PLOS | ONE

RESEARCH ARTICLE

Identification of a gene-expression predictor for diagnosis and personalized stratification of lupus patients

Yan Ding[1]☯, Hongai Li[2]☯, Xiaojie He[3], Wang Liao[4], Zhuwen Yi[3], Jia Yi[5], Zhibin Chen[6], Daniel J. Moore[7], Yajun Yi[8]*, Wei Xiang[9]*

1 Department of Dermatology, Hainan Provincial Dermatology Disease Hospital, Haikou, China, **2** Pediatrics, The Hainan Affiliated Hospital of University of South China, Haikou, China, **3** Department of Nephropathy, Children's Medical Center, The Second Xiangya Hospital, Central South University, Changsha, China, **4** Department of Cardiology, Hainan General Hospital, Haikou, China, **5** Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, NC, United States of America, **6** Department of Microbiology and Immunology, University of Miami Miller School of Medicine, Miami, FL, United States of America, **7** Departments of Pediatrics and Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, TN, United States of America, **8** Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, United States of America, **9** Department of Pediatrics, Maternal and Child Health Care Hospital of Hainan Province, Haikou, China

☯ These authors contributed equally to this work.
* 13907533529@163.com (WX); Yajun.Yi@vanderbilt.edu (YY)

Abstract

Systemic lupus erythematosus (SLE) is an autoimmune disease characterized by a wide spectrum of clinical manifestations and degrees of severity. Few genomic biomarkers for SLE have been validated and employed to inform clinical classifications and decisions. To discover and assess the gene-expression based SLE predictors in published studies, we performed a meta-analysis using our established signature database and a data similarity-driven strategy. From 13 training data sets on SLE gene-expression studies, we identified a SLE meta-signature (SLEmetaSig100) containing 100 concordant genes that are involved in DNA sensors and the IFN signaling pathway. We rigorously examined SLEmetaSig100 with both retrospective and prospective validation in two independent data sets. Using unsupervised clustering, we retrospectively elucidated that SLEmetaSig100 could classify clinical samples into two groups that correlated with SLE disease status and disease activities. More importantly, SLEmetaSig100 enabled personalized stratification demonstrating its ability to prospectively predict SLE disease at the individual patient level. To evaluate the performance of SLEmetaSig100 in predicting SLE, we predicted 1,171 testing samples to be either non-SLE or SLE with positive predictive value (97–99%), specificity (85%-84%), and sensitivity (60–84%). Our study suggests that SLEmetaSig100 has enhanced predictive value to facilitate current SLE clinical classification and provides personalized disease activity monitoring.

OPEN ACCESS

Citation: Ding Y, Li H, He X, Liao W, Yi Z, Yi J, et al. (2018) Identification of a gene-expression predictor for diagnosis and personalized stratification of lupus patients. PLoS ONE 13(7): e0198325. https://doi.org/10.1371/journal.pone.0198325

Editor: Frederic Rieux-Laucat, IMAGINE, FRANCE

Received: December 10, 2017

Accepted: May 17, 2018

Published: July 5, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was supported by the grants (81260139, 81060073, 81560275, 61562021 to WX, 30560161 to YD) from Natural Science Foundation of China; the grant (2018CXTD350 to YD) from Hainan Natural Science Foundation Innovation Research Team Project; the grant (2015SF39 to YD, ZDYF2018103 to WX) from Hainan special projects of Social Development; the grant (201515 to YD) from Hainan Association for academic excellence Youth Science and

Technology Innovation Program; and the grant (805106 to YD) from Hainan Natural Science Foundation of China. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors declare no competing interests exist.

Introduction

Systemic lupus erythematosus (SLE) is a chronic systemic autoimmune disease that predominantly affects young women. It is characterized by heterogeneous clinical manifestations with varying degrees of severity punctuated by alternating phases of remission and flares [1]. Patients classically produce high autoantibody titers that form immune complexes that are deposited in many different organs such as the skin, joints, and kidneys causing cumulative damage over time [2].

Given its wide variety of relapsing-remitting symptoms, lupus is typically difficult to diagnose. To complicate matters, there is no laboratory test that can definitively identify the illness. As a result, it can take years for a diagnosis to be made. Disease activity (DA) is also difficult to measure. One method is using any of six validated composite scores, one of which is the SLE Disease Activity Index (SLEDAI) [3]. Because SLE is heterogeneous, not all manifestations are included in the SLEDAI, making reliable patient assessment challenging. Therefore, there is a critical need for a biomarker to detect, monitor, and stratify individual patients with SLE.

The use of gene expression microarrays in clinical research has led to the establishment of biomarker signatures. SLE patients display unique blood transcriptional signatures linked to type I interferon (IFN) and granulocytes [4–7]. Preliminary work suggests that these signatures correlate with SLE disease status and DA [6]. Most studies have focused on IFN-induced transcripts or proteins as biomarkers [7–9].

The value of this approach to discover stable disease markers has been questioned [10]. One common weakness is that the traditional approach employs single study-based signatures that are derived from small sample-size and lack cross-validation making data interpretation and application difficult. This limitation underscores the need to assess larger cohorts, to use unbiased approaches that incorporate all elements of the signature, and to account for disease heterogeneity during data interpretation.

These limitations can be overcome by combining related but independent studies into a meta-analysis forming a larger sample size with lower false discovery rates. We have developed and utilized a robust meta-analysis approach called EXALT(EXpression AnaLysis Tool) for gene expression profile studies from thousands of Gene Expression Omnibus (GEO) and published breast cancer datasets [11–13]. A gene expression signature as defined by EXALT is a set of significant genes with their corresponding statistical scores and gene expression direction codes (up or down). We have previously used this approach to discover a novel and conserved gene expression signature predictive of metastasis risk in multiple cancers [13].

The present work describes the implementation of our unique EXALT approach for the meta-analysis of blood microarray transcriptional profiles on SLE. To this end, we aimed to identify a meta-signature (SLEmetaSig100) that correlates SLE status and DA from thirteen training data sets. We then validated the SLEmetaSig100 in two independent test data sets to determine its correlation with SLE and DA and prospective predictive value of SLE disease at the individual patient level. This enabled patient stratification based on a personalized transcriptional immunomonitoring signature correlating with DA in each patient.

Materials and method

Publicly available data sets and signatures

Subjects were recruited by the individual studies. Clinical diagnosis and demographic characteristics of anonymized SLE patients were confirmed and summarized in the previously published studies (S1 Table). EXALT is a database that holds original study descriptions, sample phenotypes, curated gene expression datasets, as well as thousands of gene expression signatures extracted from the GEO and other published studies. These meta-data and gene-expression profiles are encoded in a searchable format to form the basis of our data analysis [13]. With EXALT database, we are able to search data sets based on similar sample phenotypes and study design, subsequently identifying fifteen gene-expression microarray data sets on various SLE phenotypes. These were then further divided into thirteen training sets and two testing data sets. Two test data sets were selected for their large sample size and comprehensive clinical information.

Whole blood PBMC or T cell subset samples from the training data sets (n = 1,869) were grouped by their clinical attributes and study designs. Based on the existing sample group description in the published studies, each data set had at least two groups of samples including normal healthy controls and various SLE related phenotypes, and/or molecular markers such as a lupus flare, low or high disease grade activities, TLR and IFN gene expression levels (S1 Table). Two or more groups per dataset were needed to generate statistical comparisons. A total of 167 SLE gene signatures from all possible pairwise group comparisons were generated accordingly [11].

One of the signatures (PMID: 16777955) in the training set (S1 Table) was derived from a mouse model (Low vs Overexpression of TLR7). In order to define corresponding human TLR7 signature for cross-species meta-analysis, we used NCBI Gene and NCBI HomoloGene databases to translate mouse array probesets to human homolog gene symbols as we described before[14, 15].

Identification of meta-signature

We used EXALT in an iterative manner (iterative EXALT) [13] to conduct a data similarity-driven clustering analysis of the 167 SLE gene-expression signatures and to elucidate a common transcriptional signature in SLE studies. This iterative EXALT process started with all-versus-all signature similarity searches, resulting in signature clusters. More specifically, each of the 167 SLE signatures was searched against other 166 signatures to bring homologous signatures together by their intrinsic similarities. This process "grouped" or "clustered" the thirteen signatures together based on their similarities (i.e. gene names, expression directions, and confidence levels) to form the SLE signature cluster. We focused on this cluster because their phenotypes were clearly related to known SLE disease status or pathogenesis such as SLE flare activity, IFN production, or TLR7 expression.

In the cluster, each of the thirteen signatures comprised several hundred genes with various overlapping signature genes. In order to identify a recurrent and concordant gene expression pattern in the SLE signature cluster, all signature genes were assembled together to form a synthetic signature (SLEmetaSig). The top 100 genes (SLEmetaSig100), as determined by ranking their frequency of recurrence and gene expression profile concordance, were identified using the method previously described [13].

Prospective prediction of SLE status

We constructed a centroid-based reference signature associated with known SLE status and the SLEmetaSig100 signature values from the thirteen training datasets using the method described before [16].

Two GEO datasets (GSE65391 and GSE11909) were used as validation data sets to test the predictive ability of SLEmetaSig100. The 1,171 testing SLEmetaSig100 signatures were made from 92 healthy subjects and 1,079 SLE samples (211 patients) [6, 17].

By performing a Spearman's rank correlation between the reference signature and individual gene-expression profiles in test datasets, we were able to determine SLE status of individual patients from the two test data sets.

For each testing sample, a Spearman rank correlation value between the reference SLEmetaSig100 profile and the test sample profile was calculated. The sample is considered to be 'SLE' if the correlation value with the reference was equal to or above the predefined threshold value (0.3)[18]. The sample was considered healthy otherwise.

Statistics

An EXALT built-in statistical approach was used to assess signature similarity of training data sets. To evaluate of SLEmetaSig100 retrospective classification, we used *Mann-Whitney U test* to examine the difference in mean numbers SLE subjects between two SLEmetaSig100 classified groups. We used Fisher's Exact Test to compare prospective prediction rates of SLEmetaSig100 in two test data sets.

Hierarchical clustering and Spearman's rank correlation were performed and visualized using the TIGR MeV [19]. Unsupervised hierarchical clustering based on average linkage was conducted to group the patient samples. The group assignments were based on the first bifurcation of the clustering sample dendrogram [20]. The Spearman rank correlation was used to measure the correlations between the reference signature and individual testing profiles in the two test data sets.

Prospective SLE prediction in test data sets was compared with actual clinical diagnosis. The primary predictive endpoint was SLE diagnosis or SLE disease activity (DA) for the validation cohort. The predictive performance was assessed using the derived positive predictive value (PPV), the negative predictive value (NPV), sensitivity, and specificity.

Receiver operating characteristic (ROC) analysis was performed to determine the sensitivity and specificity of SLEmetaSig100 predictions and the area under the ROC curve. ROC analyses were performed using R, version 3.3.3.

Results

Identification of human SLE meta-signature

In the past two decades, a large number of gene-expression studies have been reported and deposited in public domain including GEO, PubMed, and EXALT signature database [11]. EXALT manages signatures that are derived from all possible comparisons of each data sets including all possible experimental and disease conditions.

To avoid the weaknesses of single study-derived signatures and to better utilize the available gene expression data from independent studies, we have developed a meta-analysis strategy called EXALT. EXALT is essentially a database containing thousands of gene expression signatures extracted from published studies that enables signature comparisons [13]. We have extracted signatures from over 1,500 data sets representing over 22,367 samples from various diseases and experimental conditions collected from NCBI GEO and PubMed [11]. Searching through signature similarity sample phenotypes, and design information, we identified fifteen data sets on SLE gene expression profiling and then partitioned them into thirteen training and two testing data sets with 1,869 and 1,171 samples, respectively (S1 Table). From the thirteen training data sets, we extracted 167 gene-expression signatures.

Some of these 167 gene signatures are biologically related to SLE disease status. There are inherent limitations for any individual profiling study such as small sample size relative to the large number of potential gene probes, limitations of technological platforms, sample variation, and bioinformatics or statistical method bias. To overcome these problems, we implemented a meta-analysis approach (iterative EXALT) that combines individual transcriptional profiling signatures to deduce a common transcriptional signature across studies (SLEmetaSig100).

This conserved profile (SLEmetaSig100) was derived from 1,869 patient samples from thirteen individual SLE studies (Fig 1, S1 Table). The expression directions (up or down-regulation) and the function of the genes enriched in SLEmetaSig100 are displayed in Fig 1 and summarized in S2 Table. These genes are mechanistically involved in the pathogenies of SLE or other autoimmune diseases. Thus, SLEmetaSig100 likely represents genes involved in SLE disease pathogenesis.

Pathways analysis of SLE signature genes

To determine pathways that SLEmetaSig100 may be involved in, we used KEGG Pathway database (http://www.genome.jp/kegg) and its analysis tool[21].

The pathway analysis results suggested that SLEmetaSig100's genes involved in numerous pathways such as Toll-like receptor signaling pathway, NF-kappa B signaling pathway, and cytokine-cytokine receptor interaction network.

We categorized these pathway genes into two major functional categories (DNA sensors and cytokine genes) and constructed an innate immune DNA-sensor model of SLEmetaSig100. A cartoon depicting genes and their relationship in this model is shown in Fig 2.

DNA sensor genes such as cyclic GMP-AMP Synthase (MB21D1 or cGAS), IFI16, and Toll-like receptors (TLRs) are upstream stimulators in Toll-like receptor signaling pathway (S1 Fig), NF-kappa B signaling pathway (S2 Fig), and cytokine-cytokine receptor interaction network (S3 Fig). These pathways work synergistically to mount an immune response to either a pathogen or cellular or tissue damage. The role of TLRs in the onset of autoimmune pathologies has been effectively addressed in murine models of SLE [22]. DNA sensors have been demonstrated to be essential for inducing inflammatory genes, *e.g.* IFN-β expression [23]. IFN gene over-expression patterns have been reported in SLE patients [24]. SLE can be distinguished by a remarkably homogeneous gene expression pattern with overexpression of granulopoiesis-related and interferon (IFN)-induced genes [5]. Thus, it is possible that changes in expression of DNA sensors, TLRs, NF-kappa B genes, and cytokines are indicators for SLE pathogenesis.

To further study the characteristics of SLEmetaSig100, we compared SLEmetaSig100 with some other known SLE signatures (Table 1). Unlike many other SLE signatures[5, 8, 9, 17], SLEmetaSig100 was identified based on much larger training sets from a meta-analysis of thirteen training data sets and two independent test data sets for validation.

When signature genes were analyzed by an overlapping analysis, we found that there were very few common genes between SLEmetaSig100 and others. For example, only one common gene (1%) was found between SLEmetaSig100 and the other two signatures (IFN signature[5] and IFNr signature[9]). There is no common gene between SLEmetaSig100 and other known signatures such as plasmablast signature[17] and IFN-induced genes (IFIGs) signature[8]. Because of discrepant patient populations and signature extraction methods (Table 1), the result suggests that the five previously reported SLE studies identified different blood transcriptional signatures. Despite this small amount of overlap in gene composition, the major functional component of these signature genes are all linked to interferon (IFN) and/or IFN-induced genes (IFIGs) (Table 1).

Stratification of SLE patient by the SLE 100-gene signature

In order to validate the correlation between SLEmetaSig100 and SLE disease, we surveyed gene expression profiles of SLEmetaSig100 from two independently published transcriptional profiling studies [GSE65391and GSE11909] performed on normal versus SLE disease states (Fig 3).

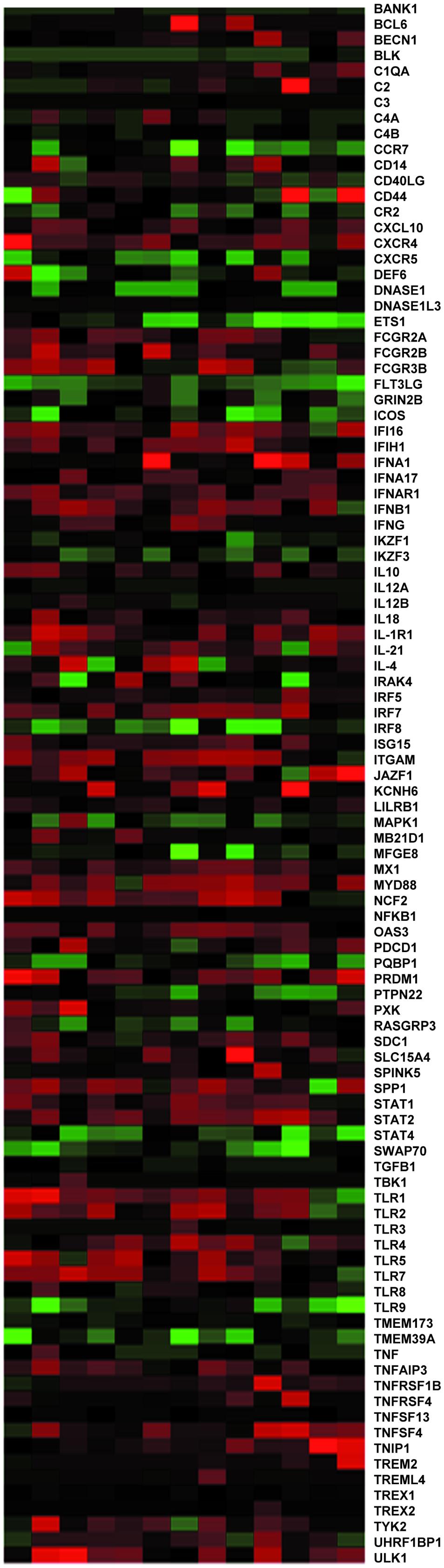

BANK1
BCL6
BECN1
BLK
C1QA
C2
C3
C4A
C4B
CCR7
CD14
CD40LG
CD44
CR2
CXCL10
CXCR4
CXCR5
DEF6
DNASE1
DNASE1L3
ETS1
FCGR2A
FCGR2B
FCGR3B
FLT3LG
GRIN2B
ICOS
IFI16
IFIH1
IFNA1
IFNA17
IFNAR1
IFNB1
IFNG
IKZF1
IKZF3
IL10
IL12A
IL12B
IL18
IL-1R1
IL-21
IL-4
IRAK4
IRF5
IRF7
IRF8
ISG15
ITGAM
JAZF1
KCNH6
LILRB1
MAPK1
MB21D1
MFGE8
MX1
MYD88
NCF2
NFKB1
OAS3
PDCD1
PQBP1
PRDM1
PTPN22
PXK
RASGRP3
SDC1
SLC15A4
SPINK5
SPP1
STAT1
STAT2
STAT4
SWAP70
TGFB1
TBK1
TLR1
TLR2
TLR3
TLR4
TLR5
TLR7
TLR8
TLR9
TMEM173
TMEM39A
TNF
TNFAIP3
TNFRSF1B
TNFRSF4
TNFSF13
TNFSF4
TNIP1
TREM2
TREML4
TREX1
TREX2
TYK2
UHRF1BP1
ULK1

Fig 1. Co-expression analysis of the 100 meta-signature genes from the SLE training data sets. Using EXALT meta-analysis, thirteen SLE signatures in columns with similar phenotypes indicated in S1 Table were displayed in a heat map with 100 genes (SLEmetaSig100) displayed in rows. The colors in the meta-heat map represent the direction of differential gene expression within a given transcriptional profile (red for up, green for down, and black for a missing match). Color intensity reflects the confidence levels of differential expression in the signatures.

https://doi.org/10.1371/journal.pone.0198325.g001

Both genes and samples were clustered by their gene-expression profiles. The sample classifications were determined based on the SLEmetaSig100 genes expression patterns (yellow bars in Fig 3A and 3B). Meta-heat maps were used to illustrate SLEmetaSig100 gene expression profiles and its capability to classify 996 (Fig 3A) and 175 human samples (Fig 3B).

Based on the meta-heat maps, we were able to classify human samples into left and right sample groups (yellow bar) in the dendrograms based on differential gene expression patterns. The associated sample phenotypes, either healthy (white bar) or SLE (black bars), were also grouped in each dataset. We observed that healthy samples were enriched in the left group of both testing data sets (Fig 3A and 3B). The results demonstrated that SLEmetaSig100 could stratify human samples into two groups: the normal healthy enriched sample groups (left) and the SLE dominant sample group (right).

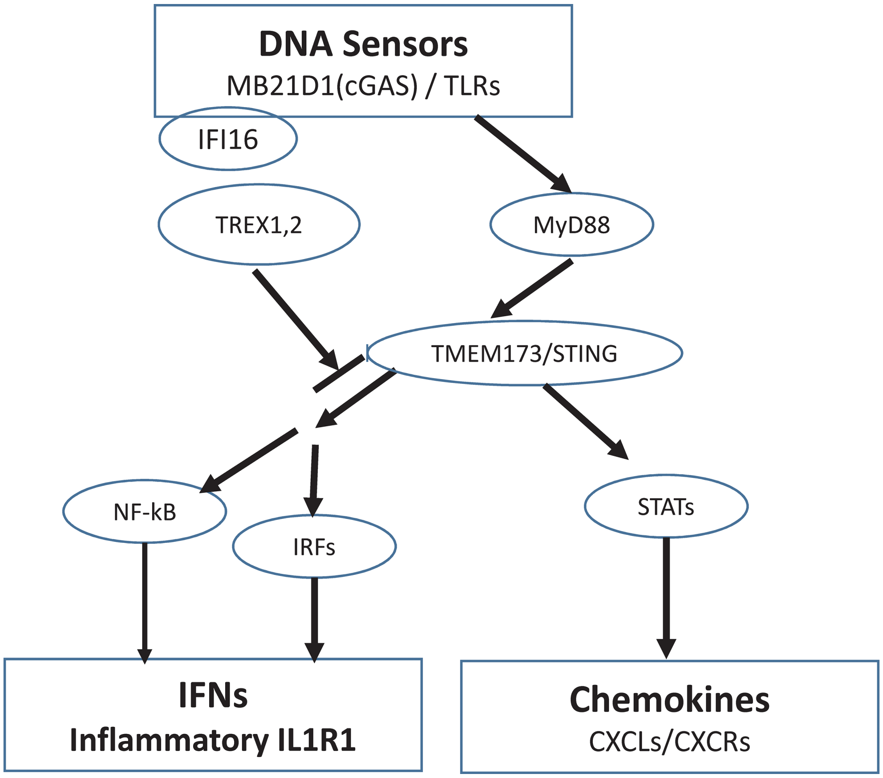


Fig 2. Pathway analysis of SLEmetaSig100 genes. Pathway analyses of SLEmetaSig100 genes identified genes involving DNA sensors and the cytokines constructing an innate immune DNA-sensor model. SLEmetaSig100 genes are marked in white circles or rectangles. DNA sensors include MB21D1(cGAS), multiple TLR genes, TMEM173/STING, and IF16 genes. In the Toll-like receptor signaling pathway, the stimulation of DNA sensor genes by microbe-derived and/or host DNA are positively regulated by MYD88 and TMEM173/STING genes and negatively regulated by TREX1 and TREX2 genes. The downstream cytokine-cytokine receptor interaction genes include NF-kappa B signaling pathway mediated IFNs, inflammatory cytokines (e.g. IL1R1), and STATs mediated chemokines (CXCL and CXCR genes).

https://doi.org/10.1371/journal.pone.0198325.g002

Table 1. SLE signature comparison.

Signature	SLEmetaSig100	Plasmablast	IFN	IFIGs	IFNr
Gene Number	100	9	10	3	3
Overlapping Genes	100	0	1	0	1
Training Set Size	1869	649	39	NA	NA
Test Set Size	1171	12	0	127	93
Method	meta-analysis	modular	Hochberg	scores	scores
SLE Association	DA	DA	DA	renal	No
Retrospective Stratification	Yes	Yes	Yes	Yes	No
Prospective Prediction	Yes	ND	ND	ND	ND

NOTE: ND, not done.

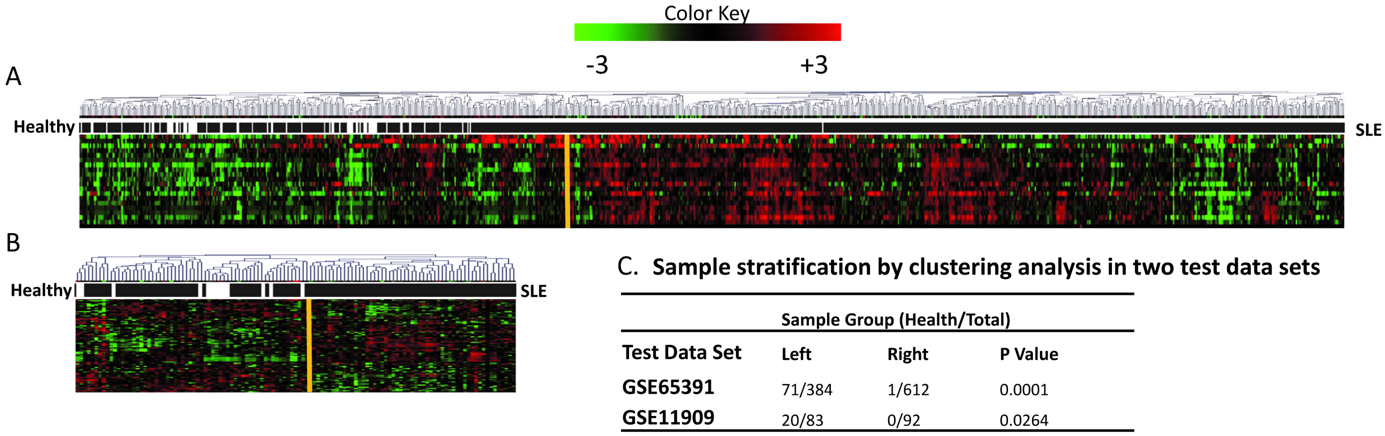
https://doi.org/10.1371/journal.pone.0198325.t001

More specifically, in the SLE dominant sample groups, there was only one (out of 72 total) normal sample in GSE65391(Fig 3A) and none (out of 20 total) in GSE11909 (Fig 3B), while remaining normal samples (71 in GSE6539 and 20 in GSE11909) were grouped in the normal sample groups. Thus, there were significantly fewer normal samples in SLE dominant sample groups than those in the normal sample groups of two test data sets (P = 0.0001 and 0.0264, Fig 3C). These results suggest that with known clinical SLE information SLEmetaSig100 is capable of stratifying testing samples into a normal health group and a SLE group.

Personalized SLE prediction in individual patients using SLEmetaSig100

SLE is a heterogeneous disease that cannot be diagnosed by a single symptom or lab test. Personalized prediction of SLE status by comparing a test sample profile to a reference SLEmetaSig100 signature may provide a new method to facilitate clinical diagnosis.

We leveraged SLEmetaSig100 to determine whether it can distinguish individual SLE patients from normal healthy subjects and to uncover associated disease activity (SLEDAI) when clinical SLE information is masked in test data sets.



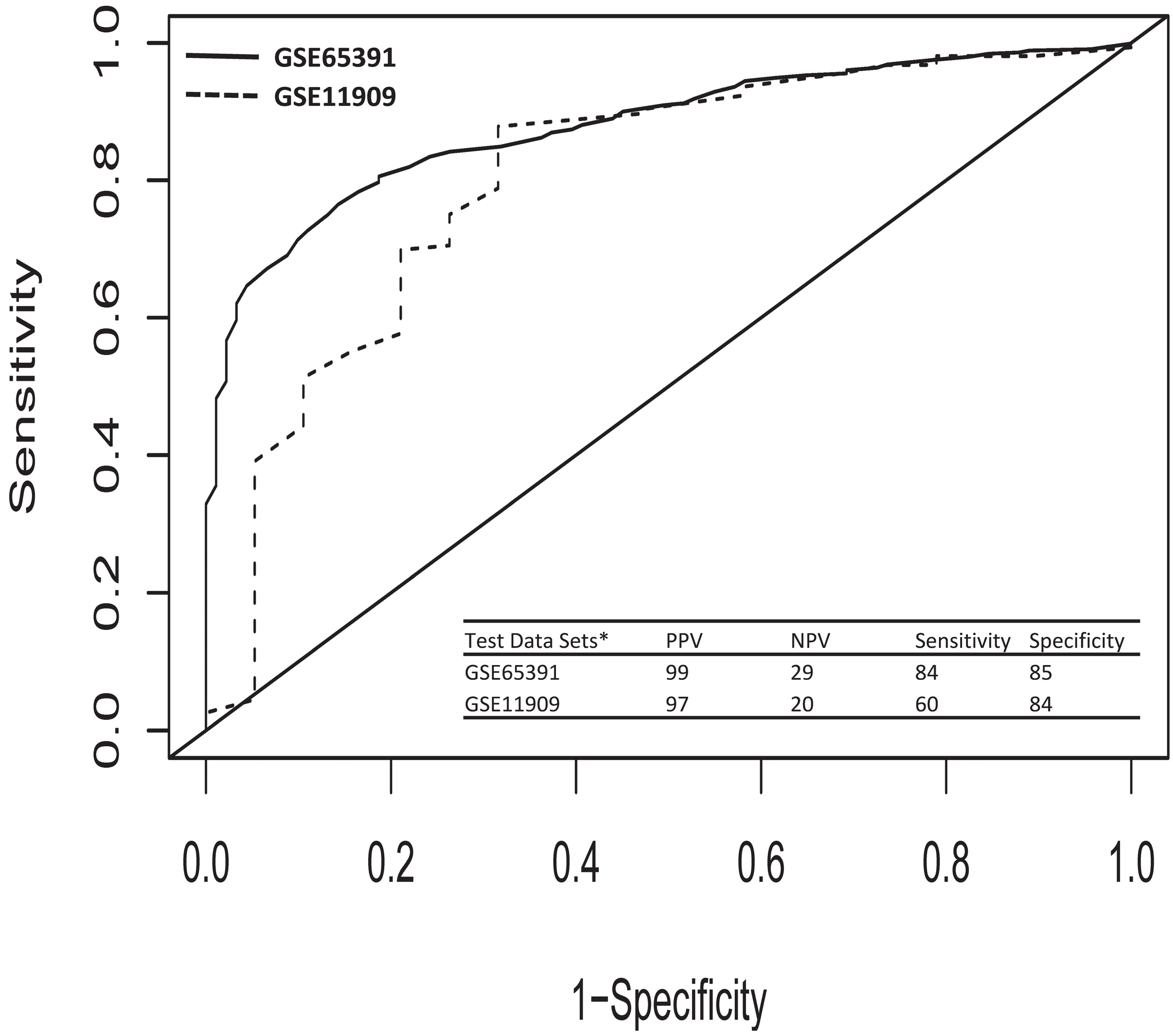
C. **Sample stratification by clustering analysis in two test data sets**

	Sample Group (Health/Total)		
Test Data Set	**Left**	**Right**	**P Value**
GSE65391	71/384	1/612	0.0001
GSE11909	20/83	0/92	0.0264

Fig 3. Stratification of human samples by clustering the SLEmetaSig100 meta-profiles. Meta-heat maps from unsupervised hierarchical clustering depict meta-profiles in two test data sets, (A) GSE65391 and (B) GSE11909. In each panel, the gene expression patterns from one given test data set are represented in rows and samples are clustered in columns. The colors in the heat map represent the direction of differential gene expression within a given transcriptional profile (red for up, green for down, and black for a missing match). Color intensity reflects the confidence levels of differential expression. Sample groups in columns are determined by the top hierarchy nodes of dendrograms (yellow bar) into left and right sample groups. The sample phenotype patterns underneath each sample dendrogram panel are indicated by black (SLE) and white (healthy) bars. The classification of healthy samples from total samples (healthy/total) by SLEmetaSig100 profiles between two sample groups (left and right) and statistic tests (Mann-Whitney U test) results are listed in a table (C).

https://doi.org/10.1371/journal.pone.0198325.g003

To evaluate the ability of SLEmetaSig100 to predict SLE, we divided 1,171 testing samples to be either predicted healthy or predicted SLE. Those results were then compared to their actual clinic diagnoses. The derived positive predictive value (PPV), the negative predictive value (NPV), sensitivity, and specificity were then calculated and compared accordingly. While assessing the SLEmetaSig100 prospective prediction results with the actual clinic diagnoses, we found that SLEmetaSig100 could significantly correctly predict SLE in two independent cohorts (sub-Table in Fig 4, P = 1.48E-36). The prospective predictions using the SLEmetaSig100 centroid model showed comparable results to those obtained using the unsupervised clustering-based retrospective prediction at group level (Fig 3). However, the centroid model can further prospectively apply to individual patients with high PPV (97%-99%), specificity (85%-84%), and sensitivity (60–84%) (sub-Table in Fig 4).



Test Data Sets*	PPV	NPV	Sensitivity	Specificity
GSE65391	99	29	84	85
GSE11909	97	20	60	84

Fig 4. Receiver operating characteristic (ROC) curves for SLEmetaSig100. Area under receiver operating characteristic curve (AUC) for performance of SLEmetaSig100 were calculated in two testing cohorts, GSE65391(solid line) and GSE11909(dash line), and SLEmetaSig100 significantly outperforms the random prediction of SLE disease (AUC, 0.89 in GSE65391 and 0.85 in GSE11909). The sub-table shows SLEmetaSig100 prediction performance in two test datasets. *Note: SLE prediction by SLEmetaSig100 in two test data sets was examined by Fisher Exact test (P value = 1.48E-36).

https://doi.org/10.1371/journal.pone.0198325.g004

SLEmetaSig100's ability to discriminate between normal healthy subjects and SLE patients was also evaluated by ROC analysis in two test data sets (GSE65391 and GSE11909). The ROC results are comparable displaying areas under the curve (AUC) of 0.89 (GSE65391) and 0.85 (GSE11909), respectively, indicating that SLEmetaSig100 has reliable performance in heterogeneous patient populations. When SLEmetaSig100 prediction was implemented with centroid model, the GSE65391 test resulted in a specificity of 85% and a sensitivity of 84%. The GSE11909 test displayed a specificity of 84% and a sensitivity of 60% (sub-Table in Fig 4).

To determine the association between SLE DA and SLEmetaSig100 prediction, we used the SLE samples from the training set with a known SLE disease activity index (SLEDAI). These samples were classified into three DA group scores: S1 (SLEDAI: 0–2), S2 (SLEDAI:3–7), or S3 (SLEDAI > 7). From this, we constructed three DA reference signatures (DA1, DA2, and DA3). Using the same SLE prediction process, we computed three correlation scores between each test sample and three reference DA signatures. The predicted SLE DA group for each sample was determined by the highest correlation score of the three correlations calculated. We then compared the average actual DA group scores (1.49, 2.01, 2.21 in Table 2) with each predicted GA group (S1, S2, and S3, respectively). We found that there were significant differences among these three average DA group scores (S1 vs S2, S1 vs S3, and S2 vs S3). Results suggested that metaSLEsig100 predicted DA groups correlated with the actual clinical SLEDAI group scores. Those groups with higher predicted SLE DA showed statistically significantly higher actual SLE DA group scores than those predicted to be lower SLE DA group ($P \leq 0.002$). Thus, the SLEmetaSig100 centroid model can serve as a SLE biomarker that can improve SLE patient diagnosis and disease activity prediction.

Discussion

SLE is a heterogeneous disease characterized by a wide spectrum of clinical manifestations and degrees of severity. A single gene-expression profiling study on SLE cannot capture the full molecular heterogeneity of SLE. Few studies on SLE gene expression signatures have progressed beyond the discovery phase because the retrospective validation of these signatures often focused on the association of DA at the group level within the same study [5, 8, 17].

We analyzed fifteen source datasets in this study from microarray platforms. In a typical transcriptional study of a disease model, a differential gene list is usually generated from a training set and is then validated in a test set. With both training and test sets from the same patient cohorts, a microarray study is often underpowered. Other common negative factors of microarray platform such as low data quality and high background noise interference can be introduced at different experimental and analysis stages.

In response to these challenges, we developed the EXALT method by combining related but independent studies into a meta-analysis to create a larger sample size and to produce a lower false discovery rate. EXALT is essentially a database containing thousands of gene expression signatures extracted from public gene expression database (e.g. GEO) and published studies that enable signature comparisons. We previously used this robust meta-analysis of gene

Table 2. Association between SLE group prediction and actual SLE disease activities.

Predicated GSE65391 Group	S1	S2	S3
SLEDAI	0–2	3–7	>7
Group DA scores	1	2	3
Actual Average Group DA Score	1.49	2.01	2.21
SD	0.67	0.63	0.75
P values	≤0.002	≤0.002	≤0.002

https://doi.org/10.1371/journal.pone.0198325.t002 expression profiles from hundreds of breast cancer datasets [11–13]. By using this approach, we discovered a novel and conserved gene expression signature that predicts metastasis risk in multiple cancers (breast, lung, and prostate cancer) [13]. Furthermore, we identified a new cancer metastasis-suppressor gene [25].

In this study, we applied our established high-throughput *in silico* method (EXALT) to identify and validate the meta-signature (SLEmetaSig100) based on fifteen available published studies. We demonstrated that heterogeneous signatures from thirteen training datasets containing 1,869 samples from heterogeneous whole blood PBMC or T cells could be systematically organized by their common data elements (i.e., intrinsic similarities and disease phenotypes) and assembled into a new signature data type called a meta-signature (S1 Table and Fig 1). We identified a meta-signature representing a common SLE gene expression profile consisting of 100 genes (SLEmetaSig100) that is capable of identifying SLE in 1,171 testing human samples from two independent datasets (Figs 3 and 4 and Table 2). These findings illustrate that there is a fundamental gene expression pattern conserved across immune cell types in SLE patients.

Besides SLEmetaSig100, several studies have shown that gene-expression signatures related to SLE DA [5, 8, 17]. We compared SLEmetaSig100 with a few representative signatures on SLE (Table 1). The main difference is that SLEmetaSig100 was identified based on much larger training sets from a unique meta-analysis (EXALT) and could stratify SLE at the individual patient level. The meta-signature represents heterogeneous cell populations that might result in weak overlap with any known SLE signatures (Table 1). Although the overlap in gene composition between SLEmetaSig100 and other published SLE signatures is small, they all contain a major functional component of the signature genes related to interferon (IFN) and/or IFN-induced genes that are included in the network of cytokine-cytokine receptor interaction. The gene functional groups and pathways of the meta-signature (SLEmetaSig100) are commonly known for their roles as seen in other SLE signatures or SLE studies (S2 Table).

The network of cytokine-cytokine receptor Interaction contains 18 SLEmetaSig100 genes (CCR7, CXCL10, CXCR4, CXCR5, CD40LG, TNFRSF1B, TNFRSF4, FLT3LG, IFNA1, IFNA17, IFNAR1, IFNB1, IL10, IL12A, IL12B, IL18, TNFSF13, TNFSF4) (S3 Fig, S2 Table). These cytokine genes often play direct effective roles in SLE pathogenesis through the regulation of systemic inflammation, local tissue damage, and immune modulation[26]. We confirmed a set of prevalent IFN-regulated transcripts in SLEmetaSig100 that are highly correlated with inflammation and IFN signaling pathways such as IFI16 IFIH1, IFNA1, IFNA17, IFNAR1, IFNB1, IFNG, IL10, IL12A, IL12B, IL18, IL-1R1, and IL-21. The dysregulation of IFN family genes (IFNA1, IFNA17, IFNAR1, IFNB1) is dominantly pervasive, and their protein and gene expression profiles may serve as markers of disease activity and severity [26–29].

Besides the overexpressed IFN-inducible genes, SLEmetaSig100 also contains DNA sensor genes as described in the innate immunity DNA-sensing model (Fig 2) such as cyclic GMP-AMP Synthase (MB21D1 or cGAS), IFI16, and Toll-like receptors (TLR)s that are required for cell proliferation and for mounting an appropriate immune response to either a pathogen or cellular/tissue damage[22]. TLR3, TLR5 and TLR7/8/9 have been reported as facilitating SLE pathogenesis [30] (S1 Fig, S2 Table). These different TLRs provide distinct or synergistic contributions. For example, the expression levels of TLR2 and TLR4 mRNAs in SLE patients' PBMCs are much higher than those in healthy subjects [31], and the expression of TLR3 mRNA increases with the progression of lupus nephritis [32, 33] while downregulation of TLR2 or TLR4 can decrease the production of autoantibodies and attenuates the development of renal injury in lpr mutation-induced murine lupus[34].

Other DNA sensor genes in the SLEmetaSig100 signature are key enzymes involved in breakdown of DNA including nucleases such as DNASE1, DNASE1lL3, TREX1, and TREX2. Importantly, a loss-of-function variant of DNASE1L3 causes a familial form of SLE. Mutations in TREX1 are associated with familial chilblain lupus and are also associated with the inflammatory disorder Aicardi-Goutieres syndrome. The SLEmetasig100 emphasizes the importance of including DNA processing pathways, which may capture the contributions of proteostasis and ER stress to SLE pathogenesis.

Lupus nephritis is a frequently seen complication in patients with SLE and is known to significantly reduce the survival of SLE patients. A hallmark of lupus nephritis is the renal inflammation caused by deposition of autoimmune complexes to kidney glomeruli. There are four SLEmetaSig100 genes (NFAIP3, IRAK4, MYD88, TLR4) in NF-kappa B signaling pathway (S2 Fig, S2 Table) that have been implicated in the pathogenesis of lupus nephritis[35] coupled with upregulation of inflammatory cytokines [36, 37].

Previous SLE signatures have been essentially equivalent in correlation to SLE DA except for the INFγ signature (Table 1). However, SLEmetaSig100 not only correlates with SLE DA but also provides a prospective prediction method that can improve SLE patient diagnosis, a capability that has not been demonstrated in any other SLE signature (Table 1, Table 2). Our result suggests that SLEmetaSig100 is capable of prospectively applying to individual patients with high PPV (97%-99%), specificity (85%-84%), and sensitivity (60–84%) (Fig 4). In the meantime, we also observed that SLEmetaSig100 has a low NPV rate (20%, Fig 4) which is the percentage of patients with a true negative test result who do not have the disease (SLE), suggesting that SLEmetaSig100 may not be suitable to predict healthy status.

Our approach may provide a new SLE biomarker for clinical diagnosis, classification and monitoring. Previous signatures have correlated with DA at the cohort level (retrospective stratification), such as IFN or plasmablast signatures (Table 1); however, they did not demonstrate the capability to predict SLE status and DA association in individual SLE patients.

A limitation of the current study is that we only tested SLEmetSig100 in samples of healthy and SLE subjects. Heterogeneous cell types with limited clinical attributes and follow-up information could hamper the training process and interpretation of our meta-analyses. There is no test data set available showing the correlation between the SLEmetaSig100 profile and a manifestation of SLE or another autoimmune disease. Treatment status was not accessible in the meta-data of the training and testing sets. Therefore, it may be possible that the SLE patients with signatures that were most similar to healthy controls were responding well to therapy. The capacity for SLEmetasig100 to predict early treatment response will be an important future application as well as its ability to distinguish SLE from other autoimmune disorders that may overlap in clinical presentation.

Future studies on SLEmetaSig100 with consecutive blood sampling from the same patient would allow us to better measure SLEmetSig100's performance by tracking disease activity and response to treatment over time. More importantly, a more controlled training set would allow us to improve our meta-signature's predictive ability to distinguish SLE profiles from those of other autoimmune disorders.

In summary, our finding supports the potential application of SLEmetSig100 as a promising biomarkers in clinical practice with an acceptable specificity and sensitivity. Biomarkers that can prospectively predict occurrence and frequency of flares will be of great clinical value in clinical practice [38]. The data mining nature of this study provides a foundation to further identify and validate more flare predictors. Additionally, the SLEmetaSig100 may also inform future study design to identify novel genes in SLE pathogenesis, classifiers, and early predictors of DA scores.

Supporting information

S1 Fig. Toll-like receptor signaling pathway. SLEmetaSig100 includes eight Toll-like receptors (TLR) genes (TLR1, TLR2, TLR3, TLR4, TLR5, TLR7, TLR8, and TLR9). Most of TLRs

are up-regulated (TLR1, TLR2, TLR4, TLR5, TLR7, and TLR8) while two TLRs (TLR3 and TLR9) are down-regulated in SLE disease conditions.
(TIF)

S2 Fig. NF-kappa B signaling pathway. NF-kappa B signaling pathway. There are four up-regulated SLEmetaSig100 genes (NFAIP3, NFIRAK4, MYD88, TLR4) and one down-regulated gene (NFKB1) in the NF-kappa B signaling pathway that are also present in the TLR signaling pathway as expected. As a negative regulator protein, NFKB1 is controlled by various mechanisms of post-translational modification and subcellular compartmentalization as well as by interactions with other cofactors or co-repressors.
(TIF)

S3 Fig. Cytokine-cytokine receptor interaction. The network of cytokine-cytokine receptor Interaction contains 18 SLEmetaSig100 genes (CCR7, CXCL10, CXCR4, CXCR5, CD40LG, TNFRSF1B, TNFRSF4, FLT3LG, IFNA1, IFNA17, IFNAR1, IFNB1, IL10, IL12A, IL12B, IL18, TNFSF13, TNFSF4). Most cytokine genes are up-regulated (S1 Table) like IFN, IFN responsive genes, or chemokines except four down-regulated genes (CCR7, CXCR5, FLT3LG, and IL12A).
(TIF)

S1 Table. Overview of SLE data sets on gene-expression profiles.
(XLSX)

S2 Table. Gene annotation for the SLE signature.
(XLSX)

Acknowledgments

We thank PengCheng Lu for help in statistical analysis and the many colleagues who contributed to data collection and data analysis of clinical samples.

Author Contributions

Conceptualization: Wang Liao, Jia Yi, Zhibin Chen, Daniel J. Moore, Yajun Yi, Wei Xiang.

Data curation: Yan Ding, Hongai Li, Xiaojie He, Wang Liao, Zhuwen Yi, Daniel J. Moore, Yajun Yi.

Formal analysis: Yan Ding, Hongai Li, Yajun Yi.

Funding acquisition: Yan Ding, Wei Xiang.

Investigation: Yan Ding, Wang Liao, Jia Yi, Yajun Yi, Wei Xiang.

Methodology: Yan Ding, Hongai Li.

Project administration: Yan Ding, Yajun Yi, Wei Xiang.

Resources: Yan Ding, Hongai Li, Xiaojie He, Wang Liao, Zhuwen Yi, Zhibin Chen, Yajun Yi, Wei Xiang.

Software: Yajun Yi.

Supervision: Yan Ding, Yajun Yi, Wei Xiang.

Validation: Yan Ding, Hongai Li, Xiaojie He, Wang Liao, Zhuwen Yi, Jia Yi, Daniel J. Moore, Yajun Yi.

Visualization: Hongai Li, Xiaojie He, Yajun Yi.

Writing – original draft: Yan Ding, Hongai Li, Daniel J. Moore, Yajun Yi, Wei Xiang.

Writing – review & editing: Jia Yi, Zhibin Chen, Daniel J. Moore, Yajun Yi, Wei Xiang.

References

1. Agmon-Levin N, Mosca M, Petri M, Shoenfeld Y. Systemic lupus erythematosus one disease or many? Autoimmunity reviews. 2012; 11(8):593–5. Epub 2011/11/02. https://doi.org/10.1016/j.autrev.2011.10.020 PMID: 22041578.
2. Tsokos GC. Systemic lupus erythematosus. The New England journal of medicine. 2011; 365 (22):2110–21. Epub 2011/12/02. https://doi.org/10.1056/NEJMra1100359 PMID: 22129255.
3. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. Arthritis and rheumatism. 1992; 35(6):630–40. Epub 1992/06/01. PMID: 1599520.
4. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100(5):2610–5. Epub 2003/02/27. https://doi.org/10.1073/pnas.0337679100 PMID: 12604793; PubMed Central PMCID: PMCPMC151388.
5. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. The Journal of experimental medicine. 2003; 197 (6):711–23. Epub 2003/03/19. https://doi.org/10.1084/jem.20021553 PMID: 12642603; PubMed Central PMCID: PMCPMC2193846.
6. Chaussabel D, Quinn C, Shen J, Patel P, Glaser C, Baldwin N, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. Immunity. 2008; 29(1):150–64. Epub 2008/07/18. https://doi.org/10.1016/j.immuni.2008.05.012 PMID: 18631455; PubMed Central PMCID: PMCPMC2727981.
7. Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. Arthritis & rheumatology (Hoboken, NJ). 2014; 66(6):1583–95. Epub 2014/03/20. https://doi.org/10.1002/art.38628 PMID: 24644022; PubMed Central PMCID: PMCPMC4157826.
8. Kirou KA, Lee C, George S, Louca K, Peterson MG, Crow MK. Activation of the interferon-alpha pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. Arthritis and rheumatism. 2005; 52(5):1491–503. Epub 2005/05/10. https://doi.org/10.1002/art.21031 PMID: 15880830.
9. Petri M, Singh S, Tesfasyone H, Dedrick R, Fry K, Lal P, et al. Longitudinal expression of type I interferon responsive genes in systemic lupus erythematosus. Lupus. 2009; 18(11):980–9. Epub 2009/09/19. https://doi.org/10.1177/0961203309105529 PMID: 19762399; PubMed Central PMCID: PMCPMC4752166.
10. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. Lancet (London, England). 2005; 365(9458):488–92. Epub 2005/02/12. https://doi.org/10.1016/s0140-6736(05)17866-0 PMID: 15705458.
11. Yi Y, Li C, Miller C, George AL Jr. Strategy for encoding and comparison of gene expression signatures. Genome biology. 2007; 8(7):R133. Epub 2007/07/07. https://doi.org/10.1186/gb-2007-8-7-r133 PMID: 17612401; PubMed Central PMCID: PMCPMC2323223.
12. Wu J, Qiu Q, Xie L, Fullerton J, Yu J, Shyr Y, et al. Web-based interrogation of gene expression signatures using EXALT. BMC bioinformatics. 2009; 10:420. Epub 2009/12/17. https://doi.org/10.1186/1471-2105-10-420 PMID: 20003458; PubMed Central PMCID: PMCPMC2799423.
13. Qiu Q, Lu P, Xiang Y, Shyr Y, Chen X, Lehmann BD, et al. A data similarity-based strategy for meta-analysis of transcriptional profiles in cancer. PloS one. 2013; 8(1):e54979. Epub 2013/02/06. https://doi.org/10.1371/journal.pone.0054979 PMID: 23383020; PubMed Central PMCID: PMCPMC3558433.
14. Jin R, Yi Y, Yull FE, Blackwell TS, Clark PE, Koyama T, et al. NF-kappaB gene signature predicts prostate cancer progression. Cancer research. 2014; 74(10):2763–72. https://doi.org/10.1158/0008-5472.CAN-13-2543 PMID: 24686169; PubMed Central PMCID: PMCPMC4024337.
15. Yi Y, Polosukhina D, Love HD, Hembd A, Pickup M, Moses HL, et al. A Murine Model of K-RAS and beta-Catenin Induced Renal Tumors Expresses High Levels of E2F1 and Resembles Human Wilms Tumor. J Urol. 2015; 194(6):1762–70. https://doi.org/10.1016/j.juro.2015.04.090 PMID: 25934441; PubMed Central PMCID: PMCPMC4782590.

16. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2009; 27(8):1160–7. Epub 2009/02/11. https://doi.org/10.1200/jco.2008.18.1370 PMID: 19204204; PubMed Central PMCID: PMCPMC2667820.

17. Banchereau R, Hong S, Cantarel B, Baldwin N, Baisch J, Edens M, et al. Personalized Immunomonitoring Uncovers Molecular Networks that Stratify Lupus Patients. Cell. 2016; 165(3):551–65. Epub 2016/04/05. https://doi.org/10.1016/j.cell.2016.03.008 PMID: 27040498.

18. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002; 415(6871):530–6. https://doi.org/10.1038/415530a PMID: 11823860.

19. Margolin AA, Greshock J, Naylor TL, Mosse Y, Maris JM, Bignell G, et al. CGHAnalyzer: a stand-alone software package for cancer genome analysis using array-based DNA copy number data. Bioinformatics. 2005; 21(15):3308–11. https://doi.org/10.1093/bioinformatics/bti500 PMID: 15905276.

20. Lukes L, Crawford NP, Walker R, Hunter KW. The origins of breast cancer prognostic gene expression profiles. Cancer research. 2009; 69(1):310–8. Epub 2009/01/02. https://doi.org/10.1158/0008-5472.CAN-08-3520 PMID: 19118016; PubMed Central PMCID: PMCPMC2613551.

21. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009; 4(1):44–57. https://doi.org/10.1038/nprot.2008.211 PMID: 19131956.

22. Deane JA, Pisitkun P, Barrett RS, Feigenbaum L, Town T, Ward JM, et al. Control of toll-like receptor 7 expression is essential to restrict autoimmunity and dendritic cell proliferation. Immunity. 2007; 27 (5):801–10. Epub 2007/11/13. https://doi.org/10.1016/j.immuni.2007.09.009 PMID: 17997333; PubMed Central PMCID: PMCPMC2706502.

23. Paludan SR, Bowie AG. Immune sensing of DNA. Immunity. 2013; 38(5):870–80. Epub 2013/05/28. https://doi.org/10.1016/j.immuni.2013.05.004 PMID: 23706668; PubMed Central PMCID: PMCPMC3683625.

24. Koenig KF, Groeschl I, Pesickova SS, Tesar V, Eisenberger U, Trendelenburg M. Serum cytokine profile in patients with active lupus nephritis. Cytokine. 2012; 60(2):410–6. Epub 2012/08/01. https://doi.org/10.1016/j.cyto.2012.07.004 PMID: 22846145.

25. Xiang Y, Qiu Q, Jiang M, Jin R, Lehmann BD, Strand DW, et al. SPARCL1 suppresses metastasis in prostate cancer. Molecular oncology. 2013; 7(6):1019–30. Epub 2013/08/07. https://doi.org/10.1016/j.molonc.2013.07.008 PMID: 23916135; PubMed Central PMCID: PMCPMC3838491.

26. Jacob N, Stohl W. Cytokine disturbances in systemic lupus erythematosus. Arthritis Res Ther. 2011; 13 (4):228. https://doi.org/10.1186/ar3349 PMID: 21745419; PubMed Central PMCID: PMCPMC3239336.

27. Ronnblom LE, Alm GV, Oberg KE. Possible induction of systemic lupus erythematosus by interferon-alpha treatment in a patient with a malignant carcinoid tumour. J Intern Med. 1990; 227(3):207–10. PMID: 1690258.

28. Niewold TB. Interferon alpha-induced lupus: proof of principle. J Clin Rheumatol. 2008; 14(3):131–2. https://doi.org/10.1097/RHU.0b013e318177627d PMID: 18525429; PubMed Central PMCID: PMCPMC2743115.

29. Sigurdsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G, et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet. 2005; 76(3):528–37. https://doi.org/10.1086/428480 PMID: 15657875; PubMed Central PMCID: PMCPMC1196404.

30. Wu YW, Tang W, Zuo JP. Toll-like receptors: potential targets for lupus treatment. Acta Pharmacol Sin. 2015; 36(12):1395–407. https://doi.org/10.1038/aps.2015.91 PMID: 26592511; PubMed Central PMCID: PMCPMC4816237.

31. Komatsuda A, Wakui H, Iwamoto K, Ozawa M, Togashi M, Masai R, et al. Up-regulated expression of Toll-like receptors mRNAs in peripheral blood mononuclear cells from patients with systemic lupus erythematosus. Clin Exp Immunol. 2008; 152(3):482–7. https://doi.org/10.1111/j.1365-2249.2008.03646.x PMID: 18373699; PubMed Central PMCID: PMCPMC2453201.

32. Patole PS, Pawar RD, Lech M, Zecher D, Schmidt H, Segerer S, et al. Expression and regulation of Toll-like receptors in lupus-like immune complex glomerulonephritis of MRL-Fas(lpr) mice. Nephrol Dial Transplant. 2006; 21(11):3062–73. https://doi.org/10.1093/ndt/gfl336 PMID: 16954173.

33. Patole PS, Grone HJ, Segerer S, Ciubar R, Belemezova E, Henger A, et al. Viral double-stranded RNA aggravates lupus nephritis through Toll-like receptor 3 on glomerular mesangial cells and antigen-presenting cells. J Am Soc Nephrol. 2005; 16(5):1326–38. https://doi.org/10.1681/ASN.2004100820 PMID: 15772251.

34. Lartigue A, Colliou N, Calbo S, Francois A, Jacquot S, Arnoult C, et al. Critical role of TLR2 and TLR4 in autoantibody production and glomerulonephritis in lpr mutation-induced mouse lupus. J Immunol. 2009; 183(10):6207–16. https://doi.org/10.4049/jimmunol.0803219 PMID: 19841185.

35. Zhang H, Sun SC. NF-kappaB in inflammation and renal diseases. Cell Biosci. 2015; 5:63. https://doi.org/10.1186/s13578-015-0056-4 PMID: 26579219; PubMed Central PMCID: PMCPMC4647710.

36. Zheng L, Sinniah R, Hsu SI. Pathogenic role of NF-kappaB activation in tubulointerstitial inflammatory lesions in human lupus nephritis. J Histochem Cytochem. 2008; 56(5):517–29. https://doi.org/10.1369/jhc.7A7368.2008 PMID: 18285351; PubMed Central PMCID: PMCPMC2324188.

37. Zheng L, Sinniah R, Hsu SI. In situ glomerular expression of activated NF-kappaB in human lupus nephritis and other non-proliferative proteinuric glomerulopathy. Virchows Arch. 2006; 448(2):172–83. https://doi.org/10.1007/s00428-005-0061-9 PMID: 16205945.

38. McKinney EF, Lee JC, Jayne DR, Lyons PA, Smith KG. T-cell exhaustion, co-stimulation and clinical outcome in autoimmunity and infection. Nature. 2015; 523(7562):612–6. Epub 2015/07/01. https://doi.org/10.1038/nature14468 PMID: 26123020; PubMed Central PMCID: PMCPMC4623162.