Bioinformatics analysis based on the results of LC-MS/MS 🔗

📖 **Journal of Pain Research**

Lihui Li[1]

[1]Department of Sport Medicine and Rehabilitation Center, Shanghai University of Sport, Shanghai, China; Rehabilitation Research Laboratory 2rLab, Department of Business Economics Health and Social Care, University of Applied Sciences and Arts of Southern Switzerland, Manno, Switzerland

dx.doi.org/10.17504/protocols.io.upievke

Lihui Li ⚡

THIS PROTOCOL ACCOMPANIES THE FOLLOWING PUBLICATION

PROTOCOL STATUS

Working

Hierarchical Clustering

1 Cluster3.0 and the Java Treeview software were used to study protein relative expression data performing hierarchicalclustering analysis. Euclidean distance algorithm for similarity measure and average linkage clustering algorithm for clusteringwere selected when performing hierarchical clustering. A heatmap was created as a visual aid in addition to the dendrogram.

GO and KEGG Pathway Annotation

2 The protein sequences of differentially expressed proteins were in batches retrieved fromUniProtKB database in FASTA format; then the NCBI **BLAST+** software was used to find homologue sequences and transferred the functional annotation to the studied sequences (http://www.blast2go.com), which include three major categories: biological processes (BP), molecular function (MF), and cellular components (CC). In our study, we retrieved the top 20 blast hits (E-value◻1e-3) for each query sequence and loaded them into **Blast2GO**10 (Version 3.3.5) for Gene Ontology (GO) mapping and annotation.

Regarding Kyoto Encyclopedia of Genes and Genomes (KEGG), the FASTA protein sequences of differentially altered proteins between two groups were blasted against the online KEGG database (http://www.genome.jp/kegg/) and then $KEGG_{12}$was subsequently used for mapping pathways. The significantly overrepresented pathways and GO terms were extracted and classified.

The enrichment analysis was performed to further explore the impact of differentially expressed protein in cell physiological process and discover the mechanism of internal relations between them. GO enrichment on three categories (biological process, molecular function, and cellular component) and KEGG pathway enrichment were analyzed by the Fisher' exact test based on the entire quantified protein annotations as the background dataset. To adjust derived p-values, Benjamini-Hochberg correction was further applied for multiple testing. Only functional categories and pathways with p-values$\leq$0.05 were considered as significant.

Protein-Protein Interact Network◻PPI◻

3 The protein–protein interaction information of the studied proteins was retrieved fromIntAct molecular interaction database by their genesymbols or STRING software. The results were downloaded in theXGMML format and imported into Cytoscape software to visualize and further analyze functionalprotein-protein interaction networks. Furthermore, to evaluate the importance of the protein in the PPI network, the degree of each protein wascalculated.