**Background**

Systemic Lupus Erythematosus (SLE) is frequently complicated by co-morbidities such as cardiovascular disease and stroke. Understanding shared biology between SLE and its co-morbidities could increase knowledge of the pathogenesis of SLE, guide tailored treatment decisions, and delineate novel disease biomarkers. Here, we use big data analytics to examine numerous relevant large-scale datasets.

**Methods**

We identified chronic kidney disease, stroke, cardiovascular disease, osteoporosis, lung cancer, Hodgkin lymphoma, and viral infection as relevant co-morbidities of SLE in a UK patient population (PMID: 26473719), and searched the Gene Expression Omnibus for gene expression datasets of SLE and these diseases. Datasets with over 30 human subjects in a case-control design were retained for analysis. Previously published methods (PMID: 27842596) were used to compute differentially expressed genes (DEGs) per dataset and then per disease. Next, we identified enriched biological pathways in each disease and computed which pathways are enriched across SLE and all co-morbidities contained in this analysis.

**Results**

36 datasets passed the filter criteria. These contained expression data from 4,776 individuals (3,320 cases, 1,456 controls). Number of DEGs per disease varied from a minimum of 2 (osteoporosis) to a maximum of 5,408 (lung cancer) genes (median 2,644, IQR 4,608). Subsequent numbers of enriched KEGG biological pathways per disease varied from 0 (osteoporosis) to 43 (stroke) pathways (Median 12, IQR 18). Extracting KEGG pathways that were significantly enriched across all diseases provided a set of 7 biological pathways (see table), of which ‘SLE’ (P=5.0x10-7) and ‘Antigen Processing & Presentation’ (P=5.2x10-7) showed the largest enrichment. Other enriched pathways included those representing natural killer (NK) cell and spliceosome activity.

|  |  |
| --- | --- |
| **KEGG Pathway** | **Enrichment P-Value** |
| SLE | 5.0x10-7 |
| Antigen Processing & Presentation | 5.2x10-7 |
| NK Cell Mediated Cytotoxicity  | 5.9x10-6 |
| Spliceosome | 8.4x10-6 |
| Leishmania Infection | 2.5x10-4 |
| Ribosome | 4.5x10-4 |
| Graft Versus Host Disease | 2.5x10-3 |

**Conclusions**

The significant enrichment of KEGG pathway gene sets across SLE and its co-morbidities suggest that these diseases may share an underlying molecular architecture. The detection of the ‘SLE’ KEGG pathway as enriched in these data serves as internal validation of this method. Additionally, subsets of NK cell populations exhibit unique phenotypes in patients with active SLE, and high levels of anti-spliceosomal autoantibodies are commonly found in serum from SLE patients.

Further analysis of these data may be able to guide treatment decisions when SLE is complicated by co-morbidities, or delineate gene expression signatures that could prompt focused screening of patients with SLE for relevant co-morbidities.