[BIOINFO 2024]

Genomic classification of intrapulmonary metastasis and multiple primary lung cancer

Jeongsoo Won[1], Yeon Seung Chung[2,3], Se-Young Jo[1], Jiho Park[1], Hyo Sup Shim[2,*], and Sangwoo Kim[1,*]

[1]*Department of Biomedical Systems Informatics, Yonsei University*
[2]*Department of Pathology, Severance Hospital*
[3]*Department of Pathology, Green Cross Laboratories*
**Corresponding author: HS Shim (shimhs@yuhs.ac), S Kim (swkim@yuhs.ac)*

The presentation of multifocal tumors at the time of lung cancer diagnosis exhibits either intrapulmonary metastasis (IPM) or multiple primary lung cancer (MPLC), the accurate discrimination of which is clinically important. However, the accuracy of current genome-based approaches is prone to platform diversity, which demands a more robust method for clinical use. We developed a new Bayesian probabilistic model, MeTel, that classifies IPM and MPLC based on genomic profiles. MeTel computes and compares the normalized likelihoods of IPM and MPLC from somatic mutation profiles, considering the size of sequencing panels, variant sharing, and ethnicity-specific background occurrence of mutations to enable robust and unbiased classification. When MeTel was tested on 686 multifocal lung cancers (194 IPM and 492 MPLC) from six independent cohorts, we achieved an accuracy of 98.40%, indicating that MeTel outperformed previous methods (88.05–94.60%). Only MeTel showed consistent accuracy (minimum accuracy of 93.75% vs. 0–78.95% in others) across various panel sizes (4–808 genes and whole exome). Application to 12 in-house patients showed 33% disagreement (4/12) between the initial histopathology-based and MeTel classifications, and the diagnoses of all cases were changed to MeTel prediction after re-adjudication. Our results verify the accuracy of the new sequencing-based classification algorithm of IPM and MPLC and justify its clinical utility in diagnosing multifocal lung cancer.