[BIOINFO 2024]

A Subset of Microsatellite Unstable Cancer Genomes Prone to Short Insertions over Deletions Is Associated with Elevated Anticancer Immunity

Sunmin Kim[1,2,3], Dong-Jin Han[1,2,3], Seo-Young Lee[1,2,3], Youngbeen Moon[1,2,3], Su Jung Kang[1,2,3], and Tae-Min Kim[1,2,3,4,*]

[1]*Department of Medical Informatics, College of Medicine, The Catholic University of Korea*
[2]*Cancer Research Institute, College of Medicine, The Catholic University of Korea*
[3]*Department of Biomedicine & Health Sciences, Graduate School, The Catholic University of Korea*
[4]*CMC Institute for Basic Medical Science, The Catholic Medical Center, The Catholic University of Korea*
**Corresponding author: tmkim@catholic.ac.kr*

Deficiencies in DNA mismatch repair (MMRd) leave characteristic footprints of microsatellite instability (MSI) in cancer genomes. We used data from the Cancer Genome Atlas and International Cancer Genome Consortium to conduct a comprehensive analysis of MSI-associated cancers, focusing on indel mutational signatures. We classified MSI-high genomes into two subtypes based on their indel profiles: deletion-dominant (MMRd-del) and insertion-dominant (MMRd-ins). Compared with MMRd-del genomes, MMRd-ins genomes exhibit distinct mutational and transcriptomic features, including a higher prevalence of T>C substitutions and related mutation signatures. Short insertions and deletions in MMRd-ins and MMRd-del genomes target different sets of genes, resulting in distinct indel profiles between the two subtypes. In addition, indels in the MMRd-ins genomes are enriched with subclonal alterations that provide clues about a distinct evolutionary relationship between the MMRd-ins and MMRd-del genomes. Notably, the transcriptome analysis indicated that MMRd-ins cancers upregulate immune-related genes, show a high level of immune cell infiltration, and display an elevated neoantigen burden. The genomic and transcriptomic distinctions between the two types of MMRd genomes highlight the heterogeneity of genetic mechanisms and resulting genomic footprints and transcriptomic changes in cancers, which has potential clinical implications.