Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures

Laurent Chiche (MD, PhD),[1,2]* Noémie Jourde-Chiche (MD, PhD),[1,3,4]* Elizabeth Whalen (PhD),[1] Scott Presnell (PhD),[1] Vivian Gersuk (PhD),[1] Kristen Dang (PhD),[1] Esperanza Anguiano (PhD),[3] Charlie Quinn,[1] Stéphane Burtey (MD, PhD),[4] Yvon Berland (MD),[4] Gilles Kaplanski (MD, PhD),[2] Jean-Robert Harle (MD),[2] Virginia Pascual (MD, PhD),[3] Damien Chaussabel (PhD).[1]

(1) Systems Immunology Division, Benaroya Research Institute, Seattle, WA, USA

(2) Aix-Marseille University, Department of Internal Medicine, Hôpital la Conception, AP-HM, Marseille, France

(3) Baylor Institute for Immunologic Research, Dallas, TX, USA

(4) Aix-Marseille University, Department of Nephrology, Hôpital la Conception, AP-HM, Marseille, France

Supplemental data

Methods

Immunological analyses

Antinuclear antibodies were detected by IIF on HEp-2 cells (Bio-Rad) with a screening dilution of 1/100. This method involved circular recombinant plasmid dsDNA. Single polystyrene wells were automatically filled and processed in the UNICAP 250 instrument (Phadia, Freiburg, Germany) at 37°C. Values higher than 16 IU/ml were considered positive. Quantitative determination of human complement factor C3, C4 was performed in serum samples by immunonephelometry on BN system (Siemens Diagnostics, Marburg, Germany). Normal values are the following: 0.75-1.44 g/L for C3 and 0.10-0.34 g/L for C4.

RNA Preparation and Microarray Hybridization

Total RNA was isolated from whole blood lysates using the MagMAXTM-96 Blood RNA Isolation Kit (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Following extraction, an Agilent 2100 Bioanalyzer (Agilent, Palo Alto, CA) was used to measure RNA Integrity Numbers (RIN) for each sample. All samples with RIN values > 7 were retained for further processing. RNA concentration was measured using a Nanodrop 1000 (Nanodrop Technologies, Wilmington, DE). Following RNA extraction and quality control analysis, Globin mRNA was depleted from a portion of each total RNA sample using the GLOBINclearTM-Human 96-well format kit (Ambion, Austin, TX). This was followed by another round of RIN and concentration determinations for quality control purposes. 250 ng of RNA from all samples passing quality control were then amplified and labeled using the Illumina TotalPrep-96 RNA amplification kit (Ambion, Austin, TX). 750 ng of amplified labeled RNA were hybridized overnight to Illumina HT12 V4 beadchips (Illumina, San Diego, CA). Following hybridization each chip was washed, blocked, stained, and scanned on an Illumina BeadStation 500 following the manufacturer’s protocols. Illumina GenomeStudio software was used for each sample to subtract background and scale average signal intensity to the global average signal intensity for all samples.

PCR analyses

Ct values (expression values) were exported from Fluidigm Real-Time PCR Analysis Software (Fluidigm Corporation) and processed using in-house methods written in R. Samples and genes that were not successful in at least 30% of samples/genes were removed from further analysis. Failed cases and cases with expression beyond the detectable range

were set to NA. Three housekeeping genes (EEF1A1, FAM105B, and MYL12B/MRLC2) were selected out of eight possibilities based on M values as previously described (1) and their geometric mean was subtracted from the other samples on a per-plate basis to give a deltaCt value that corrects for differing sample amounts. The average of two reference samples was subtracted from each deltaCt value on a per-plate basis to correct for plate batch effects to give a FC compared to reference samples. The reference samples used were an in-house pool and a commercial pool: FirstChoice® Human Total RNA Survey Panel (AM6000 from Life Technologies). Finally, the mean of the delta deltaCt values for control samples was subtracted from each sample and the fold change relative to controls was calculated as follows: FC = 2 ^ (-delta deltaCt).

Interferon Modules and related genes

The list of genes belonging to the 3 IFN-related modules M1.2 (http://www.biir.net/public_wikis/module_annotation/V2_Trial_8_Modules_M1.2), M3.4 (http://www.biir.net/public_wikis/module_annotation/V2_Trial_8_Modules_M3.4) and M5.12 (http://www.biir.net/public_wikis/module_annotation/V2_Trial_8_Modules_M5.12) are provided with corresponding probes and their annotations.

Public domain datasets

Publicly available blood gene-expression profiles from pediatric and adult independent SLE cohorts (Arrasapan 2011 - access to data graciously provided through ArrayTrack™ tool, FDA National Center for Toxicological Research, Jefferson, AR, USA; Berry and Pascual 2010, GSE22098) were used to validate IFN modular signatures observed in our cohort. Blood gene-expression profiles from patients infected with hepatitis C virus (HCV) treated with IFN-α (23), where samples at day 6 of treatment are compared to baseline or from patients with multiple sclerosis treated with IFN-β (22), where samples at months 3, 12 and 24 of treatment are compared to baseline (Malhotra 2011, GSE26104; Taylor 2008, GSE11342) were used to evaluate the influence of various types of IFN on the modular IFN signature.

Interferome database

The filter in Interferome was set as follows: species=human, exclusion= fetal brain, embryo, umbilical vein. Genes were defined as “IFN-related” if they were responsive to at least one

type of IFN (expression FC>2) after in vitro stimulation with IFN (compared to baseline). As the number of available experiments was not equivalent for all types of IFN, the median FC of all available experiment for the genes of interest was taken into account, and the number of available experiments for each type of IFN was represented on a Beeswarm box plot. An "IFN molecular distance to health" was defined as the number of genes with FC>2 for each sample compared to the average of healthy controls, where the genes counted only included those having evidence of IFN regulation from the Interferome database. Responsiveness to different types of IFN of the genes from the 3 IFN modules was evaluated using the Interferome database: the log2(FC) observed in each experiment for each gene after in vitro stimulation with type I or type II IFN, as well as with IFNα or β, were compared.

Two datasets included in Interferome were studied more specifically:

1) Baechler et al., corresponding to the mRNA profile of PBMCs from healthy individuals treated for 6h with INF-alpha/beta or INF-gamma. Their study identified 286 IFN-stimulated genes with more than a 2-fold change in expression and an absolute difference from baseline of more than 500 microarray units. The majority of those genes were affected by both type I and II IFNs. Data available (http://www.pnas.org/content/100/5/2610/suppl/DC1; Table S3) for 17, 19 and 18 of the genes of M1.2, M3.4 and M5.12 respectively showed a median FC type I/II ratio of 5.9, 3 and 2.2 respectively for M1.2, M3.4 and M5.12.

2) Der et al. (2), corresponding to the mRNA profiles of a human fibrosarcoma cell line treated for 6h with IFN-α, IFN-β or IFN-γ. One hundred twenty-two genes were identified as IFN-stimulated based on this group's data, with approximately one-third of those induced by all three IFN species.

Data available (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC28094/table/T2/) for 8, 11 and 10 of the genes of M1.2, M3.4 and M5.12 respectively showed a median FC alpha/beta ratio of 0.72, 0.58 and 0.9 respectively for M1.2, M3.4 and M5.12.

References

1-Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology, 3(7).

2-Der SD, Zhou A, Williams BR, Silverman RH. Identification of genesdifferentially regulated by interferon alpha, beta, or gamma usingoligonucleotide arrays. Proc Natl Acad Sci U S A 1998; 95:15623-8.

Table S1: Demographic, clinical and biological characteristics of SLE patients and healthy controls. SLE patients were split into three groups. The "at inclusion" group included all SLE patients at their first visit, irrespective of SLE disease activity at that time. The "quiescent" group included SLE patients at their first available visit with low disease activity, defined by no flare-up or treatment modifications for at least 60 days prior to the visit, and a SLEDAI of ≤4. The "longitudinal" group included SLE patients who had at least three consecutive visits during the study.

Parameters	**"At inclusion" All SLE patients n = 62**	**« Quiescent » (1st quiescent visit) n = 34**	**"Longitudinal" (Patients with ≥ 3 consecutive visits) n = 29**	**Matched Healthy Controls n = 20**
Age, median (range)	38 (18-70)	38 (18-70)	40 (20-70)	39 (18-67)
Gender Female, n (%)	53 (85)	30 (88)	24 (83)	17 (85)
Ethnicity, n (%)	.	.	.	.
White	55 (89)	29 (85)	26 (90)	19 (95)
Black	4 (6)	3 (9)	2 (7)	1 (5)
Asian	3 (5)	2 (6)	1 (3)	0
SLE duration (years) median (range)	7.8 (0-40)	9.3 (0.44-40)	7.2 (0-40)	-
anti-dsDNA positivity n (%)	39 (63)	20 (59)	21 (72)	-
anti-dsDNA titer (UI/L) median (range)	41 (0-578)	26 (0-465)	55 (0-578)	-
anti-Sm	11 (18)	3 (9)	5 (17)	-
anti-RNP	16 (26)	7 (21)	6 (21)	-
anti-SSA	20 (32)	6 (18)	12 (41)	-
anti-SSB	5 (8)	2 (6)	3 (10)	-
low C3	26 (42)	9 (26)	14 (48)	-
low C4	22 (35)	6 (18)	11 (38)	-
Lymphocyte count (/mm3) median (range)	1520 (230-5410)	1620 (460-5410)	1530 (520-2870)	-
Treatment n (%)	.	.	.	.
IV CYP	3 (5)	0	4 (14)	-
MMF	10 (16)	4 (12)	6 (21)	-
AZA	18 (29)	12 (35)	6 (21)	-
IV CS	1 (1.6)	0	2 (7)	-
Oral CS	52 (84)	30 (88)	28 (97)	-
Daily CS dose (mg/day) median (range)	8 (0-90)	8 (0-15)	10 (0-90)	-
Daily CS dose (mg/kg) median (range)	0.14 (0.03-1.13)	0.11 (0.03-0.2)	0.16 (0.08-0.95)	-
HCQ	35 (56)	20 (59)	14 (48)	
SLEDAI median (range)	5 (0-26)	2 (0-4)	8 (0-22)	-
Cutaneous flare n (%)	20 (32)	0	10 (34)	-
Articular flare n (%)	16 (26)	0	6 (21)	-
Hematological flare (%)	5 (8)	0	4 (14)	-
Renal flare n (%)	25 (40)	0	17 (59)	-

Table S2: Prevalence of IFN signature on the 157 SLE samples of our cohort according to various IFN scores from the literature.

IFN scores	SLE samples, n (%)
Yao et al.[25]	
High	20 (13)
Moderate	86 (55)
Low	51 (32)
Petri et al.[8]	
≥2	132 (84)
<2	25 (16)
Feng et al.[26]	
>0	139 (88.5)
≤0	18 (11.5)
Kirou et al.[27]	
High IFN-α	128 (82)
Low IFN-α	29 (18)
High IFN-γ	112 (71)
Low IFN-γ	45 (29)

Table S3: Characteristics of SLE patients at inclusion according to the modular IFN score classification.

Modular IFN score	ABSENT	MILD	MODERATE	STRONG	p
Patients, n (%)	9 (15)	7 (11)	24 (39)	22 (35)	-
Ethnicity, n (%)					
White	9 (100)	6 (86)	21 (88)	19 (87)	
Black	0	1 (14)	1 (4)	2 (9)	
Asian	0	0	2 (8)	1 (4)	0.87
Age (yrs), median (range)	34 (20-54)	39 (25-55)	34 (18-70)	39 (25-59)	0.75
Disease duration (yrs), median (range)	8 (0-30)	5 (0-29)	8 (0-40)	9 (0-33)	0.63
Gender Female, n (%)	7 (78)	7 (100)	21 (87.5)	18 (82)	0.66

Table S4: Characteristics of SLE samples at each visit according to the modular IFN score classification.

Modular IFN Score	ABSENT	MILD	MODERATE	STRONG	p
Visits, n (%)	26 (16)	17 (11)	69 (44)	45 (29)	-
SLEDAI, median (range)	5 (0-23)	4 (0-22)	4 (0-22)	6 (0-26)	0.18
Cutaneous flare, n (%)	2 (8)	3 (18)	15 (22)	13 (29)	**0.03**
Articular flare, n (%)	3 (12)	5 (29)	14 (20)	7 (16)	0.9
Hematological flare, n (%)	1 (4)	0	6 (9)	3 (7)	0.4
Renal flare, n (%)	11 (42)	4 (24)	24 (35)	25 (56)	0.2
Low C3/C4, n (%)	10 (38)	5 (29)	36 (52)	23 (51)	0.14
Anti-dsDNA titer, mean (SD)	28 (50)	27 (30)	81 (108)	82 (109)	**0.0004**
Lymphocyte count, mean (SD)	2469 (1703)	2201 (852)	1643 (786)	928 (583)	**<0.0001**
CS mg/day, median (range)	10 (0-60)	8 (0-60)	10 (0-60)	15 (0-90)	0.19
Antimalarial, n (%)	21 (81)	11 (65)	38 (55)	20 (44)	**0.002**
IS (CYP, MMF or AZA), n (%)	18 (69)	9 (53)	34 (49)	23 (51)	0.15
IS + Antimalarial, n (%)	15 (58)	9 (53)	20 (29)	10 (22)	**0.0006**

Table S5: Characteristics of quiescent SLE patients according to the modular IFN score classification.

Modular IFN score	ABSENT	MILD	MODERATE	STRONG	P
Patients, n (%)	5 (14)	6 (18)	17 (50)	6 (18)	-
Ethnicity, n (%)					
White	5 (100)	5 (83)	15 (88)	4 (67)	
Black	0	1 (17)	1 (6)	1 (17)	
Asian	0	0	1 (6)	1 (17)	0.59
Age (years), median (range)	48 (28-54)	45 (25-55)	32 (18-70)	31 (25-40)	0.42
Disease duration (years), median (range)	12 (2-33)	15 (4-29)	9 (0.4-40)	8 (2-18)	0.66
Gender Female, n (%)	4 (80)	6 (100)	15 (88)	5 (83)	0.89

Table S6: Characteristics of quiescent SLE samples at each visit according to the modular IFN score classification.

Modular IFN score	ABSENT	MILD	MODERATE	STRONG	p
Visits, n (%)	13 (20)	8 (13)	31 (48)	12 (19)	-
SLEDAI, median (range)	2 (0-4)	2 (0-4)	2 (0-4)	2 (0-4)	0.50
Low C3/C4, n (%)	2 (15)	2 (25)	14 (45)	4 (33)	0.058
Anti-dsDNA titer, mean (SD)	18 (18)	22 (33)	51 (49)	108 (172)	**0.007**
Lymphocyte count, mean (SD)	2015 (1298)	2254 (817)	1488 (470)	639 (283)	**< 0.0001**
CS mg/day, median (range)	5 (0-20)	5 (0-10)	10 (0-20)	10 (0-20)	0.063
Antimalarial, n (%)	8 (62)	3 (38)	16 (52)	7 (58)	0.74
IS (CYP, MMF or AZA), n (%)	10 (77)	2 (25)	14 (45)	4 (33)	0.072
IS + Antimalarial, n (%)	7 (54)	2 (25)	7 (23)	2 (17)	0.17

Table S7: Accuracy of the 9-gene IFN panel to predict the group of IFN-modular score

	Absent/Mild	Moderate/Strong	Class Error
Absent/Mild	33	3	0.083
Moderate/Strong	3	89	0.033

16 genes belonging to the 3 IFN modules and which primers were available in the lab at the time of the study were used for qPCR validation (see Figure S5). From this list of 16, a panel of 9 genes was generated by random forest to end with a list of 3 genes belonging to each of the 3 IFN modules: IFI44, XAF1 and SPATS2L for M1.2; MOV10, GALM and TIMM10 for M3.4; and LBA1, TRIM38 and TRIM56 for M5.12.

Table S8: Gene composition of IFN scores from the literature.

genes	Bennett	Baechler	Kirou	Feng	Nikpour	Landolt	Petri	Yao	Higgs	Chaussabel
n	**8**	**14**	**3**	**5**	**31**	**5**	**3**	**21**	**5**	**corresponding modules**
APOBEC1-like		x								NA
BST2					x					**M5.12**
C7orf6					x					NA
CIC		x								NA
DNAPTP6								x		M1.2
EIF2AK2			x							M3.4
EPSTI1					x			x		M1.2
FCGR1A		x								NA
FLJ20035					x					M1.2
HERC5								x		M1.2
HERC6					x					M3.4
IFI6		x						x	x	NA
IFI27							x	x	x	NA
IFI35					x					M3.4
IFI44	x		x		x		x	x	x	M1.2
IFI44L		x			x			x	x	M1.2
IFIH1					x					M3.4
IFIT1			x		x	x				M1.2
IFIT2					x			x		M3.4
IFIT3	x				x			x		M1.2
IFRG28					x					M1.2
IRF7					x					M3.4
ISG15	x	x		x	x	x		x		M1.2
LAMP3								x		M1.2
LAP3					x					M3.4
LGALS3BP		x								M3.4
LGP2					x					M3.4
LOC129607					x					M1.2
LY6E	x	x		x		x		x		M1.2
MX1	x	x		x	x	x		x		M1.2
MX2										NA
OAS1				x	x	x		x		M1.2
OAS2					x			x		M1.2
OAS3					x		x	x		M1.2
OASL	x	x		x	x					M1.2
PML					x					M3.4
PLSCR1	x	x						x		M3.4
RNASE2		x								NA
RSAD2					x			x	x	M1.2
RTP4								x		M1.2
SERPING1		x								M1.2
SIGLEC1								x		NA
SP110					x					**M5.12**
STAT1										M3.4
STAT2					x					M3.4
TAP1					x					**M5.12**
UBE2L6					x					M3.4
USP18					x			x		NA
XAF1	x	x			x					M1.2

The number and names of genes of each score is provided, as well as the IFN modules to which these genes belong. Most genes composing these scores belong to M1.2 (the most sensitive and the less variable IFN module), some belong to M3.4, but very few belong to M5.12 (the most variable module, upregulated only after the 2 others).

Table S9: Annotation of clusters of genes differentially expressed between quiescent SLE samples, ordered according to their modular IFN score, and healthy controls (See Figure S3).

Clusters	Annotation	n transcripts	KEGG pathways	Representative genes
1	B-cell related	38	B cell receptor signaling pathway B cell-related genes	CD19, CD23, CD69, CD72, CXCR5
2	Neutrophil-related	32	NA	cathepsin G, defensin alpha 1-3 and 4, myeloperoxidase, lactotransferrin
3	IFN-regulated chemotaxis	24	NOD-like receptor signaling pathway	CCL2, CCL8, C3AR1, SIGLEC-1
4	IFN early signaling-related	66	RIG-like receptor signaling pathway Cytosolic-DNA-sensing pathway	ISG15, IRF7, IFI27, IFIT3, SERPING1, XAF1, LY69E, MX1, OAS1-2-3
5	IFN downstream signaling-related	49	Toll-like receptor signaling pathway Chemokine signaling pathway	TLR2, SOCS1, STAT1, GBP1, FCGR1A-B-C, CCR1, CXCL10, TNFSF13B (BAFF)

Functional annotation of genes in networks was performed using the web-based tool DAVID (http://david.abcc.ncifcrf.gov).

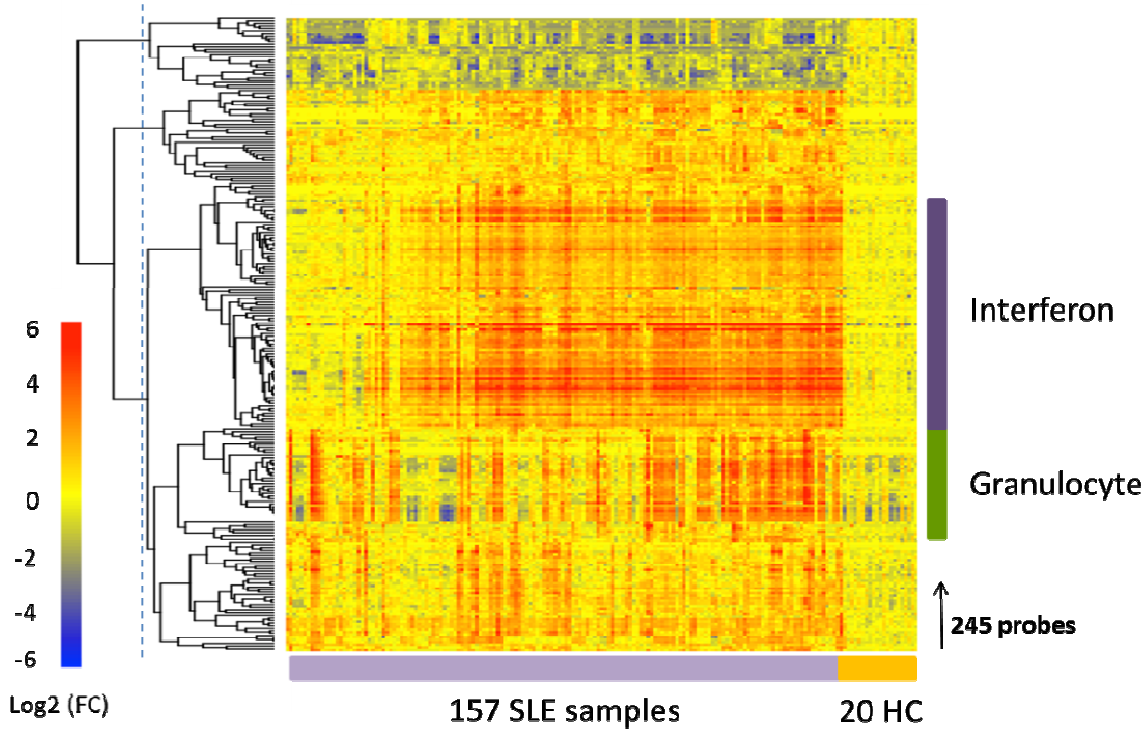

Figure S1: Gene expression in SLE patients compared to matched healthy controls. Unpaired t-tests (Benjamini–Hochberg FDR 0.05) identified 245 transcripts differentially expressed (fold change (FC $\geq$ 2 or $\leq$1/2) between SLE patients and healthy controls. Hierarchical clustering of probes was performed with Pearson correlation as the distance metric. Log2(FC) values are plotted on the heatmap (color key from -6 to 6 corresponding to FC from 1/64 to 64). Clusters of probes identified with this approach were then annotated with DAVID (see annotation on the right of the heatmap corresponding to the IFN and granulocyte clusters or signatures). During the study period, 54/62 patients (87%) had at least one sample with an IFN signature, corresponding to a significant overexpression of the gene cluster annotated "interferon". Of the 157 samples obtained from 62 SLE patients, 130 (83%) showed an IFN signature while only 27 samples (17%) from 12 patients did not show this signature.

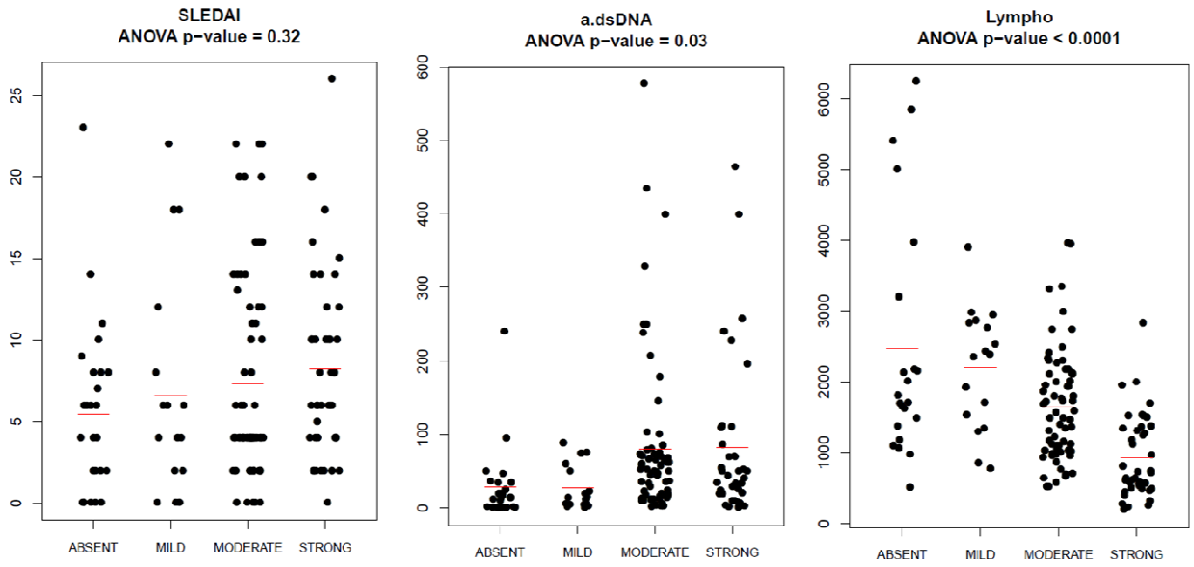

Figure S2: Clinical and biological activity of SLE at each visit according to the modular IFN score.

SLEDAI score was not significantly different between the 4 groups of IFN modular scores (ANOVA), but there was a tendency for a linear trend between groups (p=0.06). Anti-dsDNA antibodies were higher in patients with moderate and strong scores compared to patients with absent (p=0.02 vs moderate, p=0.03 vs strong) or mild (p=0.04 vs moderate, p=0.05 vs strong) scores. Lymphocyte count was lower in patients with moderate and strong IFN scores compared to patients with absent (p=0.0005 vs moderate, p<0.0001 vs strong) or mild (p=0.03 vs moderate, p<0.0001 vs strong) IFN scores, with a significant linear trend (p<0.0001).

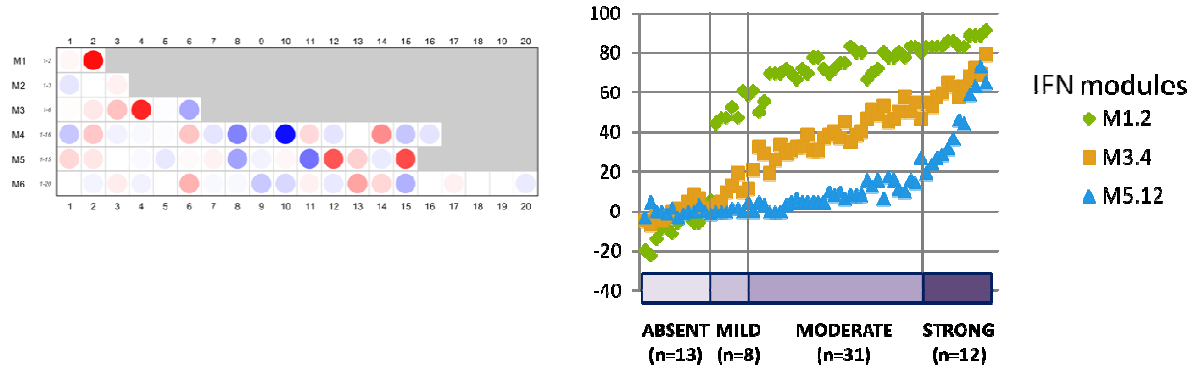

Figure S3: Group modular map of the 64 clinically quiescent SLE samples (left) and repartition of these samples according to their individual modular IFN score (right). Modular IFN signature is observed even in clinically quiescent SLE, as shown by the upregulation of IFN modules at the group level (94% probes upregulated in M1.2, 85% in M3.4 and 67% in M5.12, with no probe downregulated) as well as at the individual level (80% of samples with at least a mild modular IFN score).

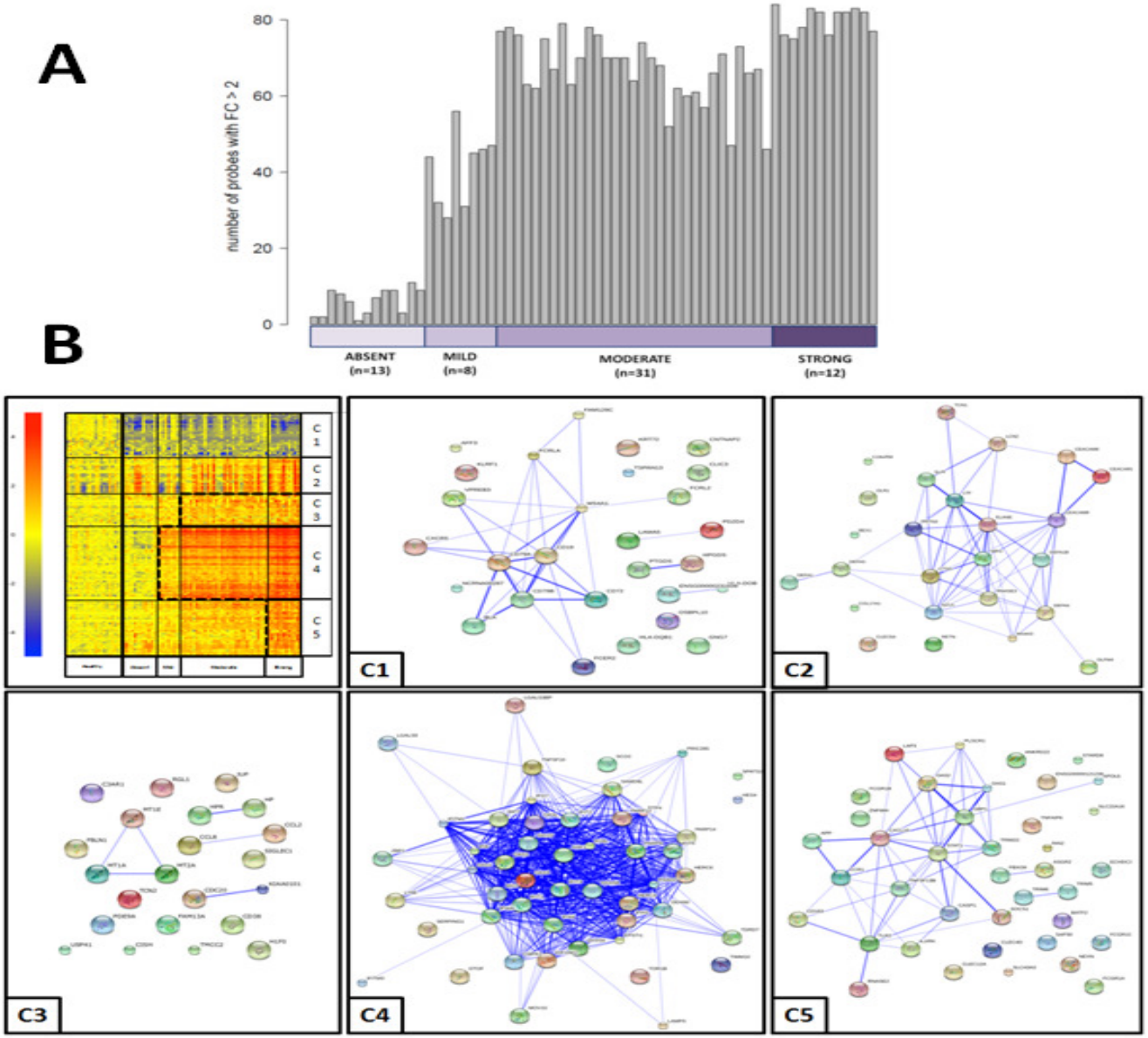
A
number of probes with FC > 2
0
20
40
60
80
ABSENT
(n=13)
MILD
(n=8)
MODERATE
(n=31)
STRONG
(n=12)
B
C1
C2
C3
C4
C5

Figure S4: Quantitative and qualitative gene expression changes in quiescent SLE samples according to the modular IFN score.

Blood gene expression data identified differentially expressed transcripts (n=209) in quiescent SLE patients compared to healthy controls (Benjamini–Hochberg FDR <0.05, FC ≥ 2).

A: A list of unique gene symbol derived from upregulated transcripts (n=171) was entered in the Interferome database and identified 75 unique "IFN-related" genes. The corresponding transcripts (n=85) were used to calculate an "IFN molecular distance to health" for each sample. In quiescent SLE patients, this "IFN molecular distance to health" was strongly correlated to the modular IFN score (Pearson correlation = 0.946, $p < 0.0001$).

B: The 64 quiescent samples were ordered according to their modular IFN score and hierarchical clustering of differentially expressed probes (SLE versus healthy controls) was performed with Pearson correlation as the distance metric. Pathway analysis revealed several clusters of up-regulated genes that were correlated with the modular classification of samples. Functional annotation of genes in networks (**see Table S9**) was performed using the web-based tool DAVID (http://david.abcc.ncifcrf.gov) and visualized using the online STRING interface (http://string-db.org).

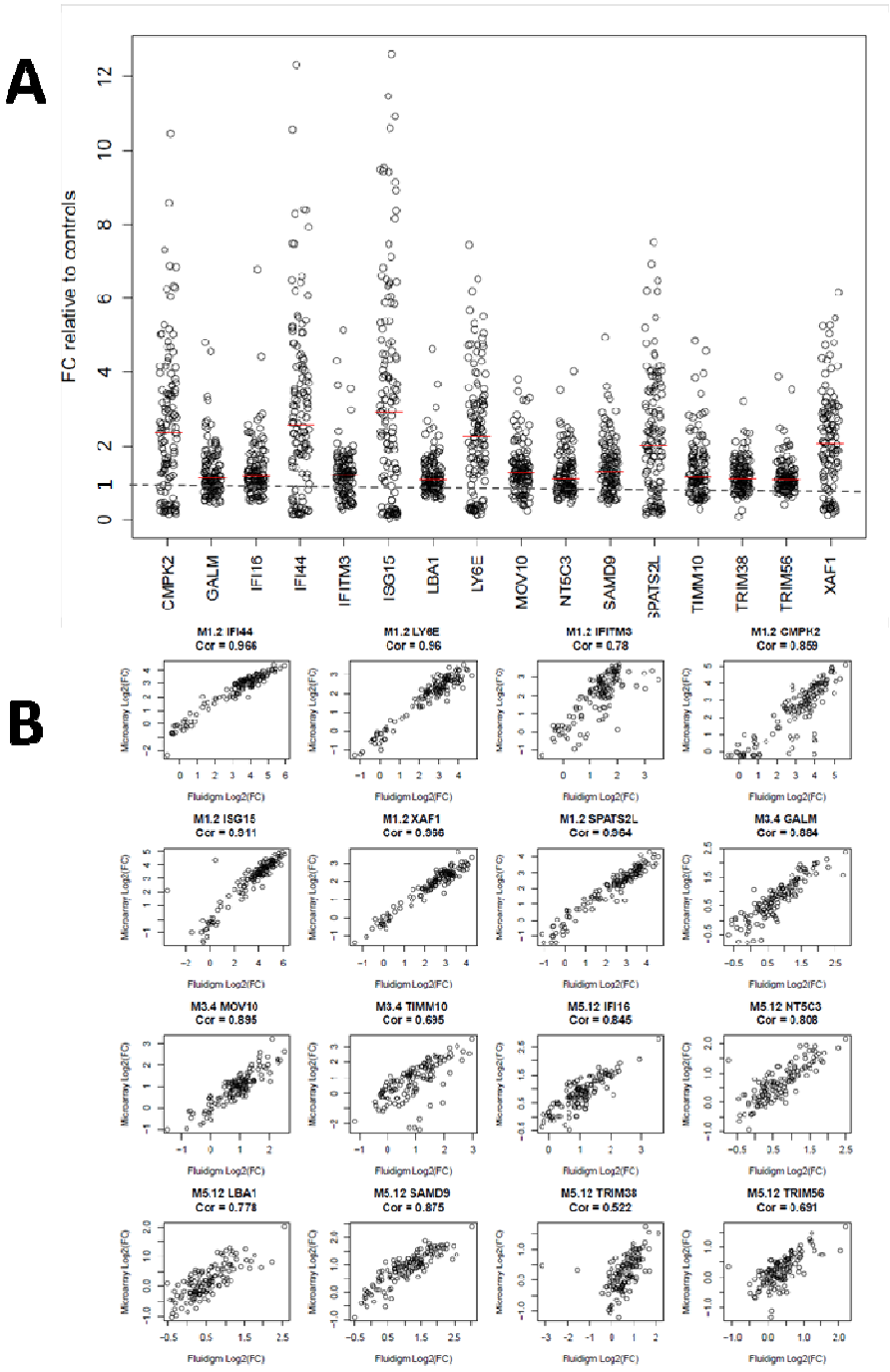

Figure S5: qPCR validation: expression of 16 IFN-inducible genes belonging to IFN modules. A: Expression of 16 IFN-inducible genes in SLE (n=128) compared to healthy controls (n=20) using TaqMan assays (median fold change). **B:** Pearson correlation of the expression of 16 genes belonging to IFN modules by qPCR (TaqMan assays, Fluidigm) and by microarray.