A Machine Learning Approach to Identify Key Tumor miRNA Expression Biomarkers Associated with Aggressive Prostate Cancer

Authors and Affiliations

Masuma Mannan[1], Hui-Yi Lin[1], PhD, Anand Paul, PhD.

[1] Louisiana State University Health Sciences Center, School of Public Health, New Orleans, LA 70112, USA.

Background

There is a range of clinical characteristics associated with prostate cancer (PCa), from aggressive disease with a high risk for metastasis to indolent tumors. Small non-coding RNAs known as microRNAs (miRNAs) control gene expression and have been linked to the aggressiveness of prostate tumors as well as the advancement of cancer. Finding miRNA signatures linked to aggressive PCa may help with risk assessment, therapy selection, and understanding of the molecular processes underlying the course of the disease.

Objective

The goal of this study was to identify key tumor miRNA expression profiles associated with aggressive PCa.

Methods

We analyzed 798 miRNA expression profiles from 320 localized prostate cancer patients (37.8% aggressive cases) from the Gene Expression Omnibus database. Data underwent quality control, normalization, and $\log_2$ transformation. We utilized the R software with the neuralnet package to implement an artificial neural network (ANN) to assess predictive power. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC). To identify significant markers, we performed univariate logistic regression for each miRNA. We screened 15 high-stringency candidates ($p < 0.01$) and applied correlation matrices to remove collinear features, resulting in a 14-miRNA subset. These features were then entered a stepwise multivariable logistic regression model to identify independent predictors associated with PCa aggressiveness.

Results

The ANN with 2 hidden layers (2 and 5 nodes) model achieved good predictive ability with an AUC of 0.718. Univariate analysis identified 61 miRNAs ($p<0.05$) significantly associated with aggressive PCa. Through the multivariable stepwise selection, we identified a significant 4-miRNA signature consisting of hsa-miR-940 (OR = 0.53, $p = 7.0 \times 10^{-3}$), hsa-miR-2278 (OR = 1.99 $p = 1.0 \times 10^{-2}$), hsa-miR-548a-5p (OR = 2.19, $p = 2.1 \times 10^{-2}$), and hsa-miR-582-5p (OR = 0.73, $p = 3.6 \times 10^{-2}$), which collectively showed strong discriminatory ability for PCa aggressiveness with an AUC of 0.710.

Conclusions

Both the standard logistic regression models and ANN support expression of specific tumor miRNAs associated with aggressive PCa. The identified signature includes miRNAs with established tumor-suppressive and oncogenic roles, providing a biological map of disease progression. These findings suggest that miRNA expression profiling is a valuable tool for predicting PCa aggressiveness, potentially enhancing personalized risk-stratification.

Recommendations

Our results support the integration of targeted miRNA profiling into clinical decision-making for the personalized management of PCa patients. Further validation in larger, diverse cohorts is recommended to facilitate clinical translation.

IRB

5543