Multi-platform genome assembly of an SHR/OlaIpcv X BN-Lx/Cub F1 rat trio

Andrea Guarracino[1], Rachel Ward[2], Flavia Villani[1], Burt Sharp[1], Farnaz Salehi[1], Denghui Chen[3], Abraham A. Palmer[3], Robert W. Williams[1], Vincenza Colonna[1], David G. Ashbrook[1], Hao Chen[2], Erik Garrison[1]

[1]Department of Genetics, Genomics and Informatics, University of Tennessee Health Science Center, Memphis, TN, USA. [2] Department of Pharmacology, Addiction Science, and Toxicology, University of Tennessee Health Science Center, Memphis, TN, USA. [3] Department of Psychiatry, University of California San Diego, San Diego, CA, USA

The Hybrid Rat Diversity Panel, comprising classic inbred strains and two recombinant inbred panels (FXLE/LEXF and HXB/BXH), offers a powerful resource for studying the genetic basis of complex traits in rats. Here, we present high-quality genome assemblies for an F1 hybrid of SHR/OlaIpcv X BN-Lx/Cub and its parental strains. This trio combines the genetic backgrounds of the HXB/BXH recombinant panel, offering potential insights into complex trait genetics and physiological mechanisms. We applied a multi-platform sequencing approach: PacBio HiFi reads for SHR/OlaIpcv (44X coverage) and BN-Lx/Cub (41X), and Oxford Nanopore Technology (ONT) reads for the F1 SHRxBN-Lx sample (81X coverage, including 36X Ultra-Long reads). All assemblies were generated using hifiasm. For SHRxBN-Lx, we explored two assembly strategies: 1) integrating parental HiFi reads with F1 ONT reads, and 2) using ONT reads only.

The assemblies show high contiguity, with contig N50 of 42.1 Mb for SHR/OlaIpcv and 35.6 Mb for BN-Lx/Cub. The F1 diploid assembly demonstrates even higher contiguity across different strategies. The ONT-only assembly achieves near-whole chromosome contigs, with contig N50 of 66.3 Mb and 131.9 Mb for haplotypes 1 and 2, respectively, surpassing the contig N50 of the latest GRCr8 reference (64.2 Mb) and approaching its scaffold N50 (137 Mb). HiFi+ONT integration yielded contig N50 of 101.7 Mb/99.6 Mb for haplotypes 1/2, respectively. BUSCO analysis (with the Glires dataset) shows high completeness: ~98% for parental HiFi-only assemblies, ~95% for ONT-only F1, and ~97% for HiFi+ONT F1. These results highlight the potential of combining multiple sequencing technologies and parental information to improve F1 assemblies. The assemblies of this trio will be made available to the research community, facilitating studies on complex trait mapping in rats.