The Identification of Significant Genes Related to Systemic Lupus Erythematosus through the Integration of the Results of a Transcriptome-Wide Association Study and an mRNA Expression Profile Analysis

Shuai Li
Department of Orthopedics, Orthopedic Research Institute, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Jiawen Xu
Department of Orthopedics, Orthopedic Research Institute, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Ye Tian
Healthy Food Evaluation Research Center, West China School of Public Health and West China Fourth Hospital, Sichuan University

Yi Zeng
Department of Orthopedics, Orthopedic Research Institute, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Haibo Si
Department of Orthopedics, Orthopedic Research Institute, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Yuangang Wu
Department of Orthopedics, Orthopedic Research Institute, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Shaoyun Zhang
Department of Orthopedics, Orthopedic Research Institute, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Bin Shen (shenbin_1971@163.com)
Department of Orthopedics, Orthopedic Research Institute, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Research Article

Keywords: Transcriptome-wide association study, Systemic lupus erythematosus, mRNA expression profiles, Gene ontology, Pathway enrichment analyses

Posted Date: November 3rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2216395/v1

Abstract

Background

Systemic lupus erythematosus (SLE) is a polygenic autoimmune connective tissue disease in which heritable components play an essential role in the pathogenesis. However, the correlation between genetic variants and pathological changes in SLE is still unclear, and it is difficult to provide insights for the early diagnosis and treatment of SLE.

Methods

We conducted a transcriptome-wide association study (TWAS) of SLE by integrating a genome‑wide association study (GWAS) summary dataset of SLE (538 diagnosed patients and 213,145 controls derived from the FinnGen consortium). To verify the results of the TWAS analysis, the significant genes were further compared with the mRNA expression profiles of SLE to screen for common genes. Finally, significant genes were analyzed using functional enrichment and annotation analysis in Metascape to examine SLE-related gene sets.

Results

The TWAS identified 30 genes with $P_{TWAS-adjusted}$ values < 1.33×10^{-6} ($0.05/37665 = 1.33\times10^{-6}$), including *HCP5* ($P_{TWAS}$ =8.74×10^{-15}) and *APOM* (P_{TWAS} = 4.57×10^{-12}). Four common genes were identified through the comparison of the TWAS results with the differentially expressed genes (DEGs) of SLE, including *APOM* (P_{TWAS} = 4.57×10^{-12}, P_{DEG} = 3.31×10^{-02}) and *C2* (P_{TWAS} = 8.04×10^{-11}, P_{DEG} = 1.54×10^{-02}). Moreover, 36 terms were detected for the enrichment results of the TWAS, including antigen processing and presentation (logP value = -4.1938). By integrating the pathway and process enrichment analysis results of DEGs, 17 terms were identified, including allograft rejection (logP value = -7.5738).

Conclusion

The study identified a group of SLE-related genes and pathways, and the findings provide novel insights for the early diagnosis and intervention of SLE.

Introduction

Systemic lupus erythematosus (SLE) is a polygenic autoimmune connective tissue disease for which the etiology has not yet been determined (1). Abnormalities in immune cell phenotypes and function cause the presence of nuclear autoantibodies, immune complex formation, and the inflammation of multiple organs in the pathogenesis of SLE (2). Multiple organ systems are affected, leading to the skin, joint, or

hematological manifestations that are observed in SLE patients (3). SLE predominantly affects young women and is among the leading causes of death in young females in the US (4). The overall global incidence of SLE ranges from 1.5 to 11 per 100,000 person-years, and the global prevalence ranges from 13 to 7,713.5 per 100,000 individuals (5). SLE patients, especially those with moderate or severe disease, require significant medical resources and incur numerous medical costs (6). However, the underlying mechanism of SLE development is still elusive because of the heterogeneity of the disease (1).

A heritable component was thought to play an essential role in the pathogenesis of SLE. In a comparative study, co-twins had a 76-fold increased risk of SLE and a 2.7-fold increased risk of any autoimmune disease (7). The heritability of SLE ranges from 44–66% (8, 9), and many genes potentially play pathogenic roles in the aberrant immunity and cellular processes of SLE (10). Although the inflammatory responses in individuals susceptible to SLE are regulated by both environmental and genetic factors, the genetic factors are not affected by environmental factors and predate the clinical manifestation of SLE (11). Gene analysis is an effective method for the early diagnosis and screening of people at high risk of SLE and the choice of the best treatment based on the patient's biological factors (12).

In previous genome-wide association studies (GWAS), a large number of SLE genetic markers were scanned across the genome, and functional pathways were found that revealed critical molecular pathways representative of each population (13, 14). It was found that the ETS1 gene was correlated to curbing the terminal Tfh2 cell differentiation process and in turn influenced disease parameters in SLE patients (15). Gene expression is a key step linking DNA sequence variation to phenotypes, GWAS have linked thousands of genomic loci to complex traits (16). However, GWAS are rarely ascertainable from identify causal genes that lead to trait changes (17). In 2018, a new omics analysis method, transcriptome-wide association studies (TWAS), emerged, which leverage expression reference panels (eQTL cohorts with expression and genotype data) to discover gene–trait associations in GWAS datasets, providing a powerful strategy that integrates GWAS results and gene expression references to identify significant expression-trait associations (18–20). Compared with GWAS, TWAS can remove most of the meaningless results obtained by GWAS, so that candidate causal genes can be screened and prioritized more accurately (18, 21). In recent years, TWASs have been widely used to identify risk genes in a variety of autoimmune diseases. For example, by using TWAS analysis, Díez-Obrero V et al. provided insight into the tissue-specific molecular processes underlying inflammatory bowel disease genetic susceptibility (17).

In the current study, by integrating an SLE GWAS summary statistics derived from the FinnGen consortium and precomputed gene expression weights of cross-tissue features (sCCA features), we conducted a TWAS analysis to identify significant genes related to SLE. To validate the TWAS results, the significant genes identified by the TWAS were further compared with the mRNA expression profiles of SLE. Finally, we reevaluated the expression of the TWAS-identified genes and performed a functional examination. The findings provide novel insights for the early diagnosis and intervention of SLE by identifying genetic variants related to pathological change

Methods And Materials

SLE GWAS summary data

Recent large-scale GWASs and meta-analyses in European populations of SLE were used here (22). The GWAS summary statistics of Europe was obtained from the FinnGen consortium (study page: https://www.finngen.fi/en/; release 5: https://r5.finngen.fi/), which was launched in Finland in 2017, including 538 diagnosed SLE patients and 213,145 controls of Finnish ancestry (22). All cases were defined by the code M13 in the International Classification of Diseases—Tenth Revision. These individuals were genotyped with Illumina and Affymetrix chip arrays (Illumina Inc, San Diego, and Thermo Fisher Scientific, Santa Clara, CA, USA) and typed at 16,962,023 variants analyzed in total. Detailed information on the participants, genotyping, imputation, and quality control can be found on the FinnGen website (22).

TWAS analysis of SLE

The cross-tissue TWAS analysis of SLE was carried out by using Functional Summary-based Imputation software (FUSION http://gusevlab.org/projects/fusion/) by integrating the SLE GWAS summary statistics and precomputed gene expression weights. FUSION can evaluate the gene expression associations between each gene and target disease (23). Specifically, the gene expression weights of sCCA features were calculated using the FUSION prediction models. Different from single-tissue TWAS analysis, sCCA features integrate eQTL data across multiple tissues such that sCCA features increase the power of cross-tissue TWAS (24) FUSION computed TWAS expression weights by using five linear models, including BLUP, BSLMM, LASSO, Elastic Net, and the top SNPs from the reference expression panels (i.e., GTExv8). When performing transcriptomic imputation, FUSION calculated an out-sample R2 using fivefold cross-validation of each model to determine the best performing prediction model for a gene. Then, the sCCA features were combined with the GWAS results to impute the association statistics between gene expression levels and target diseases. The association testing statistics between predicted gene expression and target diseases were calculated as Z TWAS = w'Z/(w'Lw)1/2, where ‘Z’ denotes the scores of SLE, ‘w’ denotes the weights, and ‘L’ denotes the SNP-correlation linkage disequilibrium (LD) matrix (23). In the present study, a TWAS P value was calculated for each gene within cross-tissue features for European populations. For the TWAS analysis, we adopted a Bonferroni-corrected p value < 1.33×10 − 06 (0.05/37665 = 1.33×10 − 06) to determine statistical significance.

mRNA expression profiles of SLE

The differentially expressed genes (DEGs) were derived from genome-wide mRNA expression profiles of SLE. The SLE mRNA expression data were downloaded from the Gene Expression Omnibus (GEO) datasets (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE181500) and the corresponding reference (25). In brief, peripheral blood samples were obtained from 6 female patients (mean age 32 ± 9.8 years, range from 24 to 45 years) diagnosed with SLE according to the classification criteria of the

American College of Rheumatology and 6 female age-matched healthy controls (25). The mRNA expression microarray was processed according to the human 4x180k long noncoding RNA array (Agilent Technologies), and quantile normalization and subsequent data processing were carried out using the GeneSpring GX v11.5.1 software package (Agilent Technologies) (25). In this study, DEGs were analyzed by the GEO2R tool. GEO2R has a simple interface that allows users to perform sophisticated R-based analysis of GEO data to help identify and visualize DEGs (26). DEGs were identified through filtering with the threshold setting of a |log fold change (LogFC) | > 1 and an adjusted P value < 0.05 by the moderated t statistic.

Functional enrichment and annotation analysis

In this study, Metascape (https://metascape.org/) was used to perform the functional enrichment and annotation analysis of the genes identified by the TWAS analysis and DEGs. Metascape was designed to allow researchers to apply powerful computational analysis pipelines to analyze and interpret large-scale datasets, facilitating functional exploration that includes Gene Ontology (GO) and pathway analysis (27). In Metascape, we searched the gene symbols of common targets by limiting the species to "*Homo sapiens*". Pathway and process enrichment analysis was carried out with the following ontology sources: GO biological processes, KEGG pathways, GO molecular functions, reactome gene sets, canonical pathways, and CORUM. Enrichment analysis was based on Fisher's exact test and the calculation of P values. Terms with P < 0.05 were considered significant.

Results

TWAS results of SLE

A total of 30 SLE-related genes were identified by cross-tissue TWAS (p value < 1.33×10^{-06}), such as *HCP5* (*P*adj = 8.70×10^{-15}), *APOM* (*P*adj = 4.57×10^{-13}), and *C4B* (*P*adj = 5.56×10^{-14}) (Fig. 1, Supplemental Table 1). Several of the most significant genes were located on chromosome 6, and the rsIDs of the most significant GWAS SNP in the locus (BEST.GWAS. ID) were SNP rs3130557 and SNP rs1270942. If these alleles are mutated, they alter the expression of downstream target genes. This study indicated that SNP rs1270942 and SNP rs3130557 were the causative genetic variants in SLE. Table 1 presents the detailed information of the top 10 significant genes identified by the TWAS, including the heritability of genes (HSQ), rsID of the most significant GWAS SNP in the locus (BEST.GWAS. ID), number of SNPs in the locus (NSNP), and TWAS P value (P TWAS).

Table 1
Top 10 genes identified by TWAS analysis

Gene	CHR	BEST.GWAS.ID	NSNP	TWAS.Z	TWAS.*P*
HCP5	6	rs1270942	272	7.7570	8.70×10^{-15}
C4B	6	rs1270942	187	-7.5180	5.56×10^{-14}
PPT2	6	rs1270942	164	-7.5180	5.56×10^{-14}
FLOT1	6	rs3130557	77	-7.4424	9.89×10^{-14}
CLIC1	6	rs1270942	240	-7.3600	1.84×10^{-13}
APOM	6	rs1270942	240	7.2375	4.57×10^{-13}
C4A	6	rs1270942	220	-7.0745	1.50×10^{-12}
SAPCD1	6	rs1270942	239	7.0313	2.05×10^{-12}
PPP1R18	6	rs3130557	77	6.9620	3.35×10^{-12}
CYP21A1P	6	rs1270942	195	-6.7817	1.19×10^{-11}

Note: Table 1 presented the detailed information of the top 10 significant genes identified by TWAS, including rsID of the most significant GWAS SNP in the locus (BEST.GWAS.ID), number of SNPs in the locus (NSNP), TWAS Z score (TWAS.Z) and TWAS *P* value (P_{TWAS}). The TWAS.*P* and TWAS.Z values were calculated by the FUSION approach (http://gusevlab.org/projects/fusion/)

TWAS: Transcriptome-Wide Association Study; GWAS: Genome-Wide Association Study

Validating the TWAS results by the SLE mRNA expression profiles

The mRNA expression profiles of SLE screened 2999 DEGs, among which 1149 were downregulated and 1850 were upregulated (Supplemental Table 2). After comparing the significant genes of the TWAS analysis and the DEGs, 4 common genes were selected and are shown in Table 2, including *APOM* ($P_{TWAS} = 4.57\times10^{-13}$ $P_{DEG} = 3.31\times10^{-02}$), *C2* ($P_{TWAS} = 8.04\times10^{-11}$, $P_{DEG} = 1.54\times10^{-02}$), *PPP1R10* ($P_{TWAS} = 4.27\times10^{-07}$, $P_{DEG} = 1.08\times10^{-04}$) and *MICB* ($P_{TWAS} = 1.05\times10^{-06}$, $P_{DEG} = 2.23\times10^{-05}$). The distribution of DEGs identified from mRNA expression profiles was visualized in the corresponding volcano plot (Fig. 2).

Table 2
The common genes identified by both TWAS and DEGs for SLE

Gene	CHR	NSNP	P_{TWAS}	P_{DEG}	logFC	Regulation
APOM	6	240	4.57×10^{-13}	3.31×10^{-02}	1.1164	UP
C2	6	224	8.04×10^{-11}	1.54×10^{-02}	-1.0707	DOWN
PPP1R10	6	173	4.27×10^{-07}	1.08×10^{-04}	2.8921	UP
MICB	6	275	1.05×10^{-06}	9.07×10^{-04}	-1.8414	DOWN

Note: Each PTWAS value was calculated by analysis of a transcriptome-wide association study (TWAS). Each PDEG value was the differentially expressed gene (DEG) derived from the published studies.

TWAS, Transcriptome-Wide Association Study; DEG, Differentially Expressed Gene; SLE, systemic lupus erythematosus; P_{TWAS}, $P_{\text{Transcriptome–Wide Association Study}}$ value; P_{DEG}, $P_{\text{Differentially Expressed Gene}}$ value;

Gene set enrichment analysis of the TWAS results

The significant genes identified by cross-tissue TWAS analysis were submitted to Metascape for functional enrichment and annotation analysis. We identified 36 significant terms enriched for the TWAS results, such as antigen processing and presentation (hsa04612, P value = 6.40×10^{-05}) and complement activation (WP545, P value = 1.35×10^{-6}). The top 10 terms enriched for the TWAS results are shown in Table 3, and all the results of the enrichment analysis are shown in supplemental table 3. The Sankey diagram and dot plot showed the top significant GO terms and related genes for the TWAS analysis (Fig. 3). By integrating the results of the enrichment analysis of DEGs, 17 terms were identified, including allograft rejection (WP2328, P value = 2.67×10^{-08}) (Supplemental Table 4).

Table 3
Significant terms identified by both TWAS analysis

Terms ID	Description	P_{value}
WP2328	Allograft rejection	2.67×10^{-08}
hsa05150	Staphylococcus aureus infection	3.91×10^{-08}
GO:2000427	positive regulation of apoptotic cell clearance	4.94×10^{-08}
R-HSA-174577	Activation of C3 and C5	4.94×10^{-08}
GO:2000425	regulation of apoptotic cell clearance	1.06×10^{-07}
hsa05322	Systemic lupus erythematosus	2.24×10^{-07}
GO:0016064	immunoglobulin mediated immune response	7.39×10^{-07}
GO:0019724	B cell mediated immunity	8.27×10^{-07}
hsa04612	Antigen processing and presentation	6.40×10^{-05}
GO:0043603	cellular amide metabolic process	7.48×10^{-03}

Note The significant genes identified by TWAS analysis were analyzed by Metascape tool (https://metascape.org/gp/index.html#/) for functional enrichment and annotation analysis.

TWAS, Transcriptome-Wide Association Study; GO, gene ontology; hsa, KEGG Pathways; WP: Wiki Pathways; R-HSA: Reactome Gene Sets

Discussion

SLE is characterized by the abnormal functioning of T and B cells, autoantibody production, and immune complex deposition, ultimately leading to multiorgan damage. A genetic component plays a significant role in the etiology of SLE (28), and the genetic basis of SLE was partially revealed by previous GWASs (29, 30). For example, through a GWAS, Cui et al. found 4 genes related to T-cell signaling, including protein phosphatase nonreceptor type 22 (*PTPN22*), implicating that these gene pathways are important in the pathogenesis of SLE [30]. In addition, the combined data from GWASs and inhibition assays implicated autophagy in SLE (30). Compared with single nucleotide polymorphism (SNP)-based GWAS, TWAS can take into account eQTLs, especially eQTLs in noncoding regions of the human genom (31), reducing the multiple-testing burden and directly implicating the gene-based mechanisms underlying complex traits (19). TWAS analysis has been widely used to identify risk genes for autoimmune diseases such as inflammatory bowel diseases, and the results have provided a better understanding of the genetic pathogeneses of these diseases (29).

By integrating the results of a TWAS analysis and an SLE mRNA expression profile analysis, we identified several common genes, such as *MICB, C2, and APOM*. Interestingly, most significant genes are located on chromosome 6, suggesting the significant role of chromosome 6 in the pathogenesis of SLE, which is consistent with the existing research (32). A variety of immune-related genes have been found on chromosome 6, such as major histocompatibility complex(*MHC*) (33), and some of them provide new ideas for the pathogenesis and early diagnosis of SLE. For example, prolactin (PRL), whose gene is near the HLA region on the short arm of chromosome 6, is a versatile hormone mainly produced in the anterior pituitary gland that has multiple functions. Hyperprolactinemia (HPRL) has been demonstrated in 20–30% of SLE patients and is related to active disease (34). The findings of this study suggest a role for hormones in the pathogenesis of SLE.

Major histocompatibility complex class I-related chains B (*MICB*) is a member of natural-killer group 2, member D ligands (*NKG2DL*), whose ligand engagement on tissue-resident effector lymphocytes promotes cell damage and inflammation (35). MICs can be shed from the cell surface to generate soluble MICs (sMICs) (36). It was found that soluble MICB (sMICB) plasma values were negatively correlated to disease activity scores in juvenile-onset SLE, suggesting clinical relevance (37). In addition, sMICs may be related to activated NK cells migrating to inflamed tissue in active SLE and dropping circulating NK cells (38). The expression of *NKG2D* ligands, including *MICB*, can be adjusted by numerous genes, such as c-Myc, and ultimately influence the function of NK cells (39). Zhang et al. found higher mRNA expression of *MICB* in B cells, monocytes, and renal biopsies from SLE patients than in those from controls in the Chinese population (40).

Complement system dysfunction plays a significant role in the pathogenesis of SLE. Complement participates in internal homeostasis and assists in the disposal of dead cells, immune complexes, and infectious microbes (41); a failure to clear autoantigens and defective waste caused by complement deficiency may be the first step of SLE (42). Complement 2 (*C2*) is an important link in the classical and mannose-binding lectin (MBL) pathways of complement activation (43) and provides defense against microbial infection and assists in the removal of immune complexes (44). *C2* defects are considered to be the most common complement deficiency and are inherited in an autosomal-recessive manner (45). Approximately 10–30% of homozygous *C2*-deficient patients develop SLE (46). SLE patients with *C2* deficiency mainly have manifestations of musculoskeletal, mucocutaneous, cutaneous, and cardiovascular-related damage (47).

The human apolipoprotein M (*APOM*) gene is located in a highly conserved segment in the major histocompatibility complex (MHC) class III locus on chromosome 6, which is close to genes related to the immune response (48). *APOM* is mainly related to HDL, which shows impaired vasculoprotective effects (49). The plasma APOM level was found to be downregulated by the inflammatory processes in active SLE, and low APOM levels were related to markers of inflammation, such as CRP and C3, which are indicators of SLE activity (50, 51). In addition, APOM is the physiological carrier of sphingosine-1-phosphate (S1P) (52). In the pathogenesis of SLE, immune complex (IC) deposition activates neutrophils (PMNs), increases vascular permeability, and leads to organ damage (53). APOM-Fc, a novel S1P

chaperone, was found to limit leukocyte escape from capillaries and protect against inflammatory injury, suggesting the therapeutic potential of APOM through attenuating tissue responses in SLE (54).

Our study also identified several significant biological pathways related to SLE. Most of these pathways were related to the immune system, such as allograft rejection and antigen processing and presentation. Allograft rejection (WP2328) is caused by recipient T-cell recognition of nonself donor alloantigens (55). All of the antibodies, T cells and complement activation were thought to be involved in the mechanism of allograft rejection, as observed in SLE. Some SLE patients may require organ transplantation due to disease progression. For example, the 5-year incidence of end-stage kidney disease (ESKD) in patients with lupus nephritis is 11% despite novel and potent therapeutic regimens (56), and the clear superiority of renal transplantation regarding prolonged survival and better quality of life for SLE patients has been demonstrated by numerous studies (57). However, the management of an SLE patient who has undergone transplantation can be more complex when immunity dysregulations coexist. In this study, we identified genes differentially expressed between SLE patients and healthy individuals involved in allogeneic immune rejection, and therapeutic measures targeting these genes may be more applicable to SLE patients who have undergone transplantation, especially considering the possibility of recurrent SLE in kidney transplant patients. Some drugs that modulate immune cells, such as the B-cell–depleting agent rituximab, have been shown to have significant therapeutic effects on SLE and immune rejection. Additionally, antigen processing and presentation (hsa04612) was identified as a pathway related to SLE. Antigen processing and presentation refers to the process by which antigens are captured and phagocytosed by antigen-presenting cells (APCs) and presented to lymphocytes in a recognizable form. The hyperactivation of APCs, including monocytes/macrophages, B cells, and dendritic cells (DCs), has been found in SLE patients and resulted in the incorrect recognition of autoantigens (58). As a result, biotherapeutic strategies targeting APCs have become a hot spot in the treatment of SLE (59). However, the development and implementation of new therapies for SLE have lagged behind those of other rheumatic diseases, and many biologic drugs cannot reach the expected therapeutic effect in clinical trials (60). In this study, the genes we identified as being differentially expressed between SLE patients and healthy controls were closely related to antigen processing and presentation. These promising molecular pathways and targets for the biotherapeutic treatment of SLE will provide new directions for future investigations.

The novelty of this study is that we used a new omics analysis method, TWAS, to explore the genetic mechanisms underlying SLE. TWAS analysis is a creative and valuable method that can integrate publicly available GWAS summary data and expression quantitative trait loci (eQTL) reference datasets to evaluate gene-trait relationships (61). In addition, the large sample size of the GWAS summary statistics ensures the accuracy of our results, and the results were further validated by integrating the results of an mRNA expression profile analysis. However, there are some limitations of our study. First, it is easy to miss causal variants without cis-gene expression effects on SLE. Second, the major significant genes we found are located on major histocompatibility complex (MHC) locus in the chromosome 6. However, the genetic variation in the MHC locus is so complicated that we should be cautioned to use this genes (such human leukocyte antigen genes) and molecular mechanisms.

Conclusion

In summary, by using a GWAS summary statistics, TWAS analysis identified significant and common susceptibility genes for SLE. Our results provide novel clues for understanding the underlying genetic mechanism of SLE, focusing on the possible roles of genes in the pathogenesis of SLE. This study also provides a new diagnostic and treatment strategy for SLE patients who have undergone organ transplantation. In addition to the specific mechanistic findings for SLE, this work outlines a systematic approach for identifying functional mediators of complex trait diseases.

Abbreviations

APCs: antigen-presenting cells

APOM: apolipoprotein M

C2: complements 2

DCs: dendritic cells

DEGs: differentially expressed genes

eQTL: expression quantitative trait loci

ESKD: end-stage kidney disease

GEO: gene expression omnibus

GO: gene ontology

GWAS: genome-wide association studies

MHC: major histocompatibility complex

MICB: major histocompatibility complex class I-related chains B

PRL: prolactin

HPRL: hyperprolacctinemia

SLE: systemic lupus erythematosus

sMICB: soluble MICB

SNP: single nucleotide polymorphism

TWAS: transcriptome-wide association study

Declarations

Ethics approval and consent to participate

The source of the data was a publicly available data base and no human participants were involved, hence ethical parameters are not applicable.

Consent for publication

Not applicable

Availability of data and supporting materials

The GWAS summary statisticss are available from the FinnGen study(https://r5.finngen.fi/). The datasets analyzed during the current study are available from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/gds) accession number: GSE181500.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (grant number 81974347 and 81802210); the Department of Science and Technology of Sichuan Province (grant number 2021YFS0122). Financial support had no impact on the outcomes of this study.

Authors' Contributions

(I) Conception and design: Author 1, Author 8

(II) Administrative support: Author 1, Author 8

(III) Provision of study materials: Author 1, Author 3, Author 4

(IV) Collection and assembly of data: Author 1, Author 4,Author 5

(V) Data analysis and interpretation: Author 1, Author 2, Author 5, Author 6, Author 7, Author 8

(VI) Manuscript writing: Author 1, Author 2

(VII) Final approval of manuscript: All authors

Acknowledgements

Not applicable

References

1. Dörner T, Furie R. Novel paradigms in systemic lupus erythematosus. Lancet 2019; 393:2344–2358.
2. Durcan L, O'Dwyer T, Petri M. Management strategies and future directions for systemic lupus erythematosus in adults. Lancet 2019; 393:2332–2343.
3. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Annals of the rheumatic diseases 2019; 78:1151–1159.
4. Yen EY, Singh RR. Brief Report: Lupus-An Unrecognized Leading Cause of Death in Young Females: A Population-Based Study Using Nationwide Death Certificates, 2000–2015. Arthritis & rheumatology (Hoboken, NJ) 2018; 70:1251–1255.
5. Barber MRW, Drenkard C, Falasinnu T, Hoi A, Mak A, Kow NY, et al. Publisher Correction: Global epidemiology of systemic lupus erythematosus. Nature reviews Rheumatology 2021; 17:642.
6. Murimi-Worstell IB, Lin DH, Kan H, Tierce J, Wang X, Nab H, et al. Healthcare Utilization and Costs of Systemic Lupus Erythematosus by Disease Severity in the United States. The Journal of rheumatology 2021; 48:385–393.
7. Ulff-Moller CJ, Svendsen AJ, Viemose LN, Jacobsen S. Concordance of autoimmune disease in a nationwide Danish systemic lupus erythematosus twin cohort. Semin Arthritis Rheum 2018; 47:538–544.
8. Kuo CF, Grainge MJ, Valdes AM, See LC, Luo SF, Yu KH, et al. Familial Aggregation of Systemic Lupus Erythematosus and Coaggregation of Autoimmune Diseases in Affected Families. JAMA Intern Med 2015; 175:1518–26.
9. Woo JMP, Parks CG, Jacobsen S, Costenbader KH, Bernatsky S. The role of environmental exposures and gene-environment interactions in the etiology of systemic lupus erythematous. J Intern Med 2022; 291:755–778.
10. Ha E, Bae SC, Kim K. Recent advances in understanding the genetic basis of systemic lupus erythematosus. Semin Immunopathol 2022; 44:29–46.
11. Kwon YC, Chun S, Kim K, Mak A. Update on the Genetics of Systemic Lupus Erythematosus: Genome-Wide Association Studies and Beyond. Cells 2019; 8.
12. Goulielmos GN, Zervou MI, Vazgiourakis VM, Ghodke-Puranik Y, Garyfallos A, Niewold TB. The genetics and molecular pathogenesis of systemic lupus erythematosus (SLE) in populations of different ancestry. Gene 2018; 668:59–72.
13. Owen KA, Grammer AC, Lipsky PE. Deconvoluting the heterogeneity of SLE: The contribution of ancestry. J Allergy Clin Immun 2022; 149:12–23.
14. Owen KA, Price A, Ainsworth H, Aidukaitis BN, Bachali P, Catalina MD, et al. Analysis of Trans-Ancestral SLE Risk Loci Identifies Unique Biologic Networks and Drug Targets in African and European Ancestries. American Journal of Human Genetics 2020; 107:864–881.

15. Kim CJ, Lee CG, Jung JY, Ghosh A, Hasan SN, Hwang SM, et al. The Transcription Factor Ets1 Suppresses T Follicular Helper Type 2 Cell Differentiation to Halt the Onset of Systemic Lupus Erythematosus. 2019; 50.
16. Gallagher MD, Chen-Plotkin AS. The Post-GWAS Era: From Association to Function. American journal of human genetics 2018; 102:717–730.
17. Liao C, Vuokila V, Catoire H, Akcimen F, Ross JP, Bourassa CV, et al. Transcriptome-wide association study reveals increased neuronal FLT3 expression is associated with Tourette's syndrome. Commun Biol 2022; 5:289.
18. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, et al. Opportunities and challenges for transcriptome-wide association studies. Nat Genet 2019; 51:592–599.
19. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nature genetics 2016; 48:245–52.
20. Gusev A, Mancuso N, Won H, Kousi M, Finucane HK, Reshef Y, et al. Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. Nat Genet 2018; 50:538–548.
21. Li D, Liu Q, Schnable PS. TWAS results are complementary to and less affected by linkage disequilibrium than GWAS. Plant physiology 2021; 186:1800–1811.
22. .
23. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. Nature genetics 2018; 50:1593–1599.
24. Feng H, Mancuso N, Gusev A, Majumdar A, Major M, Pasaniuc B, et al. Leveraging expression from multiple tissues using sparse canonical correlation analysis and aggregate tests improves the power of transcriptome-wide association studies. PLoS Genet 2021; 17:e1008973.
25. Liu X, Lin J, Wu H, Wang Y, Xie L, Wu J, et al. A Novel Long Noncoding RNA lincRNA00892 Activates CD4(+) T Cells in Systemic Lupus Erythematosus by Regulating CD40L. Front Pharmacol 2021; 12:733902.
26. Yuan J, Jia J, Wu T, Liu X, Hu S, Zhang J, et al. Comprehensive evaluation of differential long non-coding RNA and gene expression in patients with cartilaginous endplate degeneration of cervical vertebra. Experimental and therapeutic medicine 2020; 20:260.
27. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature communications 2019; 10:1523.
28. Wang YF, Zhang Y, Lin Z, Zhang H, Wang TY, Cao Y, et al. Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups. Nat Commun 2021; 12:772.
29. Cui Y, Sheng Y, Zhang X. Genetic susceptibility to SLE: recent progress from GWAS. J Autoimmun 2013; 41:25–33.

30. Molineros JE, Yang W, Zhou XJ, Sun C, Okada Y, Zhang H, et al. Confirmation of five novel susceptibility loci for systemic lupus erythematosus (SLE) and integrated network analysis of 82 SLE susceptibility loci. Hum Mol Genet 2017; 26:1205–1216.
31. Doke T, Huang SZ, Qiu CX, Liu HB, Guan YT, Hu HL, et al. Transcriptome-wide association analysis identifies DACH1 as a kidney disease risk gene that contributes to fibrosis. J Clin Invest 2021; 131.
32. Kunz M, Konig IR, Schillert A, Kruppa J, Ziegler A, Grallert H, et al. Genome-wide association study identifies new susceptibility loci for cutaneous lupus erythematosus. Exp Dermatol 2015; 24:510–5.
33. Dilthey AT. State-of-the-art genome inference in the human MHC. Int J Biochem Cell Biol 2021; 131:105882.
34. Jacobi AM, Rohde W, Ventz M, Riemekasten G, Burmester GR, Hiepe F. Enhanced serum prolactin (PRL) in patients with systemic lupus erythematosus: PRL levels are related to the disease activity. Lupus 2001; 10:554–61.
35. Eagle RA, Trowsdale J. Promiscuity and the single receptor: NKG2D. Nat Rev Immunol 2007; 7:737–44.
36. Groh V, Wu J, Yee C, Spies T. Tumour-derived soluble MIC ligands impair expression of NKG2D and T-cell activation. Nature 2002; 419:734–8.
37. Hamada S, Caballero-Benitez A, Duran KL, Stevens AM, Spies T, Groh V. Soluble MICB in Plasma and Urine Explains Population Expansions of NKG2D(+)CD4 T Cells Inpatients with Juvenile-Onset Systemic Lupus Erythematosus. Open J Immunol 2017; 7:1–17.
38. Hervier B, Ribon M, Tarantino N, Mussard J, Breckler M, Vieillard V, et al. Increased Concentrations of Circulating Soluble MHC Class I-Related Chain A (sMICA) and sMICB and Modulation of Plasma Membrane MICA Expression: Potential Mechanisms and Correlation With Natural Killer Cell Activity in Systemic Lupus Erythematosus. Front Immunol 2021; 12:633658.
39. Lee YS, Heo W, Son CH, Kang CD, Park YS, Bae J. Upregulation of Myc promotes the evasion of NK cellmediated immunity through suppression of NKG2D ligands in K562 cells. Mol Med Rep 2019; 20:3301–3307.
40. Zhang YM, Zhou XJ, Cheng FJ, Qi YY, Hou P, Zhao MH, et al. Polymorphism rs3828903 within MICB Is Associated with Susceptibility to Systemic Lupus Erythematosus in a Northern Han Chinese Population. Journal of immunology research 2016; 2016:1343760.
41. Ricklin D, Hajishengallis G, Yang K, Lambris JD. Complement: a key system for immune surveillance and homeostasis. Nat Immunol 2010; 11:785–97.
42. Walport MJ. Complement. Second of two parts. N Engl J Med 2001; 344:1140–4.
43. Tange CE, Johnson-Brett B, Cook A, Stordeur P, Brohet F, Jolles S, et al. Quantification of human complement C2 protein using an automated turbidimetric immunoassay. Clin Chem Lab Med 2018; 56:1498–1506.
44. Lintner KE, Wu YL, Yang Y, Spencer CH, Hauptmann G, Hebert LA, et al. Early Components of the Complement Classical Activation Pathway in Human Systemic Autoimmune Diseases. Front Immunol 2016; 7:36.

45. D Cu, Lochlainn DJM, Bateman E, Misbah SA. Lessons of the month 2: Meningococcal epiglottitis and connective tissue disease associated with C2 deficiency. Clin Med (Lond) 2021; 21:e539-e540.
46. Macedo AC, Isaac L. Systemic Lupus Erythematosus and Deficiencies of Early Components of the Complement Classical Pathway. Front Immunol 2016; 7:55.
47. Sharma M, Vignesh P, Tiewsoh K, Rawat A. Revisiting the complement system in systemic lupus erythematosus. Expert Rev Clin Immunol 2020; 16:397–408.
48. Luo G, Xu N. Apolipoprotein M: Research Progress and Clinical Perspective. Adv Exp Med Biol 2020; 1276:85–103.
49. Kobayashi T, Kurano M, Nanya M, Shimizu T, Ohkawa R, Tozuka M, et al. Glycation of HDL Polymerizes Apolipoprotein M and Attenuates Its Capacity to Bind to Sphingosine 1-Phosphate. J Atheroscler Thromb 2021; 28:730–741.
50. Tyden H, Lood C, Jonsen A, Gullstrand B, Kahn R, Linge P, et al. Low plasma concentrations of apolipoprotein M are associated with disease activity and endothelial dysfunction in systemic lupus erythematosus. Arthritis Res Ther 2019; 21:110.
51. Du W, Shen T, Li H, Liu Y, He L, Tan L, et al. Low apolipoprotein M serum levels correlate with Systemic lupus erythematosus disease activity and apolipoprotein M gene polymorphisms with Lupus. Lipids Health Dis 2017; 16:88.
52. Christoffersen C, Obinata H, Kumaraswamy SB, Galvani S, Ahnstrom J, Sevvana M, et al. Endothelium-protective sphingosine-1-phosphate provided by HDL-associated apolipoprotein M. Proc Natl Acad Sci U S A 2011; 108:9613–8.
53. Atehortua L, Rojas M, Vasquez GM, Castano D. Endothelial Alterations in Systemic Lupus Erythematosus and Rheumatoid Arthritis: Potential Effect of Monocyte Interaction. Mediators Inflamm 2017; 2017:9680729.
54. Burg N, Swendeman S, Worgall S, Hla T, Salmon JE. Sphingosine 1-Phosphate Receptor 1 Signaling Maintains Endothelial Cell Barrier Function and Protects Against Immune Complex-Induced Vascular Injury. Arthritis Rheumatol 2018; 70:1879–1889.
55. Ronca V, Wootton G, Milani C, Cain O. The Immunological Basis of Liver Allograft Rejection. Front Immunol 2020; 11:2155.
56. Tektonidou MG, Dasgupta A, Ward MM. Risk of End-Stage Renal Disease in Patients With Lupus Nephritis, 1971–2015: A Systematic Review and Bayesian Meta-Analysis. Arthritis Rheumatol 2016; 68:1432–41.
57. Katalinic L, Eliasson E, Bubic-Filipi L, Kes P, Anic B, Basic-Jukic N. [Renal transplantation in patients with lupus nephritis]. Lijec Vjesn 2014; 136:219–23.
58. Honarpisheh M, Kohler P, von Rauchhaupt E, Lech M. The Involvement of MicroRNAs in Modulation of Innate and Adaptive Immunity in Systemic Lupus Erythematosus and Lupus Nephritis. J Immunol Res 2018; 2018:4126106.
59. Gordon RE, Nemeth JF, Singh S, Lingham RB, Grewal IS. Harnessing SLE Autoantibodies for Intracellular Delivery of Biologic Therapeutics. Trends Biotechnol 2021; 39:298–310.

60. Murphy G, Isenberg DA. New therapies for systemic lupus erythematosus - past imperfect, future tense. Nat Rev Rheumatol 2019; 15:403–412.
61. Veturi Y, Ritchie MD. How powerful are summary-based methods for identifying expression-trait associations under different genetic architectures? Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing 2018; 23:228–239.

Figures

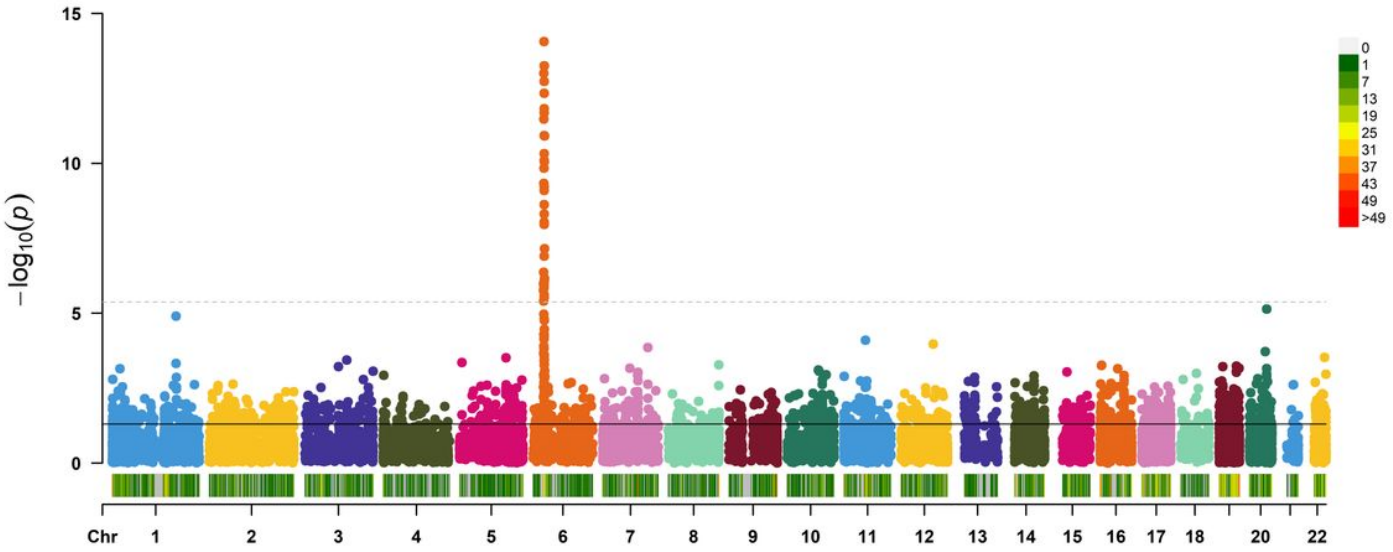

Figure 1

Manhattan plot showing TWAS-identified genes

Note: Manhattan plot showing TWAS-identified genes and significantly expressed genes related to systemic lupus erythematosus (SLE; annotated points). Each point represents a single gene, and the physical position (chromosome localization) is plotted on the x-axis, while the -log10 (*P*value) of the association between gene and SLE is plotted on the y-axis.

TWAS: Transcriptome-wide association study;

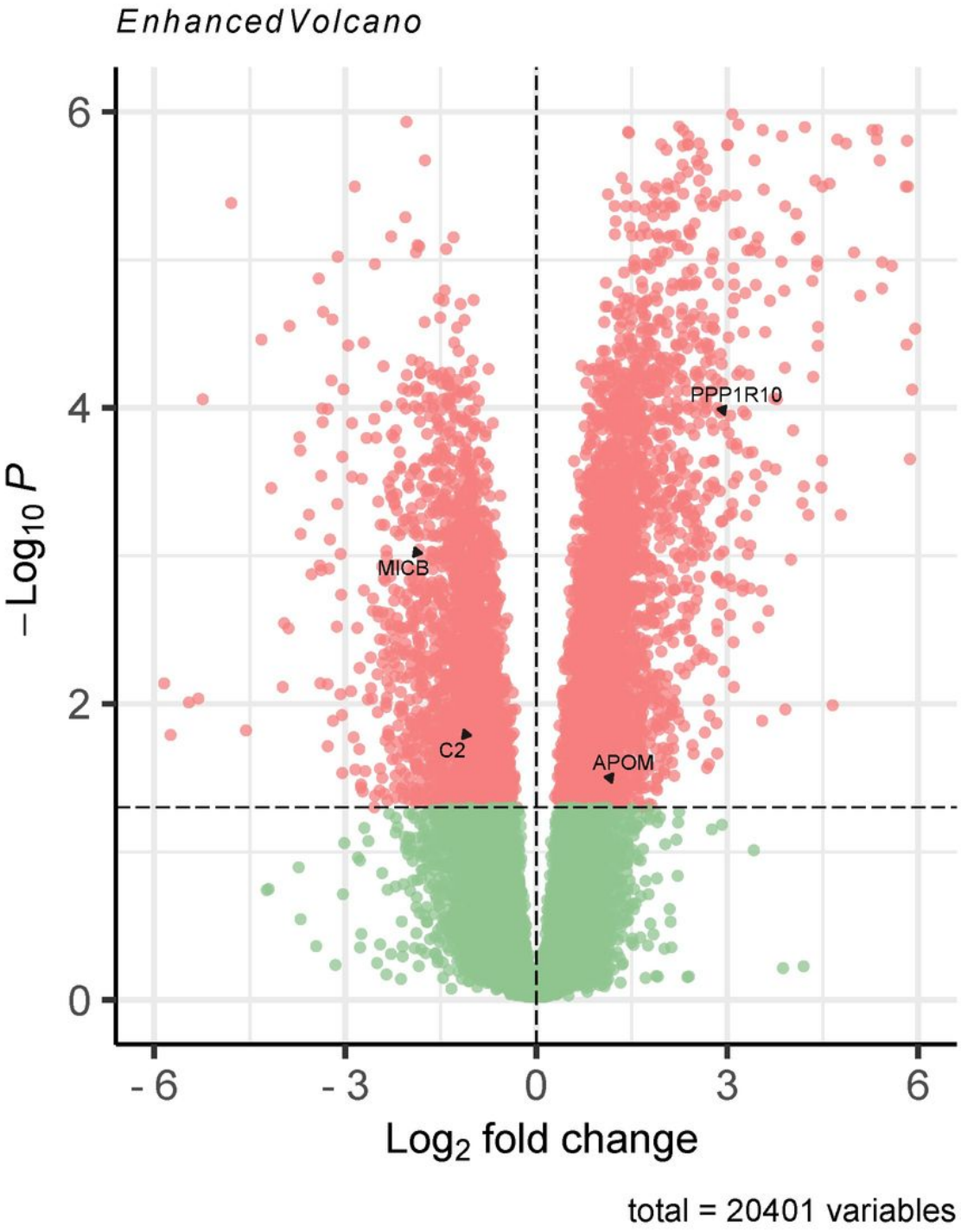

Figure 2

The volcano plot of mRNA expression profiles for systemic lupus erythematosus (SLE)

Note: The results of mRNA expression profiles were output to the volcano map. Genes were marked in red point as differentially expressed when the following two conditions were met: adjusted *P*-value of < 0.05 by the moderated t statistic and |logFC| > 1.

DEG: differentially expressed gene; SLE: systemic lupus erythematosus

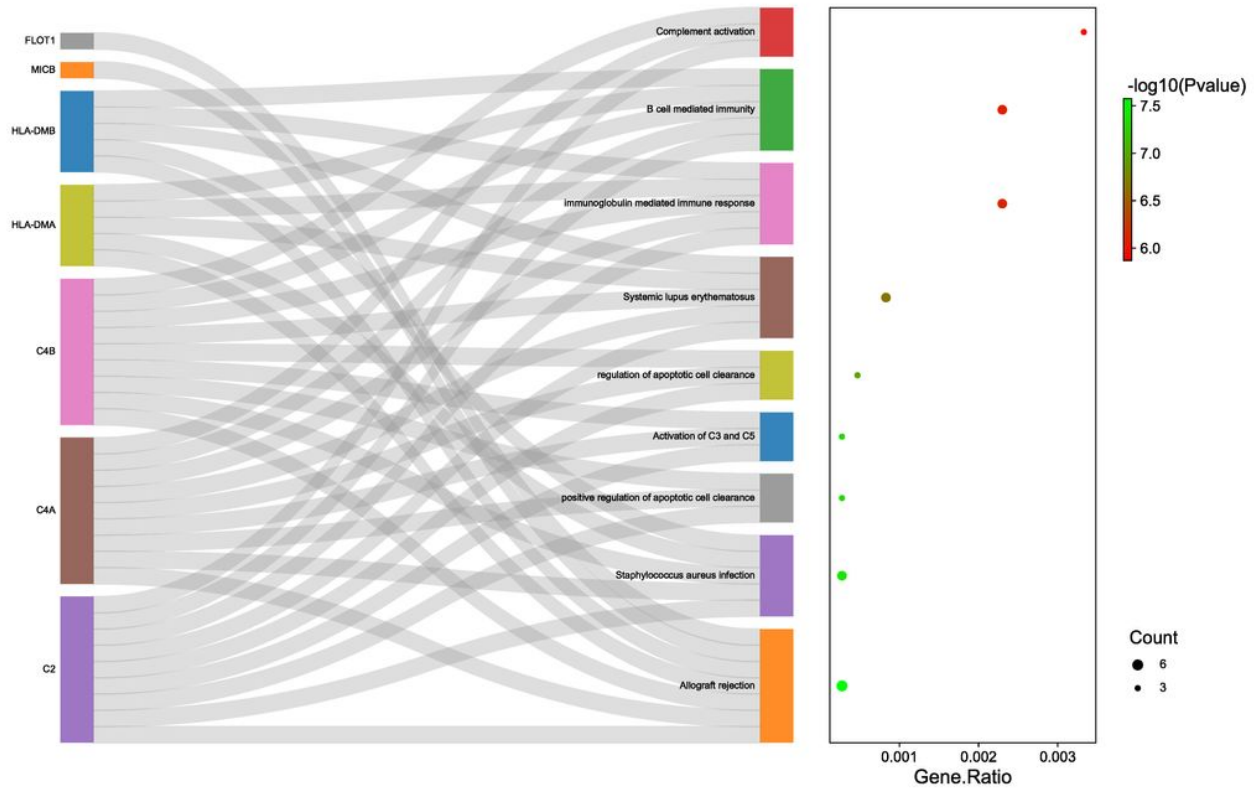

Figure 3

The significant GO term and related genes

Sankey diagram showed the relationship between the genes and top significant GO terms. The dot plot showed the ratio between the genes identified involved in GO terms and the total number of genes included in each GO terms (FDR $P \leq 0.05$).

GO: Gene Ontology

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementtable.xlsx