Name: Panjun Kim
Email: pkim11@uthsc.edu
PI Name: Hao Chen
PI email: hchen3@uthsc.edu
Presentation preference: Poster presentation

De novo genome assembly for HXB/BXH recombinant inbred strains using Hi-C data

Panjun Kim[1], Hao Chen[2]
[1]Department of Genetics, Genomics and Informatics, University of Tennessee Health Science Center; [2]Department of Pharmacology, Addiction Science and Toxicology, University of Tennessee Health Science Center

Background: High-quality genome assemblies are crucial in genetic and genomic studies because they not only empower accurate identification of strain-specific variants, but also benefit pangenome research. HXB/BXH recombinant inbred (**RI**) strains have beneficial phenotypes being investigated in behavioral or physiological research, and the identification of causal variants for the phenotypes of interest can be facilitated by improving their quality of genomes.

Rationale: Hi-C is an assay to examine chromatin contact frequencies across a genome to identify three-dimensional (**3D**) genome architecture. However, the frequencies were also found to be applicable to *de novo* genome assembly as long-range linking information, and it enhances the genome assemblies at chromosome scale.

Results: We built 30 genome assemblies for the RI strains including their parent strains, BN-Lx and SHR/Ola, on the Supernova pipeline using 60x coverage of linked-reads from 10X Genomics. The genome assemblies achieved an average of 34,088 kb of contig N50 which is a unit used for genome contiguity. Among them, we improved scaffold N50 for the genomes of BN-Lx and SHR/Ola from 6.8 Mb to 36 Mb and from 2.2 Mb to 26 Mb using Hi-C data generated from the prefrontal cortex tissue of the strains.

Discussion: These enhanced genome assemblies can be used for base assemblies in pangenome and enable the pangenome to have the more complete dispensable genome as well as genetic diversity. The diverse genetic variations in the pangenome can contribute to finding novel causal variants for phenotypes of the RI strains.