Multiplex matrix network analysis of protein complexes in the human TCR signalosome

Stephen E. P. Smith, Steven C. Neier, Brendan K. Reed, Tessa R. Davis, Jason P. Sinnwell, Jeanette E. Eckel-Passow, Gabriel F. Sciallis, Carilyn N. Wieland, Rochelle R. Torgerson, Diana Gil, Claudia Neuhauser and Adam G. Schrum (August 2, 2016)

Science Signaling **9** (439), rs7. [doi: 10.1126/scisignal.aad7279]

Science Signaling (ISSN 1937-9145) is published weekly, except the last December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005.

Science Signaling
AAAS

www.sciencesignaling.org/cgi/content/full/9/439/rs7/DC1

Supplementary Materials for

Multiplex matrix network analysis of protein complexes in the human TCR signalosome

Stephen E. P. Smith, Steven C. Neier, Brendan K. Reed, Tessa R. Davis, Jason P. Sinnwell, Jeanette E. Eckel-Passow, Gabriel F. Sciallis, Carilyn N. Wieland, Rochelle R. Torgerson, Diana Gil, Claudia Neuhauser,* Adam G. Schrum*

*Corresponding author. Email: schrum.adam@mayo.edu (A.G.S.); neuha001@umn.edu (C.N.)

Published 2 August 2016, *Sci. Signal.* **9**, rs7 (2016)
DOI: 10.1126/scisignal.aad7279

This PDF file includes:

Equations

Equation S1. Linear regression model.

$$\log_2(y_{ijk} + 0.5) = \mu + \alpha_i + \epsilon_{ijk}$$

Variable	Description
i	# of stimulation conditions
j	# of intra-experimental replicates for stimulation condition i
k	# of fluorescent bead measurements in replicate j for condition i
y_{ijk}	PE fluorescence intensity measurement for bead k in replicate j of condition i
μ	Overall mean
α_i	Treatment effect for stimulation condition i
ϵ_{ijk}	Error in the model

Equation S2. ANC α-cutoff calculation. The α-cutoff required to determine statistical significance per experiment (α_κ) was calculated to maintain an overall type I error of 0.05 adjusted for multiple hypothesis testing with Bonferroni correction. The α-cutoffs were calculated as follows:

i) Let $\varkappa = 0.7$

ii) Define $C = \left\{C_j = (C_{jS+}, C_{jNS}, C_{jS-}) \in \mathbb{N}_0^3 \quad \text{such that} \quad C_{jS+} + C_{jNS} + C_{jS-} = N_{exp}\right\}$

iii) For every $\kappa \in \{1, 2, \ldots, N_{exp}\}$ that satisfies $\frac{\kappa}{N_{exp}} \geq \varkappa$, we find all ordered sets $C_j \in C$ with,

$$C_{jS^+} = \kappa \quad \text{or} \quad C_{jS^-} = \kappa \qquad (\star)$$

iv) Let $J_\kappa^\star = \{j \text{ such that } C_j \text{ satisfies } (\star)\}$

v) Solve for α_κ, where

$$\frac{\alpha_{bc}}{N_\kappa} = \sum_{j \in J_\kappa^\star} \left(\frac{\alpha_\kappa}{2}\right)^{C_{jS^+}} * (1-\alpha_\kappa)^{C_{jNS}} * \left(\frac{\alpha_\kappa}{2}\right)^{C_{jS^-}} * \frac{N_{exp}\,!}{C_{jS^+}\,! * C_{jNS}\,! * C_{jS^-}\,!}$$

Variable	Description
$\varkappa$	Consistency cutoff value (set at 70% for this dataset)
α_{bc}	0.05 corrected for multiple hypotheses using the Bonferroni correction method
κ	# of experiments a PiSCES is consistent (defined as statistically significant with fold change in the same direction)
N_{exp}	# of inter-experimental replicates
C_{S^+}	# of experiments a PiSCES has Log2 fold change ≥ 0 & is statistically significant
C_{S^-}	# of experiments a PiSCES has Log2 fold change < 0 & is statistically significant
C_{NS}	# of experiments a PiSCES is not significantly changed
C	The set of ordered tuples, $C_j = (C_{jS+}, C_{jNS}, C_{jS-}) \in \mathbb{N}_0^3$ such that $C_{jS+} + C_{jNS} + C_{jS-} = N_{exp}$
J	$J = \{1, 2, \ldots, \lvert C\rvert\}$, where $\lvert C\rvert$ = size of C
j	An index used to represent a given combination in the set C, where $C_j = \{C_{jS^+}, C_{jNS}, C_{jS^-}\}$ indicates the jth combination
$J_\kappa^\star$	Values j such that C_j satisfy $(\star)$, i.e. $C_{jS^+} = \kappa$ or $C_{jS^-} = \kappa$
N_κ	# of $\kappa \in \{1, 2, \ldots, N_{exp}\}$ that satisfies $\frac{\kappa}{N_{exp}} \geq \varkappa$
α_κ	α-cutoff when $C_{jS^+} = \kappa$ or $C_{jS^-} = \kappa$, which is then empirically corrected for technical error using intra-assay duplicates (see Methods)

Equation S3. Adaptive ANC α-cutoff calculation example for four experiments and 70% consistency.

i) Let $\varkappa = 0.7$

ii) Let $N_{exp} = 4$, and so $C = \{(4,0,0),(0,0,4),(3,1,0),(3,0,1),(0,1,3),(1,0,3),(2,2,0),(2,1,1),(2,0,2),(0,2,2),(1,1,2),(1,3,0),(1,2,1),(0,3,1)\}$

iii) There are 2 values $\kappa \in \{1,2,3,4\}$, which satisfy $\frac{\kappa}{N_{exp}} \geq \varkappa$. Thus $N_\kappa = 2$.

iv) First select $\kappa_4 = 4$

v) Let $j \in \{1,2\}$ where $C_{j=1} = (4,0,0)$ and $C_{j=2} = (0,0,4)$ which satisfy either $C_{jS^+} = \kappa_4 = 4$ or $C_{jS^-} = \kappa_4 = 4$

vi) Solve for α_{κ_4}, where

$$\frac{\alpha_{bc}}{N_\kappa} = \sum_{j\in\{1,2\}} \left(\frac{\alpha_{\kappa_4}}{2}\right)^{C_{jS^+}} * (1-\alpha_{\kappa_4})^{C_{jNS}} * \left(\frac{\alpha_{\kappa_4}}{2}\right)^{C_{jS^-}} * \frac{N_{exp}\,!}{C_{jS^+}\,! * C_{jNS}\,! * C_{jS^-}\,!}$$

which simplifies to,

$$\frac{\alpha_{bc}}{2} = \left[\left(\frac{\alpha_{\kappa_4}}{2}\right)^4 * 1\right] * 2$$

vii) Next select $\kappa_3 = 3$, which also satisfies $\frac{\kappa}{N_{exp}} \geq \varkappa$

viii) Let $j \in \{3,4,5,6\}$ where $C_{j=3} = (3,1,0)$, $C_{j=4} = (3,0,1)$, $C_{j=5} = (0,1,3)$, and $C_{j=6} = (1,0,3)$ which satisfy either $C_{jS^+} = \kappa_3 = 3$ or $C_{jS^-} = \kappa_3 = 3$

ix) Solve for α_{κ_3}, where

$$\frac{\alpha_{bc}}{N_\kappa} = \sum_{j\in\{3,4,5,6\}} \left(\frac{\alpha_{\kappa_3}}{2}\right)^{C_{jS^+}} * (1-\alpha_{\kappa_3})^{C_{jNS}} * \left(\frac{\alpha_{\kappa_3}}{2}\right)^{C_{jS^-}} * \frac{N_{exp}\,!}{C_{jS^+}\,! * C_{jNS}\,! * C_{jS^-}\,!}$$

which simplifies to,

$$\frac{\alpha_{bc}}{2} = \left[\left(\frac{\alpha_{\kappa_3}}{2}\right)^3 * (1-\alpha_{\kappa_3}) * 4 + \left(\frac{\alpha_{\kappa_3}}{2}\right)^3 * \frac{\alpha_{\kappa_3}}{2} * 4\right] * 2$$

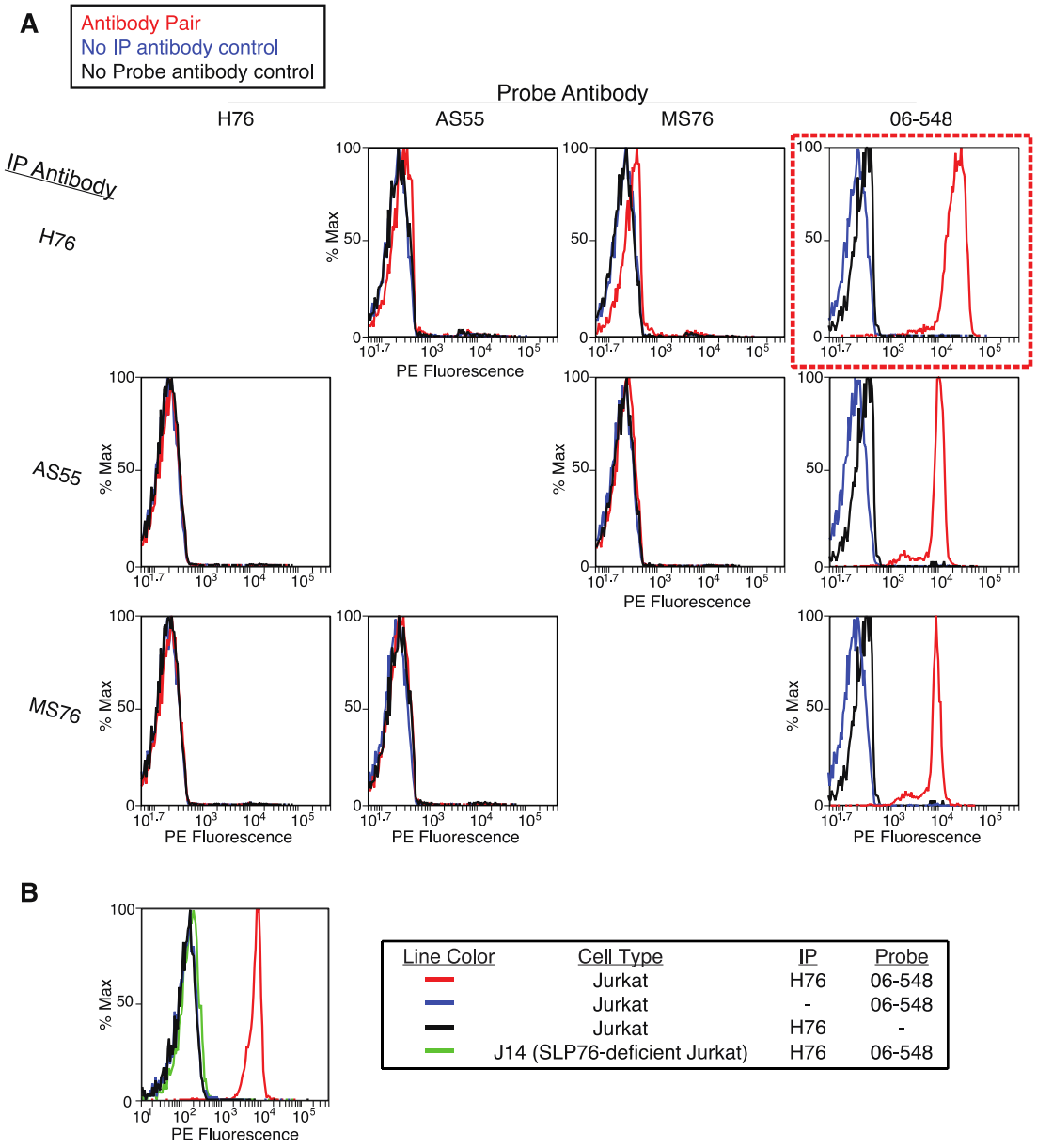

Fig. S1. Screening strategy for multiplex panel antibodies, using SLP-76 as an example. (**A**) For each of 20 target proteins, we first screened three to five antibodies for their ability to specifically bind to the target in postnuclear lysates prepared from Jurkat cells. For SLP-76, four antibodies were screened as immunoprecipitation (IP) or detection (probe) reagents. Capture antibodies were exclusively monoclonal, whereas polyclonal antibodies were sometimes included as probes. The red dashed box indicates the capture-probe antibody combination that was selected for further use because of its high signal-to-noise ratio. (**B**) The selected antibody pair was then screened for specificity with a cell line known to lack the target protein. When possible, mutant Jurkat cell lines in which the target protein was specifically deleted were used, as shown here for the SLP-76–deficient mutant Jurkat cell line J14. If the MFI was reduced by >90% in the mutant cell line compared to that in the wild-type Jurkat cells, the antibody pair was considered validated. For a complete list of validated antibody pairs, see table S1. Primary data for the other target proteins are available upon request.

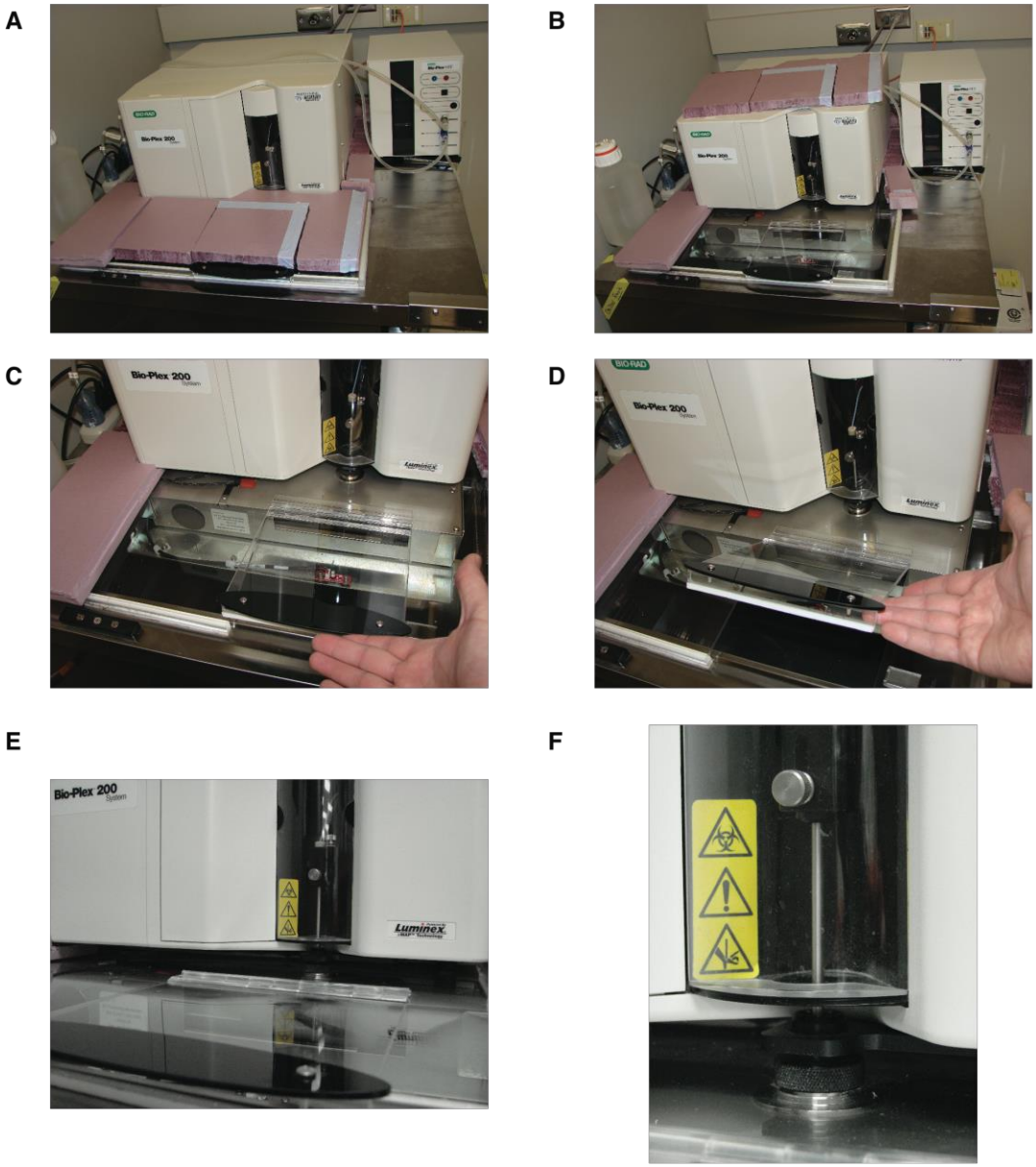

Fig. S2. Instrument setup for optimal analysis of protein complexes. Preservation of protein complexes during processing for analysis requires that they be kept cold. Our solution to refrigerating samples on plates during data acquisition on the Bioplex 200 instrument is presented here. The lower plate carrier unit of the instrument was placed in a commercial sandwich-prep refrigerator (Norlake Inc.). The upper flow cytometer portion remained at room temperature, but resided on an insulated acrylic divider placed on top of the refrigerator, and the sip-needle of the flow cytometer accessed samples from the refrigerated plate carrier through a bored hole in the acrylic. (**A**) Top view of the instrument setup. Purple insulation is placed on top of the acrylic divider that covers the refrigerator, which serves to eliminate potential water condensation problems at the temperature differential interface. (**B**) Removing the insulation panels reveals the acrylic divider upon which the flow cytometer resides. Through the acrylic, the lower plate carrier portion of the Bioplex 200 can be seen. (**C** and **D**) Built into the acrylic divider is a flip-lid access point, which is lifted to insert a 96-well plate. (**E** and **F**) Close-up views of the sip needle that accesses samples through a bored hole in the acrylic.

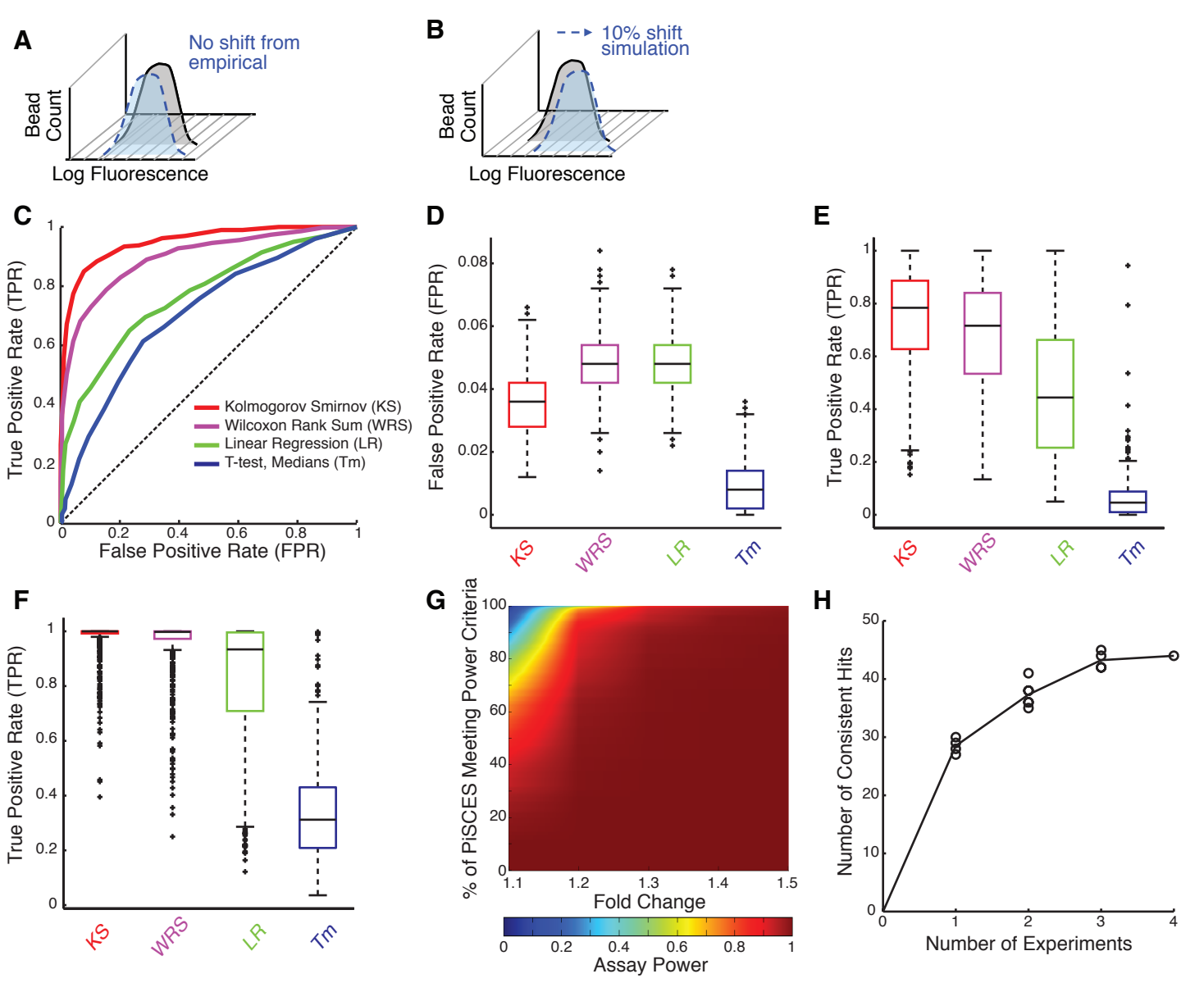
A
No shift from empirical
Bead Count
Log Fluorescence
B
10% shift simulation
Bead Count
Log Fluorescence
C
True Positive Rate (TPR)
False Positive Rate (FPR)
Kolmogorov Smirnov (KS)
Wilcoxon Rank Sum (WRS)
Linear Regression (LR)
T-test, Medians (Tm)
D
False Positive Rate (FPR)
KS
WRS
LR
Tm
E
True Positive Rate (TPR)
KS
WRS
LR
Tm
F
True Positive Rate (TPR)
KS
WRS
LR
Tm
G
% of PiSCES Meeting Power Criteria
Fold Change
Assay Power
H
Number of Consistent Hits
Number of Experiments

Fig. S3. Development and evaluation of ANC analysis. (**A**) FPRs for various statistical tests were evaluated by comparing bootstrapping-inspired resamplings that did not impose any shift in the distribution from that of the original empirical data. (**B**) TPRs were evaluated similarly, but resamplings were performed with imposed distribution shifts of at least 10% to assess detection of a known shift in the data. (**C**) Given a 10% shift, FPR versus TPR was visualized for the different statistical tests, performed with many different type I and II error thresholds (type I error risk, FPR, x-axis; type II error risk, β, where 1 - β = TPR, y-axis). Colors indicate the different statistical methods tested, whereas the dotted line represents the expectation for random guessing. (**D**) Boxplot of the FPR for each protein pair measured with each of the statistical tests using an α-cutoff = 0.05. (**E** and **F**) Boxplot of the TPR for each protein pair measured with each of the statistical tests using an α-cutoff = 0.05 for empirical shifts of 10% (E) or 20% (F). As the magnitude of the shift increased, all tests increased their respective TPR. The lower TPR for Tm suggests that the use of summary MFI values reduced the ability to detect shifts. (**G**) Power calculation for consistently significant hits across four experiments (using the α-cutoff applied to the SEE stimulation experiment). For a consistent shift of 1.2-fold, we predict there is at least 90% power in 90% of the PiSCES measurements. (**H**) Using experimental data from the SEE stimulation data set, the number of consistent hits observed in all experiments was plotted with an increasing number of inter-experimental replicates. Dots indicate each combination for the given number of experiments, and the line intersects through the mean number of consistent hits. Because the number of hits using ANC analysis plateaued after three or four inter-experimental replicates, we adopted the practice of performing three or more experiments.

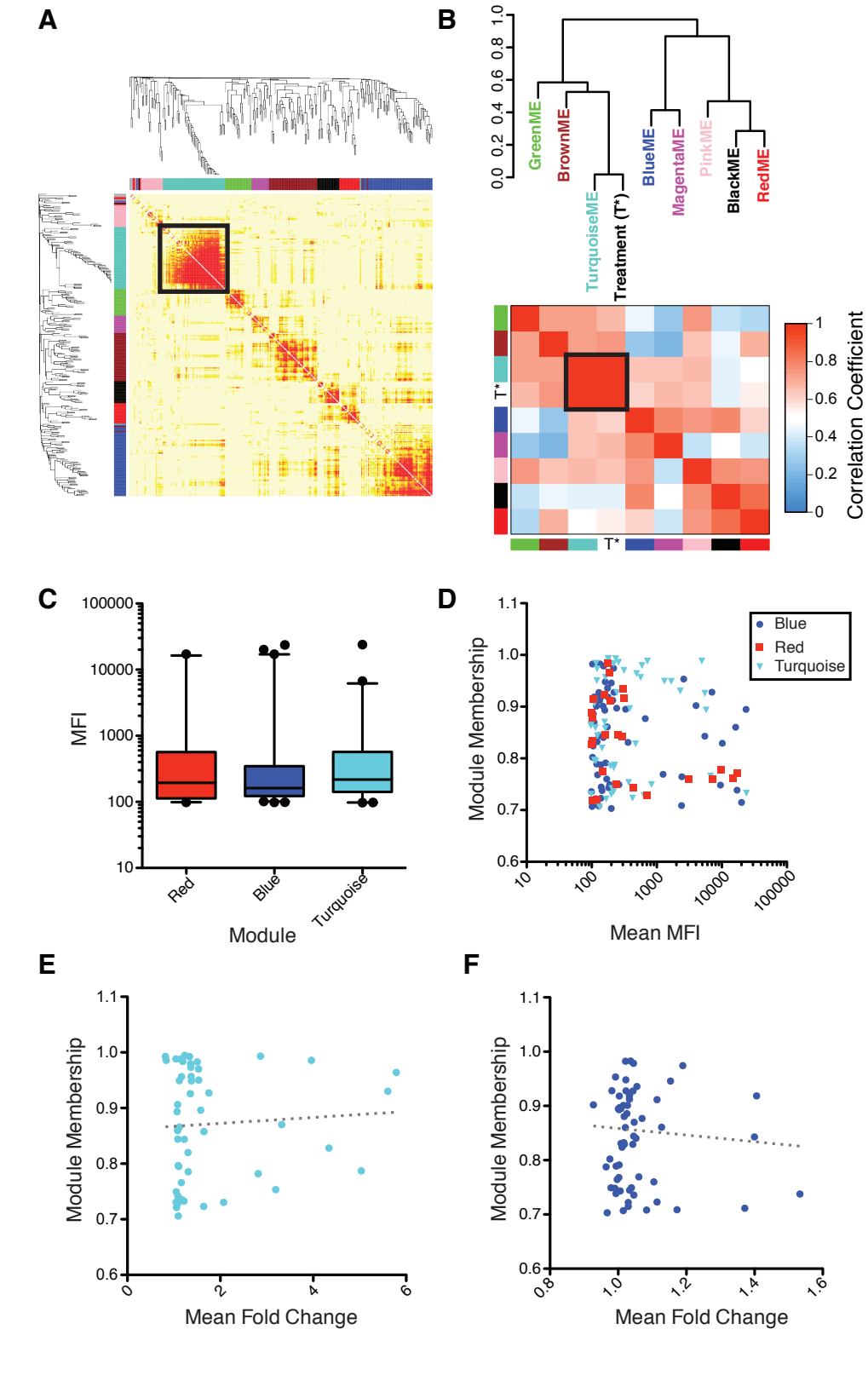
A
B
0.0 0.2 0.4 0.6 0.8 1.0
GreenME
BrownME
TurquoiseME
Treatment (T*)
BlueME
MagentaME
PinkME
BlackME
RedME
T*
T*
1
0.8
0.6
0.4
0.2
0
Correlation Coefficient
C
100000
10000
1000
100
10
MFI
Red
Blue
Turquoise
Module
D
Blue
Red
Turquoise
Module Membership
1.1
1.0
0.9
0.8
0.7
0.6
10
100
1000
10000
100000
Mean MFI
E
Module Membership
1.1
1.0
0.9
0.8
0.7
0.6
0
2
4
6
Mean Fold Change
F
Module Membership
1.1
1.0
0.9
0.8
0.7
0.6
0.8
1.0
1.2
1.4
1.6
Mean Fold Change

Fig. S4. Application and evaluation of WCNA analysis for the SEE-stimulated and unstimulated Jurkat cell data. (**A**) The WGCNA program for R generated a topological overlap matrix (TOM) to visualize the weighted correlation network, which depicts the correlation (dark colors indicate high correlation) between each PiSCES (each represented as a row and column). The dendrogram represents the results of network clustering, in which the differently colored blocks represent groups of correlated PiSCES, named modules (from four independent experiments). (**B**) Each module was assigned an "eigenvector" value that summarized the overall behavior of that module. The relationship between the eigenvector and treatment (T*, unstimulated versus SEE-stimulated) is shown as a hierarchical clustering dendrogram of the eigenvectors (top) or as a heat map representing eigenvector adjacency (bottom). The "turquoise" module and SEE treatment status are statistically significantly correlated ($P < 0.00001$). The turquoise module and treatment are outlined by a black box. (**C**) The box-and-whiskers plot shows that MFI was similar in selected modules (ANOVA: $F_{2,133}=0.64$, $P = 0.52$), indicating that MFI did not define module membership. (**D**) Within modules, MFI did not correlate with module membership (Blue: $r^2 = 0.01$, NS; Red: $r^2 = 0.12$, NS; Turquoise: $r^2 = 0.02$, NS). Each point represents one PiSCES, whereas colors indicate modules. (**E** and **F**) Fold-change in response to stimulation with SEE did not correlate with module membership, even in the turquoise module that was associated with stimulation (Blue: $r^2 = 0.01$, NS; Turquoise: $r^2 = 0.01$, NS). NS, not significant. Data are from four independent SEE treatment experiments.

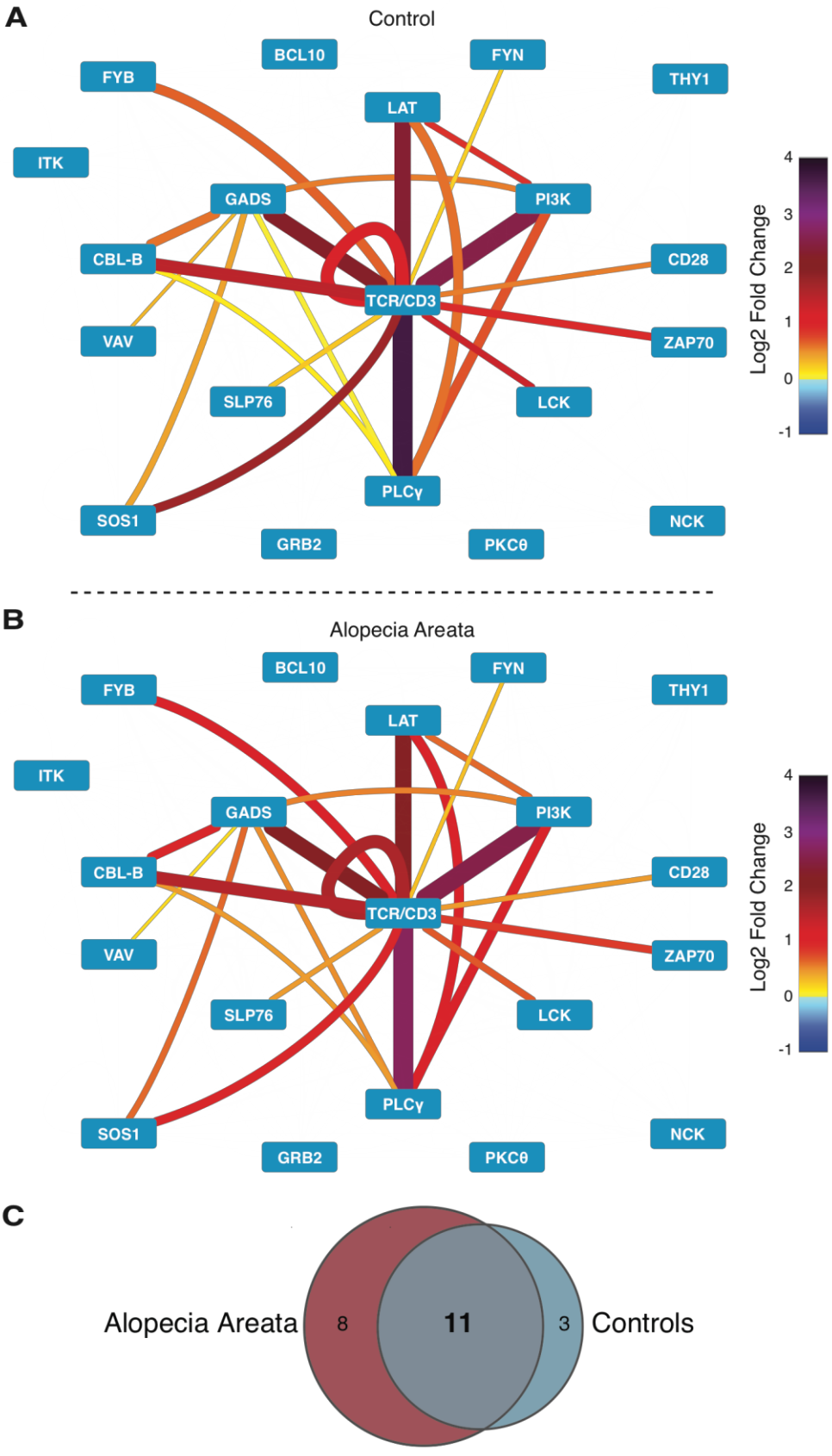
A
Control
BCL10
FYN
FYB
THY1
LAT
ITK
GADS
PI3K
CBL-B
CD28
TCR/CD3
VAV
ZAP70
SLP76
LCK
PLCγ
SOS1
NCK
GRB2
PKCθ
Log2 Fold Change
4
3
2
1
0
-1
B
Alopecia Areata
BCL10
FYN
FYB
THY1
LAT
ITK
GADS
PI3K
CBL-B
CD28
TCR/CD3
VAV
ZAP70
SLP76
LCK
PLCγ
SOS1
NCK
GRB2
PKCθ
Log2 Fold Change
4
3
2
1
0
-1
C
Alopecia Areata
8
11
3
Controls

Fig. S5. Stimulation-induced PiSCES network is similar between control and alopecia areata patient groups. (**A** and **B**) Visualization of the stimulation-induced PiSCES signature consistent in ANC and WCNA (ANC ∩ WCNA) for hits identified as statistically significant in either experimental group. Edge color and thickness correspond to mean $\log_2$ fold-change (color legend at right) for stimulation-responsive PiSCES in control (A) or alopecia areata patients (B). (**C**) Venn diagram of hits that were statistically significantly different in controls, alopecia areata patients, or both (n = 7 alopecia areata patients; n = 5 for control patients).

Table S1. Validated antibody pairs used to identify each target in Jurkat cells. Negative control cells for specificity are listed in the right-most column. Where possible, target-negative Jurkat mutant cell lines were used. When these were not available, GeneAtlas RNA expression profiles were used to select a cell type that lacked the protein, and these were used as controls. For widely expressed targets, RNAi was used.

	Target	Capture anitbody	Probe antibody	Cell specificity control
1a	TCR	IP-26 (eBioscience)	JOVI-1 (in-house)	JRT3 (Jurkat-negative mutant)
1b	CD3z	H146 (in-house)	6B10 (eBioscience)	X63 (myeloma)
2	LAT	10-17 (eBioscience)	661002 (R&D Systems)	Anj (Jurkat-negative mutant)
3	ZAP70	D1C10E Cell Signaling)	1E7 (eBioscience)	P116 (Jurkat-negative mutant)
4	SLP76	H76 (Biolegend)	06-548 (Millipore)	J-14 (Jurkat-negative mutant)
5	PLCγ	10/PLC (BD Biosciences)	NBP1-61254 (Novus)	Jgamma-1 (Jurkat-negative mutant)
6	PI3K p85	U5 (Thermo)	AB6 (Millipore)	X63 (myeloma)
7	VAV	9C1 (Novus)	05-219 (Millipore)	J-VAV (Jurkat-negative mutant)
8	LCK	73A5 (Cell Signaling)	3A5 (Santa Cruz)	Jcam-1 (Jurkat-negative mutant)
9	CD28	CD28.2 (eBioscience)	AF-342-PB (R&D Systems)	Mouse T cells, and Renca cells
10	GRB2	3F2 (Millipore)	SAB4501290 (Sigma)	Mouse T cells
11	SOS1	SOS-01 (AbCam)	07-337 (Millipore)	RNAi knockdown in Jurkat
12	NCK	Y531 (AbCam)	06-288 (Millipore)	Mouse cerebellum
13	FYN	FYN59 (Biolegend)	Fyn15 (Santa Cruz)	Renca
14	FYB	460107 (R&D Systems)	6348 (AbD Serotec)	Mouse T cells, and Renca cells
15	ITK	Y402 (AbCam)	2F12 (BD Biosciences)	Renca
16	GRAP2	UW40 (Novus)	1G12 (AbNova)	Renca
17	Cbl-b	246C5A (AbCam)	B-5 (Santa Cruz)	NIH3T3
18	BCL10	EPR3174 (AbCam)	4F8 (Thermo)	NIH3T3
19	PKCθ	MAB4368 (R&D Systems)	NBP1-00985 (Novus)	X63 (myeloma)
20	Thy1	1A1 (R&D Systems)	5E10-PE (Biolegend)	Mouse T cells, and Renca cells

Table S2. Phenotypic characteristics of alopecia areata patients and controls and their T cell populations. Isolated cells were analyzed by flow cytometry and gated on CD3. SP denotes single-positive populations for either CD4 or CD8. Patient ages are given in years. M, male; F, female.

Subject	T cells ($\times 10^6$)	% CD4 SP	% CD8 SP	% CLA^+CCR4^+	Age	Sex
aa1	2	20.9	56.9	12.9	29	M
aa2	3	65.4	28	74.2	50	F
aa3	8.4	14.7	76.3	17.4	50	F
aa4	3.87	88.9	8.7	44.5	81	F
aa5	1.32	21.6	59.3	21.7	39	F
aa6	1.58	85.1	7.7	71.5	41	F
aa7	1.58	69.3	25	16.6	59	M
cntl1	1.5	82.9	6.1	24.4	51	M
cntl2	1.6	78	17.9	50.5	77	M
cntl3	6.8	91.8	5.1	84.2	51	M
cntl4	6.2	98.1	0.3	79.1	76	M
cntl5	2.8	85.9	12.3	59	89	F