Diana Lin
Mariia Radaeva

Connection Assignment

Field Review: Single cell technologies in sequence assembly and genome construction
Current Paper: Obtaining high-quality draft genomes from uncultured microbes by cleaning and co-assembly of single-cell amplified genomes

The advent of next-generation sequencing (NGS) has revolutionized the field of genome biology. Using NGS, scientists sequenced full genomes of a wide variety of species including microorganisms. In particular, it became possible to study microbial genomes recovered directly from environmental samples (metagenomics). However, NGS applied to metagenomics was inherently multi-cell sequencing, where samples were composed of millions of cells. In this sense, multi-cell sequencing yields averaged data sets, where the data does not necessarily represent each cell. This limitation in metagenomics studies hides within species heterogeneity and makes it challenging to link specific genes to organisms of origin.

With the rise of single-cell sequencing, detailed and comprehensive genome analysis of individual cells became possible. Combined with the distinctive advantage of metagenomics approach to sequence unculturable bacteria single-cell sequencing allowed to study bacterial diversity on the cell-to-cell level. However, single-cell sequencing has its drawbacks. The limited amount of genetic material in a single cell requires the use of whole-genome amplification (WGA), a process that induces technical noise and errors[1]. For example, chimeric artifacts arise due to the linkage of noncontiguous sequences[2]. Moreover, single-cell sequencing usually yields incomplete genomes with a significant proportion of contamination.

To overcome complications of single-cell sequencing authors developed a novel genome assembly workflow called ccSAG (Cleaning and Co-assembly of a Single-Cell Amplified Genome)[3]. ccSAG removes chimeric sequences and then co-assembles multiple single-cell amplified genomes (SAG) of related cells into nearly complete genomes. The investigation of mice gut microbiota using ccSAG revealed that the overall quality of obtained genomes was roughly equal to the quality of genomes assembled from conventional bulk DNA sequencing. The superiority of ccSAG to conventional chimaera removal tool jackknifing was presented in the paper. Also, the authors were able to capture a SNP within a single bacterial strain. They claim that this was the first time when heterogeneity was directly observed within an unculturable bacterial strain. Nevertheless, the authors acknowledge that to achieve high coverage and quality the computational workflow was enforced by the upstream single-cell sequencing technique called single-droplet MDA (multiple displacement amplification).

In summary, single-cell sequencing allows exploring the genetic diversity between individual cells. Implications include, but are not limited to, the fields of metagenomics and cancer research where tumour heterogeneity plays a significant role in the disease progression[4]. Complications that arise during single-cell genome assembly have been addressed and partially resolved with the newly developed computational pipeline called ccSAG. The pipeline combined with the novel experimental method called single-droplet MDA enables to achieve high-coverage and accuracy genomes from single cells.

Diana Lin
Mariia Radaeva

Connection Assignment

References

1. Wang, Y. & Navin, N. E. Advances and Applications of Single-Cell Sequencing Technologies. *Molecular Cell* vol. 58 598–609 (2015).
2. Lasken, R. S. & Stockwell, T. B. Mechanism of chimera formation during the Multiple Displacement Amplification reaction. *BMC Biotechnology* vol. 7 19 (2007).
3. Kogawa, M., Hosokawa, M., Nishikawa, Y., Mori, K. & Takeyama, H. Obtaining high-quality draft genomes from uncultured microbes by cleaning and co-assembly of single-cell amplified genomes. *Sci. Rep.* **8**, 2059 (2018).
4. Navin, N. E. Cancer genomics: one cell at a time. *Genome Biol.* **15**, 452 (2014).