PAPER NAME

Manuscript plagarism check.docx

AUTHOR

Premanand Adaikalasamy

WORD COUNT

4784 Words

CHARACTER COUNT

28636 Characters

PAGE COUNT

15 Pages

FILE SIZE

110.2KB

SUBMISSION DATE

Aug 19, 2024 9:41 PM GMT+5:30

REPORT DATE

Aug 19, 2024 9:42 PM GMT+5:30

● **18% Overall Similarity**

The combined total of all matches, including overlapping sources, for each database.

- 15% Internet database
- Crossref database
- 0% Submitted Works database

- 14% Publications database
- Crossref Posted Content database

● **Excluded from Similarity Report**

- Bibliographic material

Integrative Bioinformatics Analysis of transcriptomic data from CD8+ T cells in Systemic Lupus Erythematosus

Premanand Adaikalasamy[1], Shruti Kumari[1], Sona Elizabeth Jacob[1], Sudhriti Bhuiyan[1], Reena Rajkumari Baskaran[1*], Shobana Sampath[2], and Sandhanasamy Devanesan[3]

Abstract

Introduction: Systemic Lupus Erythematosus (SLE) is a complex, multisystem autoimmune disorder characterized by extensive inflammation that affects nearly all organ systems in the body. It is primarily mediated by auto-antibodies and immune complexes, and it predominantly affects women more than men. This study employs an in-silico approach to identify key genes potentially involved in the pathogenesis of SLE.

Objectives: To identify key genes potentially involved in SLE pathogenesis using in-silico approach.

Methods: High-throughput sequencing dataset GSE97264, from the Gene Expression Omnibus (GEO) database, which contains RNA transcriptome data from CD8+ T-cells of 18 SLE patients and 14 healthy controls was utilized for the analysis. Differentially expressed genes (DEGs) were identified using the Bioconductor DESeq2 package in R platform. Gene Ontology (GO) and pathway enrichment analyses were performed using the ToppGene suite. Motif analysis of the genes' promoter regions was conducted using HOMER software. Protein-protein interaction (PPI) and Reactome functional interaction (FI) networks were created using Cytoscape plugins StringApp and ReactomeFIViz, and analysed to identify hub genes.

Results: Our analysis identified 931 DEGs, with 577 upregulated and 354 downregulated. GO and pathway enrichment analyses indicated that upregulated genes were associated with immune responses, including cytokine production and receptor activation. Motif analysis identified key regulatory motifs linked to immune regulation in upregulated genes and T-cell activation in downregulated genes. PPI and FI networks analyses revealed 29 cell cycle-associated hub genes, with 10 genes—CDK1, TPX2, BIRC5, CCNA2, BUB1, BUB1B, AURKA, KIF2C, PLK1, and CDCA8—common to both biological networks, suggesting their crucial role in SLE pathogenesis.

Conclusion: This study suggests that dysregulation of the identified 10 genes may impact immune responses and contribute to the autoimmune-like conditions observed in SLE. Several

of these genes are also implicated in other autoimmune diseases, highlighting their potential as SLE biomarkers. Despite their known roles in other immune-related diseases involving CD8+ T cells, their direct association with SLE had not been previously established. This novel finding underscores the potential of these genes as therapeutic targets and may contribute to the development of diagnostic tools.

Keywords: Systemic Lupus Erythematosus (SLE), CD8+ T cells Transcriptomics, Gene Expression Analysis, Biological Networks, Biomarkers for SLE

Introduction

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune disease characterized by the immune system's attack on its own tissues, which can affect both sexes. This leads to inflammation and, in some cases, permanent tissue damage affecting various body parts including the skin, joints, heart, lungs, kidneys, and brain [1]. The exact causes of SLE remain largely unknown; however, scientists have identified several potential triggers, such as genetic, environmental, and inflammatory factors [2]. It is postulated that an impairment in the normal clearance of apoptotic and necrotic cells may result in immune system dysregulation, whereby the immune system is misled into fighting itself, leading to SLE [2]. The course of disease progression is not linear, often characterized by relapse and remission phases, leading to considerable variability among patients, which presents significant difficulties in diagnosis and therapeutic interventions [1].

The pathophysiology of SLE is mediated by autoantibodies and immune complexes. A deficiency in complement proteins (C1, C2, C3, C4) impairs the efficient removal of apoptotic cells by macrophages, resulting in the exposure of nuclear material and internal proteins (antigens). These antigens are captured by Antigen-Presenting Cells (APCs) and presented to naive T-helper cells, triggering their activation and differentiation. In SLE, IL-4 secretion favors T-helper 2 cell maturation, which in turn promotes antibody-mediated immune responses. Autoantibody production, particularly anti-nuclear antibodies (ANA) and anti-double-stranded DNA (anti-dsDNA) antibodies, is attributed to activated B-cells that evolve into plasma cells, which remember the antigen and secrete more autoantibodies. [2].

When autoantibodies bind to nuclear proteins, they trigger inflammation through various mechanisms. Immune complexes may deposit on organs or bind to Fc receptors on immune

cells, leading to complement activation and the release of pro-inflammatory cytokines like interferons and TNF-alpha [2]. After surrounding immune cells become sensitized to the antigens, they release cytokines upon recognizing the nuclear antigen through the TOLL-like receptor, exacerbating the inflammatory response and further damaging nearby cells. The inflammation eventually subsides due to the immune system's innate regulation [2].

CD8+ cytotoxic T cells, part of the host defence mechanism, are activated by APCs presenting antigens via MHC-I molecule. Several studies have reported varied functions of CD8+ cells in SLE. In SLE patients, systemic CD8+ T cells demonstrate diminished effector functions and cytolytic activity, which compromises their ability to efficiently clear infections and autoreactive B-cells. These cells also demonstrate increased infiltration capacity, especially in the kidneys, contributing to organ damage [3]. Despite various assumptions about the role of CD8+ T cells in SLE, the precise scenario remains unclear, warranting further exploration.

In this study, we have utilised the dataset from Buang et al., deposited in the Gene Expression Omnivus Database, containing mRNA samples from blood CD8+ T cells of SLE patients and healthy controls [4]. We identified Differentially Expressed Genes (DEGs) between the SLE and control samples using the DESeq tool and categorized the upregulated and downregulated genes. To understand the functional and biological significance of these genes, we performed Gene Ontology (GO), Gene Family, and Reactome Pathway Enrichment analysis using the ToppGene Suite server. We also built a Reactome Functional Interactome Network and a STRING protein-protein interaction (PPI) network to identify hub genes in both networks. This study aims to unveil previously unexplored pathogenic genes, enhancing understanding of SLE's genetic-environmental interplay, aiding in diagnosis and treatment.

Methods and Materials

RNA-Seq Data from SLE Patients and Healthy Controls

The raw data for this study were sourced from the RNA transcriptome deposited by Buang et al., [4] in the Gene Expression Omnibus (GEO) under entry GSE97264. The transcriptome was obtained from the blood CD8+ T cells of SLE patients and healthy controls. The study included 16 SLE patients with active disease, 18 with less active disease, and 14 healthy controls. The British Isles Lupus Assessment Group (BILAG) and Systemic Lupus Erythematosus Disease Activity Index (SLEDAI score) were used to classify the patients. For this study, we selected data from patients with active SLE and healthy controls from the GSE97264 dataset.

Differential Gene Expression Analysis

The differential gene expression (DGE) analysis was conducted using DESeq2, a Bioconductor package, in R platform. Initially the count data was converted into "DESeqDataSet" object using the function 'DESeqDataSetFromMatrix' from DESeq2 package v 1.40.2. Then the DGE analysis based on the Negative Binomial (a.k.a. Gamma-Poisson) distribution was performed using the 'DESeq' function. The function followed a default workflow of estimating size factors, estimating dispersions, fitting model and testing (replacing outliers and refitting genes). The Wald test was employed to evaluate the significance of DEGs. The resulting p-values were then adjusted for multiple testing using the Benjamini-Hochberg (BH) method to control for the false discovery rate (FDR). Finally, the genes that have log2FC > |1| and padj (FDR) < 0.05 were considered as DEGs.

Gene Ontology and Pathway Enrichment Analyses

GO functional and pathway analyses was performed using the ToppGene Suite server. GO analyses (biological process, molecular function, cellular component), gene family, and pathway enrichment analyses were executed using the ToppFun function. The probability distribution function was selected for the p-value method, and a cut-off criterion of Gene count < 2 and FDR B&H q value < 0.05 was applied in the ToppFun function for these analyses [5].

Motif Analysis

The Homer v4.11 software was used to identify gene-based motifs in the promoter regions of the DEGs. For this study, the promoter regions were defined as 2,000 bp upstream and 200 bp downstream of the transcriptional start sites based on RefSeq genes (Hg38)). Motifs with a maximum length of 12 bases were probed, with a Benjamini-Hochberg-corrected p-value threshold of < 0.05 [5].

STRING protein-protein network analysis

The STRING database was used to construct the PPI network for the identified DEGs. The Cytoscape plugin stringApp v2.0.1 was employed to generate the PPI interaction network, setting the highest confidence interaction score at 0.900 to refine the network [5].

Reactome functional interaction network analysis

ReactomeFIViz v8.0.6, a Cytoscape tool, was used to create networks using the Reactome functional interaction (FI) network. The tool utilizes Reactome, a biological pathway database,

to derive interaction information. The FI network for our DEGs was built using the 2022 version of the Reactome FI network [5].

Network analysis and hub gene identification

Hub genes in the PPI and FI networks were identified using the Cytohubba v0.1 plugin. They were ranked according to six topological centralities and algorithms, including Maximal Clique Centrality (MCC), Maximum Neighbourhood Component (MNC), Density of Maximum Neighbourhood Component (DMNC), Degree, Closeness, and Betweenness [5].

Results

Identification of DEGs in SLE CD8+ T cells

The DGE analysis identified 931 genes as DEGs with, 354 genes downregulated and 577 upregulated in CD8+ T cells from active SLE patients compared to healthy controls (Supplementary I). The top 10 upregulated and downregulated genes are shown in Figure 1. Additionally, the DEGs identified in our analysis are statistically depicted using a volcano plot in Figure 1.

Gene Ontology and Pathway enrichment Analyses

Gene family enrichment analysis revealed that the upregulated DEGs are associated with immunoglobulins and CD molecules, while down-regulated DEGs are linked to keratins, collagens, fibronectin (Supplementary II-1). GO functional enrichment analysis (Figure 2, Supplementary II-2) showed distinct profiles for DEGs in SLE. Upregulated genes were predominantly associated with inflammation-related GO terms across three categories: molecular function (MF), biological process (BP), and cellular component (CC). Specifically, the enriched GO MF terms included antigen binding and toll-like receptor binding. In the CC category, terms like the IgA and IgG immunoglobulin complexes were prominent. Meanwhile, BP terms highlighted were adaptive immune response, chronic inflammatory response, cytokine production, innate immune response, interferon-mediated signaling pathway, and leukocyte migration involved in the inflammatory response (Figure 2, Supplementary II-2). These terms collectively suggest a heightened inflammatory activity consistent with the clinical manifestations typically observed in SLE [6].

Conversely, downregulated genes were enriched with GO MF terms such as oxidoreductase activity, and BP terms like G protein-coupled receptor signaling pathway, Wnt signaling

pathway, and cell-cell signaling by Wnt (Figure 2, Supplementary II-2). The enrichment of these GO terms for downregulated DEGs reflects disrupted immune responses characterized by altered chemotaxis, impaired adhesion, and modified CD8+ T cell function, a known feature of SLE pathology [6].

Pathway enrichment analysis enriched Reactome pathways such as FCGR3A mediated IL-10 synthesis, activation of B cell receptor, and several cell cycle associated pathways for upregulated DEGs. On the other hand, pathways related to matrisome, basement membrane and collagen were enriched for the downregulated DEGs (Supplementary II-3).

Motif Analysis of dysregulated genes in SLE

The motif enrichment analysis on the promoter regions of DEGs provided insights on the potential transcription factors that might play a significant role in SLE pathogenesis (Figure 3, Supplementary II- 4&5). Downregulated genes had motifs for transcription factors such as MyoG, ASCL1, LRF, Gata1, EGR1, E2A, and NFAT (Figure 3). MyoG, a basic helix-loop-helix (bHLH) family member, is involved in MEF2's immune response in T lymphocytes [7]. ASCL1, also a bHLH factor, plays a role in T-cell development and the Wnt/β-catenin pathway [8]. LRF, a Kruppel family member, is crucial for T cell differentiation [9], while GATA1, a zinc finger transcription factor, is linked with T-cell differentiation [10]. EGR1, another zinc finger transcription factor, activates macrophage transcription and is associated with immune response genes, including Tumor Necrosis Factor [11]. E2A, a bHLH member, is essential for early B and T cell development [28]. NFAT, a redox-dependent factor, affects CD8+ T cell cytotoxicity and metabolism [12]. The enrichment of these transcription factor motifs in the down-regulated genes suggests immune system aberration in SLE.

Upregulated genes were enriched with motifs for IRF4, T1SRE, E2F7, E2F3, E2F1, IRF1, ISRE, IRF3, and IRF2 (Figure 3). The Interferon Regulatory Factor family members (IRF1, IRF2, IRF3, and IRF4) were notably enriched, suggesting their involvement in SLE's pro-inflammatory processes [13]. IRF1 is involved in pro-inflammatory transcription, characteristic of SLE pathology [13], while IRF2 responds to persistent IFN signaling, leading to CD8+ T cell exhaustion [14]. IRF3 activation is associated with increased type I interferon expression in dendritic cells [15] , and IRF4 overexpression is linked to CD8+ T cell exhaustion [16]. The presence of T1ISRE hints at increased type 1 interferon gene expression. E2F1's role in T cell apoptosis and its association with T cell cycle progression [17], alongside E2F7 and E2F3's links to CD8+ T cell infiltration, further reveal SLE's complex cellular mechanisms

[18]. ISRE motifs in cytokine genes contribute to heightened type 1 interferon production, reflecting altered mitochondrial state of CD8+ T cells in SLE patients [4].

Construction and Investigation of STRING protein-protein interaction network for the identification of hub genes

The primary PPI network, after discarding the disconnected nodes, constructed by the StringApp had 249 nodes (215 upregulated genes, 26 downregulated genes, and 8 linker genes) and 1478 edges (Figure 4). Topological analysis revealed a clustering coefficient of 0.531, network diameter of 12, network density of 0.048, characteristic path length of 4.302, and an average of 11.871 neighbors per node. The nodes were further subjected to functional enrichment analysis by utilizing the STRING plugin available within the StringApp. This analysis enriched pathway terms related to cell cycle checkpoints, interferon alpha/beta signaling, cytokine signaling, IL-24 signaling and type II interferon signaling (Supplementary II-6). While GO analysis linked the PPI network genes to processes associated with cell cycle and immune system, such as apoptosis, chromosome segregation, DNA damage response, cytokine-mediated signaling pathway, and interleukin 27 signaling (Supplementary II-6).

To identify central or 'hub' genes in the PPI network, we applied four topological analysis methods- MCC, DMNC, MNC, and Degree - along with two measures of centrality, Closeness and Betweenness [5]. The top 20 hub nodes (genes) identified for each category in the PPI network are listed in Table 1. Genes appearing in at least three of categories are considered significant and they are CDK1, KIF11, BUB1, KIF20A, TOP2A, KIF2C, BUB1B, CCNB1, DLGAP5, CDC20, CDCA8, UBE2C, TPX2, CENPF, CCNA2, NUSAP1, CCNB2 and BIRC5.

Construction and Investigation of Reactome Functional Interaction Network for the Identification of Hub genes

The primary FI network, after discarding the disconnected nodes, generated using the ReactimeFIViz tool comprised of 361 nodes (278 up-regulated and 83 downregulated genes) and 4081 edges (Figure 5). Topological analysis of this FI network revealed an average node count of 22.609, a network diameter of 14, a characteristic path length of 3.958, a clustering coefficient of 0.481, and an average network density of 0.0653 neighbors. Pathway enrichment and Gene Ontology biological process analyses were conducted using the inbuilt network function plugin in ReactomeFIViz. This analysis highlighted terms predominantly related to the cell cycle, such as mitotic cell cycle, cell division, chromosome segregation, cell cycle checkpoint, mitotic prometaphase and G1/S transition (Supplementary II-7).

Similar to the PPI network, the top 20 hub nodes (genes) for each category in the FI network were identified using four topological analysis methods and two centralities. These hub nodes are detailed in Table 2. Genes appearing in at least three categories are considered significant and they are CDK1, CCNA2, AURKA, CDCA5, NDC80, CDCA8, CENPE, ZWINT, CCNB2, SPC25, BUB1, BUB1B, PLK1, EXO1, FOXM1, NEK2, TPX2, KIF2C and BIRC5.

Prospective Genetic Markers of SLE pathology

A total of 29 hub genes were identified from the analysis of PPI and FI networks, (Supplementary II-8). All these hub genes were found to be up-regulated. Among these, 10 genes - CDK1, TPX2, BIRC5, CCNA2, BUB1, BUB1B, AURKA, KIF2C, PLK1 and CDCA8 were found to be hub genes in both networks and are thus considered prospective genetic markers of SLE pathology. GO biological process and pathway enrichment analyses implicated these genes in several cell-cycle related processes (Supplementary II-9).

CDK1, a serine-threonine kinase and part of the M phase-promoting factor, plays a pivotal role in the cell cycle, particularly at the G1-S and G2-M checkpoints. It is instrumental in the IFN type I induced phosphorylation of STAT-1 and elevating ISG expression, a process central to SLE prognosis and inflammation [19]. Inhibition of CDK1 has been linked to reduced expression of pro-inflammatory genes [20]. TPX2, a key protein in mitotic spindle development and function, is associated with increased expression of pro-inflammatory cytokines. Studies in CRS-affected mice suggest that silencing TPX2 diminishes inflammation, indicated by changes in GSK3β, IL-10, TNF-α, IL-6, and IL-8 levels [21]. BIRC5 which encodes the protein survivin, significant for immune system maintenance, shows enrichment in leukocytes accumulated in inflamed tissues, a common feature in autoimmune disorders. It particularly affects the activity threshold of cytotoxic T lymphocytes against other survivin-expressing cells [22]. In rheumatoid arthritis, BUB1 is associated with abnormal cell proliferation, migration, invasion, PI3K/Akt pathway disruption, and pro-inflammatory cytokine release [23]. AURKA, found in psoriasis patients, promotes inflammation by impeding autophagy-mediated AIM2 inflammasome suppression and activating the Akt/mTOR pathway. It has also been linked to increased TNF- α expression in gastric mucosa of mice with gastrointestinal cancer [24]. KIF2C, coding for Mitotic centromere-associated kinesin, influences cell motility and migration by affecting the actin-MT cytoskeleton and FA turnover [25]. BUB1B, linked to exhausted T-cell signature and inflammatory CD8+ cells, has been shown to reduce invasion and proliferation in RCC cell lines upon knockdown [26].

PLK1, increases the activity of the NLRP3 inflammasome, a critical component in SLE pathophysiology. It affects the microtubule-organizing centre architecture and activation of NLRP3 causes cell damage and death in lupus nephritis [27]. CDCA8, essential for mitosis and chromosome segregation, has shown a significant correlation with tumor purity and immune cell infiltration levels in a study [28]. CDCA5, AURKB, CCNB1, and FOXM1, enriched in the transcriptomes of SLE patients, correlate with autoantibody titers [29]. NDC80 and its constituent SPC25, were linked to increased levels pro-inflammatory cytokines and chemokines in RA and are involved in regulating cell proliferation through the PK/AKT/Notch1 signaling pathway [30]. CENPE, an autoantigen in systemic sclerosis, has been found to trigger the formation of specific autoantibodies [31]. ZWINT, significant for kinetochore and mitotic checkpoint function, is implicated in enhanced CD8+ T cell infiltration [32]. NEK2 overexpression affects IL-22 mediated inflammation and cytokine production in psoriasis patients [33]. KIF20A and DLGAP5, correlated with elevated pro-inflammatory cytokines in RA, along with CDC20, a hub gene in the disease, are integral to inflammatory processes [30]. CCNB2, positively correlated with various immune cells, plays a role in immune infiltration [34]. NUSAP1, causing aberrant mitosis, is positively correlated with inflammatory gene expression [35]. UBE2C, associated with regulating the expression of several immunological checkpoints including chemokines [36]. CENPF expression is closely tied with CD8+ T cell immunological infiltration [37]. Exo1, KIF11, and TOP2A, enriched in SLE and other autoimmune conditions, are critical in the disease's pathology [38].

Since most of the hub genes are associated with cell-cycle functions, a single-sample Gene Set Enrichment Analysis (ssGSEA) of 16 gene sets related to cell-cycle processes and diseases (Supplementary II-10) obtained from the GSEA-MSigDB webserver was conducted on the samples using the GSVA v1.52.3 and GSEABase v1.66.0 Bioconductor packages. The ssGSEA heatmap (Figure 6) revealed that these 16 gene sets are up-regulated in SLE-active patients compared to healthy controls. This upregulation of cell-cycle-related gene sets suggests that the hub genes involved in cell-cycle functions likely play a significant role in SLE pathology. Therefore, these hub genes are proposed as potential biomarkers and therapeutic targets for further investigation in SLE.

Discussion

Our study identified 931 genes with differential expression between control and diseased samples, including 577 upregulated and 354 downregulated genes (Supplementary I). To elucidate the biological significance of these DEGs, we utilized the ToppGene Suite server.

The analysis revealed that the upregulated genes predominantly belong to immune-related gene families, such as immunoglobulins, C-type lectin domain family, and kinesins, among others (Supplementary II-1). These genes are crucial in various immune responses, including toll-like receptor binding and cytokine production, as indicated by Gene Ontology (GO) enrichment analysis (Figure 2). Pathway enrichment analysis further underscored their involvement in pathways such as FCGR3A-mediated IL-10 synthesis, B cell receptor activation, and several cell cycle-related pathways (Supplementary II-3). Conversely, the downregulated genes are associated with gene families including keratins, type II collagens, fibronectin type III domain-containing proteins, receptor tyrosine kinases, FERM domain-containing proteins, low-density lipoprotein receptors, and fibulins (Supplementary II-1). These downregulated genes are linked to signaling pathways such as G protein-coupled receptors and Wnt signaling (Figure 4). Pathways related to the matrisome, collagen, basement membrane, and collagen were particularly enriched for the downregulated genes in pathway enrichment analysis (Supplementary II-3).

Our investigation also identified several motifs in the promoter regions of both downregulated and upregulated genes. Downregulated genes exhibited motifs for transcription factors involved in T- cell activation (MyoG), T-cell differentiation (LRF, GATA1), regulation of T-cell development (ASCL1), transcriptional activation, TNF signalling and maturation of immune cells (EGR1), B and T-cell development (E2A) and cell cytotoxicity (NFAT). Notably, GATA1 is linked to transcriptional regulation of dendritic cell development and lineage separation of dendritic cells and macrophages [39]. In contrast, upregulated genes had motifs for IRF1, IRF3, and IRF4 (immune regulatory transcription factors), as well as T1ISRE and ISRE. These motifs are associated with increased type 1 interferon production and expression. The binding of IRFs to ISRE motifs in the promoter regions of cytokine genes results in heightened type 1 interferon production and continuous exposure to IFNα in SLE patients. This may lead to alterations in mitochondrial metabolism of CD8+ T cells, contributing to their apoptosis [4].

The analyses of two biological networks namely the String PPI and Reactome FI networks, identified 29 important genes. Among these, 10 genes—CDK1, TPX2, BIRC5, CCNA2, BUB1, BUB1B, AURKA, KIF2C, PLK1, and CDCA8—were found to be enriched in both networks. These common hub genes play significant roles in the abnormalities observed in CD8+ T cells in SLE.

CDK1 facilitates the phosphorylation of STAT-1, upregulates ISG expression and enhances type 1 IFN signalling, which has been linked to accelerated disease activity in SLE patients [19]. TPX2 wass associated with increased production of pro-inflammatory cytokines and could thus be contributing to the exaggerated inflammation seen in SLE patients [21]. BIRC5 was associated with CD8+ T cell production, innate and adaptive immune responses and increased accumulation of defective CD8+ T cells in tissues. Moreover, cytotoxic T lymphocytes expressing this gene were found to have a lowered activity threshold against survivin expressing cells, implying that these cytotoxic T cells became increasingly active and attacked healthy cells expressing survivin, leading to autoimmune-like conditions [22].

Pathways such as E2F targets, interferon alpha response, G2M checkpoint, IL6-JAK-STAT signalling, interferon gamma response, and inflammatory response were all regulated by CCNA2, which also increases CD8+ T cell invasiveness [40]. BUB1 facilitates cytokine release, regulates the PI3K/Akt signaling pathway associated with autoimmune development, and enhances tissue infiltration [23]. AURKA promotes inflammation, activates the Akt/mTOR and is associated with increased TNF-α expression which may have an impact on inflammation, CD8+ T cell apoptosis, and the production of autoantigens [24]. KIF2C enrichment suggests a potential dysregulation of motility and migratory functions of the CD8+ T cells [25]. BUB1B is linked to exhausted T cell signature, inflammatory CD8+ and IFN-gamma signature, cell proliferation and invasion [26]. The activation of the NLRP3 inflammasome is crucial in autoimmune development by triggering the production of pro-inflammatory cytokines and inducing pyroptosis through GSMD. PLK1 enriched in our study is a major player in NLRP3 inflammasome activation suggesting a central role of this gene in SLE pathogenesis [27]. Finally, CDCA8 may be involved in determining the immunological environment linked to SLE pathology because of the positive correlation found between CDCA8 expression and immune cell infiltration in SLE, including B cells, CD8+ T cells, CD4+ T cells, and macrophages [28].

Our study aligns with the previously established role of CD8+ T cell in the clinical manifestation of SLE which indicated that these cells exhibit improper cytotoxic activity, increased infiltration in the sites of tissue damage and increased cytokine production [3]. Although this aspect of the disease had been previously explored, the specific genes responsible of these abnormalities were not well understood. In this study, we utilised computational methods to investigate the genes underlying these aberrations by analysing the RNA transcriptome from SLE patients. We identified genes such as CDK1, TPX2, BIRC5, CCNA2,

BUB1, BUB1B, AURKA, KIF2C, PLK1 and CDCA8 as biomarkers of dysfunctional CD8+ T cells in SLE patients. These genes may serve as potential therapeutic targets and warrant further investigation to elucidate their roles more comprehensively.

Conclusion

SLE represents a highly complex multigenic disorder characterized by its multifaceted nature and the diverse array of factors influencing its symptomatology. Despite extensive research, the precise mechanisms underlying its pathophysiology remain elusive. Our study leveraged bioinformatics tools to pinpoint genes potentially instrumental in the aberrant functioning of CD8+ T cells in SLE patients. We identified ten key genes - CDK1, TPX2, BIRC5, CCNA2, BUB1, BUB1B, AURKA, KIF2C, and PLK1, along with CDCA8 - as critical in this context. Although these genes are primarily recognized for their roles in cell cycle regulation, our analysis suggests they may also have a broader impact. Their dysregulation could influence immune responses, potentially contributing to the autoimmune-like conditions observed in SLE and similar disorders. Importantly, some of these genes have been implicated in other autoimmune diseases, indicating their potential as biomarkers for SLE. However, to fully understand their relevance and mechanism in SLE pathology, further targeted studies are necessary. This research represents a significant step toward unraveling the genetic underpinnings of SLE, and could guide future diagnostic and therapeutic strategies.

Table 1. Top 20 hub genes determined using various topological methods and centrality measures in the String's protein-protein interaction network

Topological Algorithms				Centralities	
MCC	**MNC**	**DMNC**	**DEGREE**	**BETWEENNESS**	**CLOSENESS**
CDK1	CDK1	PBK	CDK1	FN1	CDK1
KIF11	CCNA2	NEK2	CCNA2	AURKA	CCNA2
BUB1	CCNB1	AURKA	CCNB1	BRCA1	CCNB1
KIF20A	CDC20	MKI67	CDC20	STAT1	BUB1B
TOP2A	BUB1	PTTG1	BUB1	CXCL10	CDC20
KIF2C	BUB1B	TTK	KIF11	ISG15	BUB1
BUB1B	KIF11	CEP55	BUB1B	PCNA	KIF11
CCNB1	KIF20A	CENPE	KIF20A	LCN2	TOP2A
DLGAP5	TOP2A	NUSAP1	TOP2A	EGF	KIF20A
CDC20	DLGAP5	NCAPG	DLGAP5	IL10	DLGAP5
CDCA8	KIF2C	PRC1	KIF2C	CCNA2	PLK1
UBE2C	CENPF	CDCA8	CENPF	CDK1	TPX2
TPX2	CCNB2	ASPM	CCNB2	LTF	CCNB2
CENPF	TPX2	MELK	TPX2	CENPA	UBE2C
CCNA2	CDCA8	KIF15	PLK1	ITGA2B	AURKA
CEP55	AURKB	BIRC5	CDCA8	BUB1B	KIF2C
MELK	PLK1	CENPF	BIRC5	TYMS	CENPF
NUSAP1	BIRC5	CDCA3	AURKB	RAD51	AURKB
TTK	NUSAP1	GBP4	NUSAP1	CD40LG	CDCA8
CCNB2	UBE2C	UBE2C	ASPM	MMP9	BIRC5

The colour code designates the presence of the highlighted genes in more than two columns—red indicates presence in five columns, violet indicates presence in four columns, blue indicates presence in three columns, and green indicates presence in two columns

Table 2. Top 20 hub genes determined using various topological methods and centrality measures with in the FI network.

Topological Algorithms				Centralities	
MCC	**MNC**	**DMNC**	**DEGREE**	**BETWEENESS**	**CLOSENESS**
ZWINT	CDK1	SGO1	CDK1	STAT1	CDK1
CDK1	CCNA2	SPC24	CCNA2	BRCA1	CCNA2
CCNB2	PLK1	KNL1	PLK1	FCGR1A	PLK1
SKA1	CDCA8	CDCA3	CDCA8	BIRC5	AURKA
SPC25	EXO1	CENPN	EXO1	TEAD3	EXO1
NDC80	CCNB2	IGHV2-26	AURKA	GLIS1	FOXM1
SKA3	AURKA	IGHV1-24	CCNB2	HLA-DRA	NEK2
CCNA2	FOXM1	IGHV1-69D	FOXM1	GLIS2	STAT1
AURKA	NEK2	IGHV3-43	NEK2	KIT	CDCA8
BUB1	NDC80	IGHV5-51	NDC80	E2F1	CENPE
BUB1B	ZWINT	IGHV6-1	ZWINT	CHEK1	NDC80
CDCA5	CDCA5	OIP5	CDCA5	PLCB1	BRCA1
CCNB1	CENPE	IGKV1-27	CENPE	CDK1	CDCA5
CDC20	BUB1B	IGKV1-8	BUB1B	CTBP2	KIF11
SGO1	SPC25	IGKV1-6	SPC25	CDKN1A	KIF2C
AURKB	TPX2	CENPM	TPX2	CCR5	CENPF
CDC25C	KIF2C	CKS2	KIF2C	MET	CCNB1
CENPE	BUB1	PTTG1	BUB1	VEGFA	TPX2
CDCA8	CENPA	SKA3	CENPA	EGF	CDC20

The colour code designates the presence of the highlighted genes in more than two columns—red indicates presence in five columns, violet indicates presence in four columns, blue indicates presence in three columns, and green indicates presence in two columns.

Figure 1. Volcano plot depicting the distribution of the DEGs identified in the SLE samples with the cut-off values log2FC>|1| and q-value < 0.05 and top 10 Up-Regulated and Down-Regulated DEGs Identified in the Study .

Figure 2. Gene Ontology analysis result of DEGs (FDR< 0.05) using ToppGene Suite

Figure 3. Motif enrichment analysis result

Figure 4. Protein-Protein interaction network built using stringApp

The circles are designated for the genes (nodes) while the lines represent the edges. The downregulated genes are represented using the red nodes while the upregulated genes are represented using the green nodes. Linker genes have been represented using blue nodes.

Figure 5. Reactome functional interaction network created using ReactomeFIViz

The circles are designated for the genes (nodes) while the lines represent the edges. The downregulated genes are represented using the red nodes while the upregulated genes are represented using the green nodes.

Figure 6. single sample Gene Set Enrichment Analysis (ssGSEA) of Cell Cycle related Gene sets on the samples

● **18% Overall Similarity**

Top sources found in the following databases:

- 15% Internet database
- Crossref database
- 0% Submitted Works database

- 14% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	**ncbi.nlm.nih.gov** Internet	**6%**
2	**frontiersin.org** Internet	**2%**
3	**nature.com** Internet	**<1%**
4	**biorxiv.org** Internet	**<1%**
5	**Yao Cai, Fei Ma, LiuHong Qu, Binqing Liu, Hui Xiong, Yanmei Ma, Sitao ...** Crossref	**<1%**
6	**Claudio Carini, Mark Fidock, Alain van Gool. "Handbook of Biomarkers ...** Publication	**<1%**
7	**aging-us.com** Internet	**<1%**
8	**science.gov** Internet	**<1%**

33   Victor R. Preedy, Ross Hunter. "Cytokines", CRC Press, 2019   **<1%**
Publication

34   Hans-Uwe Simon. "CRC Desk Reference for Allergy and Asthma", CRC ...   **<1%**
Publication

35   Li, Hongyan, Xiaonan Zhao, Jing Wang, Minru Zong, and Hailing Yang. "...   **<1%**
Crossref

36   Thomas F. Kresina. "Immune Modulating Agents", CRC Press, 2020   **<1%**
Publication