[BIOINFO 2023]

Trajectory of lymph node metastasis characterizes molecular features of esophageal squamous cell carcinoma

Jiho Park[1], Seong Yong Park[2], Ha Eun Kim[3], Se-Young Jo[1], Jeongsoo Won[1],
Dae Joon Kim[3,*], Sangwoo Kim[1,*]

[1] Department of Biomedical Systems Informatics, Yonsei University College of Medicine
[2] Department of Thoracic and Cardiovascular Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine
[3] Department of Thoracic and Cardiovascular Surgery, Yonsei University College of Medicine
* Corresponding authors: swkim@yuhs.ac, kdjcool@yuhs.ac

Esophageal squamous cell carcinoma (ESCC) frequently accompanies lymph node metastasis to the neck, chest, and abdomen. The relationship between genomic and molecular characteristics and the trajectory of lymph node metastasis remains unknown, despite its significance as a key prognostic factor. We performed whole-exome sequencing (WES) and transcriptome sequencing (RNA-seq) from a cohort of 10 ESCC patients with 45 multi-regional samples (10 primary tumors, 10 normal esophageal tissues, and 25 lymph node tumors) to characterize the underlying patterns in lymph node metastasis and their associations with genomic characteristics. Phylogenetic analysis of genomic mutations in primary and lymph nodes reconstructed the temporal trajectory of invasion. Somatic mutations that were acquired before metastasis included major driver mutations including *TP53* and *KMT2D*, and exhibited a mutational process associated with alcohol consumption (SBS16, p=0.006 (Mann-Whitney U-test)). On the other hand, post-metastasis mutations were sporadic and associated with clock-like (SBS5, p=0.04 (Mann-Whitney U-test)) and defective DNA mismatch repair (SBS6, p=0.0022 (Mann-Whitney U-test)) signatures. The tumors within later metastasized lymph nodes acquired mutations at a faster rate (Spearman's $\rho = 0.79$, $p = 2.76 \times 10^{-6}$), indicative of higher genomic instability. We observed genomic patterns of nodal skip metastasis (NSM) in three of the nine patients (33%), two of which were only identifiable with trajectory-level information, emphasizing the possibility of misclassification when determining NSM. Transcriptome analysis revealed enrichment of epithelial-mesenchymal transition (EMT) ($FDR < 0.001$) and KRAS signaling ($FDR < 0.003$) pathways in NSM tumors, suggesting the underlying genomic characteristic for the poor prognosis. Our study provides a molecular understanding of lymph node metastasis and the potential usage of node skipping patterns in staging ESCC.