**Original Research**

# Meta-analysis of gene expression profiles of peripheral blood cells in systemic lupus erythematosus

Sang-Cheol Bae[1], Young Ho Lee[2]*

[1] Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Korea
[2] Department of Rheumatology, Korea University College of Medicine, Seoul, Korea

**Correspondence to:** lyhcgh@korea.ac.kr

**Abstract:** Our purpose was to identify differentially expressed (DE) genes and biological processes associated with gene expression changes in systemic lupus erythematosus (SLE). We performed a meta-analysis using the INMEX program (integrative meta-analysis of expression data) on publicly available microarray Genetic Expression Omnibus (GEO) datasets of peripheral blood mononuclear cells (PBMCs) of SLE patients and healthy controls. We performed Gene Ontology (GO) enrichment analysis by using hypergeometric tests. In total, five comparisons (2 B cells, 2 CD4 T cells, and 1 myeloid cell) from two GEO datasets containing 51 cases and 46 controls were included in the meta-analysis. We identified 483 genes consistently differentially expressed in SLE (260 up-regulated and 223 down-regulated). The up-regulated gene with the lowest P-value (P-value = 7.33E-10) was TAP1 (transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)). The up-regulated gene with the largest effect size (ES = 2.7799, P-value = 8.28E-06) was STAT1 (signal transducer and activator of transcription 1, 91 kDa). The down-regulated gene with the lowest P-value (P-value = 2.53E-06) was EIF3F (eukaryotic translation initiation factor 3, subunit F), and the down-regulated gene with the largest ES (ES = −1.8543, P-value = 8.56E-06) was FBL (fibrillarin). The most significant enrichment among 317 GO categories was the type I interferon (IFN)-mediated signaling pathway with P = 1.93E-20. Other significant GO categories included cellular response to type I IFN (P = 1.93E-20) and response to type I IFN (P = 2.86E-20). Our meta-analysis demonstrates that up-regulated genes mediate IFN-regulated and cytokine-mediate signaling pathways, innate immune response, and antigen processing and presentation. These results provide insights into molecular mechanisms associated to the pathophysiology of SLE.

*Key words:* SLE; Gene expression; Meta-analysis.

## Introduction

Systemic lupus erythematosus (SLE) is the prototype of systemic autoimmune diseases, and it is characterized by autoantibody production and immune complex formation, both of which lead to intense inflammation and multiple organ damage. Although the etiopathology of SLE remains largely unknown, strong genetic components that influence disease susceptibility and modify its clinical expression are well established(1). A combination of affected genes and biological pathways, all with small contributions, is believed to cause SLE expression. Autoimmunity is considered the main etiology of SLE, but the biological mechanisms associated with the disease remain unclear.

High-throughput genomic technologies such as microarrays have been developed to improve our understanding of complex interactions and networks during disease development. Microarrays measure gene expression on a genome-wide scale(2). Alterations in gene profiles can be correlated to altered gene function and abnormal biochemical activities. Microarrays are a powerful tool that have become one of the most rapidly growing techniques in medical research(3).

Identification of differentially regulated genes in SLE may lead to potential biomarkers for the disease, and provide insights into its pathogenesis. Identification of gene expression signatures that differentiate disease from health depends on the size and quality of the samples, and has generally yielded heterogeneous datasets(4). Although many microarray studies have generated lists of differentially expressed (DE) genes, these tend to be inconsistent among studies, because most are limited by small sample sizes (5).

To address these challenges, meta-analysis has been applied to publically available genome-wide gene expression datasets of diseases (6,7). Meta-analysis can enhance reliability and generalizability of the results and obtain a more precise estimate of gene expression(8). Meta-analysis enhances statistical power that allows the identification of more robust and reliable gene signatures (9-11). Recently, a new user-friendly microarray meta-analysis tool, integrative meta-analysis of expression data (INMEX), has been developed to support meta-analysis of multiple gene-expression datasets, as well as to enable integration of datasets from gene expression to biological pathways(12).

In order to overcome the limitations posed by individual studies, to resolve inconsistencies, and to reduce the rate of false-positive and false-negative associations generated by random errors, we turned to meta-analysis. The objective of this study was to identify DE genes and biological processes associated with gene expression changes in SLE by using a microarray meta-analysis approach.

## Materials and Methods

### Identification of eligible gene expression datasets of SLE

A search for microarray datasets that examined DE genes between SLE and healthy controls was conducted. We used NCBI GEO (Gene Expression Omnibus) database (http://www.ncbi.nlm.nih.gov/geo/) to identify microarray datasets through July 2013(13). The following keyword was used for the search: "SLE". Studies were included in the analysis if: (1) they were case-controlled, (2) contained gene expression profiling of peripheral blood mononuclear cells (PBMCs), and (3) included patients diagnosed with SLE based on the SLE classification criteria. We excluded the following: (1) animal studies and (2) studies in which the microarray data could not be ascertained for meta-analysis. Data were extracted from the original studies by two independent researchers. Discrepancies between researchers were resolved by consensus or by a third party. The following information was extracted from each selected study: (1) GEO accession, (2) sample type, (3) platform, (4) numbers of cases and controls, (5) references, and (6) gene expression data.

### Meta-analysis of microarray datasets

All available microarray datasets of SLE were downloaded from the NCBI GEO database. We made data tables containing gene expression or relative expression values with genes/probes in rows and samples/experiments delineated in individual columns. After we uploaded the datasets in INMEX (http://www.inmex.ca/INMEX), we annotated them by converting different gene or probe IDs to Entrez IDs. We inspected the data to determine whether the class labels were consistently used across different datasets. For each probeset, intensity values were log-transformed and normalized to zero mean and unit variance(14). When all datasets were uploaded, processed, and annotated, we performed a data integrity check before proceeding to the meta-analysis stage. Standardized difference, also known as effect size (ES), is the difference between the means of two groups divided by the standard deviation of the data (Cohen's 'd'). Standardized differences are considered combinable across studies. There are two methods to do this, namely, the fixed- and the random-effect models(15). The fixed-effect model assumes that a genetic factor has a similar effect on disease susceptibility across all investigated studies and that observed variations among studies are caused by chance alone(16). The random-effect model assumes that different studies show substantial diversity and assesses both within-study sampling error and between-study variance(17). When study groups are homogeneous, the two models behave similarly. If the study groups lack homogeneity, the random-effect model usually provides wider confidence intervals (CIs) than the fixed-effect model. Cochran's Q test is commonly used to evaluate the homogeneity of the datasets. We used a Q–Q plot to evaluate the homogeneity of the datasets and to choose the appropriate model. When the estimated Q values approximate a chi-squared distribution, the fixed-effect model assumption is more appropriate; otherwise, the random-effect model should be used. The random-effect model deals better with the heterogeneous nature encountered during meta-analysis of microarray data, when there is significant heterogeneity between studies(17). Statistical analyses were performed using the INMEX program.

### Functional analysis

INMEX's functional analysis module is designed to generate new hypotheses by taking advantage of inherent characteristics of the DE gene list identified in the previous meta-analysis (12). 'Pattern extractor' visualizes their expression profiles across different datasets/conditions as heatmaps. To further understand the functions of the gene list, we performed GO enrichment analysis using hypergeometric tests to detect enriched functional attributes based on gene-associated GO terms (http://www.geneontology.org/)(18).

## Results

### Studies included in the meta-analysis

Two human datasets met the inclusion criteria(19,20). The studies contained data from three different groups (CD 19 B, CD4 T, myeloid cells)(19), and two different groups (B, CD4 T cells)(20), respectively. These groups were treated independently and, in total, included 51 cases and 46 controls. Selected details of the individual studies are summarized in Table 1. The first study (GEO Accession Number GSE10325) was designed to determine whether specific lymphocyte and myeloid subsets isolated from the blood of SLE patients. The samples were hybridized on Affymetrix Human Genome U133A Array. In the second study (GEO Accession Number GSE4588), CD4 T and B cells were sorted by flow cytometry from PBMCs from SLE patients or healthy controls. Affymetrix Human Genome U133 Plus 2.0 Array was hybridized in monoplicates. The heatmap of the differential expression analysis of individual dataset identified a subset of genes across three studies (Figure 1).

**Table 1.** Characteristics of the individual studies included in the meta-analysis.

| Study [Ref] | GEO Accession | Numbers | | Sample | Platform |
| | | SLE | Control | | |
| --- | --- | --- | --- | --- | --- |
| Datatet2-1(19) | GSE10325 | 14 | 9 | CD 19 B cells | GPL96 Affymetrix Human Genome U133A Array |
| Datatet2-2(19) | GSE10325 | 11 | 8 | CD4 T cells | GPL96 Affymetrix Human Genome U133A Array |
| Datatet2-3(19) | GSE10325 | 11 | 10 | Myeloid cells | GPL96 Affymetrix Human Genome U133A Array |
| Datatet3-1(20) | GSE4588 | 7 | 9 | B cells | GPL570 Affymetrix Human Genome U133 Plus 2.0 Array |
| Datatet3-2(20) | GSE4588 | 8 | 10 | CD4 T cells | GPL570 Affymetrix Human Genome U133 Plus 2.0 Array |

*Ref*, reference; *GEO*, Gene Expression Omnibus; *GSE*, Gene Expression Series; *GPL*, Gene Platform; *SS*, SLE.

**Figure 1.** Heatmap visualization of the patterns of change for a particular gene across different datasets (row-wise comparison). It was generated by re-scaling individual dataset in order to prevent the view from being dominated by study-specific effects. Datatset2 = GSE10325, datatset3 = GSE4588.

## Meta-analysis of gene expressions in PBMCs from SLE

We used a random-effect model on effect size measurements to integrate gene expression patterns and to incorporate between-study heterogeneities because the estimated Q value did no approximate a chi-squared distribution. We identified 372 gained genes and 207 lost genes in this meta-analysis. Gained genes are those DE genes exclusively identified through the meta-analysis(12). These genes show relatively weak but consistent expression profiles across different datasets. They benefited from analyzing more samples; thus, we could confidently declare them as DE genes. Lost genes are those identified as DE genes in any individual analysis, but not in the meta-analysis. These genes either show conflicting changes of expression profile, or very big variations across different studies (i.e. batch effect or system bias due to different platforms).

## Identifying genes differentially expressed in PBMCs from SLE

We identified 483 genes across the studies, which showed a consistently differential expression with SS (Supplementary Table 1). Among the 483 DE genes, 260 genes were up-regulated and 223 genes were down-regulated. The top-30 lists of the up- and down-regulated genes are presented in Table 2 and 3, respectively. The up-regulated gene with the lowest P-value (P-value = 7.33E-10) was TAP1 (transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)). TAP1 protein is

involved in the pumping of degraded cytosolic peptides across the endoplasmic reticulum into the membrane-bound compartment where class I molecules assemble. TAP1 plays a key role in innate immune response(21). The up-regulated gene with the largest ES (ES = 2.7799, P-value = 8.28E-06) was STAT1 (signal transducer and activator of transcription 1, 91 kDa), which product is a member of the STAT protein family. In response to cytokines and growth factors, STAT family members are phosphorylated by receptor-associated kinases and form homo- or hetero-dimers that translocate to the cell nucleus where they act as transcription activators. This protein mediates the expression of a variety of genes, and it is thought to be important for cell viability in response to different cellular stimuli and pathogens. STAT1 is involved in the innate immune response, and in type I interferon (IFN)-regulated and cytokine-mediated signaling pathways(22,23). Many genes that were consistently differentially expressed across the datasets participate in the IFN pathway activation, such as IFI35, STAT1, OASL, GBP1, IFI6, IFIT5, and IFIH1 (Table 2). The down-regulated gene with the lowest P-value (P-value = 2.53E-06) was EIF3F (eukaryotic translation initiation factor 3, subunit F). EIF3F is required during translational initiation, as well as in protein metabolism. It also positively regulates Notch signaling, by deubiquitinating activated NOTCH1 and promoting its nuclear translocation. The down-regulated gene with the largest ES (ES = −1.8543, P-value = 8.56E-06) was FBL (fibrillarin). FBL protein is a component of a nucleolar

**Table 2.** The 30 most up-regulated genes in SLE.

| Entrez ID | Gene, symbol | Combined ES | P-value | Gene name |
|---|---|---|---|---|
| 6890 | TAP1 | 1.8588 | 7.33E-10 | transporter 1, ATP-binding cassette, sub-family B (MDR/TAP) |
| 5721 | PSME2 | 1.8773 | 7.33E-10 | proteasome (prosome, macropain) activator subunit 2 (PA28 beta) |
| 3430 | IFI35 | 1.9045 | 7.33E-10 | interferon-induced protein 35 |
| 2633 | GBP1 | 1.9587 | 5.96E-08 | guanylate binding protein 1, interferon-inducible |
| 4496 | MT1H | 1.5585 | 2.58E-07 | metallothionein 1H |
| 5898 | RALA | 1.4835 | 2.30E-06 | v-ral simian leukemia viral oncogene homolog A (ras related) |
| 51056 | LAP3 | 2.169 | 2.30E-06 | leucine aminopeptidase 3 |
| 6772 | STAT1 | 2.7799 | 8.28E-06 | signal transducer and activator of transcription 1, 91 kDa |
| 5687 | PSMA6 | 1.3328 | 9.58E-06 | proteasome (prosome, macropain) subunit, alpha type, 6 |
| 9997 | SCO2 | 1.3264 | 9.80E-06 | SCO2 cytochrome c oxidase assembly protein |
| 51513 | ETV7 | 1.3273 | 1.02E-05 | ets variant 7 |
| 8638 | OASL | 2.2618 | 1.02E-05 | 2'-5'-oligoadenylate synthetase-like |
| 4493 | MT1E | 1.3164 | 1.06E-05 | metallothionein 1E |
| 4501 | MT1X | 1.7944 | 1.14E-05 | metallothionein 1X |
| 80055 | PGAP1 | 1.3043 | 1.27E-05 | post-GPI attachment to proteins 1 |
| 4494 | MT1F | 1.2881 | 1.57E-05 | metallothionein 1F |
| 57097 | PARP11 | 1.4243 | 1.70E-05 | poly (ADP-ribose) polymerase family, member 11 |
| 644 | BLVRA | 1.2671 | 1.84E-05 | biliverdin reductase A |
| 2537 | IFI6 | 2.1271 | 1.88E-05 | interferon, alpha-inducible protein 6 |
| 9830 | TRIM14 | 1.3265 | 2.26E-05 | tripartite motif containing 14 |
| 5359 | PLSCR1 | 2.0912 | 3.01E-05 | phospholipid scramblase 1 |
| 11274 | USP18 | 2.3386 | 3.97E-05 | ubiquitin specific peptidase 18 |
| 5292 | PIM1 | 1.2117 | 4.59E-05 | pim-1 oncogene |
| 24138 | IFIT5 | 1.3789 | 4.63E-05 | interferon-induced protein with tetratricopeptide repeats 5 |
| 840 | CASP7 | 1.2844 | 6.59E-05 | caspase 7, apoptosis-related cysteine peptidase |
| 7295 | TXN | 1.1814 | 7.06E-05 | thioredoxin |
| 9246 | UBE2L6 | 2.1228 | 8.76E-05 | ubiquitin-conjugating enzyme E2L 6 |
| 5214 | PFKP | 1.271 | 9.35E-05 | phosphofructokinase, platelet |
| 64135 | IFIH1 | 1.8569 | 0.000147 | interferon induced with helicase C domain 1 |
| 1890 | TYMP | 1.1492 | 0.000154 | thymidine phosphorylase |

ES effect size.

small nuclear ribonucleoprotein (snRNP) particle that is thought to participate in the first step during preribosomal RNA processing (Table 3). FBL is also involved in macromolecule metabolism. Although interpreting these results is not straightforward, further studies are encouraged because DE genes could play a role in the pathogenesis of SLE.

### Functional analysis

To identify the biological processes associated with gene expression changes in SLE, we performed GO analysis of the DE genes. We identified 317 significant enrichments among the DE genes (Supplementary Table 2). The top-30 GO list is presented in Table 4. The GO term that is most significantly enriched is type I IFN-mediated signaling pathway (P = 1.93E-20). Other significant GO categories include cellular response to type I IFN (P = 1.93E-20) and response to type I IFN (P = 2.86E-20) (Table 4). In addition, enrichment analysis using GO terms suggests that innate immune response, cytokine-mediated signaling pathway, immune response, antigen processing and presentation, regulation of cell cycle, and p53-mediated DNA-damage response might be involved in SLE pathogenesis (Table 4).

### Discussion

Many genes tend to be differentially expressed in autoimmune diseases including SLE. The challenge is to identify those that would help us to understand the molecular and cellular events during SLE pathogenesis(3).

To identify biological functions relevant to the pathogenesis of the disease in the peripheral blood cells of SLE patients, we have performed meta-analysis of 5 groups of microarray datasets of PBMCs. We identified, across studies, 483 consistently DE genes in SLE: 260 up-regulated and 223 down-regulated. The up-regulated gene with the lowest P-value (P-value = 7.33E-10) was TAP1. The up-regulated gene with the largest ES (ES = 2.7799, P-value = 8.28E-06) was STAT1. TAP1 transporter is involved in the pumping of degraded cytosolic peptides across the endoplasmic reticulum into the membrane-bound compartment. TAP1 plays a key role in the innate immune response and antigen processing and presentation. STAT1 mediates the expression of important genes for cell viability in response to different cellular stimuli and pathogens. A TAP1 polymorphism(24), and polymorphisms affecting STAT1 function(25) or near the STAT1 gene have been asso-

**Table 3.** The 30 most down-regulated genes in SLE.

| Entrez ID | Gene, symbol | Combined ES | P-value | Gene name |
|---|---|---|---|---|
| 8665 | EIF3F | -1.4124 | 2.53E-06 | eukaryotic translation initiation factor 3, subunit F |
| 2091 | FBL | -1.8543 | 8.56E-06 | fibrillarin |
| 8664 | EIF3D | -1.3483 | 9.10E-06 | eukaryotic translation initiation factor 3, subunit D |
| 3757 | KCNH2 | -1.3255 | 1.02E-05 | potassium voltage-gated channel, subfamily H (eag-related), member 2 |
| 293 | SLC25A6 | -1.5699 | 1.20E-05 | solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 6 |
| 51611 | DPH5 | -1.2487 | 2.76E-05 | diphthamide biosynthesis 5 |
| 1975 | EIF4B | -1.6128 | 3.26E-05 | eukaryotic translation initiation factor 4B |
| 6138 | RPL15 | -1.1742 | 8.61E-05 | ribosomal protein L15 |
| 738 | C11orf2 | -1.424 | 0.000429 | chromosome 5 open reading frame, human **C11orf2** |
| 6125 | RPL5 | -1.0706 | 0.000429 | ribosomal protein L5 |
| 64174 | DPEP2 | -1.0697 | 0.000448 | dipeptidase 2 |
| 1397 | CRIP2 | -1.0527 | 0.000529 | cysteine-rich protein 2 |
| 4048 | LTA4H | -1.0342 | 0.000747 | leukotriene A4 hydrolase |
| 5217 | PFN2 | -1.0306 | 0.000747 | profilin 2 |
| 79891 | ZNF671 | -1.023 | 0.000934 | zinc finger protein 671 |
| 56172 | ANKH | -1.0342 | 0.001064 | **ANKH** inorganic pyrophosphate transport regulator |
| 23492 | CBX7 | -0.99968 | 0.001142 | chromobox homolog 7 |
| 4673 | NAP1L1 | -0.99508 | 0.001353 | nucleosome assembly protein 1-like 1 |
| 26751 | SH3YL1 | -0.99016 | 0.001363 | SH3 domain containing, Ysc84-like 1 (*S. cerevisiae*) |
| 10399 | GNB2L1 | -0.98967 | 0.001363 | guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1 |
| 4306 | NR3C2 | -1.035 | 0.001406 | nuclear receptor subfamily 3, group C, member 2 |
| 11331 | PHB2 | -0.98485 | 0.00149 | prohibitin 2 |
| 6432 | SRSF7 | -0.973 | 0.001573 | serine/arginine-rich splicing factor 7 |
| 4216 | MAP3K4 | -0.96369 | 0.001984 | mitogen-activated protein kinase kinase kinase 4 |
| 3590 | IL11RA | -0.95531 | 0.00225 | interleukin 11 receptor, alpha |
| 23387 | SIK3 | -0.94896 | 0.002406 | SIK family kinase 3 |
| 26053 | AUTS2 | -0.94702 | 0.002406 | autism susceptibility candidate 2 |
| 8409 | UXT | -0.94166 | 0.002406 | ubiquitously-expressed, prefoldin-like chaperone |
| 6134 | RPL10 | -0.93524 | 0.002704 | ribosomal protein L10 |
| 79913 | ACTR5 | -0.93582 | 0.002905 | ARP5 actin-related protein 5 homolog (yeast) |

ES, effect size.

ciated with SLE(26). STAT1 is involved in the innate immune response, and in cytokine-mediated and type I IFN-regulated pathways(27). Many genes involved in the IFN signaling pathway, such as IFI35, STAT1, OASL, GBP1, IFI6, IFIT5, and IFIH1, were also identified in our meta-analysis. This overexpression of IFN-regulated genes in PBMCs of SLE patients is in agreement with previous studies(28). The down-regulated gene with the lowest P-value (P-value = 2.53E-06) was EIF3F, and the down-regulated gene with the largest ES (E`S = −1.8543, P-value = 8.56E-06) was FBL. EIF3F and FBL are involved in translational initiation and protein metabolism, and macromolecule metabolism, respectively. Although interpreting these results is not straightforward, further studies are encouraged because DE genes could play a role in SLE pathogenesis.

The roles of all DE genes associated with SLE have not been elucidated yet. However, a large proportion of the identified genes are involved in IFN regulation, innate immune response, signaling pathway, antigen processing and presentation, and cell cycle regulation. These results could provide insights into molecular me-chanisms relevant to the pathophysiology of SLE.

We identified 317 significant GO enrichments among the DE genes. The 3 GO categories that were most significantly enriched were type I IFN-mediated signaling pathway, cellular response to type I IFN, and response to type I IFN. All of them are related to IFN pathway. In addition, GO enrichment analysis suggests that immune response mechanisms, cytokine-mediated signaling pathway, antigen processing and presentation, regulation of cell cycle, and p53-mediated DNA-damage response are all involved in SLE pathogenesis. The GO categories we identified deserve further investigation, despite that at this moment it is impossible for us to discuss all the significant functional categories differentially expressed in SLE.

Our observations are in agreement with previous studies that demonstrated that the innate immune reaction and IFN pathways are major up-regulated pathways in SLE(29,30). The increased expression of genes involved in the IFN pathway suggests that it is a major factor in the biological mechanism of SLE(28,29). Because this conclusion was drawn from pathways that were enriched

**Table 4.** The top 30 enriched GO terms among the DE genes in SLE

| GO ID | Term | P-value | Genes |
|---|---|---|---|
| GO:0060337 | type I interferon-mediated signaling pathway | 1.93E-20 | IFI35; STAT1; OASL; IFI6; USP18; ISG20; IRF7; ISG15; OAS3; IFI27; MX1; IFIT3; IFITM1; XAF1; GBP2; IFITM2; HLA-B; HLA-A; IFIT1; OAS2; HLA-C; IFITM3; ADAR; MX2; OAS1 |
| GO:0071357 | cellular response to type I interferon | 1.93E-20 | IFI35; STAT1; OASL; IFI6; USP18; ISG20; IRF7; ISG15; OAS3; IFI27; MX1; IFIT3; IFITM1; XAF1; GBP2; IFITM2; HLA-B; HLA-A; IFIT1; OAS2; HLA-C; IFITM3; ADAR; MX2; OAS1 |
| GO:0034340 | response to type I interferon | 2.86E-20 | IFI35; STAT1; OASL; IFI6; USP18; ISG20; IRF7; ISG15; OAS3; IFI27; MX1; IFIT3; IFITM1; XAF1; GBP2; IFITM2; HLA-B; HLA-A; IFIT1; OAS2; HLA-C; IFITM3; ADAR; MX2; OAS1 |
| GO:0009615 | response to virus | 2.61E-19 | IFI35; STAT1; OASL; PLSCR1; IFIH1; ISG20; RSAD2; IRF7; ISG15; OAS3; HERC5; DHX58; IFI44; IFI44L; MX1; APOBEC3A; PML; IFITM1; DDX60; TRIM5; PCBP2; EIF2AK2; IFITM2; BST2; APOBEC3G; HLA-A; IFIT1; EEF1G; IFI16; BCL2L1; OAS2; DDX58; CXCL10; IFITM3; TRIM22; CREB3; CFL1; MX2; TBK1; MYD88; OAS1 |
| GO:0045087 | innate immune response | 2.62E-17 | TAP1; IFI35; GBP1; STAT1; OASL; IFI6; PLSCR1; USP18; TXN; UBE2L6; IFIH1; MT2A; ISG20; IRF7; ISG15; OAS3; HERC5; DHX58; RAB27A; IFI27; MX1; APOBEC3A; IFIT3; PML; IFITM1; DDX60; LY96; NLRP1; NFKBIB; PCBP2; GCH1; XAF1; GBP2; IFITM2; HLA-B; BST2; CASP1; APOBEC3G; HLA-A; SERPING1; IFIT1; TRAFD1; ZBP1; BCL2L1; CD44; OAS2; DDX58; NFKBIA; HLA-C; IFITM3; PANX1; ADAR; MX2; TBK1; MYD88; CEBPG; OAS1 |
| GO:0051707 | response to other organism | 4.53E-13 | IFI35; STAT1; OASL; SLC25A6; PLSCR1; IFIH1; ISG20; RSAD2; IRF7; ISG15; OAS3; HERC5; HIST1H2BC; DHX58; IFI44; IFI44L; MX1; APOBEC3A; PML; IFITM1; DDX60; LY96; TRIM5; NLRP1; PCBP2; GCH1; EIF2AK2; IFITM2; HLA-B; BST2; ABCC5; APOBEC3G; HLA-A; HNRNPA0; IFIT1; EEF1G; IFI16; BCL2L1; OAS2; DDX58; NFKBIA; CXCL10; IFITM3; FAS; TRIM22; CREB3; CFL1; MX2; TBK1; MYD88; ANXA3; OAS1 |
| GO:0019221 | cytokine-mediated signaling pathway | 6.66E-13 | IFI35; GBP1; STAT1; OASL; IFI6; USP18; UBE2L6; MT2A; ISG20; IRF7; ISG15; OAS3; HERC5; IL15RA; LRP8; IFI27; MX1; IFIT3; PML; IFITM1; XAF1; GBP2; IFITM2; HLA-B; HLA-A; IFIT1; KIT; CD44; OAS2; HLA-C; IFITM3; ADAR; NUP37; NUP62; STAT3; MX2; MYD88; OAS1; VRK2 |
| GO:0006952 | defense response | 8.55E-13 | TAP1; IFI35; GBP1; STAT1; PSMA6; OASL; SLC25A6; IFI6; PLSCR1; USP18; TXN; UBE2L6; IFIH1; NMI; MT2A; ISG20; RSAD2; IRF7; LTA4H; ISG15; OAS3; HERC5; HIST1H2BC; DHX58; RAB27A; IFI27; MX1; APOBEC3A; IFIT3; PML; IFITM1; DDX60; LY96; NLRP1; NFKBIB; PCBP2; GCH1; EIF2AK2; XAF1; GBP2; LGALS3BP; IFITM2; HLA-B; BST2; CASP1; APOBEC3G; HLA-A; HNRNPA0; SERPING1; IFIT1; KIT; TRAFD1; KCNN4; JAM3; ZBP1; BCL2L1; CD44; OAS2; DDX58; NFKBIA; CXCL10; HLA-C; IFITM3; PANX1; ADAR; STAT3; CREB3; MX2; TBK1; MYD88; SIGLEC1; CEBPG; ANXA3; ADRB2; OAS1; APOL2 |
| GO:0009607 | response to biotic stimulus | 3.38E-12 | IFI35; STAT1; OASL; SLC25A6; PLSCR1; IFIH1; ISG20; RSAD2; IRF7; ISG15; OAS3; HERC5; HIST1H2BC; DHX58; IFI44; IFI44L; MX1; APOBEC3A; PML; IFITM1; DDX60; LY96; TRIM5; NLRP1; PCBP2; GCH1; EIF2AK2; IFITM2; HLA-B; BST2; ABCC5; APOBEC3G; HLA-A; HNRNPA0; IFIT1; EEF1G; IFI16; BCL2L1; OAS2; DDX58; NFKBIA; CXCL10; IFITM3; FAS; TRIM22; CREB3; CFL1; MX2; TBK1; MYD88; ANXA3; OAS1 |
| GO:0016032 | viral reproduction | 5.26E-12 | PSME2; TAP1; PSMA6; OASL; SLC25A6; PLSCR1; RPL5; RPL5; RSAD2; IRF7; PSMB9; OAS3; PSMA4; PSME1; RPL10; RPL7; APOBEC3A; RPL18; RPL4; EIF2AK2; RPL35A; RPS8; ZSCAN18; APOBEC3G; HLA-A; PSMA3; PSMD8; IFIT1; PSMB10; BCL2L1; PSME4; PSMA5; RPL3; NUP37; NUP62; CREB3; RBX1; RPL8; RPS21; RPL13; OAS1; ZNF274 |
| GO:0006955 | immune response | 1.01E-11 | TAP1; IFI35; GBP1; STAT1; OASL; IFI6; PLSCR1; USP18; TXN; UBE2L6; IFIH1; MT2A; ISG20; RSAD2; IRF7; ISG15; OAS3; HERC5; DHX58; IFI44L; RAB27A; IFI27; MX1; APOBEC3A; IFIT3; PML; IFITM1; DDX60; LY96; INPP5D; NLRP1; NFKBIB; PCBP2; GCH1; PRKCB; XAF1; GBP2; HLX; IFITM2; HLA-B; BST2; CASP1; APOBEC3G; HLA-A; CRIP1; SERPING1; IFIT1; PSMB10; KIT; TRAFD1; JAM3; ZBP1; BCL2L1; CD44; OAS2; ICAM3; DDX58; NFKBIA; CXCL10; HLA-C; IFITM3; PANX1; ADAR; CD38; FAS; TRIM22; MX2; TBK1; MYD88; CEBPG; ANXA3; OAS1 |
| GO:0051704 | multi-organism process | 1.60E-11 | TAP1; IFI35; RALA; STAT1; OASL; SLC25A6; PLSCR1; IFIH1; ISG20; RSAD2; IRF7; PSMB9; ISG15; OAS3; PSMA4; HERC5; GNB2L1; HIST1H2BC; DHX58; IFI44; LDLR; IFI44L; MX1; APOBEC3A; PML; IFITM1; DDX60; LY96; TRIM5; NLRP1; PCBP2; GCH1; EIF2AK2; IFITM2; HLA-B; BST2; ABCC5; APOBEC3G; HLA-A; PSMA3; PSMB2; HNRNPA0; IFIT1; PSMB10; EEF1G; IFI16; BCL2L1; OAS2; BUB1; DDX58; GADD45GIP1; NFKBIA; CXCL10; HLA-C; IFITM3; CFLAR; CD38; FAS; TRIM22; STAT3; CREB3; CFL1; MX2; TBK1; RBX1; MYD88; ANXA3; OAS1; TOP1; NXF1; VRK2; APOL2 |

| GO ID | Term | p-value | Genes |
|---|---|---|---|
| GO:0006950 | response to stress | 3.52E-11 | PSME2; TAP1; IFI35; GBP1; STAT1; PSMA6; OASL; SLC25A6; IFI6; PLSCR1; USP18; TXN; UBE2L6; IFIH1; NMI; MTHFD2; MT2A; ISG20; RSAD2; IRF7; LTA4H; PSMB9; ISG15; OAS3; PSMA4; HERC5; HIST1H2BC; DHX58; MAP3K4; BAG1; ATOX1; LRP8; PSME1; RAB27A; ACTR5; IFI27; MX1; APOBEC3A; CENPE; IFIT3; PHF17; PML; IFITM1; DDX60; LY96; FANCL; INPP5D; NLRP1; BRCA1; CCNG1; NFKBIB; TSC22D3; PCBP2; GCH1; PRKCB; EIF2AK2; XAF1; GBP2; CHEK1; PRDX1; LGALS3BP; IFITM2; F5; HLA-B; BST2; CASP1; APOBEC3G; HLA-A; PSMA3; KIF4A; PSMB2; PSMD8; HNRNPA0; SERPING1; IFIT1; BRCA2; PSMD14; PSMB10; ANXA2; KIT; GPX4; TRAFD1; KCNN4; JAM3; BARD1; IFI16; ZBP1; BCL2L1; CD44; OAS2; PSME4; ATG16L1; MANF; PSMA5; DTL; DDX58; NFKBIA; CXCL10; HLA-C; ZBTB40; IFITM3; MCTS1; SEPP1; PKD2; PANX1; ADAR; FANCI; CD38; DGKD; FAS; SLC7A6; TRIM28; SESN1; STAT3; CREB3; CFL1; MX2; TBK1; RBX1; MYD88; MMS19; FADS1; SIGLEC1; CEBPG; NHEJ1; ANXA3; ATG3; ADRB2; SLC16A1; OAS1; SDC1; ERO1L; CALM3; TMEM204; VRK2; GMPR; APOL2; RASGRP2; ADD1; ECT2; MKI67; RACGAP1 |
| GO:0002376 | immune system process | 4.71E-11 | PSME2; TAP1; IFI35; GBP1; STAT1; PSMA6; OASL; IFI6; PLSCR1; USP18; PSMA4; ISG20; RSAD2; CRIP2; IRF7; PSMB9; ISG15; OAS3; PSMA4; HERC5; DHX58; IL15RA; PSME1; IFI44L; RAB27A; IFI27; MX1; APOBEC3A; IFIT3; PML; IFITM1; DDX60; LY96; IMPDH2; INPP5D; NLRP1; NFKBIB; PCBP2; GCH1; PRKCB; XAF1; GBP2; ITPKB; HLX; IFITM2; HLA-B; BST2; CASP1; APOBEC3G; HLA-A; CRIP1; PSMA3; PSMB2; PSMD8; SERPING1; IFIT1; BRCA2; PSMD14; PSMB10; KIT; TRAFD1; KCNN4; JAM3; IFI16; ZBP1; BCL2L1; CD44; OAS2; PSME4; ICAM3; PSMA5; DDX58; NFKBIA; CXCL10; HLA-C; PATZ1; IFITM3; CXCR5; PANX1; ADAR; CD38; FAS; SLC7A6; TRIM22; CREB3; MX2; TBK1; SEC24A; MYD88; TGFBR2; CEBPG; NHEJ1; ANXA3; SLC16A1; OAS1 |
| GO:0071345 | cellular response to cytokine stimulus | 5.58E-11 | IFI35; GBP1; STAT1; OASL; IFI6; IFI6; PLSCR1; USP18; UBE2L6; MT2A; ISG20; IRF7; ISG15; OAS3; HERC5; IL15RA; LRP8; IFI27; MX1; IFIT3; PML; IFITM1; XAF1; GBP2; IFITM2; HLA-B; HLA-A; IFIT1; KIT; CD44; OAS2; HLA-C; IFITM3; ADAR; NUP62; STAT3; MX2; MYD88; OAS1; VRK2 |
| GO:0044419 | interspecies interaction between organisms | 5.92E-11 | TAP1; RALA; STAT1; SLC25A6; IFIH1; RSAD2; IRF7; PSMB9; ISG15; PSMA4; GNB2L1; DHX58; LDLR; PML; TRIM5; EIF2AK2; HLA-B; APOBEC3G; HLA-A; PSMA3; PSMB2; PSMD8; IFIT1; PSMB10; BCL2L1; BUB1; DDX58; GADD45GIP1; NFKBIA; HLA-C; CFLAR; TRIM22; STAT3; CREB3; TBK1; RBX1; MYD88; TOP1; NXF1; VRK2 |
| GO:0034097 | response to cytokine stimulus | 6.85E-11 | IFI35; GBP1; STAT1; OASL; MT1X; IFI6; IFI6; USP18; UBE2L6; IFI6; PLSCR1; USP18; UBE2L6; MT2A; ISG20; IRF7; ISG15; OAS3; HERC5; IL15RA; LRP8; IFI27; MX1; IFIT3; PML; IFITM1; GCH1; XAF1; GBP2; IFITM2; HLA-B; HLA-A; IFIT1; KIT; TRAFD1; BCL2L1; CD44; OAS2; HLA-C; IFITM3; ADAR; CD38; NUP37; NUP62; STAT3; MX2; MYD88; OAS1; VRK2 |
| GO:0002479 | antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent | 5.79E-10 | PSME2; TAP1; PSMA6; PSMB9; PSMA4; PSME1; HLA-B; HLA-A; PSMA3; PSMB2; PSMD8; PSMB10; PSME4; PSMA5; HLA-C |
| GO:0051607 | defense response to virus | 1.05E-09 | PLSCR1; IFIH1; RSAD2; IRF7; OAS3; HERC5; DHX58; APOBEC3A; PML; DDX60; PCBP2; BST2; APOBEC3G; HLA-A; IFIT1; DDX58; CXCL10; CREB3; OAS1 |
| GO:0042590 | antigen processing and presentation of exogenous peptide antigen via MHC class I | 1.31E-09 | PSME2; TAP1; PSMA6; PSMB9; PSMA4; PSME1; HLA-B; HLA-A; PSMA3; PSMB2; PSMD8; PSMB10; PSME4; PSMA5; HLA-C |
| GO:0002478 | antigen processing and presentation of exogenous peptide antigen | 2.81E-09 | PSME2; TAP1; PSMA6; PSMB9; PSMA4; PSME1; HLA-B; HLA-A; PSMA3; PSMB2; PSMD8; PSMB10; PSME4; PSMA5; HLA-C |
| GO:0019884 | antigen processing and presentation of exogenous antigen | 4.04E-09 | PSME2; TAP1; PSMA6; PSMB9; PSMA4; PSME1; HLA-B; HLA-A; PSMA3; PSMB2; PSMD8; PSMB10; PSME4; PSMA5; HLA-C |

| GO ID | Term | p-value | Genes |
| --- | --- | --- | --- |
| GO:0002474 | antigen processing and presentation of peptide antigen via MHC class I | 5.73E-09 | PSME2; TAP1; PSMA6; PSMB9; PSMA4; PSME1; HLA-B; HLA-A; PSMA3; PSMB2; PSMD8; PSMD14; PSMB10; PSME4; PSMA5; HLA-C; SEC24A |
| GO:0000216 | M/G1 transition of mitotic cell cycle | 1.10E-08 | PSME2; PSMA6; PSMB9; PSMA4; PSME1; GMNN; MCM10; PSMA3; CDC45; PSMB2; PSMD8; PSMD14; PSMB10; PSME4; PSMA5 |
| GO:0071158 | positive regulation of cell cycle arrest | 1.31E-08 | PSME2; PSMA6; PSMB9; PSMA4; PSME1; PML; BRCA1; PSMA3; PSMB2; PSMD8; PSMD14; PSMB10; PSME4; PSMA5; PKD2 |
| GO:2000602 | regulation of interphase of mitotic cell cycle | 1.68E-08 | PSME2; PSMA6; PSMB9; PSMA4; PSME1; PML; FHL1; CCNG1; PSMA3; CDC45; PSMB2; PSMD8; BRCA2; PSMD14; PSMB10; BCL2L1; PSME4; PSMA5; PKD2; UBE2L3 |
| GO:0048002 | antigen processing and presentation of peptide antigen | 2.23E-08 | PSME2; TAP1; PSMA6; PSMB9; PSMA4; PSME1; HLA-B; HLA-A; PSMA3; PSMB2; PSMD8; PSMD14; PSMB10; PSME4; PSMA5; HLA-C; SEC24A |
| GO:0090068 | positive regulation of cell cycle process | 2.68E-08 | PSME2; PSMA6; PSMB9; PSMA4; PSME1; PML; ESPL1; BRCA1; PSMA3; PSMB2; PSMD8; PSMD14; PSMB10; PSME4; PSMA5; AURKA; PKD2; DLGAP5; ECT2; RACGAP1 |
| GO:0007346 | regulation of mitotic cell cycle | 3.93E-08 | PSME2; PSMA6; PSMB9; PSMA4; PSME1; PML; FHL1; ESPL1; CCNG1; CHEK1; TMEM8B; PSMA3; CDC45; PSMB2; PSMD8; BRCA2; PSMD14; PSMB10; BCL2L1; PSME4; BUB1; PSMA5; AURKA; TPX2; PKD2; DLGAP5; UBE2L3; NEK2; ZWINT |
| GO:0006977 | DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest | 6.67E-08 | PSME2; PSMA6; PSMB9; PSMA4; PSME1; PML; PSMA3; PSMB2; PSMD8; PSMD14; PSMB10; PSME4; PSMA5 |

across all of the datasets, it might be highly representative of biological pathways associated to SLE. Because SLE is a systemic disease that affects multiple organs, other biological and cellular abnormalities present in the disease are important to be investigated. Our analysis also revealed novel sets of genes and pathways not previously implicated with SLE. Some of these could have interesting functions that could potentially regulate the immune response during SLE.

The present study has some limitations that require consideration. First, heterogeneity and confounding factors may have distorted the analysis. Clinical samples could be heterogeneous with respect to clinical activity, severity, or gender. Second, there are differences in gene expression between blood cells such as B and T cells that were not taken into account here. We could not do subgroup analysis due to the limited data. However, having integrated samples from different tissues during our meta-analysis might have enabled us to detect genes that we would have otherwise missed in subgroup analysis.

In conclusion, the meta-analysis of gene expression profiles of PBMCs provided a global overview of differential gene expression in SLE: 483 genes were differentially expressed, of which 260 were up-regulated and 223 were down-regulated. Integrated pathway analysis of the DE genes indicated that these belong to the type I IFN-mediated and cytokine-mediated signaling pathways, immune response, antigen processing and presentation, regulation of cell cycle, and p53-mediated DNA-damage response. Our meta-analysis reveals previously unknown transcriptional changes in SLE. Further functional studies may provide additional insights into role of the differentially regulated genes in the pathophysiology of SLE.

### Acknowledgements

### Conflict of interest statement
The authors have no financial or non-financial conflict of interest to declare.

### References

1. Lee YH & Nath SK. Systemic lupus erythematosus susceptibility loci defined by genome scan meta-analysis. Hum Genet 2005; 118:434-43.

2. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999; 286:531-7.

3. Jones C, Simpson P, Mackay A & Lakhani SR. Expression profiling using cDNA microarrays. Methods Mol Med 2006; 120:403-14.

4. Ramasamy A, Mondry A, Holmes CC & Altman DG. Key issues in conducting a meta-analysis of gene expression microarray datasets. PLoS Med 2008; 5:e184.

5. Siddiqui AS, Delaney AD, Schnerch A, Griffith OL, Jones SJ & Marra MA. Sequence biases in large scale gene expression profiling data. Nucleic Acids Res 2006; 34:e83.

6. Cahan P, Rovegno F, Mooney D, Newman JC, St Laurent G, 3rd & McCaffrey TA. Meta-analysis of microarray results: challenges, opportunities, and recommendations for standardization. Gene

2007; 401:12-8.

7. Rung J & Brazma A. Reuse of public genome-wide gene expression data. Nat Rev Genet 2013; 14:89-99.

8. Griffith OL, Melck A, Jones SJ & Wiseman SM. Meta-analysis and meta-review of thyroid cancer gene expression profiling studies identifies important diagnostic biomarkers. J Clin Oncol 2006; 24:5043-51.

9. Lee YH. Meta-analysis of genetic association studies. Annals of laboratory medicine 2015; 35:283-7.

10. Lee YH, Lee HS, Choi SJ, Ji JD & Song GG. Associations between TLR polymorphisms and systemic lupus erythematosus: a systematic review and meta-analysis. Clin Exp Rheumatol 2012; 30:262-5.

11. Lee YH, Bae S-C, Choi SJ, Ji JD & Song GG. Associations between interleukin-10 polymorphisms and susceptibility to rheumatoid arthritis: a meta-analysis. Molecular biology reports 2012; 39:81-7.

12. Xia J, Fjell CD, Mayer ML, Pena OM, Wishart DS & Hancock RE. INMEX--a web-based tool for integrative meta-analysis of expression data. Nucleic Acids Res 2013; 41:W63-W70.

13. Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for functional genomics data sets--10 years on. Nucleic Acids Res 2011; 39:D1005-10.

14. Bolstad BM, Irizarry RA, Astrand M & Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003; 19:185-93.

15. Choi JK, Yu U, Kim S & Yoo OJ. Combining multiple microarray studies and modeling interstudy variation. Bioinformatics 2003; 19 Suppl 1:i84-90.

16. Egger M, Smith GD & Phillips AN. Meta-analysis: principles and procedures. BMJ 1997; 315:1533-7.

17. DerSimonian R & Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7:177-88.

18. Falcon S & Gentleman R. Using GOstats to test gene lists for GO term association. Bioinformatics 2007; 23:257-8.

19. Lee HM, Sugino H, Aoki C & Nishimoto N. Underexpression of mitochondrial-DNA encoded ATP synthesis-related genes and DNA repair genes in systemic lupus erythematosus. Arthritis Res Ther 2011; 13:R63.

20. Garaud JC, Schickel JN, Blaison G, Knapp AM, Dembele D, Ruer-Laventie J, et al. B cell signature during inactive systemic lupus is heterogeneous: toward a biological dissection of lupus. PLoS One 2011; 6:e23900.

21. Gromme M & Neefjes J. Antigen degradation or presentation by MHC class I molecules via classical and non-classical pathways. Mol Immunol 2002; 39:181-202.

22. Hale MB, Krutzik PO, Samra SS, Crane JM & Nolan GP. Stage dependent aberrant regulation of cytokine-STAT signaling in murine systemic lupus erythematosus. PLoS One 2009; 4:e6756.

23. Pelzel C, Begitt A, Wenta N & Vinkemeier U. Evidence against a role for beta-arrestin1 in STAT1 dephosphorylation and the inhibition of interferon-gamma signaling. Mol Cell 2013; 50:149-56.

24. Huang CM, Hang LW, Chen CL, Wu JY & Tsai FJ. Polymorphisms of TAP1 transporter genes in Chinese patients with systemic lupus erythematosus in Taiwan. Rheumatol Int 2004; 24:130-2.

25. Kanemitsu S, Ihara K, Saifddin A, Otsuka T, Takeuchi T, Nagayama J, et al. A functional polymorphism in fas (CD95/APO-1) gene promoter associated with systemic lupus erythematosus. J Rheumatol 2002; 29:1183-8.

26. Namjou B, Sestak AL, Armstrong DL, Zidovetzki R, Kelly JA, Jacob N, et al. High-density genotyping of STAT4 reveals multiple haplotypic associations with systemic lupus erythematosus in different racial groups. Arthritis Rheum 2009; 60:1085-95.

27. Bachmann M, Ulziibat S, Hardle L, Pfeilschifter J & Muhl H.

IFNalpha converts IL-22 into a cytokine efficiently activating STAT1 and its downstream targets. Biochem Pharmacol 2013; 85:396-403.

28. Arasappan D, Tong W, Mummaneni P, Fang H & Amur S. Meta-analysis of microarray data using a pathway-based approach identifies a 37-gene expression signature for systemic lupus erythematosus in human peripheral blood mononuclear cells. BMC Med 2011;

9:65.

29. Qing X & Putterman C. Gene expression profiling in the study of the pathogenesis of systemic lupus erythematosus. Autoimmun Rev 2004; 3:505-9.

30. Aringer M, Gunther C & Lee-Kirsch MA. Innate immune processes in lupus erythematosus. Clin Immunol 2013; 147:216-22.