Unsupervised co-clustering unveils metabolic pathway aberrations linked with patient prognosis in head and neck squamous cell carcinoma

Giulia Toti1 , Kevin Ramzi Nasir1 , Vito Carlo Alberto Caponio2 , Le Minh Thao Doan3 , Gennaro Musella1 , Lorenzo Lo Muzio1 , Annalisa Occhipinti3 , Pier Paolo Claudio4
1Università degli studi di Foggia, Foggia, Italy, 2Link Campus University, Roma, Italy, 3Teesside University, Middlesbrough, United Kingdom, 4University Of Mississippi Medical Center, Jackson, United States

Background: Tumor heterogeneity and clinical outcomes in head and neck squamous cell carcinoma (HNSCC) are influenced by metabolic reprogramming. This study aimed to identify patient subgroups based on metabolic gene expression using biclustering and to develop a reproducible prognostic score. Method: Metabolic pathways were retrieved from the Kyoto Encyclopedia of Genes and Genomes (KEGG), and annotated genes were used to filter TCGA gene expression data from HNSCC patients. Spectral biclustering was performed in Python to identify patient clusters based on gene expression. These clusters were compared for Disease-Free Survival (DFS), Progression-Free Survival (PFS), Overall Survival (OS), and Disease-Specific Survival (DSS) using Kaplan–Meier analysis and Cox regression in IBM SPSS Statistics. A cluster score was calculated from five upregulated and five downregulated genes in cluster B versus A, ranked by |logFC| × −log10(adjusted p-value). The score was validated in independent CPTAC and GSE41613 datasets. Differentially expressed genes were identified using the R package limma. Results: Analysis of 500 patients and 418 genes identified two clusters (A = 274; B = 226). Cluster A was associated with worse DFS (HR = 1.45; p = 0.02), PFS (HR = 1.55; p = 0.003), and DSS (HR = 1.42; p = 0.05), whereas OS was not significant (HR = 1.14; p = 0.33). The highest-ranked upregulated genes were HSD11B1, IL4I1, CYP27A1, KMO, and ATP6V0D2, whereas the highest ranked downregulated genes were COX2, GAPDH, COX3, ENO1, and COX8A. The score discriminated the clusters with an absolute threshold > 1. Validation using independent datasets confirmed poorer PFS (HR = 4.39; p = 0.07) and OS (HR = 3.97; p = 0.02) among patients with expression profiles similar to cluster A. Conclusion: Unsupervised biclustering identified a patient group with poorer prognosis, and 10 genes may potentially serve as biomarkers or therapeutic targets in HNSCC.