

Analysis and Validation of Diagnostic Biomarkers and Immune Cell Infiltration Characteristics in Osteoarthritis by Integrating Bioinformatics and Machine Learning

Tianyang Li¹, Jinpeng Wei¹, Hua Wu², Chen Chen^{2, *}

Department of Orthopedics, Third Hospital of Shanxi Medical University, Taiyuan, China

Email address:

litianyang@sxmu.edu.cn (Tianyang Li), wjpsspine@163.com (Jinpeng Wei), wuhua@hust.edu.cn (Hua Wu), 13835120261@163.com (Chen Chen)

*Corresponding author

Abstract

Osteoarthritis (OA) is one of the most common musculoskeletal diseases, characterized by degeneration of articular cartilage accompanied by subchondral bone remodeling, synovial inflammation, and osteophyte formation. Current diagnostic methods rely heavily on clinical functional assessment and imaging techniques, which do not allow for early diagnosis of OA and may lead to diagnostic discrepancies due to subjective judgments. Exosomes, small extracellular vesicles of 30 – 150 nanometers in diameter secreted by various cell types, facilitate intercellular communication by transferring proteins, lipids, and RNA between cells and have been implicated in inflammation, chondrocyte activity, and cartilage degradation in OA; however, the specific mechanisms of exosome action in OA remain unclear. The aim of this study was to identify potential exosome-related biomarkers in OA and to explore the role and mechanism of exosome-related genes in articular cartilage. Gene expression profile microarrays of normal and OA cartilage samples were obtained from the Gene Expression Omnibus database, and exosome-related genes were retrieved from the GeneCards database. Differential gene expression analysis was performed using the limma package, and exosome-related differential genes (ERDEGs) were identified by intersecting differentially expressed genes with exosome-related genes. Multiple enrichment analyses were conducted to reveal potential mechanisms, and machine learning algorithms—including LASSO logistic regression, support vector machine-recursive feature elimination, and Random Forest—were employed to screen hub ERDEGs. A nomogram was constructed to assess diagnostic ability, and single-sample gene set enrichment analysis (ssGSEA) was used to evaluate immune cell infiltration characteristics. The results identified 105 ERDEGs, from which seven hub genes were screened: TOLLIP, ALB, HP, RHOBTB3, GSTM2, S100A8, and AKR1B1. Among these, AKR1B1, S100A8, and HP were significantly upregulated and RHOBTB3 was significantly downregulated in OA cartilage samples, as validated by qRT-PCR. The nomogram demonstrated high predictive ability for OA diagnosis, and ROC curve analysis confirmed that all seven hub genes exhibited high diagnostic value (AUC > 0.75). ssGSEA analysis revealed that eight types of immune cell infiltration and five types of immune cell activation were significantly increased in OA samples, and these hub genes were significantly correlated with multiple immune cell types and immune functions. In conclusion, this study identified seven exosome-related hub genes with excellent diagnostic potential for OA, which may serve as novel biomarkers and therapeutic targets; however, further experimental validation through large-scale cohorts and in vivo models is needed to support these findings.

Keywords

Osteoarthritis, Exosome, Machine Learning, Immune Infiltration, Biomarkers