Haplotype-based analyses of phylogeny and regional genome diversity in laboratory rats

Yanchao Pan[1], Hao Chen[2], Jun Z. Li[1]

[1] Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA. [2] Department of Pharmacology, Addiction Science, and Toxicology, University of Tennessee Health Science Center, Memphis, TN, USA.

The common rat (Rattus norvegicus) is a key species in biomedical research. We developed methods to study ancestral haplotypes and regional genome diversity in laboratory rats using whole-genome sequencing (WGS) data. As reported previously, an international team of researchers performed linked-read WGS and joint variants calling for 163 samples representing most commonly used laboratory rat strains, including classic inbred lines and several RI panels. We selected a subset of 33 samples, representing most of the distinct strains in the dataset, for haplotype analysis.

We confirmed that all 33 samples are effectively inbred, with nearly completely homozygous genomes. We generated an ancestral recombination graph using tsinfer, followed by our in-house algorithms for vertical and horizontal merging to remove the branches driven by recent mutation, and vertical renumbering to recreate the most likely haplotype genealogy. This resulted in a database of ancestral haplotypes detailing 1) haplotype block boundaries corresponding to historical recombination, 2) ancestral haplotype assignment for each sample at each block, and 3) the likely genealogy among the observed haplotypes.

While the panel represents 33 rat strains, most regions contain only 2 to 6 distinct ancestral haplotypes, with the five most common haplotypes comprising over 98% of the genomes in the 33 samples. This pattern reflects extensive sharing of ancestral genome materials when the modern-day strains were created. Our findings provide insights into the evolution of rat strains and inform regional-specific diversity to support haplotype-based genetic mapping studies. The methodologies we developed do not require a reference founder panel and should have wilder adaptability.