Efficient high-throughput sequencing for quantitative immune profiling using unique molecular identifiers

Genomics

Bryan Bell (bryan_bell@takarabio.com), Takara Bio USA, **Qiang Li**, Takara Bio USA, **Andrew Farmer**, Takara Bio USA

Next-generation sequencing (NGS) for immune repertoire profiling has become a powerful tool for understanding the role of the adaptive immune system in health and disease. Additionally, unique molecular identifiers (UMIs) have become a vital aspect of this approach and are used for preserving quantitative information (i.e., accurate clonotype counts) of the repertoire by removing PCR/sequencing errors and duplicates. Here, we integrated UMIs with our SMART technology to detect genuine low-frequency events in full-length variable regions of B-cell receptor (BCR) genes and T-cell receptor (TCR) genes. For each library, >90% of reads were on-target, and the most highly represented clonotypes remained consistent among the technical duplicates in the range of 10ng–1µg of input RNA or 50–10,000 cells. A sensitivity assay demonstrated RNA transcripts corresponding to multiple UMIs could be detected when spiked in to input RNA at a relative concentration of 0.001%. We also developed software to analyze multiple sequence data with UMIs to generate detailed stats for reads, clonotypes, UMIs, and mapping rate as output. The updated human TCR profiling kit, SMARTer Human TCRa/b Profiling Kitv2 (TCRv2), is designed to be compatible with any Illumina platform using 2x150bp reads. Moreover, unique dual indexes (UDIs) were incorporated to avoid crossover contamination caused by index hopping. Thus, our immune profiling technology can be used to observe clonal selection and hypermutation events in rare clonotypes found in blood and tumor tissues. These methods could also serve as a basis for the discovery of antibody-based therapeutics.