics data, which I’m sure will require a lot of work due to their large data size. With that, I expect to operate tools like Biosample, metaSPAdes, and many others in the future.