Supporting Information

Materials and Methods

Study subjects

For the CNV-GWAS analysis, 100 RA patients (76 females and 24 males, mean age 54±11 years), who met the 1997 American College of Rheumatology criteria (1), were recruited from St. Vincent Hospital (Suwon, Korea) and Eulji University Hospital (Daejeon, Korea). As a normal control group, 400 healthy individuals (198 females and 202 males, mean age 49±8 years) were randomly selected from the Korean Genome Epidemiology Study (KoGES). For independent replication of the significant CNV, 764 RA patients and 1224 healthy individuals were recruited from Korea (St. Vincent Hospital) and the USA (Section of Rheumatology, Yale University School of Medicine, CT06520, USA); 599 RA cases (421 females and 178 males, mean age 54±12 years) and 966 controls (931 females and 35 males, mean age 41±13 years) from Korea, and 165 RA cases (105 females and 60 males, mean age 63±9) and 258 controls (69 females and 74 males, mean age 47±15) from the USA (***SI Appendix*, Table S1**). Genomic DNA was extracted from whole blood by using a DNeasy Blood & Tissue Kit (Qiagen). This study was performed with the approval of the institutional review board (CUMC09U034).

Defining CNV regions and statistical analyses

After CNV calling, we defined CNV regions (CNVR) as the union of overlapping CNVs from the 500 individuals using CNVRuler software (2). To avoid potential overestimation, regions of very low density of overlapping CNVs (<10% of the total

CNVs consisting of each CNVR) were not merged into CNVRs as described previously (3). After constructing CNVRs, logistic regression analysis was conducted by adjusting for age and sex. The false discovery rate (FDR) method was used for multiple comparison correction. CNVRs with FDR < 0.01 were considered to be potentially significant.

Quantitative PCR for LSP1 DNA

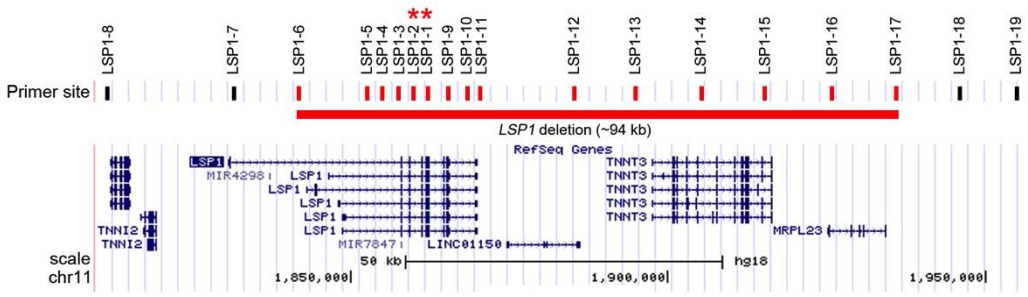
To validate the existence and boundaries of the deletion encompassing *LSP1* locus, we performed genomic qPCRs for multiple points within and outside the expected CNV breakpoints estimated by SNP microarray. For this, we designed 19 target specific primer pairs (see a **Table** below). Through the multiple point genomic qPCR, the *LSP1* deletion breakpoint was clearly defined (see a **Figure** below). The length of the deletion confirmed by genomic qPCR (~94 kb) was longer than the length estimated by SNP microarray (~43 kb). In addition to the *LSP1*, there is another coding gene *TNNT3* in this deletion-CNVR. TNNT3 is a member of the Tropomyosin gene family. We did not perform replication and functional analyses for *TNNT3*, because *TNNT3* is known to involve in congenital muscle disorders (4) but never reported to be associated with RA or other inflammatory disorders.

Next, we performed genomic qPCR for the replication cohort using two *LSP1* specific primer pairs located in the *LSP1* gene. Primer sequences specific for the *LSP1* gene consisted of a forward primer (5′-GATGGAGGGTGGGCTTTAC-3′) and a reverse primer (5′-CTTCTCTATGGAGCGGTTTAGG-3′) for *LSP1-1*, and a forward primer (5'-TCCGTAATGGAGGCTGAGA-3') and a reverse primer (5'-GGCATGCACATCTGGAAATAC-3') for *LSP1-2*. As a diploid internal reference gene

for genomic quantitative PCR, HS6ST3 gene was used. The forward primer of HS6ST3 (5′-CGCTACCACCACACCAAGCAG-3′) and a reverse primer (5′-CCACCTGGCTGTTGTAGTCCTC-3′) were used as described elsewhere (5). Genomic quantitative PCR was performed using the Viia 7 system (Life Technologies, Carlsbad, CA). In brief, 10 μl of reaction mixtures containing 10 ng of genomic DNA, THUNDERBIRD TMSYBR® quantitative PCR Mix (TOYOBO, Osaka, Japan), and 0.6 pmole of each primer were used. Thermal cycling conditions consisted of one cycle of 60 seconds (sec) at 95℃, followed by 40 cycles of 5 sec at 95℃, 10 sec at 61℃, and 20 sec at 72℃. After PCR reactions, the melting curve analysis was performed using the ViiATM 7 Softrware ver.1.1 (Applied Biosystem). Copy number of each target was defined by using as $2^{-\Delta\Delta CT}$ equation, where ΔCt is the difference in threshold cycles for the sample in question normalized against reference gene and expressed relative to the value obtained by calibrator DNA, NA10851 (individual/calibrator). We set the median value of qPCR ratio of *LSP1* (individual/NA10851) in the healthy control group (0.92) to that of a diploid as described elsewhere (6). Then, the measured copy number ratio values were rounded off to the nearest integer as described elsewhere (3). Once the results from both *LSP1-1* and *LSP1-2* qPCRs were consistent, we defined the copy number status of the individual. If *LSP1*-2 qPCR was not available due to limitations of genomic DNA, we used the *LSP1-1* qPCR result for defining the copy number status.

Primers for LSP1 genomic qPCR

Primer	Forward	Reverse
LSP1-1	GATGGAGGGTGGGCTTTAC	CTTCTCTATGGAGCGGTTTAGG
LSP1-2	TCCGTAATGGAGGCTGAGA	GGCATGCACATCTGGAAATAC
LSP1-3	CTGGTTCTACTTAGGGAGGGA	GACTCCAAGGCCCATTCAA
LSP1-4	CAGACACAGGAAATCGCCTAAT	TGTCAGGACTCTGGGATGG
LSP1-5	CTGAACCCTGGGTAGTGAGA	CCCTAGGTGCTTCTACTCAGA
LSP1-6	GGGAACAGAGAGGGTTCTTG	GCTCCTGGTTCTCCACTTT
LSP1-7	ATGGCAGAGGTTGTGTGTAG	CCATTGGTACCACAGAGTACAG
LSP1-8	CCCTTTCTGTCTGTCTCCTTG	GCCATCCCATCGGGAATAA
LSP1-9	CCTTCACTCTTGGTCTCCTTTC	CTTCCACAAGCACCTTCTCATA
LSP1-10	CAAACGCAAGTTGTCCTTCC	AGGACACAGTCATGCTCATATC
LSP1-11	CAGTATAAGACGCCAGGACATAA	GGGTAGGGATATCACAGACATTC
LSP1-12	AAATGGTGAAGTGGTGAAATGG	GCTGCACACTTATCTGCTTTC
LSP1-13	GCAGAACGAGTCAATGAAACTG	GACACAGCACTTGGGAACT
LSP1-14	GTCAGAGAGCTGCACATTCA	GCTCGAGTCACGCAAGAG
LSP1-15	GGGACCTAGGACAGGGAATTA	GATGAGGTTTCACCATCCTGAG
LSP1-16	TTGGGTTTGAGTGGGAAAGG	GGAATCCAGAGAACATGCAGAG
LSP1-17	CATGCGAATGCTCAGAAGAAC	TGCCCAAGCAGAAGGAAA
LSP1-18	TGCCCTTTATCTGAGGGTATCT	TCCCATCTCCCTGAGCTATTT
LSP1-19	TTCAGTGCTGGCTTGAAGAG	GAACCACCTGGACCTCATTATT



Breakpoints of the deletion-CNVR in 11p15.5. On the primer site track (19 target specific primer pairs in 11p15.5), red vertical bares represent copy number loss and black bars represent diploid copy number. The asterisks represent the primers that used in genomic qPCR for replication cohort.

Microarray experiments

Total RNA was isolated from splenic T cells of *Lsp1*-deficient and wild-type mice, which were stimulated with anti-CD3/CD28 Abs for 6 hours. The total RNA integrity was evaluated by Bioanalyzer 2100 (Agilent, Santa Clara, CA). The RNA integrity number in all the samples was larger than 9.5. The RNA was reverse-transcribed and amplified according to standard Agilent protocols. Then, it was hybridized to an array chip (SurePrint G3 Mouse GE 8x60K Microarray, Agilent) containing 62,976 probes for 24,241 annotated genes.

Identification of differentially expressed genes (DEGs)

Log_2-probe-intensities in individual samples were normalized using the quantile normalization method (7). An integrative statistical method previously reported was then applied to the normalized log_2-probe-intensities (8). In short, for each gene, Student's *t*-test and log_2-median-ratio test were performed to compute *T* values and log_2-median-ratios between *Lsp1* (-/-) and *Lsp1* (+/+) T cells. Empirical null distributions of both *T* values and log_2-median-ratios were generated by random permutations of the samples 1,000 times and then by applying Gaussian kernel density estimation method (9) to the *T* values and log_2-median-ratios from the random permutations. For each gene, adjusted *P* values of the two tests were obtained by two-tailed tests for the observed T value and log_2-median-ratio using the empirical distribution, and the adjusted *P* values of the two tests were combined to compute an overall *P* value using Stouffer's method (10). The DEGs were identified with the following criteria; the overall *P* value less than 0.05 and the absolute log_2-median-ratio larger than the cutoff. The cutoff was determined as the mean absolute values of 2.5th

and 97.5th percentiles of the empirical null distribution.

Functional enrichment analysis

The enrichment analysis of GOBPs and KEGG pathways was conducted for up- or down-regulated genes in *Lsp1* (-/-) T cells, compared to *Lsp1* (+/+) T cells, in the absence and presence of anti-CD3/28 Abs using DAVID software (8). GOBPs and KEGG pathways with their *P* values < 0.1 were identified as the ones significantly represented by the up- or down-regulated genes in each condition. Migration-related genes were identified as 1359 genes annotated with cell motility (GO:0048870), cell migration (GO:0016477), leukocyte migration (GO:0050900), and chemotaxis (GO:0006935). The list of the genes annotated with each GOBP term was obtained from AmiGO (11).

Significance of migration-related DEGs being regulated by ERK-downstream TFs

The significance that the DEGs in *LSP1*-deficeint T cells in the presence of anti-CD3/28 Abs and migration-related genes are shared was computed. The same number of the genes with the DEGs was randomly sampled 10^6 times from the list of the genes spotted on the microarray, and then the shared genes with the migration-related genes were counted after each random sampling. An empirical null distribution for the number of shared genes was generated. An adjusted *P* value was obtained by the one-tailed test for the observed number of the shared genes (n=67) using the empirical distribution.

For the shared migration-related DEGs, a list of ERK downstream TFs was first obtained from the KEGG pathway database (12) and previous literatures (13, 14). For each TF, the number of target migration-related DEGs regulated by the TF was counted.

For the number of target DEGs regulated by each TF, the *P* value was computed based on random sampling as described above. After randomly sampling 67 genes from the list of genes spotted on the microarray, the number of target genes from the 67 randomly sampled genes was counted for each TF, based on TF-target relationships in GeneGO MetaCore^{TM} (15). This procedure was repeated 10,000 times. Using the numbers of the target DEGs from the random sampling, an empirical distribution for the number of target DEGs was generated, and a *P* value for the observed target DEGs was computed using the empirical distribution.

Culture of mononuclear cells and Jurkat T cells

Heparinized peripheral blood was aseptically collected from healthy controls and RA patients who met 1987 the American College of Rheumatology criteria (1). Peripheral blood mononuclear cells (PBMC) were isolated by density-gradient centrifugation on Ficoll-Hypaque. The cells were resuspended in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS; Gibco BRL, Gaithersburg, MD). Cells were then incubated at 37℃ in a 5% CO_2 atmosphere. Jurkat T cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA), and were maintained in RPMI 1640 supplemented with 10% FBS (Gibco BRL), L-glutamine, and penicillin/streptomycin.

Cloning and overexpression of LSP1 gene in Jurkat cells

LPS1 full-length cDNA-inserted plasmid (pOTB7) was purchased from the NIH mammalian gene collection (Open Biosystems, Huntsville, AL), and the gene was subcloned into a mammalian expression vector, EGFP-C1 (Clontech**,** Mountain View,

CA), at *EcoRI* sites. The DNA sequence was verified by sequencing (Cosmo Genetech, Seoul, Korea). The *LSP1* plasmid DNA was tagged with *GFP* (20 μg) and was then transfected into Jurkat cells ($3x10^6$ cells) by electroporation (Invitrogen, Carlsbad, CA). After confirming LSP-1 expression in transfected cells, cell lines where the LSP1 gene was stably overexpressed were then generated by the limiting dilution method, as previously described (16). The cells were subsequently maintained in 2 mg/ml G418 (Gibco BRL).

Mice

Mice genetically deficient in *Lsp1* (*Lsp1* –/– mice) were backcrossed onto the C57BL/6 background (generation N10) (17). Age and sex-matched wild-type C57BL/6 mice (*Lsp1*+/+) were used as controls. All mice were 8-12 weeks of age unless specified otherwise. The mice were maintained in specific pathogen-free conditions and were used according to guidelines of the Institutional Animal Care Committee.

Induction of delayed type hypersensitivity reaction

Evaluation of DTH responses was performed as previously described (18). Briefly, mice were immunized subcutaneously with 200 μl of 1 mg/ml methylated BSA (mBSA; Sigma, St Louis, MO) emulsified with 100 μg of complete Freund's adjuvant (CFA; Chondrex, Redmond, WA). Seven days after immunization, mice were challenged again subcutaneously into one footpad with 10 μl of 10 mg/ml mBSA in phosphate-buffered saline (PBS); the opposite foot pad was injected with an equal volume of PBS alone as a control. At the indicated times after the challenge, footpad thickness was measured with a dial caliper (Kori Seiki MFG Co Ltd, Tokyo, Japan).

The magnitude of the DTH response was determined as follows: [footpad swelling (mm)] = [footpad thickness of mBSA-injected footpad (mm)] - [footpad thickness of PBS injected footpad (mm)]. At days 8, spleen and lymph node cells were isolated from the mice

Induction of Antigen-induced arthritis (AIA)

Antigen-induced arthritis was induced as described previously (19). Briefly, *Lsp1*-deficient and wild-type mice were immunized subcutaneously with 100 μg of mBSA (Sigma) emulsified with CFA (Chondrex) 14 and 21 days before AIA induction. Arthritis was induced by intra-articular inoculation of 100 μg of mBSA in the ankle joints (day 0). Severn days after intra-articular injection, the animals were sacrificed and their ankle joints were harvested for histology. The sections were stained with haematoxylin and eosin and evaluated in a blinded manner according to a histological scoring system, ranging from 0 to 3 (0: no, 1: mild, 2: moderate, 3: severe alterations).

For collagen-induced arthritis (CIA), male DBA/1 mice (Jackson Laboratories, Bar Harbor, ME) were immunized with bovine type II collagen (Chondrex) at 8 to 12 weeks of age, as described previously (20, 21).

Cell viability: MTT assay

T cells viability was determined by an MTT assay as described previously (21).

T cell migration assays

T cells were isolated from human PBMC or from spleen and lymph node of mice using commercially available magnetic beads (Miltenyi Biotec, Auburn, CA). Primary

T cells and transfected Jurkat cells were cultured in migration medium (RPMI 1640 containing 1% FBS, 10 mM HEPES buffer, pH 6.9) overnight. Migration assays were performed in transwell chambers with 5 μm polycarbonate membrane (Costar, Cambridge, MA). Human SDF1 (R&D Systems, Minneapolis, MN), MCP-1 (R&D Systems), and IL-6 (R&D Systems) were diluted to appropriate concentrations in migration medium and added to the lower chamber of the transwells. The cells were loaded into the upper chamber in migration medium, and allowed to migrate for the indicated times at 37℃ in 5% CO_2. The migrated cells were counted manually or determined by flow cytometry.

Flow cytometry analysis of LSP1

Expressions of CD4, CD8, and LSP1 in PBMC were determined by flow cytometry analysis. Briefly, mononuclear cells ($1x10^6$) were stained with phycoerythrin (PE)-labeled anti-CD4 monoclonal (IgG2b; eBioscience, San Diego, CA) or allophycocyanin (APC)-labeled anti-CD8 monoclonal Ab (IgG2a; eBioscience) for 30 minutes at 4^oC in the dark. APC-labeled IgG2a and PE-labeled IgG2b isotype Abs (eBioscience) were used as controls. After permeabilization using a cytoperm kit (BD Biosciences, San Jose, CA), PBMC were stained with FITC-conjugated anti-LSP1 Ab (IgG1; BD Biosciences). FITC-labeled IgG1 isotype Ab (eBioscience) was used as a control. Fluorescence intensity was measured by the FACS Canto II system (BD Biosciences).

Western blot analysis for LSP1 and ERK

LSP1 expressions in primary T cells and Jurkat cells were also measured by

western blot. Final protein concentrations were calculated using a Bradford assays (BioRad, Hercules, CA). Electrophoresis was carried out by sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis (SDS-PAGE), and blots were transferred to a nitrocellulose membrane. Membranes were incubated with Abs to LSP1 (BD Biosciences), p-ERK1/2 (Cell signaling, Beverly, MA), GFP (Santa Cruz Biotechnology, Santa Crcuz, CA), and β-actin (Sigma), followed by horseradish peroxidase-conjugated anti-mouse or anti-rabbit IgG (Santa Cruz Biotechnology). Binding of the Ab on the blot was detected and visualized using a western blot detection kit (GE Healthcare, Amersham, UK).

Immunoprecipitation

After pre-clearing with goat normal IgG (Sigma), samples were incubated with a pERK Ab with shaking at 4^{o}C overnight and followed by further incubation with an agarose-conjugated protein G (Invitrogen) for 1 hour at 4 ^{o}C. The immune complexes were precipitated by centrifugation and washed with cold PBS three times. The complex of LSP1 and pERK was detected by western blot analysis using anti-LSP1 Ab.

Immunohistochemistry and immunofluorescence staining

For immunohistochemistry, sections (5 μm) of paraffin-embedded RA synovia were treated with pepsin for 30 minutes and blocked with bovine serum albumin (BSA) for 30 minutes at room temperature. Tissue sections were then incubated with anti-LSP1 Ab (BD Biosciences) overnight at 4℃. Isotype control Ab (Sigma) was used as a negative control. The LSP1 positive cells were detected by using peroxidase-conjugated streptavidin (Vector Laboratories) followed by 3’3-diamino-benzidine

tetrahydrochloride (DAB; Vector Laboratories). The slides were counterstained with haematoxylin.

For immunofluorescence staining, RA synovia were fixed with cold acetone for 10 minutes at -20℃ and blocked with 1% BSA for 30 minutes at room temperature. Tissue sections were then incubated with anti-CD3 Ab (Santa Cruz Biotechnology, Santa Cruz, CA) plus anti-LSP1 Ab (Santa Cruz Biotechnology) overnight at 4℃. Each slide was washed three times in PBS and incubated with Cy3-conjugated anti-IgG for CD3 (Abm, Richmond, Canada) and Alexa Fluor 488-conjugated anti-IgG for LSP1 (Invitrogen, Carlsbad, CA) in a dark humid chamber for 40 minutes. After washing with PBS, the coverslips were mounted on the glass slides with ProLong Antifade solution (Invitrogen) and examined under a confocal laser scanning microscope (LSM 510; Carl Zeiss, Göttingen, Germany).

Real-time PCR analysis

Cells were cultured with or without TCR triggering for 24 hours. RNA was isolated from transfected Jurkat and T cells of mouse spleen or lymph nodes using the RNeasy® Mini kit (Qiagen, Hilden, Germany). Real-time PCR was performed using SYBR Green Supermix with an iCycler thermal cycle (Bio-Rad). PCR primers used in this study were purchased from Bioneer (Daejeon, Korea) and are listed in ***SI Appendix,* Table S6**. The data were collected and analyzed using the comparative Ct (threshold cycle) method using GADPH as the reference gene.

Measurement of serum anti-mBSA IgG

Anti-mBSA IgG levels in the sera were measured by enzyme-linked

immunosorbent assay (ELISA). Briefly, 96-well plates were coated overnight with 100 μl of a 5 μg/ ml mBSA solution at 4^{o}C and then blocked by incubation with 2% casein (Sigma) in PBS for 1 hour at room temperature. After washing three times in PBS, 100 μl of each mouse serum, diluted in PBS, was added for 1 hour. After extensive washing, 100 μl of HRP-conjugated goat anti-mouse IgG (ZYMED, San Francisco, CA) was added. The HRP activities were measured using *p*-nitrophenylphosphate (Sigma) as a substrate.

Statistical analysis

Data are expressed as the mean ± standard deviation (SD). Comparisons of the numerical data between groups were performed by the paired or unpaired Mann-Whitney U-test. *P* values less than 0.05 were considered statistically significant.

Figures

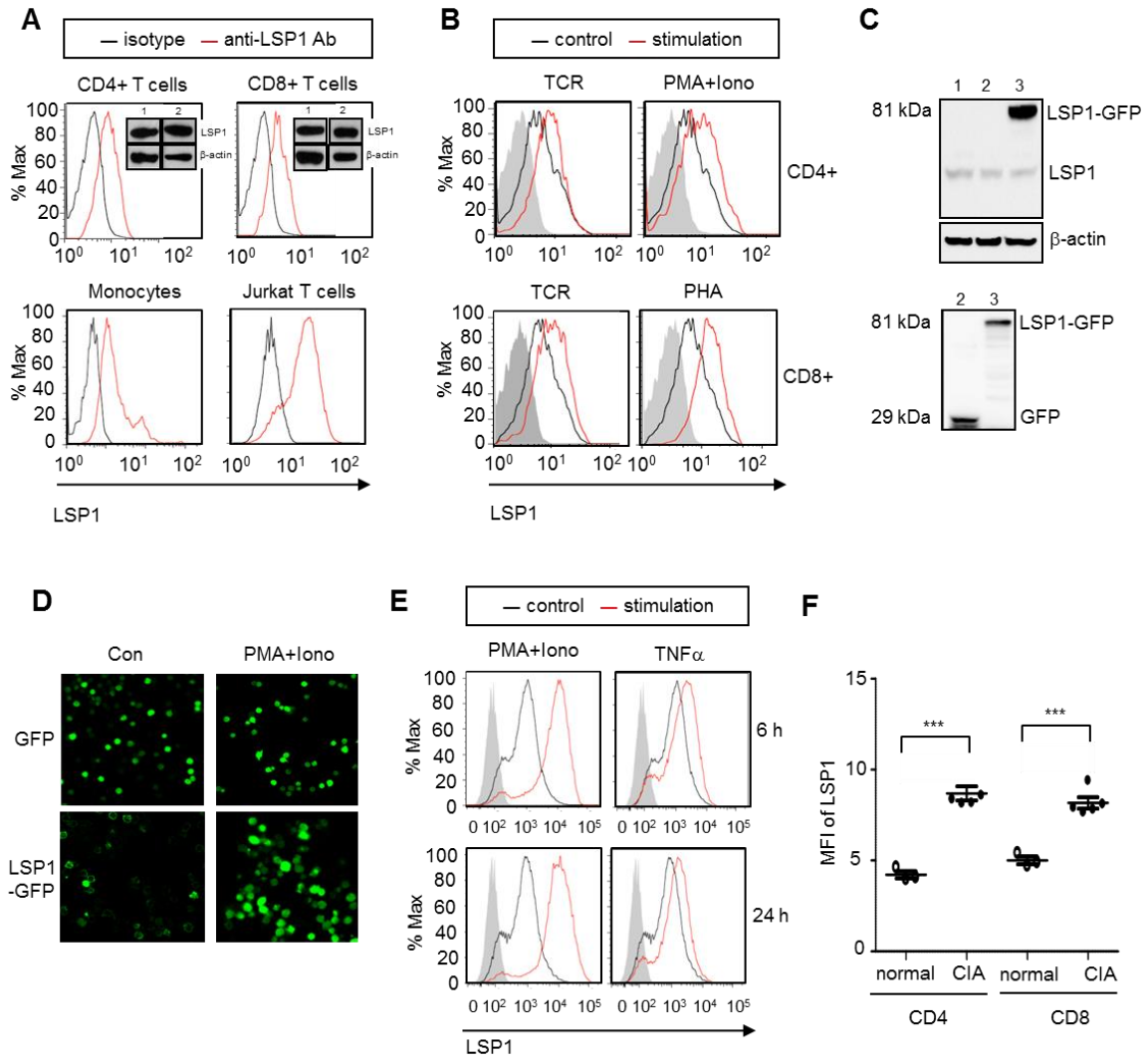


Fig. S1. Expression of LSP1 in primary human T cells and Jurkat T cells. (A) Expression of LSP1 in peripheral blood mononuclear cells (PBMC) of healthy controls as determined by flow cytometry using FITC-conjugated anti-LSP1 Ab. The protein expression levels for LSP1 (52 kD) were also determined by western blot; lane 1 and 2 indicate healthy controls 1 and 2, respectively. **(B)** Upregulation of LSP1 expression in primary T cells by various stimuli. Normal PBMC were stimulated with anti-CD3 (1 µg/ml) plus anti-CD28 (2 µg/ml) for 24 hours (TCR), PMA (50 ng/ml) plus ionomycin (Iono, 1 µg/ml) for 6 hours or PHA (5 µg/ml) for 12 hours. Cells were stained with PE-

labeled anti-CD4 Ab, APC-conjugated anti-CD8 Ab, and FITC-conjugated anti-LSP1 Ab. Data are representative of three independent experiments. Gray-colored histogram indicates the isotype control. **(C)** Increase in LSP1 expression level after the stable transfection of *LSP1-GFP* fusion gene (LSP-GFP). The LSP1-GFP fusion protein expression in Jurkat T cells was determined by western blot analysis using anti-LSP1 Ab (upper panel) or anti-GFP Ab (lower panel); lane 1: untransfected, lane 2: *GFP* only, and lane 3: *LSP1-GFP* transfection. **(D)** Induction of LSP1 expression in *LSP1-GFP*-transfected cells by PMA plus ionomycin. Cells were stimulated with PMA plus ionomycin for 4 hours, and then LSP1 expression was examined by confocal microscopy. Representative photographs are shown. **(E)** Up-regulation of LSP1 expression in *LSP1-GFP*-transfected Jurkat cells by Ca^{2+} activator or TNF-α (10 ng/ml). LSP1 expression was analyzed by flow cytometry. Gray-colored histogram indicates the isotype control. **(F)** LSP1 expression in T cells in spleen and lymph node of mice with collagen-induced arthritis (CIA) versus non-arthritic normal mice, as determined by flow cytometry. *** $P < 0.001$.

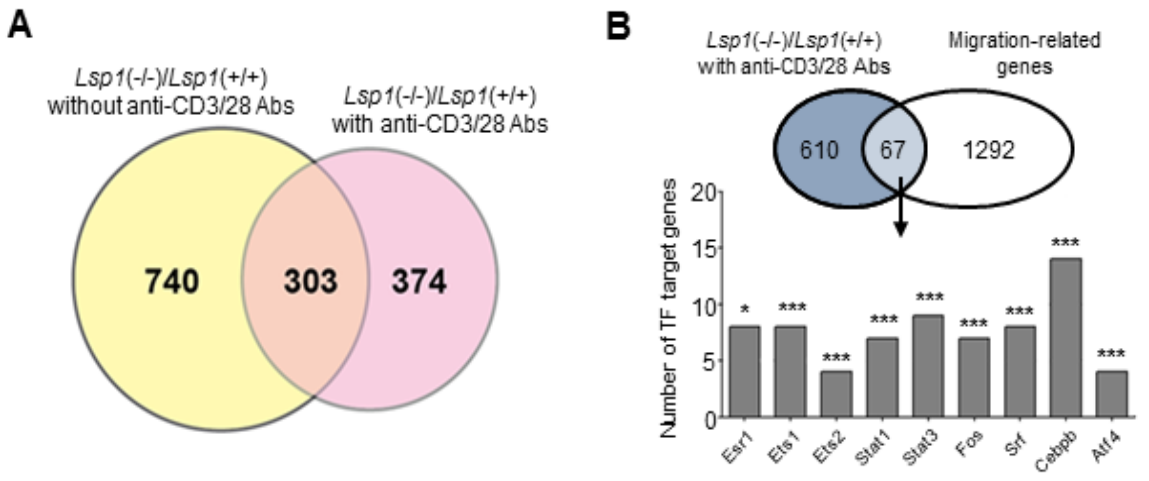


Fig. S2. (A) Venn diagram showing the relationship of the two sets of DEGs identified from the two comparisons, respectively: 1) *Lsp1* (-/-)/*Lsp1* (+/+) without anti-CD3/28 Abs (1043 DEGs) and 2) *Lsp1* (-/-)/*Lsp1*(+/+) with anti-CD3/28 Abs (677 DEGs). Of the DEGs, 303 were shared between the two sets. **(B)** ERK-downstream transcription factor (TF) related to cell migration. Venn diagram showing the relationships between migration-related genes and the DEGs in *Lsp1* (-/-) T cells, compared to *Lsp1* (+/+) T cells, in the presence of anti-CD3/28 Abs. The bar graph represents the number of DEGs regulated by ERK-downstream TFs out of 67 shared DEGs. * $P < 0.05$ and *** $P < 0.001$.

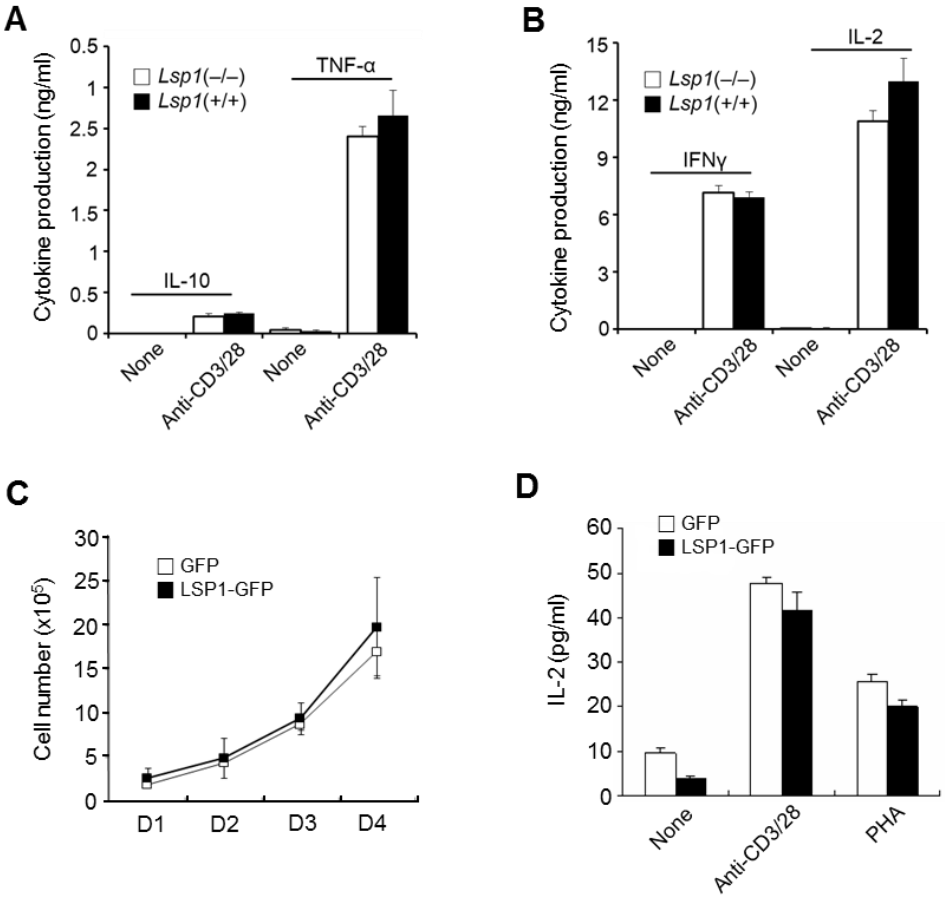


Fig. S3. (A and B) Cytokine production by splenic T cells of *Lsp1*-deficient (n=6) and wild-type mice (n=6). Splenic T cells ($2x10^5$) were sorted by anti-CD3 magnetic beads and were stimulated with anti-CD3 (1 µg/ml) plus anti-CD28 (1 µg/ml) for 24 hours. Levels of IL-10, TNF-α, IFNγ, and IL-2 were determined in the culture supernatants by ELISA. **(C)** No effect of LSP1 on the proliferation of Jurkat T cells. The number of T cells stably transfected with *LSP1-GFP* fusion gene (LSP-GFP) or *GFP* only was counted daily using trypan blue exclusion over the 4 days of culture. **(D)** IL-2 production by Jurkat T cells ($1x10^6$) transfected with *LSP1-GFP* fusion gene (LSP-GFP) or *GFP* only. The cells were stimulated with anti-CD3 (2 µg/ml) plus anti-CD28 (2 µg/ml) or PHA (PHA (5 µg/ml) for 24 hours, and IL-2 concentration was determined in the culture supernatants by ELISA.

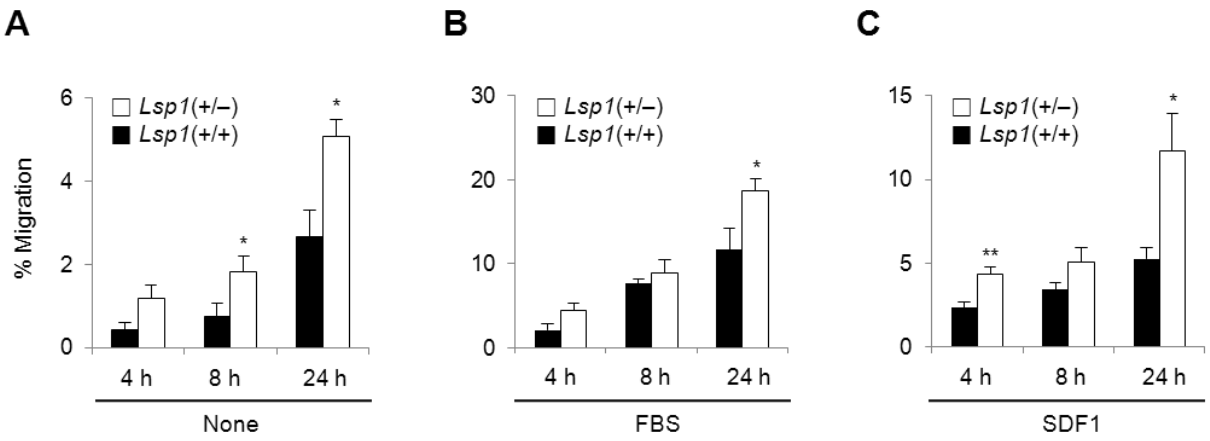


Fig. S4. Increase in T cell migration by *Lsp1*-haplo-insufficieny. In vitro migration of CD4+ T cells ($5x10^5$ cells) obtained from the spleens of *Lsp1*-haplo-insufficient (+/–) mice (n=6) and wild-type mice (n=6). Migration assays were performed in transwell chambers in the presence or absence of SDF1α (100 ng/ml) or 10% FBS. The results are mean ± SD. * $P < 0.05$, ** $P < 0.01$ versus *Lsp1* (+/+) T cells.

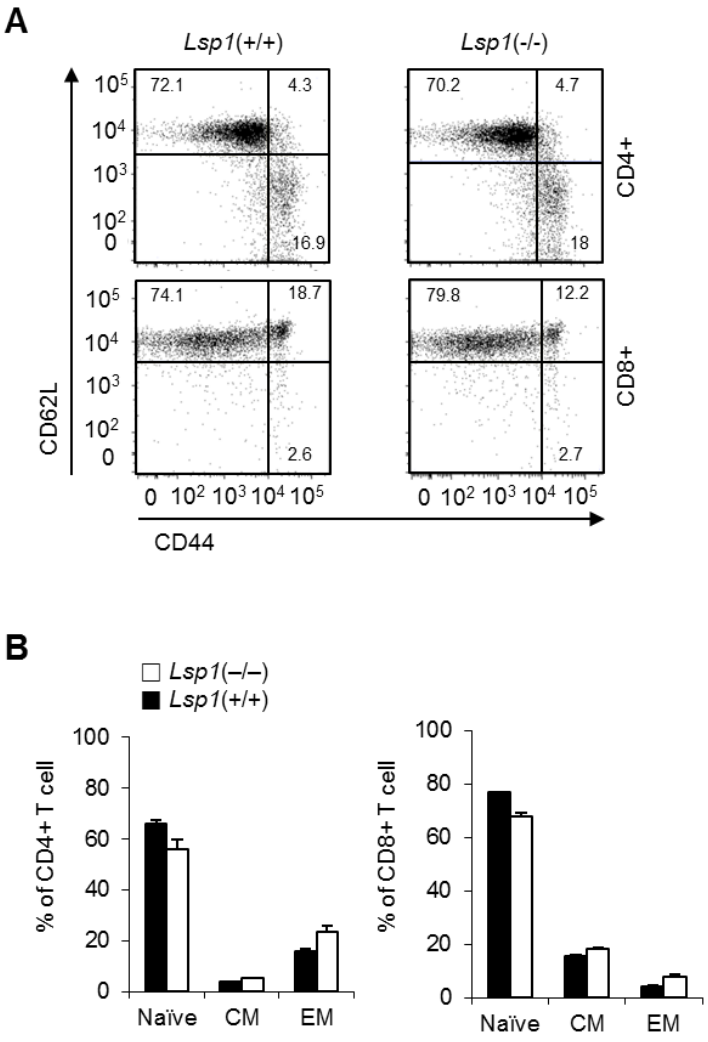


Fig. S5. T cell subpopulation in the spleen of *Lsp1*-deficient and wild-type mice. Splenocytes were isolated from *Lsp1*-deficient (n=4) and wild-type mice (n=4) and then stained with PE-labeled anti-CD44 Ab, APC-labeled anti-CD62L Ab, and FITC-conjugated anti-CD4 or anti-CD8 Ab. The percentage of naïve ($CD44^-CD62L^+$), central memory (CM, $CD44^+CD62L^+$), and effector memory T cells (EM, $CD44^+CD62L^-$) was determined by flow cytometry analysis. Representative dot plots are shown in **Fig. S5*A***. Data in **Fig. S5*B*** are mean ± SD of four mice for each group.

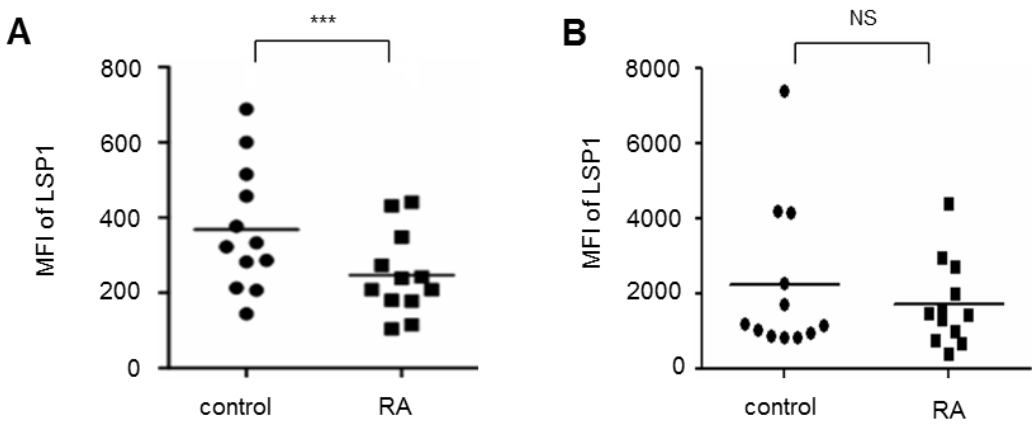


Fig. S6. LSP1 expression in total T cells (A) and monocytes (B). LSP1 expression was determined in freshly isolated peripheral blood mononuclear cells of rheumatoid arthritis (RA) patients and healthy controls by flow cytometry. *** P <0.001. NS=not significant.

Tables

Table S1. Characteristics of the CNV and study population.

Table S1-1. Characteristics of the CNVs identified in this study

Parameters	RA patients (n=100)	Healthy controls (n=400)	Total (n=500)
Total number of CNVs	5,037	26,336	31,373
Average CNVs per sample (Median)	50.4 (51)	65.8 (43)	62.7(45)
Gain	30.0 (30)	22.0 (21)	23.6 (22.5)
Loss	20.4 (20)	43.8 (20)	39.1 (20)
Average of size (kb) (Median)	66.0 (17.1)	61.7 (18.7)	62.4 (18.6)
Ratio (Loss/Gain)	0.7	2.0	1.7

Table S1-2. General characteristics of the discovery and replication cohorts

Parameter	Discovery set		Replication set (Korean)		Replication set (Caucasian)	
	RA patients (N=100)	Controls (N=400)	RA patients (N=599)	Controls (N=966)	RA patients (N=165)	Controls* (N=258)
Age	54±11	49±8	54±12	41±13	63±9	47±15
Gender						
Male	24(24%)	202(51%)	178(30%)	35(4%)	60(36%)	74(52%)
Female	76(76%)	198(49%)	421(70%)	931(96%)	105(64%)	69(48%)

*115 of the 258 control individuals were not available for age and sex information

Table S2. Copy number variation regions (CNVRs) associated with RA risk as identified in the discovery set.

CNVR Location	Start (kb)†	End (kb)†	Length (kb)	Type	Genes	aCGH Discovery (100 RA versus 400 controls)		
						P	FDR‡	OR (95% CI)
4q35.2	189998	190006	9.1	Gain	-	$4.7x10^{-6}$	$9.2x10^{-4}$	5.49 (2.65-11.37)
4q13.2	70168	70264	96.0	Gain/Loss	*UGT2B28*	$8.7x10^{-6}$	$1.6x10^{-3}$	4.70 (2.38-9.29)
9p21.1	31117	31146	29.5	Gain	-	$1.8x10^{-5}$	$3.2x10^{-3}$	118.6 (13.39-1049.53)
5q21.1	100608	100618	10.1	Gain	-	$2.5x10^{-5}$	$4.4x10^{-3}$	41.36 (7.31-233.95)
11p15.5	1857	1901	43.4	Loss	*LSP1*, *TNNT3*	$4.6x10^{-5}$	$7.6x10^{-3}$	5.10 (2.32-11.06)
1q21.1	147377	147383	5.3	Gain	-	$6.38x10^{-5}$	$9.9x10^{-3}$	5.12 (2.29-11.41)
1q31.3	194664	194697	33.0	Gain/Loss	-	$6.53x10^{-5}$	$9.9x10^{-3}$	16.41 (4.16-64.81)

Logistic regression analysis was conducted to adjust for the age and sex. 95% CI = 95% confidence interval.

† hg18

‡ False discovery rate (FDR) was calculated with the *P* values for all 3936 CNV regions.

Table S3. Eight clusters of the DEGs based on their up-, and down-regulation patterns. Eight clusters include all possible combination of up- and down-regulation of DEGs in the two comparisons: 1) *Lsp1* (–/–) versus *Lsp1* (+/+) T cells without anti-CD3/CD28 Abs and 2) *Lsp1* (–/–) versus *Lsp1* (+/+) T cells with anti-CD3/CD28 Abs. Colors represent up- (U; red) and down-regulation (D; green) of DEGs in the corresponding comparisons. The count indicates the number of DEGs in each cluster.

Cluster	*Lsp1* (–/–)/*Lsp1* (+/+) without anti-CD3/CD28 Abs	*Lsp1* (–/–)/*Lsp1* (+/+) with anti-CD3/CD28 Abs	Count
1	U	-	327
2	U	U	196
3	-	U	164
4	D	U	20
5	D	-	413
6	D	D	86
7	-	D	210
8	U	D	1

Table S4. GOBP and KEGG terms enriched by the DEG in *Lsp1*-deficient T cells. GOBPs and KEGG pathways significantly represented by the DEGs in *Lsp1*-deficient T cells, compared to wild-type T cells, in the absence and presence of anti-CD3/28 Abs for 6 hours. Data are presented as Z scores, significance measures of $N^{-1}(1-P)$ in which P is p-value computed by DAVID and $N^{-1}(\cdot)$ is the inverse standard normal distribution (22). Blue colored boxes indicate the Z scores with statistical significance.

Terms	media		anti-CD3/28	
	up	down	up	down
cytokine-cytokine receptor interaction	1.767	1.869	3.289	1.291
immune response	7.039	0.813	5.199	3.104
response to wounding	2.153	-1.609	3.568	0.129
inflammatory response	2.497	-1.335	4.196	-0.127
innate immune response	3.140	-0.162	3.203	-3.303
leukocyte migration	2.116	-2.571	3.694	-3.303
cell migration	0.471	-2.208	3.350	-1.688
cell motility	0.439	-1.174	3.866	-1.946
chemotaxis	1.230	0.523	3.615	-3.303
calcium ion homeostasis	0.309	-0.899	3.427	1.178
cell adhesion	1.255	-0.758	1.673	0.937
phagocytosis	1.256	-2.571	2.276	0.088
proteolysis	2.356	0.110	0.962	-0.427
glycolysis / gluconeogenesis	1.359	-2.571	-5.326	-3.303
RIG-I-like receptor signaling pathway	1.359	0.595	0.361	-3.303
chemokine signaling pathway	0.072	1.579	-0.412	-3.303
regulation of transcription	-1.102	2.112	-1.508	-2.093
natural killer cell mediated cytotoxicity	-0.218	2.321	-1.280	2.578
Jak-STAT signaling pathway	-1.256	0.847	-0.763	1.666
regulation of apoptosis	0.241	-0.363	0.234	2.614
Fc epsilon RI signaling pathway	-0.272	-0.574	-0.791	1.332
complement and coagulation cascades	1.781	-0.480	0.992	1.445

Table S5. DEGs related to cell migration in addition to regulated by ERK-downstream TFs in TCR-activated T cells

Entrez ID	Symbol	Description
Up-regulated		
16176	Il1β	interleukin 1 beta
12978	Csf1r	colony stimulating factor 1 receptor
19225	Ptgs2	prostaglandin-endoperoxide synthase 2
20296	Ccl2	chemokine (C-C motif) ligand 2
24047	Ccl19	chemokine (C-C motif) ligand 19
17329	Cxcl9	chemokine (C-X-C motif) ligand 9
Down-regulated		
16985	LSP1	lymphocyte specific 1
20682	Sox9	SRY-box containing gene 9
16189	Il4	interleukin 4
14634	Gli3	GLI-Kruppel family member GLI3

Table S6. Sequences of primers used in quantitative real-time PCR.

Gene name	Forward primer sequence	Reverse primer sequence
hs-Gapdh	5'-TGTGGGCATCAATGGATTTGG-3'	5'-ACACCATGTATTCCGGGTCAAT-3'
hs-Il-1β	5'-ATGATGGCTTATTACAGTGGCAA-3'	5'-GTCGGAGATTCGTAGCTGGA-3'
hs-Csf1r	5'-GGGAATCCCAGTGATAGAGCC-3'	5'-TTGGAAGGTAGCGTTGTTGGT-3'
hs-Ptgs2	5'-TAAGTGCGATTGTACCCGGAC-3'	5'-TTTGTAGCCATAGTCAGCATTGT-3'
hs-Ccl2	5'-CATCTCCTACACCCCACGAAG-3'	5'-GGGTTGGCACAGAAACGTC-3'
hs-Ccl19	5'-TACATCGTGAGGAACTTCCACT-3'	5'-CTGGATGATGCGTTCTACCCA-3'
hs-Cxcl9	5'-CCAGTAGTGAGAAAGGGTCGC-3'	5'-AGGGCTTGGGGCAAATTGTT-3'
hs-LSP1	5'-GGAGCACCAGAAATGTCAGCA-3'	5'-TCGGTCCTGTCGATGAGTTTG-3'
hs-Sox9	5'-AGCGAACGCACATCAAGAC-3'	5'-CTGTAGGCGATCTGTTGGGG-3'
hs-Il-4	5'-GCCAAGACCCCTTCGAGAAAT-3'	5'-CCGATCCTGTTATCTGCCTCC-3'
hs-Gli3	5'-GAAGTGCTCCACTCGAACAGA-3'	5'-GTGGCTGCATAGTGATTGCG-3'
mu-Gapdh	5'-TGACGTGCCGCCTGGAGAAA-3'	5'-AGTGTAGCCCAAGATGCCCTTCAG-3'
mu-Il-1β	5'-GTGGCTGTGGAGAAGCTGTG-3'	5'-GAAGGTCCACGGGAAAGACAC-3'
mu-Csf1r	5'-CATGGCCTTCCTTGCTTCTAAA-3'	5'-CAGCACGTTTCGAGCTGCTA-3'
mu-Ptgs2	5'-*CCCCCACAGTCAAAGACACT*-3'	5'-GGTTCTCAGGGATGTGAGGA-3'
mu-Ccl2	5'-GTTGGCTCAGCCAGATGCA-3'	5'-AGCCTACTCATTGGGATCATCTTG-3'
mu-Ccl19	5'-GACCTTCCCAGCCCCAACT-3'	5'-CGGAAGGCTTTCACGATGTT-3'
mu-Cxcl9	5'-*TCTGCCATGAAGTCCGCTG* -3'	5'-*CAGGAGCATCGTGCATTCCT*-3'
mu-Sox9	5'-GACAAGCGGAGGCCGAA-3'	5'-CCAGCTTGCACGTCGGTT--3'
mu-Il-4	5'-CCTCAC AGCAACGAAGAACA -3'	5'-TGGACTCATTCATGGTGCAG-3'
mu-Gli3	5'-GCTCCAACATTTCCAACAC-3'	5'-TGTGGGCTTGCTCTGTGAGG-3'

Dataset

Dataset S1. DEGs between *Lsp1* (-/-) and *Lsp1* (+/+) T cells in the presence and absence of anti-CD3/28 Abs. For each DEG, the $\log_2$-fold-change and the overall *P* value computed from the integrative statistical method are shown in the two comparisons: 1) *Lsp1* (-/-)/*Lsp1* (+/+) without anti-CD3/28 Abs and 2) *Lsp1* (-/-)/*Lsp1* (+/+) with anti-CD3/28 Abs. Also, the probe name in Agilent microarray, gene symbol, Entrez ID, and gene description are included.

References

1. Arnett FC, *et al.* (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 31(3):315-324.
2. Kim JH, *et al.* (2012) CNVRuler: a copy number variation-based case-control association analysis tool. *Bioinformatics* 28(13):1790-1792.
3. Kim JH, *et al.* (2013) Deletion variants of RABGAP1L, 10q21.3, and C4 are associated with the risk of systemic lupus erythematosus in Korean women. *Arthritis Rheum* 65(4):1055-1063.
4. Kee AJ & Hardeman EC (2008) Tropomyosins in skeletal muscle diseases. *Adv Exp Med Biol* 644:143-157.
5. Jung SH, *et al.* (2014) Genome-wide copy number variation analysis identifies deletion variants associated with ankylosing spondylitis. *Arthritis Rheumatol* 66(8):2103-2112.
6. Yim SH, *et al.* (2011) The potential role of VPREB1 gene copy number variation in susceptibility to rheumatoid arthritis. *Mol Immunol* 48(11):1338-1343.
7. Bolstad BM, Irizarry RA, Astrand M, & Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19(2):185-193.
8. Huang da W, Sherman BT, & Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1):44-57.
9. Bowman AW & Azzalini A (1997) Applied Smoothing Techniques for Data

Analysis: The Kernel Approach with S-Plus Illustrations (*Oxford University Press*, New York,).

10. Hwang D, *et al.* (2005) A data integration methodology for systems biology. *Proc Natl Acad Sci USA* 102(48):17296-17301.
11. Carbon S, *et al.* (2009) AmiGO: online access to ontology and annotation data. *Bioinformatics* 25(2):288-289.
12. Kanehisa M, Goto S, Sato Y, Furumichi M, & Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40 (Database issue):D109-114.
13. Mendoza MC, Er EE, & Blenis J (2011) The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem Sci* 36(6):320-328.
14. Chang F, *et al.* (2003) Signal transduction mediated by the Ras/Raf/MEK/ERK pathway from cytokine receptors to transcription factors: potential targeting for therapeutic intervention. *Leukemia* 17(7):1263-1293.
15. Chung J, Uchida E, Grammer TC, & Blenis J (1997) STAT3 serine phosphorylation by ERK-dependent and -independent pathways negatively modulates its tyrosine phosphorylation. *Mol Cell Biol* 17(11):6508-6516.
16. Kivens WJ, *et al.* (1998) Identification of a proline-rich sequence in the CD2 cytoplasmic domain critical for regulation of integrin-mediated adhesion and activation of phosphoinositide 3-kinase. *Mol Cell Biol* 18(9):5291-5307.
17. Jongstra-Bilen J, *et al.* (2000) LSP1 modulates leukocyte populations in resting and inflamed peritoneum. *Blood* 96(5):1827-1835.
18. Allen IC (2013) Delayed-type hypersensitivity models in mice. *Methods Mol Biol* 1031:101-107.

19. van den Berg WB, Joosten LA, & van Lent PL (2007) Murine antigen-induced arthritis. *Methods Mol Med* 136:243-253.

20. Trentham DE, Townes AS, & Kang AH (1977) Autoimmunity to type II collagen an experimental model of arthritis. *J Exp Med* 146(3):857-868.

21. Kong JS, et al. (2010) Inhibition of synovial hyperplasia, rheumatoid T cell activation, and experimental arthritis in mice by sulforaphane, a naturally occurring isothiocyanate. *Arthritis Rheum* 62(1):159-170.

22. Ekins S, Nikolsky Y, Bugrim A, Kirillov E, & Nikolskaya T (2007) Pathway mapping tools for analysis of high content data. *Methods Mol Biol* 356:319-350.