



OPEN ACCESS

EDITED BY

Wesley H. Brooks,
University of South Florida, United States

REVIEWED BY

Christopher Fenton,
UiT The Arctic University of Norway,
Norway
Celine C. Berthier,
University of Michigan, United States
Simon Helfgott,
Harvard Medical School, United States

*CORRESPONDENCE

Anastasia Fila
✉ afila@bioacademy.gr

†These authors have contributed
equally to this work and share
last authorship

RECEIVED 11 April 2023

ACCEPTED 15 June 2023

PUBLISHED 05 July 2023

CITATION

Sentis G, Loukogiannaki C, Malissovass N,
Nikolopoulos D, Manolakou T, Flouda S,
Grigoriou M, Banos A, Boumpas DT and
Fila A (2023) A network-based approach
reveals long non-coding RNAs associated
with disease activity in lupus nephritis: key
pathways for flare and potential biomarkers
to be used as liquid biopsies.
Front. Immunol. 14:1203848.
doi: 10.3389/fimmu.2023.1203848

COPYRIGHT

© 2023 Sentis, Loukogiannaki, Malissovass,
Nikolopoulos, Manolakou, Flouda, Grigoriou,
Banos, Boumpas and Fila. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

A network-based approach reveals long non-coding RNAs associated with disease activity in lupus nephritis: key pathways for flare and potential biomarkers to be used as liquid biopsies

George Sentis¹, Catherine Loukogiannaki¹, Nikos Malissovass¹,
Dionysis Nikolopoulos², Theodora Manolakou¹, Sofia Flouda²,
Maria Grigoriou^{1,3}, Aggelos Banos¹, Dimitrios T. Boumpas^{1,2†}
and Anastasia Fila^{1,3*†}

¹Laboratory of Autoimmunity and Inflammation, Center of Clinical, Experimental Surgery and Translational Research, Biomedical Research Foundation Academy of Athens, Athens, Greece, ²4th Department of Internal Medicine, Attikon University Hospital, National and Kapodistrian University of Athens Medical School, Athens, Greece, ³1st Department of Internal Medicine, University Hospital of Alexandroupolis, Democritus University of Thrace, Alexandroupolis, Greece

Objective: A blood-based biomarker is needed to assess lupus nephritis (LN) disease activity, minimizing the need for invasive kidney biopsies. Long non-coding RNAs (lncRNAs) are known to regulate gene expression, appear to be stable in human plasma, and can serve as non-invasive biomarkers.

Methods: Transcriptomic data of whole blood samples from 74 LN patients and 20 healthy subjects (HC) were analyzed to identify differentially expressed (DE) lncRNAs associated with quiescent disease and flares. Weighted gene co-expression network analysis (WGCNA) was performed to uncover lncRNAs with a central role (hub lncRNAs) in regulating key biological processes that drive LN disease activity. The association of hub lncRNAs with disease activity was validated using RT-qPCR on an independent cohort of 15 LN patients and 9 HC. cis- and trans-targets of validated lncRNAs were explored *in silico* to examine potential mechanisms of their action.

Results: There were 444 DE lncRNAs associated with quiescent disease and 6 DE lncRNAs associated with flares (FDR <0.05). WGCNA highlighted IFN signaling and B-cell activity/adaptive immunity as the most significant processes contributing to nephritis activity. Four disease-activity-associated lncRNAs, namely, NRIR, KLHDC7B-DT, MIR600HG, and FAM30A, were detected as hub genes and validated in an independent cohort. NRIR and KLHDC7B-DT emerged as potential key regulators of IFN-mediated processes. Network analysis suggests that FAM30A and MIR600HG are likely to play a central role in the regulation of B-cells in LN through cis-regulation effects and a competing

endogenous RNA mechanism affecting immunoglobulin gene expression and the IFN- λ pathway.

Conclusions: The expression of lncRNAs NRIR, KLHDC7B-DT, FAM30A, and MIR600HG were associated with disease activity and could be further explored as blood-based biomarkers and potential liquid biopsy on LN.

KEYWORDS

lupus nephritis, long non-coding RNAs, WGCNA, disease activity, ceRNA, blood-based biomarker, RNA-sequencing, SLEDAI-2K

Introduction

Systemic lupus erythematosus (SLE) is a chronic autoimmune disorder with manifestations of variable severity in multiple organs, predominantly affecting women of reproductive age (1). Lupus nephritis (LN), the renal manifestation of SLE, affects a significant proportion of patients and is accompanied by permanent organ damage and increased morbidity and mortality rates (2, 3). Initial diagnosis and monitoring of LN rely on invasive kidney biopsy (4). Current therapies are unable to suppress flares in more than half of LN patients, and residual inflammation is detected in cases of repeat biopsies of clinically inactive patients (5) with long-term disease leading to damage accrual (4). Considering the above and the fact that low disease activity is related to more favorable outcomes for SLE patients (6), it is essential to identify a non-invasive blood-based biomarker to reliably quantify disease activity in LN patients without the need for kidney biopsy.

Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nucleotides either without coding potential or featuring small open reading frames (ORFs) that translate in peptides of insignificant length (7, 8). They have been known to partake in transcriptional regulation, affecting the expression of genes in their vicinity (*cis*-) or distant targets (*trans*-acting lncRNAs) (9), exerting their effect at a transcriptional, posttranscriptional, or chromatin modification level with repressive or inducing aftermaths (10). lncRNAs have emerged in the last years as promising biomarker molecules of prognostic or diagnostic value, mainly in the field of cancer research (11–13). Although their role is not yet clear in SLE, studies in autoimmunity have shown lncRNAs to be stable in human plasma samples; thus, they are suitable candidates for non-invasive biomarkers (14).

Previous studies have reported that lncRNAs are differentially expressed in SLE patients compared with healthy controls. In particular, by sequencing total RNA isolated from peripheral blood mononuclear cells (PBMCs), researchers identified over 1000 DE lncRNAs, and using qPCR experiments, they validated the results of seven transcripts located near lupus susceptibility loci (15). A second study profiled the expression of lncRNAs in kidney tissue samples of a murine model with LN identifying and validating DE lncRNAs (i.e., Neat1, Lincpint) and DE mRNAs (Tgfb1, Riok3). The same study evaluated the co-expression of

lncRNAs with neighboring mRNAs and validated five gene co-expression pairs, namely, Gm26601-Dip2c, 2500002B13Rik-Hmgb2, Gm26556-Ppp1r9a, 1700020N18Rik-Hes6, and Gm20513-'H2-Aa' (16). A third study identified three lncRNAs (GAS5, linc0949, and lnc-DC) with stable plasma levels that are differentially expressed between healthy controls and LN patients. Interestingly, when comparing lnc-DC expression in non-nephritis SLE patients to healthy controls, they found the lncRNA to be upregulated, with contrasting results emerging in the LN patients vs. healthy controls comparison where the gene was found to be downregulated (17). This observation hints the existence of a biomarker, possibly based on the expression of an lncRNA, able to distinguish LN patients from non-nephritis SLE patients. Another group investigated the transcriptional landscape of non-coding RNAs in SLE showing the repressive capabilities of lncRNA ENSG00000236525 on the gene of C-C chemokine receptor type 7 (CCR7), ultimately affecting the differentiation of follicular helper T cells with an impact on autoimmunity (18). Finally, a study focusing specifically on lncRNA NEAT1 and its association with SLE revealed an upregulation of its expression in peripheral blood monocytes and an active role in the inflammation signaling of TLR4 (19).

Our study focuses on the association of lncRNAs with disease activity in LN. In this report, we describe the involvement of lncRNAs in flares of LN and propose lncRNAs with the potential to be used as liquid biopsy. We also suggest potential mechanisms for their involvement in this significant organ-specific manifestation of SLE.

Materials and methods

Sample collection

Samples from 74 LN patients and 20 healthy subjects (HC) served as our Discovery Cohort (Figure 1, Supplementary Table 1) (20). Patient samples either had active LN at the time of sampling ($n = 34$) or had a previous manifestation of LN over the course of the disease ($n = 40$). Active LN was defined as previously described (5). An independent cohort of 15 LN patients and 9 HC were used as the Validation Cohort (Supplementary Table 2). LN samples were split in

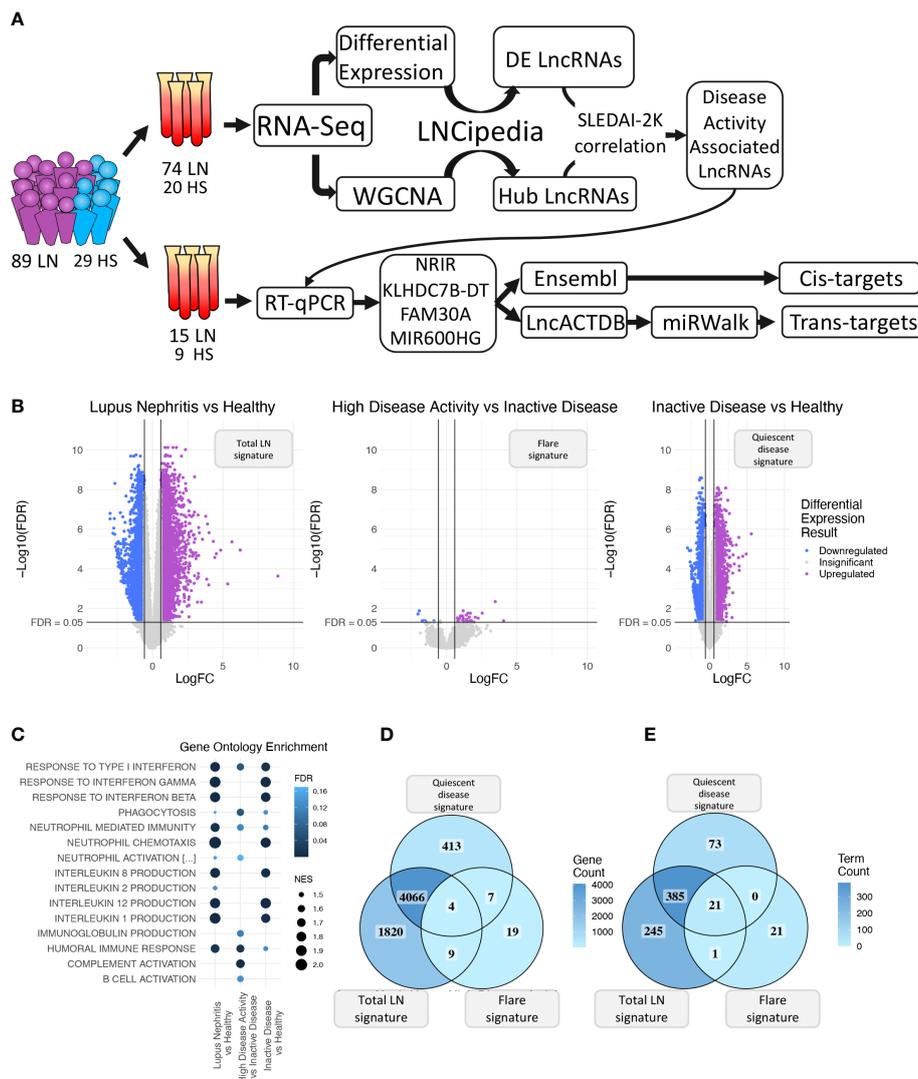


FIGURE 1

(A) Graphical overview of our research steps. (B) Volcano plot of the three differential expression analyses (DEA) performed comparing (left to right) Lupus Nephritis patients vs. Healthy controls, High Disease Activity vs. Inactive Disease patients, and Inactive Disease patients vs. Healthy controls. Upregulated genes are colored violet, and downregulated genes are colored blue. Genes not reaching our significance thresholds ($|\log_2FC| > 0.58$ and $FDR < 0.05$) are shown in gray. (C) Bubble plot showing inflammation-related Gene Ontology terms found as significantly enriched in each of the three DEA when performing gene set enrichment analysis (GSEA). Color represents FDR values whereas bubble size represents the Normalized Enrichment Score of each term. (D) Venn diagram comparing the DE genes of each DEA. Color gradient corresponds to the gene count in each compartment. (E) Venn diagram comparing the enriched terms of each GSEA. Color gradient corresponds to the term count in each compartment.

three groups according to their clinical SLE disease activity index (SLEDAI-2K) (21). The groups were defined as Inactive Disease (InaD) (SLEDAI-2K: 0-4), Intermediate Disease Activity (IDA) (SLEDAI-2K:5-11), and High Disease Activity (HDA) (SLEDAI-2K:12+) groups.

Library construction

Whole-blood extracts of the Discovery Cohort were total RNA-sequenced. PAXgene Blood RNA Kit IVD (#762174, Qiagen) and TempusTM Spin RNA Isolation Kit (#4380204, Thermo Fisher Scientific) were used for RNA isolation. Library construction was performed using NEBNext[®] rRNA Depletion Kit v2 (#E7400, New

England Biolabs) and NEBNext[®] UltraTM II Directional RNA Library Prep with Sample Purification Beads Kit (#E7765 New England Biolabs). Library quality was assessed using a 2100 Bioanalyzer (Agilent), and a Qubit 4 Fluorometer with dsDNA HS assay kit (#Q32854, Thermo Fisher Scientific) was used for quantitation of libraries. 100-bp paired-end sequencing was performed on an Illumina Nova-Seq 6000 System.

Sequencing QC and analysis

Quality of sequencing data was assessed using FastQC software (version:0.11.9, RRID : SCR_014583) (22). Adapter sequences and low-quality bases ($Q < 30$) of the 3' end were trimmed using Cutadapt

(v:1.18, RRID : SCR_011841) (23), and trimmed reads were aligned to the human reference genome (v:hg38) with GENCODE annotation (v:39, RRID : SCR_014966) (24) using STAR (v:2.6.1b, RRID : SCR_004463) (25). Samtools (v:1.9, RRID : SCR_002105) (26) was used to sort bam files, and HTSeq (v:0.11.0, RRID : SCR_005514) (27) was used to extract gene expression counts.

Differential expression analysis

Raw counts were normalized and analyzed using edgeR (28) package (v:3.38.1, RRID : SCR_012802) in R (29) (v:4.2.0, RRID : SCR_001905) to identify mRNAs and lncRNAs that are differentially expressed (DE) between (a) all LN patients and healthy controls (HC) (DEmRNAs, DElncRNAs), (b) HDA and InaD groups, and (c) InaD and HC. Genes were considered DE when $|FC| > 1.5$ and $FDR < 0.05$. Results were visualized using ggplot2 (v3.4.1, RRID : SCR_014601) (30).

Gene set enrichment analysis

Preranked gene-set enrichment analysis (GSEA) against Gene Ontology Biological Process (GO : BP MSigDB (31) v2022.1.Hs, RRID : SCR_016863) terms was performed using GSEA (32) software (v:4.2.2, RRID : SCR_003199) along with log2FC and FDR values from each DE analysis. Genes were ranked according to the product of $-\log_{10}(p \text{ value})$ multiplied by $\log_2(\text{FoldChange})$ in descending order.

Weighted gene co-expression network analysis

Weighted gene co-expression network analysis was performed using the R package WGCNA (33) (v:1.71, RRID : SCR_003302) to identify groups (modules) of co-expressed genes using the gene expression data of the 74 SLE patients. Identified modules were correlated with the patients' clinical SLEDAI-2K, and significant modules ($p < 0.05$) were tested for functional enrichment using g:Profiler (RRID : SCR_006809) (34). Genes with a central role (Hub genes) in each of the significant modules were determined using the connectivity measure of Module Membership ($MM > 0.8$).

lncRNA annotation and cis-gene identification

The LNCipedia (35) database (v:5.2) was used to identify lncRNAs in the DE gene list and hub gene list. cis-Genes (-10 kb upstream of gene start position, +10 kb downstream of gene end position) of the hub lncRNAs were extracted from the Ensembl (36) database (version 105, RRID : SCR_002344) using the R package biomaRt (37) (v:2.52.0, RRID : SCR_019214). cis-Elements were visualized using the packages Gviz (38) and karyoploteR (RRID : SCR_021824) (39).

Gene-level correlation and RT-qPCR validation

The correlation of clinical SLEDAI-2K with the expression of the identified hub lncRNAs and of lncRNAs belonging to both 'quiescent disease' and 'flare' signature was tested in R using the Spearman coefficient and the function 'cor.test' with a significance threshold of $p < 0.05$. Healthy samples were assigned a SLEDAI-2K score of -1. The expression of hub lncRNAs significantly associated with SLEDAI-2K ($p < 0.05$) was validated in an independent cohort of 15 patients and 9 HC (Validation Cohort) using quantitative reverse transcription-polymerase chain reaction (RT-qPCR). RNA was isolated as mentioned above, and cDNA was created using PrimeScript RT-PCR Kit (#RR037A, Takara). Patient samples of the validation cohort were equally distributed in each activity group: five InaD, five IDA, and five HDA patients. GAPDH gene expression was used as baseline reference for calculating the relative expression of target genes. Experiments were performed on an Applied Biosystems QuantStudio 5 Real-Time PCR System using the KAPA SYBR Fast Universal Kit (#KK4602, Kapa Biosystems). Primer sequences are available in [Supplementary Table 3](#). Spearman coefficient and function 'cor.test' with a significance threshold of $p < 0.05$ were used to evaluate clinical SLEDAI-2K correlation with ΔCt values of each gene in the validation cohort.

Competing endogenous RNA network construction

lncACTdb 3.0 (40) was utilized to discover kidney-related ceRNA effects of lncRNAs validated by qRT-PCR. A list of mRNA and miRNA targets was retrieved for lncRNAs present in the database. To narrow down the list of potential mRNA targets of each lncRNA, only mRNAs which were part of the same module as their corresponding lncRNA were kept. To expand the list of possible ceRNA events, the miRNA lists ([Supplementary Table 4](#)) were used as input to miRWalk (RRID : SCR_016509) (41). Putative mRNA targets for all regions (3'UTR, 5'UTR, and CDS) were queried, and results were filtered for experimentally supported interactions as indicated by miTarBase (42) ([Supplementary Table 5](#)). Finally, the new list of potential mRNA targets was filtered, keeping only those present in the *Lightgreen* and *Lightyellow* modules and only triplets containing downregulated mRNAs ($\log_{2}FC < 0$ for LN vs. healthy differential expression analysis) ([Supplementary Table 6](#)). The network was visualized using packages igraph (RRID : SCR_019225) (43), qgraph (44), and ggraph (RRID : SCR_021239) (45) in R.

Results

Blood transcriptome profile of LN during quiescence and flare

Initially, we explored the transcriptome (both coding and non-coding) of LN patients and healthy controls (HC) using

transcriptomic data from our discovery cohort and identified 5,899 DE genes between LN and HC (total LN signature), 39 DE genes between HDA vs. InaD patients (flare signature), and 4,490 DE genes between InaD vs. HC (quiescent disease signature) (Figure 1B). Using GSEA, we identified enriched biological processes and pathways, well established in SLE, such as interferon (IFN) type I response and interleukin (IL) production (i.e., IL1, IL2, IL8, and IL12) (46, 47). It was noted that response to type I IFN signaling is deregulated in quiescent disease and further contributes to flares. However, IL production seems to be involved only in quiescent disease and not in flares. It was also shown that complement activation, B-cell activation, and immunoglobulin (Ig) production processes were deregulated in flares only (Figure 1C). A higher number of genes and pathways (4,070 genes, 406 GO : BP terms) were found to be common between the “total LN” signature and the “quiescent disease” signature compared with genes and pathways shared between “total LN” and “flare” signature (13 genes, 22 GO : BP terms, Figures 1D, E). Overall, these results indicate that most molecules and signaling cascades are deregulated at quiescent disease, and only few pathways, such as IFN, complement, B-cell activation, and Ig production contribute to flares.

LncRNA expression during quiescence and flare: biomarkers for flares

Next, we focused specifically on lncRNAs to explore how this transcript population is implicated in disease activity. We pinpointed 816 DE lncRNAs in the “total LN” signature, 6 DE lncRNAs in the “flare” signature, and 444 DE lncRNAs in the “quiescent disease” signature (Figure 2A). The expression of DE lncRNAs for each signature is shown in Figures 2B–D, and lncRNAs that have previously been associated with SLE such as NEAT1 and ENSG00000236525 are highlighted. Accordingly with the total transcriptome comparison, we observed more genes to be common between the “total LN” and “quiescent disease” signatures compared with the “total LN” and “flare” signatures (412 DE lncRNAs vs. 1 DE lncRNA, ENSG00000283064) (Figure 2E). This large number of deregulated lncRNAs in the “quiescent disease” signature suggests that lncRNAs may play a major role at LN onset. In contrast, the “flare” signature involves only six lncRNAs, suggesting that only few lncRNAs are involved in the transition from quiescence to flare. Noticeably, lncRNAs TCL6 and ENSG00000257275 were part of both “quiescent disease” and “flare” signatures, indicating that these genes undergo expression alterations both at the establishment of the disease and during flares. These data suggest that TCL6 and ENSG00000257275 are potential biomarkers able to discern between individuals under flare, quiescent disease and healthy state.

LncRNAs play a central role in disease activity-associated pathways

To further explore the transcriptomic landscape related to disease activity, we performed a weighted gene co-expression network analysis (WGCNA) using patient transcriptomic data of

the discovery cohort. WGCNA analysis identified 35 modules of co-expressed genes, six of which were significantly correlated with clinical SLEDAI-2K values (Figure 3A) and thus were labeled as disease-activity-related. This was followed by enrichment analysis to identify deregulated pathways in each module. In this analysis, the top 3 modules, *Salmon*, *Lightyellow*, and *Lightgreen*, were associated with IFN signaling, adaptive immune response, and B-cell signaling, respectively (Figure 3B). The last positively correlated module, *Darkgrey*, was enriched in DNA-protein binding, and the two negatively correlated modules, *Violet* and *Steelblue*, were enriched in transmembrane or electron transfer activity and calcium channel functions, respectively (Supplementary Figure 1A). To identify hub lncRNAs, we used the measure of module membership (MM), a measure that reflects the similarity in expression patterns between a gene and the rest of the genes in a module. A total of 18 hub lncRNAs were identified in *Salmon* (n=7), *Lightgreen* (n=10) and *Lightyellow* (n=1) modules (Figure 3C). *Darkgrey*, *Violet*, and *Steelblue* had no hub lncRNAs and were not further investigated (Supplementary Figure 1B). The lack of hub lncRNA in the last three modules suggests that lncRNAs are not involved in all disease activity-related processes, although their presence in the IFN- and adaptive-immunity-related modules highlights their pivotal position regulating key pathways involved in flares.

Assessment of the biomarker potential of hub lncRNAs

To determine if each hub lncRNA could be useful as a potential biomarker of disease activity, we performed correlation analysis of lncRNA expression levels with SLEDAI-2K values using both patient samples and HC of the discovery cohort. The large number of DE genes and lncRNAs in the quiescent disease signature highlights the different state of patients at quiescence and HC. This difference should be reflected in our measure of disease activity in order to assess whether the biomarker can not only determine disease activity but also distinguish between HC and quiescent state patients. Therefore, we assigned HC a SLEDAI-2K value of -1 to differentiate them from the InaD patient group. We also included lncRNA ENSG00000257275 in the correlation analysis because it appeared along with hub lncRNA TCL6 in both of the “quiescent disease” and “flare” signatures, indicating a discriminatory potential that covers the complete spectrum of disease activity. Correlation at the gene level was significant for 9 out of 19 tested lncRNAs (Figure 3D). Six hub lncRNAs, NRIR, KLHDC7B-DT, ENSG00000233785, BISPR, ENSG00000280007, and LINC02574 (hub lncRNAs of the *Salmon* module), were positively correlated with disease activity, whereas FAM30A, MIR600HG, and LINC00494 (hub lncRNAs of the *Lightgreen* module) had a negative correlation. RT-qPCR experiments measuring gene expression in an independent validation cohort validated the positive correlation of NRIR and KLHDC7B-DT (Figure 3E), thus suggesting an inducing effect on flares. The negative correlation of MIR600HG and FAM30A was also validated (Figure 3E), implying that these lncRNAs may have some flare-inhibitory properties.

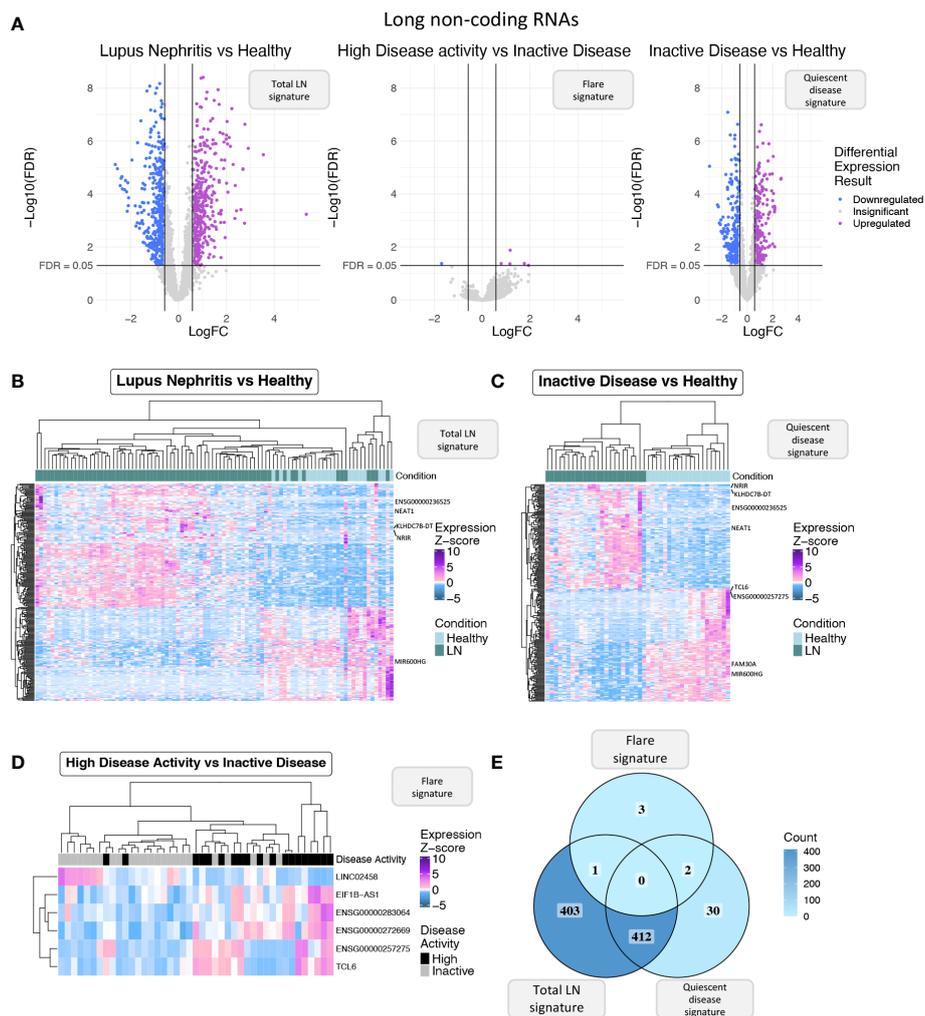


FIGURE 2
(A) Volcano plot of the long non-coding RNAs (lncRNAs) in each of the three differential expression analyses (DEA) performed comparing (left to right) LN patients vs. HC, HDA vs. InaD patients, and InaD patients vs. HC. Upregulated lncRNAs are colored violet, and downregulated lncRNAs are colored blue. lncRNAs not reaching our significance thresholds ($|\log_2FC| > 0.58$ and $FDR < 0.05$) are shown in gray. **(B)** Heatmap showing the expression profile of the top 250 lncRNAs with the highest absolute log₂ fold change value found as DE between LN and HC. Expression values were z-score normalized. Top annotation row shows the condition of each sample, colored green for LN patients and light blue for HC. **(C)** Heatmap showing the expression profile of the top 250 lncRNAs with the highest absolute log₂ fold change value found as DE between InaD patients and HC. Expression values were z-score normalized. Color scale of top annotation is the same as **(B)**. **(D)** Heatmap showing the expression profile of the six lncRNAs found as DE between HDA and InaD patients. Expression values were z-score normalized. Top annotation row shows the disease activity group of each sample with black representing HDA patients and gray representing InaD patients. **(E)** Venn diagram comparing the DE lncRNAs of each DEA. Color gradient corresponds to the term count in each compartment.

Identification of cis-targets of significant lncRNAs

Following the validation of NRIR, KLHDC7B-DT, MIR600HG, and FAM30A expression relative to disease activity, we investigated how these lncRNAs may exert their regulatory actions. We identified one cis-gene for NRIR, CMPK2, which was co-expressed with NRIR (belongs to the *Salmon* module) and three cis-genes for KLHDC7B-DT; SYCE3, ODF3B, and ENSG00000273272 (**Supplementary Figures 2A, B**). ODF3B and ENSG00000273272 belong to the *Salmon* module, whereas SYCE3 was not found in our dataset. We also identified two cis-genes in the vicinity of MIR600HG, STRBP, and MIR600 (**Supplementary**

Figure 2C). Both MIR600HG and STRBP are members of the *Lightgreen* module. MIR600 is a miRNA and was not detected in our dataset. Finally, we identified 29 cis-genes ± 10 kb of FAM30A, including the previously identified hub lncRNA ENSG00000244620 ($MM = 0.835$), a novel transcript ENSG00000288730, which is a hub RNA of the *Lightgreen* module ($MM = 0.878$), is not present in LNCipedia, and has no associated function, and 27 immunoglobulin heavy chain genes (IGH-) (**Figure 4A**). The presence of a high number of IGH-genes in the vicinity of FAM30A (**Figure 4A**), combined with the facts that a) the *Lightyellow* and *Lightgreen* modules contain 74% and 7% of immunoglobulin genes expressed in the dataset, respectively (**Figure 4B**), b) FAM30A has a high module membership in both

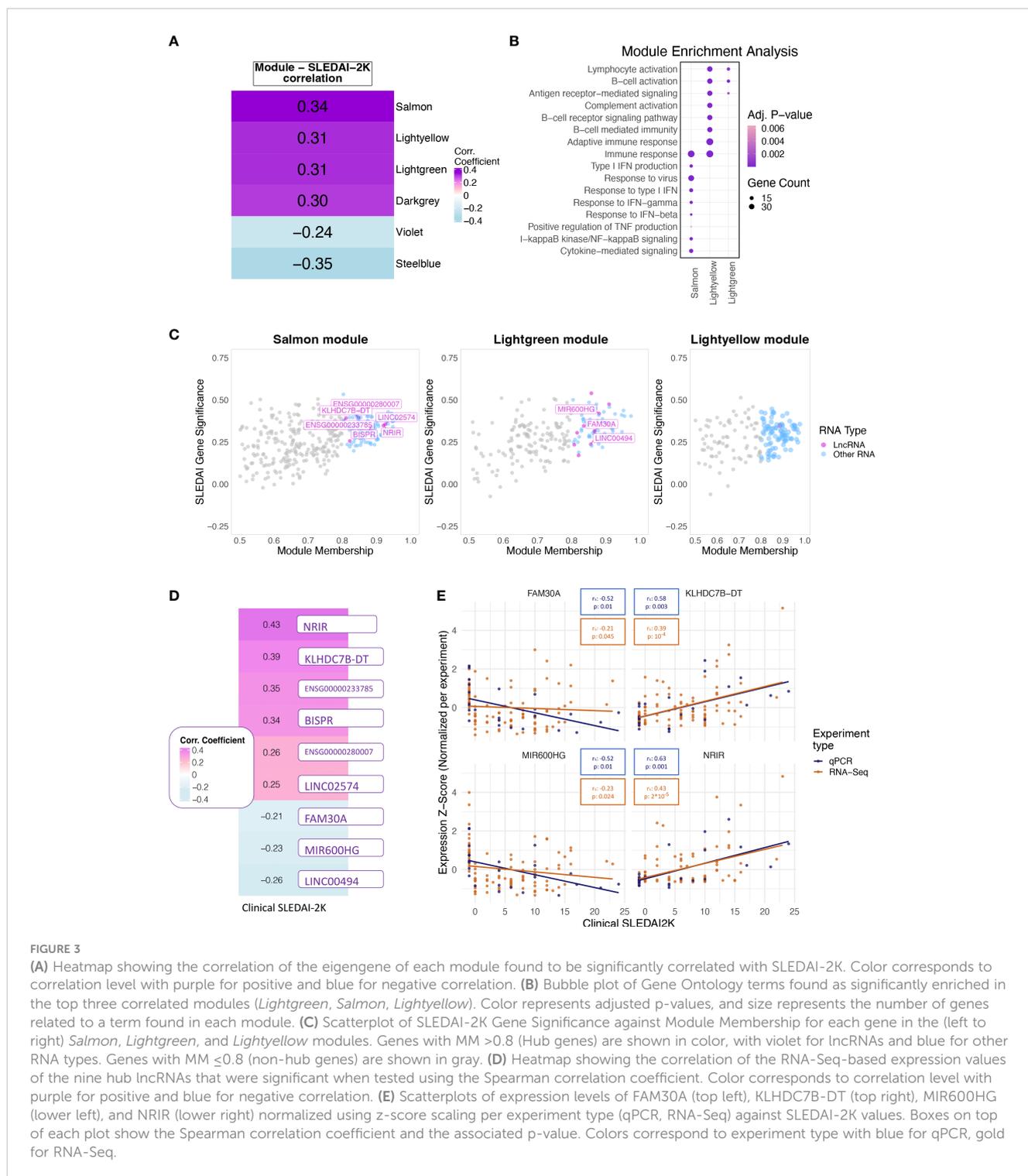


FIGURE 3

(A) Heatmap showing the correlation of the eigengene of each module found to be significantly correlated with SLEDAI-2K. Color corresponds to correlation level with purple for positive and blue for negative correlation. (B) Bubble plot of Gene Ontology terms found as significantly enriched in the top three correlated modules (*Lightgreen*, *Salmon*, *Lightyellow*). Color represents adjusted p-values, and size represents the number of genes related to a term found in each module. (C) Scatterplot of SLEDAI-2K Gene Significance against Module Membership for each gene in the (left to right) *Salmon*, *Lightgreen*, and *Lightyellow* modules. Genes with $MM > 0.8$ (Hub genes) are shown in color, with violet for lncRNAs and blue for other RNA types. Genes with $MM \leq 0.8$ (non-hub genes) are shown in gray. (D) Heatmap showing the correlation of the RNA-Seq-based expression values of the nine hub lncRNAs that were significant when tested using the Spearman correlation coefficient. Color corresponds to correlation level with purple for positive and blue for negative correlation. (E) Scatterplots of expression levels of FAM30A (top left), KLHDC7B-DT (top right), MIR600HG (lower left), and NRIR (lower right) normalized using z-score scaling per experiment type (qPCR, RNA-Seq) against SLEDAI-2K values. Boxes on top of each plot show the Spearman correlation coefficient and the associated p-value. Colors correspond to experiment type with blue for qPCR, gold for RNA-Seq.

Lightgreen ($MM = 0.867$) and *Lightyellow* ($MM = 0.784$) modules, and c) pathway enrichment of both modules is closely related to adaptive immunity, concludes that FAM30A is a gene with a significant role in both modules. Therefore, whereas most IGH-genes (81%) are assigned to the *Lightyellow* module, they have a strong probability of being cis-targets of FAM30A, rendering FAM30A a potential key regulator of immunoglobulin gene expression and antibody formation.

Delineation of trans-effects of MIR600HG and FAM30A

lncRNAs can also have trans-regulatory effects, regulating the expression of distant genes. The competing endogenous RNA (ceRNA) hypothesis (48) describes such trans-regulation events where lncRNAs bind to miRNA and inhibit miRNA-guided degradation of mRNAs. We searched LncActDB, a database of

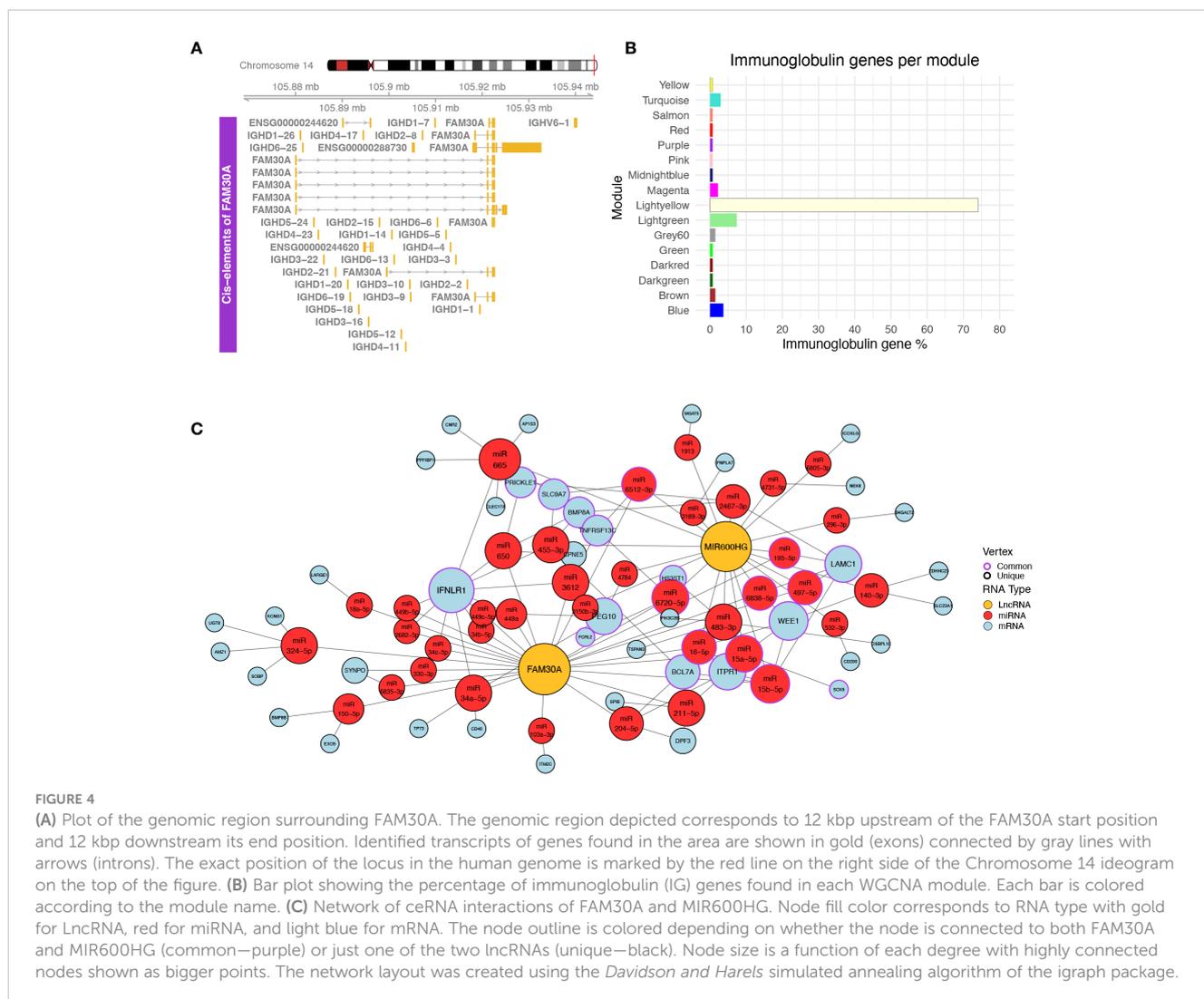


FIGURE 4

(A) Plot of the genomic region surrounding FAM30A. The genomic region depicted corresponds to 12 kbp upstream of the FAM30A start position and 12 kbp downstream its end position. Identified transcripts of genes found in the area are shown in gold (exons) connected by gray lines with arrows (introns). The exact position of the locus in the human genome is marked by the red line on the right side of the Chromosome 14 ideogram on the top of the figure. (B) Bar plot showing the percentage of immunoglobulin (IG) genes found in each WGCNA module. Each bar is colored according to the module name. (C) Network of ceRNA interactions of FAM30A and MIR600HG. Node fill color corresponds to RNA type with gold for LncRNA, red for miRNA, and light blue for mRNA. The node outline is colored depending on whether the node is connected to both FAM30A and MIR600HG (common—purple) or just one of the two lncRNAs (unique—black). Node size is a function of each degree with highly connected nodes shown as bigger points. The network layout was created using the Davidson and Harel's simulated annealing algorithm of the igraph package.

experimentally validated ceRNA interactions, to identify ceRNA events NRIR, KLHDC7B-DT, MIR600HG, and FAM30A participate in. Data were available only for MIR600HG and FAM30A. Because LncActDB provides ceRNA interactions stratified by tissue, we focused on the 15 and 116 events identified respectively for MIR600HG and FAM30A in kidney tissue. Each event is defined by an affected mRNA and a list of miRNAs interacting with both the lncRNA and the mRNA. To verify the presence of these interactions, we used our WGCNA data and intersected the mRNA targets with the *Lightyellow* and *Lightgreen* modules. We identified only one common gene, B4GALT2, between LncActDB and the *Lightyellow* module. Thus, we further expanded the analysis *in silico* to identify putative ceRNA interactions based on our data. To achieve this, we used the lists of miRNAs that interact with MIR600HG (MIR600HG miRNAs, $n = 53$) and FAM30A (FAM30A miRNAs, $n = 43$) provided by LncActDB and used them as input to the miRNA target prediction tool miRWalk. After filtering using miTarBase information to include only experimentally supported interactions, miRWalk identified 3,071 interactions for MIR600HG miRNAs and 2,589 interactions for FAM30A miRNAs. We further refined the mRNA target list by removing mRNAs not

included in the *Lightgreen* and *Lightyellow* modules. Finally, since MIR600HG and FAM30A are downregulated in LN patients, thus their miRNA-”sponging” effect would not be observed and miRNAs would be able to bind to their mRNA targets lowering their expression. To take this into account, we included only downregulated mRNAs of our “total LN” signature. Our final network consists of FAM30A, MIR600HG, 38 miRNAs, and 42 mRNAs of which 32 belong to the *Lightgreen* module and 10 belong to the *Lightyellow* module (Figure 4C). Network analysis revealed IFNLR1, an interferon-related mRNA, as the gene influenced the most by being targeted by 11 miRNAs. Interestingly, FAM30A and MIR600HG do not have a common way of regulating this mRNA. FAM30A can regulate IFNLR1 by interacting with nine different miRNAs (hsa-miR-34a-5p, hsa-miR-3612, hsa-miR-449a, hsa-miR-650, hsa-miR-2682-5p, hsa-miR-34b-5p, hsa-miR-34c-5p, hsa-miR-449b-5p, hsa-miR-449c-5p), whereas MIR600HG regulates IFNLR1 through two different miRNAs (hsa-miR-455-3p hsa-miR-665). This analysis reveals the impact of FAM30A and MIR600HG on multiple targets involved in adaptive immunity and provides a mechanism of regulation that features multiple parallel ways of targeting the same mRNA.

Discussion

In this study, we used whole-blood RNA samples of LN patients and HC to explore the transcriptomic landscape of LN focusing on lncRNAs, the role of which is largely unexplored. DE and enrichment analyses confirmed the deregulation of pathways such as IFN response and IL-2 production, which have previously been reported to be associated with SLE (46, 47). Similar to previous studies, we define distinct LN-related signatures for quiescent disease and flares (49, 50). We focused on lncRNAs and the relationship of their gene expression levels with disease activity as measured using SLEDAI-2K and identified six DE lncRNAs in the “flare” signature, hinting that changes in lncRNA expression may be subtle during periods of exacerbated symptoms. This led us to use a different approach in order to determine lncRNAs with a key role in LN flares.

To this end, WGCNA was performed and six modules associated with disease activity were identified. Previous studies have reported an IFN-related module as the most significantly associated with disease status both in renal tissue and blood samples (51, 52). Consistent with these results, our findings show the direct positive correlation of the IFN-module with SLEDAI-2K. We also report that the adaptive immunity and B-cell receptor pathways appear to be significant contributors to high disease activity through the correlation of the relevant modules with SLEDAI-2K. There were 18 lncRNAs with a high module membership and a possible central role in regulating the disease-activity-related processes discovered, with nine of them showing a significant correlation of their expression levels with SLEDAI-2K. This correlation is an indication that hub lncRNAs could be utilized as blood-based biomarkers. Importantly, assigning a pseudo-SLEDAI-2K value of -1 to HC to distinguish them from InAD patients is a useful step to better evaluate their use as a biomarker, since a large number of DE lncRNAs were found in the quiescent disease signature supporting the previously reported observation that patients with achieved clinical remission may not always have achieved histological remission (53). Thus, with our approach, we also evaluate their potential use as biomarker to detect LN molecular activity present in clinical remission cases. Quantitative PCR experiments validated the significant association observed in the transcriptomic data for four of these lncRNAs, namely, NRIR (positive), KLHDC7B-DT (positive), FAM30A (negative), and MIR600HG (negative).

NRIR has been extensively investigated in previous studies in autoimmunity. In systemic sclerosis (SSc) (54), it was associated with the IFN score of SSc patients; in primary Sjogren’s syndrome (pSS), it was found to correlate with pSS disease activity levels (55); and in SLE, it was observed with substantially higher expression in SLE patients compared with HC (56). In the last study, researchers also correlated the expression levels of two IFN-stimulated genes, RSAD2 and USP18, with the SLEDAI-2K score of LN samples of their SLE cohort, although the correlation could not reach statistical significance for NRIR. We corroborate and expand their findings showing the significant relationship of NRIR expression levels with SLEDAI-2K. The second lncRNA, KLHDC7B-DT, has been studied

previously in pancreatic ductal adenocarcinoma (57), where it was found to bind to the IL6 promoter region, and in psoriasis (58), where researchers showed a relationship between the lncRNA and ILF2, a T-cell-associated enhancer of the IL2 gene, with both genes being upregulated in lesional skin. Our novel finding suggests a central role of KLHDC7B-DT in LN disease activity. The third lncRNA, FAM30A, has previously been detected to interact with hub genes in a WGCNA analysis of a rheumatoid arthritis study (59), with our study revealing a similar central role of FAM30A in LN for the first time. Lastly, MIR600HG has been shown in a recent study to be differentially expressed between SLE patients and HC (60). The same study showed its co-expression with CD40LG, a gene encoding the CD40 ligand which is expressed on the surface of T cells and interacts with the CD40 antigen on the surface of B cells to signal B-cell activation. In light of another study (61) which has shown a direct positive correlation of CD40 expression and SLEDAI index in pediatric SLE patients, MIR600HG appears as a valid candidate for a blood-based biomarker of disease activity.

We also investigated the cis-effects of NRIR, KLHDC7B-DT, MIR600HG, and FAM30A, to identify a potential mechanism of action through which they influence disease activity. cis-Targets of NRIR and KLHDC7B-DT were defined as genes 10 kb upstream or downstream of the gene belonging to the IFN-module (Salmon). We detected CMPK2, a gene participating in IFN-dependent and IFN-independent antiviral immunity (62), as a potential cis-target of NRIR. CMPK2 has been reported as part of the SLE “flare” signature in a previous study (63), further linking NRIR’s locus to disease activity. For the KLHDC7B-DT gene, we identified three potential cis-targets, ENSG00000273272, ODF3B, and KLHDC7B. ENSG00000273272 has not been studied yet and no function is associated with this gene, ODF3B has been associated with SSc (64) and multiple sclerosis (65), and KLHDC7B has been linked to the IFN signaling pathway (66). cis-Targets of MIR600HG had to satisfy the 10-kb distance threshold and belong to the Lightgreen module. We identified STRBP, which a previous study using machine learning on SLE patient samples has shown to correlate with the expression levels of CD22, a surface marker of mature B cells (67, 68). The membership of MIR600HG and STRBP in a B-cell-related module and the previous association of STRBP with CD22 indicate the possibility of a B-cell-specific cis-regulation mechanism. The last validated lncRNA, FAM30A, is a hub gene of the *Lightgreen* module. We used the same distance threshold when determining cis-targets of FAM30A; however, due to the high module membership of FAM30A in both adaptive immunity (*Lightyellow*) and B-cell receptor (*Lightgreen*) modules, we considered genes as cis-targets if they belonged to either module. Excluding two hub RNAs (ENSG00000244620, ENSG00000288730) with no known function, cis-targets of FAM30A were 27 immunoglobulin heavy (IGH-) chain genes, members of the *Lightyellow* module. The high MM of FAM30A in the *Lightyellow* module suggests that FAM30A expression levels fluctuate with the same pattern as the IGH-genes’ expression levels, agreeing with a previous study (69) which uncovered a positive relationship of antibody titers with FAM30A expression levels after immune response to vaccination. Moreover, a recent tool

investigating human gene co-expression using transcriptomic data from the GTEx consortium clearly illustrates the fact that FAM30A is highly co-expressed with immunoglobulin genes (70, 71). This evidence suggests a role of FAM30A in the regulation of IGH-genes and possibly antibody production.

Furthermore, we explored the trans-effects of FAM30A and MIR600HG through the *in silico* resources of LncActDB and miRwalk. Using the experimentally validated FAM30A-miRNA and MIR600HG-miRNA interactions of LncActDB and the experimentally validated miRNA-mRNA interactions of miRwalk, we created a ceRNA network of both lncRNAs. Intriguingly, their most prominent trans-target is interferon lambda (IFN λ) receptor 1 (IFNLR1), whose deficiency lowers the activation of immune cells and reduces organ damage in kidneys without affecting production of antibodies in murine lupus (72). Furthermore, another study marks IFN λ as an overlooked factor driving aberrancies in B cells in SLE and associates IFNLR1 with the expansion of the CD11c⁺CD21⁻ B-cell subset (73). Taking into account data from the ABIS Gene Viewer tool (74), which shows a B-cell- and plasmablast-specific expression of FAM30A (Supplementary Figure 3), the possibility of FAM30A exerting ceRNA effects on IFNLR1 and, thus, affecting B cells in LN becomes even more probable. This is an important finding considering the relationship of B-cell activity and antibody production with disease activity in LN. Given the negative correlation of FAM30A expression with SLEDAI-2K levels, it would be interesting to study the effect of FAM30A overexpression on disease activity, as it may reveal inhibitory effects of FAM30A on Ig gene transcription, resulting in lower antibody production and averting B-cell-activity-induced flares. Further research is needed to elucidate the exact mechanisms of FAM30A involvement in disease activity and assess the therapeutic potential of its overexpression.

This study elucidates the landscape of non-coding transcriptome in autoimmunity, by using total RNA sequencing to investigate the association of lncRNAs with disease activity in LN. Subsequent *in silico* analysis focusing on significant lncRNAs suggests the potential way of action of these lncRNAs. A potential limitation of our data is the use of SLEDAI-2K as a measure of disease activity as this index is not specific to renal activity of SLE. A validated index for lupus nephritis is not available. Additionally, the treatments the recruited patients received are another potential limitation, yet no pattern of medication effect was observed during exploratory analysis of our transcriptomic data, and sequencing results were experimentally replicated in an independent cohort.

In summary, using transcriptomic data from blood samples of LN patients and a network-based approach, we emphasize the key role of IFN pathway and reveal the importance of B-cell activity and antibody production in LN flares. Furthermore, our findings identified four lncRNAs, NRIR, KLHDC7B-DT (IFN-related), and MIR600HG, FAM30A (B-cell-related), as potential biomarkers of disease activity and central components of IFN signaling and adaptive immunity with a possible cis- and trans-effect on genes

of the same pathways. Finally, we provide a network of trans-targets of FAM30A and MIR600HG, emphasizing the most prominent one, IFNLR1, and further corroborating the regulatory involvement of FAM30A in B-cell signaling, immunoglobulin gene expression, antibody production, and, consequently, in LN disease activity.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://ega-archive.org/>, EGAS00001007117.

Ethics statement

The studies involving human participants were reviewed and approved by Attikon University Hospital Research Ethic Committee, Athens, Greece. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

GS and AF designed the experiments and analysis. GS performed the experiments, analyzed and interpreted data, generated figures, and wrote the manuscript. CL analyzed and interpreted data. NM participated in the library construction. DN, SF, and AB performed clinical evaluation of patients and provided human specimens. TM participated in the design of validation experiments and the interpretation of data. MG participated in sample preparation. DN, TM, MG, DB, and AF also participated in the editing of the manuscript. DB and AF acquired the grants and supervised the study, analysis of data, and the writing of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by a research grant from the State Scholarships Foundation (IKY - MIS-5033021) to AF and research grants from the EU (SYSCID grant agreement number 733100) and European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (GAN:742390) to DB. Computational time was granted from the National Infrastructures for Research and Technology S.A. (GRNET S.A.) in the National HPC facility - ARIS - under project ID pr011006_taskp-BIOSLE.

Acknowledgments

We thank Dr. Ioannis Vatsellas from the Greek Genome Center (BRFAA) for the next-generation sequencing service. We thank the patients and their referring physicians and nurses at the Attikon Rheumatology Unit. We also thank Maria Karvouni, MSc, for valuable discussions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Christou EAA, Banos A, Kosmara D, Bertias GK, Boumpas DT. Sexual dimorphism in SLE: above and beyond sex hormones. *Lupus* (2019) 28(1):3–10. doi: 10.1177/0961203318815768
- Kandane-Rathnayake R, Kent JR, Louthrenoo W, Luo SF, Wu YJ, Lateef A, et al. Longitudinal associations of active renal disease with irreversible organ damage accrual in systemic lupus erythematosus. *Lupus* (2019) 28(14):1669–77. doi: 10.1177/0961203319887799
- Fanourakis A, Kostopoulou M, Cheema K, Anders HJ, Aringer M, Bajema I, et al. 2019 Update of the joint European league against rheumatism and European renal association-European dialysis and transplant association (EULAR/ERA-EDTA) recommendations for the management of lupus nephritis. *Ann Rheum Dis* (2020) 79(6):S713–23. doi: 10.1136/annrheumdis-2020-216924
- Fanourakis A, Tziolos N, Bertias G, Boumpas DT. Update in the diagnosis and management of systemic lupus erythematosus. *Ann Rheumatic Dis* (2021) 80:14–25. doi: 10.1136/annrheumdis-2020-218272
- Frangou E, Garantziotis P, Grigoriou M, Banos A, Nikolopoulos D, Pieta A, et al. Cross-species transcriptome analysis for early detection and specific therapeutic targeting of human lupus nephritis. *Ann Rheum Dis* (2022) 81(10):1409–19. doi: 10.1136/annrheumdis-2021-222069
- Ugarte-Gil MF, Mendoza-Pinto C, Reátegui-Sokolova C, Pons-Estel GJ, Van Vollenhoven RF, Bertias G, et al. Achieving remission or low disease activity is associated with better outcomes in patients with systemic lupus erythematosus: a systematic literature review. *Lupus Sci Med* (2021) 8. doi: 10.1136/lupus-2021-000542
- Fitzgerald KA, Caffrey DR. Long noncoding RNAs in innate and adaptive immunity. *Curr Opin Immunol* (2014) 26:140–6. doi: 10.1016/j.coi.2013.12.001
- Bocchetti M, Scrima M, Melisi F, Luce A, Sperlongano R, Caraglia M, et al. LncRNAs and immunity: coding the immune system with noncoding oligonucleotides. *Int J Mol Sci* (2021) 22:1–30. doi: 10.3390/ijms22041741
- Kopp F, Mendell JT. Functional classification and experimental dissection of long noncoding RNAs. *Cell* (2018) 172:393–407. doi: 10.1016/j.cell.2018.01.011
- Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet* (2009) 10:155–9. doi: 10.1038/nrg2521
- Zhou RS, Zhang EX, Sun QF, Ye ZJ, Liu JW, Zhou DH, et al. Integrated analysis of lncRNA-miRNA-mRNA ceRNA network in squamous cell carcinoma of tongue. *BMC Cancer* (2019) 19(1). doi: 10.1186/s12885-019-5983-8
- Li Y, Jiang T, Zhou W, Li J, Li X, Wang Q, et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. *Nat Commun* (2020) 11(1). doi: 10.1038/s41467-020-14802-2
- Fu Y, Wei X, Han Q, Le J, Ma Y, Lin X, et al. Identification and characterization of a 25-lncRNA prognostic signature for early recurrence in hepatocellular carcinoma. *BMC Cancer* (2021) 21(1). doi: 10.1186/s12885-021-08827-z
- Ye DQ, Wu GC, Hu Y, Guan SY, Pan HF. Differential plasma expression profiles of long non-coding RNAs reveal potential biomarkers for systemic lupus erythematosus. *Biomolecules* (2019) 9(6). doi: 10.3390/biom9060206
- Ye H, Wang X, Wang L, Chu X, Hu X, Sun L, et al. Full high-throughput sequencing analysis of differences in expression profiles of long noncoding RNAs and their mechanisms of action in systemic lupus erythematosus. *Arthritis Res Ther* (2019) 21(1). doi: 10.1186/s13075-019-1853-7
- Wang J, Wu X, Tu Y, Dang J, Cai Z, Liao W, et al. An integrated analysis of lncRNA and mRNA expression profiles in the kidneys of mice with lupus nephritis. *PeerJ* (2021) 9. doi: 10.7717/peerj.10668
- Wu GC, Li J, Leng RX, Li XP, Li XM, Wang DG, et al. Identification of long non-coding RNAs *GAS5*, *linc0597* and *linc-DC* in plasma as novel biomarkers for systemic lupus erythematosus. Available at: www.impactjournals.com/oncotarget.
- You Y, Zhao X, Wu Y, Mao J, Ge L, Guo J, et al. Integrated transcriptome profiling revealed that elevated long non-coding RNA-AC007278.2 expression repressed CCR7 transcription in systemic lupus erythematosus. *Front Immunol* (2021) 12:615859. doi: 10.3389/fimmu.2021.615859
- Zhang F, Wu L, Qian J, Qu B, Xia S, La T, et al. Identification of the long noncoding RNA NEAT1 as a novel inflammatory regulator acting through MAPK pathway in human lupus. *J Autoimmun* (2016) 75:96–104. doi: 10.1016/j.jaut.2016.07.012
- Nikolopoulos D, Kostopoulou M, Pieta A, Karageorgas T, Tseronis D, Chavatzas K, et al. Evolving phenotype of systemic lupus erythematosus in caucasians: low incidence of lupus nephritis, high burden of neuropsychiatric disease and increased rates of late-onset lupus in the 'Attikon' cohort. *Lupus* (2020) 29(5):514–22. doi: 10.1177/0961203320908932
- Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol* (2002) 29(2):288–91.
- Andrews S. *FastQC - a quality control tool for high throughput sequence data*. Available at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* (2011) 17(1):10–2. doi: 10.14806/ej.17.1.200
- Frankish A, Diekhans M, Jungreis I, Lagarde J, Loveland JE, Mudge JM, et al. GENCODE 2021. *Nucleic Acids Res* (2021) 49(D1):D916–23. doi: 10.1093/nar/gkaa1087

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1203848/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

(A) Bubble plot of Gene Ontology terms found as significantly enriched in the *Darkgrey*, *Violet* and *Steelblue* modules. Color represents adjusted p-values and size represents the number of genes related to a term found in each module. (B) Scatterplot of SLEDAI-2K Gene Significance against Module Membership for each gene in the (Left to right) *Darkgrey*, *Violet* and *Steelblue* module. Genes with MM>0.8 (Hub genes) are shown in color, with and blue for Other (Non-lncRNA) RNA types. Genes with MM ≤ 0.8 (Non-hub genes) are shown in grey.

SUPPLEMENTARY FIGURE 2

Plot of the genomic region surrounding NRIR (A), KLHDC7B-DT (B) and MIR600HG (C). The genomic region depicted corresponds to 12Kbp upstream of gene start position and 12Kbp downstream its end position. Identified transcripts of genes found in the area are shown in gold (exons) connected by grey lines with arrows (introns). The exact position of the locus in the human genome is marked by the red line on the Chromosome ideogram on the top of the figure.

SUPPLEMENTARY FIGURE 3

Cell-type-specific expression of FAM30A as indicated by ABIS (ABsolute Immune Signal) tool.

25. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* (2013) 29(1):15–21. doi: 10.1093/bioinformatics/bts635
26. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience* (2021) 10(2). doi: 10.1093/gigascience/giab008
27. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* (2015) 31(2):166–9. doi: 10.1093/bioinformatics/btu638
28. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* (2010) 26(1):139–40. doi: 10.1093/bioinformatics/btp616
29. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (2022).
30. Wickham H. *ggplot2*. New York, NY: Springer New York (2009).
31. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. *Cell Syst* (2015) 1(6):417–25. doi: 10.1016/j.cels.2015.12.004
32. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* (2005) 102(43):15545–50. doi: 10.1073/pnas.0506580102
33. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf* (2008) 9. doi: 10.1186/1471-2105-9-559
34. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. G:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res* (2007) 35(SUPPL.2). doi: 10.1093/nar/gkm226
35. Volders PJ, Anckaert J, Verheggen K, Nuytens J, Martens L, Mestdagh P, et al. Lncipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Res* (2019) 47(D1):D135–9. doi: 10.1093/nar/gky1031
36. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2022. *Nucleic Acids Res* (2022) 50(D1):D988–95. doi: 10.1093/nar/gkab1049
37. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* (2005) 21(16):3439–40. doi: 10.1093/bioinformatics/bti525
38. Hahne F, Ivanek R. Visualizing genomic data using gviz and bioconductor. *Methods Mol Biol* (2016) 1418:335–51. doi: 10.1007/978-1-4939-3578-9_16
39. Gel B, Serra E. KaryoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* (2017) 33(19):3088–90. doi: 10.1093/bioinformatics/btx346
40. Wang P, Guo Q, Qi Y, Hao Y, Gao Y, Zhi H, et al. LncACTdb 3.0: an updated database of experimentally supported ceRNA interactions and personalized networks contributing to precision medicine. *Nucleic Acids Res* (2022) 50(D1):D183–9. doi: 10.1093/nar/gkab1092
41. Sticht C, de la Torre C, Parveen A, Gretz N. Mirwalk: an online resource for prediction of microRNA binding sites. *PLoS One* (2018) 13(10). doi: 10.1371/journal.pone.0206239
42. Huang HY, Lin YCD, Li J, Huang KY, Shrestha S, Hong HC, et al. MiRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res* (2020) 48(D1):D148–54. doi: 10.1093/nar/gkz896
43. Csárdi G, Nepusz T. The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695. (2006) Available at: <https://igraph.org>.
44. Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: Network visualizations of relationships in psychometric data. *J Stat Softw* 48(4):1–18. doi: 10.18637/jss.v048.i04
45. Thomas M, Pedersen L. *Type package title an implementation of grammar of graphics for graphs and networks version 2.1.0* (2022). Available at: <https://orcid.org/0000-0002-5147-4711>.
46. Rönnblom L, Leonard D. Interferon pathway in SLE: one key to unlocking the mystery of the disease. *Lupus Sci Med* (2019) 6(1). doi: 10.1136/lupus-2018-000270
47. Manolakou T, Nikolopoulos D, Gkikas D, Filia A, Samiotaki M, Stamatakis G, et al. ATR-mediated DNA damage responses underlie aberrant B cell activity in systemic lupus erythematosus. *Sci Adv* (2022) 8. doi: 10.1126/sciadv.abo5840
48. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* (2011) 146(3):353–8. doi: 10.1016/j.cell.2011.07.014
49. Panousis NI, Bertias GK, Ongen H, Gergianaki I, Tektonidou MG, Trachana M, et al. Combined genetic and transcriptome analysis of patients with SLE: distinct, targetable signatures for susceptibility and severity. *Ann Rheum Dis* (2019) 78(8):1079–89. doi: 10.1136/annrheumdis-2018-214379
50. Nakano M, Ota M, Takeshima Y, Iwasaki Y, Hatano H, Nagafuchi Y, et al. Distinct transcriptome architectures underlying lupus establishment and exacerbation. *Cell* (2022) 185(18):3375–3389.e21. doi: 10.1016/j.cell.2022.07.021
51. Yao M, Gao C, Zhang C, Di X, Liang W, Sun W, et al. Identification of molecular markers associated with the pathophysiology and treatment of lupus nephritis based on integrated transcriptome analysis. *Front Genet* (2020) 11. doi: 10.3389/fgene.2020.583629
52. Shen L, Lan L, Zhu T, Chen H, Gu H, Wang C, et al. Identification and validation of IFI44 as key biomarker in lupus nephritis. *Front Med (Lausanne)* (2021) 8. doi: 10.3389/fmed.2021.762848
53. Malvar A, Pirruccio P, Alberton V, Lococo B, Recalde C, Fazini B, et al. Histologic versus clinical remission in proliferative lupus nephritis. *Nephrol Dialysis Transplantation* (2017) 32(8):1338–44. doi: 10.1093/ndt/gfv296
54. Mariotti B, Servaas NH, Rossato M, Tamassia N, Cassatella MA, Cossu M, et al. The long non-coding RNA NRIR drives IFN-response in monocytes: implication for systemic sclerosis. *Front Immunol* (2019) 10(JAN). doi: 10.3389/fimmu.2019.00100
55. Peng Y, Luo X, Chen Y, Peng L, Deng C, Fei Y, et al. LncRNA and mRNA expression profile of peripheral blood mononuclear cells in primary sjögren's syndrome patients. *Sci Rep* (2020) 10(1). doi: 10.1038/s41598-020-76701-2
56. Shen M, Duan C, Xie C, Wang H, Li Z, Li B, et al. Identification of key interferon-stimulated genes for indicating the condition of patients with systemic lupus erythematosus. *Front Immunol* (2022) 13. doi: 10.3389/fimmu.2022.962393
57. Li M, Wang H, Yuan C, Ma Z, Jiang B, Li L, et al. KLHDC7B-DT aggravates pancreatic duct adenocarcinoma development via inducing cross-talk between cancer cells and macrophages. *Clin Sci* (2021) 135(4):629–49. doi: 10.1042/CS20201259
58. Yin X, Yang Z, Zhu M, Chen C, Huang S, Li X, et al. ILF2 contributes to hyperproliferation of keratinocytes and skin inflammation in a KLHDC7B-DT-Dependent manner in psoriasis. *Front Genet* (2022) 13. doi: 10.3389/fgene.2022.890624
59. Li X, Yang Y, Sun G, Dai W, Jie X, Du Y, et al. Promising targets and drugs in rheumatoid arthritis a MODULE-BASED AND CUMULATIVELY SCORING APPROACH. *Bone Joint Res* (2020) 9(8):501–14. doi: 10.1302/2046-3758.98.BJR-2019-0301.R1
60. Lv J, Chen L, Wang X, Gao Q, Zhao L. Immune-relevant genes of systemic lupus erythematosus by transcriptome profiling analysis. *Cytokine* (2022) 158. doi: 10.1016/j.cyto.2022.155975
61. Asmiyau A, Bakr AM, Shahin DA, Wahba Y. CD40 and CD72 expression and prognostic values among children with systemic lupus erythematosus: a case-control study. *Lupus* (2020) 29(10):1270–6. doi: 10.1177/096120320941931
62. Lai JH, Wu DW, Wu CH, Hung LF, Huang CY, Ka SM, et al. Mitochondrial CMPK2 mediates immunomodulatory and antiviral activities through IFN-dependent and IFN-independent pathways. *iScience* (2021) 24(6). doi: 10.1016/j.isci.2021.102498
63. Li Y, Ma C, Liao S, Qi S, Meng S, Cai W, et al. Combined proteomics and single cell RNA-sequencing analysis to identify biomarkers of disease diagnosis and disease exacerbation for systemic lupus erythematosus. *Front Immunol* (2022) 13. doi: 10.3389/fimmu.2022.969509
64. Zhu H, Zhu C, Mi W, Chen T, Zhao H, Zuo X, et al. Integration of genome-wide DNA methylation and transcription uncovered aberrant methylation-regulated genes and pathways in the peripheral blood mononuclear cells of systemic sclerosis. *Int J Rheumatol* (2018) 2018. doi: 10.1155/2018/7342472
65. Ryu J, Woo J, Shin J, Ryoo H, Kim Y, Lee C. Profile of differential promoter activity by nucleotide substitution at GWAS signals for multiple sclerosis. *Med (United States)* (2014) 93(28):e281. doi: 10.1097/MD.0000000000000281
66. Jeong G, Bae H, Jeong D, Ham J, Park S, Kim HW, et al. A kelch domain-containing KLHDC7B and a long non-coding RNA ST8SIA6-AS1 act oppositely on breast cancer cell proliferation via the interferon signaling pathway. *Sci Rep* (2018) 8(1). doi: 10.1038/s41598-018-31306-8
67. Erickson LD, Tygrett LT, Bhatia SK, Grabstein KH, Waldschmidt TJ. Differential expression of CD22 (Lyb8) on murine B cells. *Int Immunol* (1996) 8(7):1121–9. doi: 10.1093/intimm/8.7.1121
68. Le TT, Blackwood NO, Taroni JN, Fu W, Breitenstein MK. Integrated machine learning pipeline for aberrant biomarker enrichment (i-mAB): characterizing clusters of differentiation within a compendium of systemic lupus erythematosus patients. *AMIA Annu Symp Proc* (2018) 2018:1358–67.
69. de Lima DS, Cardozo LE, Maracaja-Coutinho V, Suhrbier A, Mane K, Jeffries D, et al. Long noncoding RNAs are involved in multiple immunological pathways in response to vaccination. *Proc Natl Acad Sci USA* (2019) 116(34):17121–6. doi: 10.1073/pnas.1822046116
70. The GTEx Consortium. The GTEx consortium atlas of genetic regulatory effects across human tissues. *Sci* (1979) (2020) 369:1318–30. doi: 10.1126/science.aaz1776
71. Zogopoulos VL, Malatras A, Kyriakidis K, Charalampous K, Makrygianni EA, Duguez S, et al. HGCA2.0: an RNA-seq based webtool for gene coexpression analysis in homo sapiens. *Cells* (2023) 12(3). doi: 10.3390/cells12030388
72. Goel RR, Wang X, O'Neil LJ, Nakabo S, Hasneen K, Gupta S, et al. *Interferon lambda promotes immune dysregulation and tissue inflammation in TLR7-induced lupus*. Available at: <https://www.ncbi.nlm.nih.gov/geo/>.
73. Barnas JL, Albrecht J, Meednu N, Alzamareh DF, Baker C, McDavid A, et al. B cell activation and plasma cell differentiation are promoted by IFN- λ in systemic lupus erythematosus. *J Immunol* (2021) 207(11):2660–72. doi: 10.4049/jimmunol.2100339
74. Monaco G, Lee B, Xu W, Mustafah S, Hwang YY, Carré C, et al. RNA-Seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Rep* (2019) 26(6):1627–1640.e7. doi: 10.1016/j.celrep.2019.01.041