Abstract

Structural variants (SVs) are large-scale genomic alterations, including deletions, insertions, duplications, inversions, and translocations, that change the structure of chromosomes. By altering gene dosage, disrupting coding sequences, or creating novel gene fusions, SVs are more likely to affect gene function than single nucleotide variants (SNVs). Yet the functional and evolutionary impacts of SVs remain poorly understood because they are much more difficult to identify than SNVs. The advent of third-generation sequencing technologies has enabled the assembly of reference-quality genomes for a large number of individuals belonging to the same species, thereby alleviating the limitations in SV detection at the population level. We recently generated the *S. cerevisiae* Reference Assembly Panel (ScRAP), comprising reference-quality genomes from hundreds of strains that sample the species genomic space. ScRAP has revealed over 36,000 SVs, many of which affect protein-coding genes. However, the functional consequences of these variants on gene expression and protein evolution remain largely unexplored. This doctoral research project aims to (i) characterize the impact of SVs on gene repertoire evolution and (ii) quantify their effects on gene expression at the proteomic level. By leveraging ScRAP and newly generated quantitative proteomics data from 134 strains, we will assess how SVs influence protein abundance and uncover novel SV-derived proteins, including de novo and chimeric proteins. We will also investigate the structural properties of these newly emerged protein domains to assess the contribution of SVs to the evolution of protein folds. Finally, we will explore the evolutionary and adaptive significance of SVs by associating their presence with phenotypic traits across ecological niches. This interdisciplinary project integrates genomics, proteomics and structural bioinformatics. By combining expertise in yeast genetics and computational biology, we aim to provide novel insights into the role of SVs in gene expression, highlighting how large-scale structural rearrangements drive genome evolution and molecular innovation.