The identification of combined genomic expressions as a diagnostic factor for oral squamous cell carcinoma

Ki-Yeol Kim, phD [1,*]

[1] Oral Cancer Research Institute, College of Dentistry, Yonsei University, SEOUL, 120-752, Korea

*Contact author: kky1004@yuhs.ac

Keywords: Oral squamous cell carcinoma, combined biomarker, microarray dataset

Trends in genetics are transforming in order to identify differential coexpressions of correlated gene expression rather than the significant individual gene. Moreover, it is known that a combined biomarker pattern improves the discrimination of a specific cancer. The identification of the combined biomarker is also necessary for the early detection of invasive oral squamous cell carcinoma (OSCC). To identify the combined biomarker that could improve the discrimination of OSCC, we explored an appropriate number of genes in a combined gene set in order to attain the highest level of accuracy. After detecting a significant gene set, including the pre-defined number of genes, a combined expression was identified using the weights of genes in a gene set. We used the Principal Component Analysis (PCA) for the weight calculation. In this process, we used three public microarray datasets. One dataset was used for identifying the combined biomarker, and the other two datasets were used for validation. The discrimination accuracy was measured by the out-of-bag (OOB) error. There was no relation between the significance and the discrimination accuracy in each individual gene. The identified gene set included both significant and insignificant genes. One of the most significant gene sets in the classification of normal and OSCC included *MMP1, SOCS3* and *ACOX1*. Furthermore, in the case of oral dysplasia and OSCC discrimination, two combined biomarkers were identified. The combined genomic expression achieved better performance in the discrimination of different conditions than in a single significant gene. Therefore, it could be expected that accurate diagnosis for cancer could be possible with a combined biomarker.

References

[1] A. de la Fuente (2010), From 'differential expression' to 'differential networking' - identification of dysfunctional regulatory networks in diseases, Trends Genet 26, 326-333.

[2] J.K. Choi, U. Yu, O.J. Yoo, S. Kim (2005), Differential coexpression analysis using microarray data and its application to human cancer, Bioinformatics 21, 4348-4355.