http://dx.doi.org/10.7124/bc.000AEE

Network analysis identified novel disease module in rheumatoid arthritis

V.V. Bashynska[1,2], P. Fau[3], A.S. Romanenko[3,2], A.O. Kompaniiets[3,2], P.V. Havrysh[4,2], M.V. Mykhailichenko[5,2], D.O. Nishchenko[6,2], N.M. Kasianchuk[3,2], F. Iordachi[7], O.O. Petrenko[8,2]

[1] O.M. Marzeiev Institute for Public Health, NAMS Ukraine
50, Hetman Pavlo Polubotka Str., Kyiv, Ukraine, 02094
[2] NGO Genomics UA
3, Dorogozhitska Str., Kyiv, Ukraine, 04119
[3] Adam Mickiewicz University
1, Wieniawskiego Str., Poznan, Poland, 61–712
[4] Institute of Plant Genetics, Polish Academy of Sciences
34, Strzeszynska Str., Poznan, Poland, 60–479
[5] University of Wroclaw
1, Uniwersytecki plac, Wroclaw, Poland, 50–137
[6] Taras Shevchenko National University of Kyiv
60, Volodymyrska Str., Kyiv, Ukraine, 01033
[7] Stanford University's OHS
415, Broadway, Redwood City, CA, United States, 94063
[8] Medical University of Vienna
23, Spitalgasse, Vienna, Austria, 1090
oleksandr.petrenko@meduniwien.ac.at

Background. Rheumatoid arthritis (RA) is an autoimmune disease causing chronic joint inflammation and is a leading cause of disability. Despite existing therapies, unmet needs like refractory disease remain a patients' burden. This study aims to explore molecular associations with RA to identify new biomarkers or molecular targets. **Methods.** We assessed total RNA sequencing from synovial biopsies (GEO GSE89408: 28 controls, 152 RA; ENA E-MTAB-6141: 32 controls, 122 RA) and synovial fluid LC-MS/MS proteomics (PXD000740, 20 RA). Transcriptomic data were analyzed with DESeq2, defining differentially expressed genes (DEGs) by $L_2FC \geq |2|$, $P_{adjusted} \leq 0.05$. We constructed gene co-expression network (GCN) layer between DEGs and inferred transcription factors with GENIE3 to construct gene-regulatory layer (GRN). Protein-protein interaction networks (PPIN) were inferred using STRINGdb (≥0.7 confidence) and filtered by proteomics data. The genomic links related to RA were exported from DisGeNET (75th percentile) and GWAS (Ishigaki *et al.*, Nature 2022). The drugs affecting PPIN nodes were assessed from DGIdb and added as a separate layer. The RA network module (n = 25) identification was implemented using the maximum clique centrality and density of maximum neighborhood component scores. EnrichR was used for gene set analysis of the 1st neighbors. **Results.** The multi-layer network consisted of 187120 edges (both intra- and inter-layer links) and 1235 nodes: GRN (67), GWAS (14), DisGeNET (3), GCN (758), PPIN (292), DGIdb (101). The GCN was the most intra- and inter-connected layer. Average clustering coefficient (0.2) and centralization (0.37) indicated presence of separated communities, mainly from drug-protein links, with the largest cluster in the GCN. The double-score screening identified the RA disease module, including known associations (CCR2, CTLA4, CXCR6) and novel candidates (CD2, CD80, CRTAM, LAMP3). The disease module's functional neighborhood involved the cytokine signaling, neutrophil migration, and T-cell activation. **Results** are reposited on GitHub: github.com/xander-p/eebg24-networks-project. **Conclusions.** The identified molecular associations suggest novel links in RA, offering potential targets for mechanistic studies to understand their role in the disease pathobiology.

K e y w o r d s: rheumatoid arthritis, network analysis, multiomics