RESEARCH ARTICLE SUMMARY

SYSTEMS BIOLOGY

Conditional density-based analysis of T cell signaling in single-cell data

Smita Krishnaswamy, Matthew H. Spitzer, Michael Mingueneau, Sean C. Bendall, Oren Litvin, Erica Stone, Dana Pe'er,*† Garry P. Nolan†

INTRODUCTION: Cellular circuits sense the environment, process signals, and compute decisions using networks of interacting proteins. Emerging high-dimensional single-cell technologies such as mass cytometry can measure dozens of protein epitopes simultaneously in millions of individual cells. With thousands of individual cells, each providing a point of data on co-occurring protein states, it is possible to infer and quantify the functional forms of the relationships between proteins. However, in practice these underlying relationships are typically obscured by statistical limitations of the data, hence rendering the analysis and interpretation of single-cell data challenging. We developed computational methods, tailored to single-cell data, to more completely define the function and strength of signaling relationships.

RATIONALE: We demonstrate the utility of our methods using single-cell data collected from T cells. Although T cell subpopulations are phenotypically delineated into several cell subsets—such as regulatory, effector, and memory—and are thought to have similarly wired signaling networks, their responses to activation differ in ways that are not understood.

RESULTS: We used mass cytometry to measure the abundance of 20 internal and surface protein epitopes, at 13 time points, after two different types of T cell receptor (TCR) activation in T cells of B6 mice—resulting in more than 2 million data points. To study TCR signaling, we developed conditional-Density Resampled Estimate of Mutual Information (DREMI) to quantify the strengths of the influence that a protein X has on protein Y, and conditional-Density Rescaled Visualization (DREVI) to visualize and characterize the edge-response function underlying their molecular interaction. A key conceptual shift in DREMI and DREVI is our use of the conditional probability of $Y|X$ rather than the joint probability of X and Y. We show that the consensus Y-response for each value of X is much easier to identify in the conditional density estimate, especially when the joint density is concentrated in a narrow range, which is typical of such data (Fig. 1A).

We used DREMI to characterize the rapid dynamics of signaling interactions upon TCR activation (Fig. 1B) and show that the strength of signal transfer peaks in canonical pathway order (Fig. 1C). We compared edges in naïve and antigen-exposed CD4+ T cells and identified differential signal transmission along a key signaling cascade that starts at pCD3ζ and continues through pSLP76, pERK, and pS6. At each stage in this cascade, more information (higher DREMI) is transferred downstream from one protein to another, over a longer time period, in naïve cells than in antigen-exposed cells. We validated our characterization in mice lacking the extracellular-regulated mitogen-activated protein kinase (ERK2), demonstrating stronger influence of pERK on pS6 in naive cells, as predicted.

CONCLUSION: DREMI solves a challenging problem: quantifying the strength of the underlying complex relationships between proteins from noisy data. Our approach reveals how signaling is fine-tuned between T cell subpopulations: The differences we identified between naïve and antigen-exposed T cells suggest that naïve cells more sensitively transmit upstream signaling inputs along a key signaling cascade. In contrast, trained effector or memory cells seem poised for fast responses upon repeated exposure.

DREVI and DREMI are broadly applicable across biological systems and single-cell technologies. As single-cell data become more abundant, our methods will enable the construction of quantitative models of cellular signaling and comparison between healthy and diseased cells. ∎
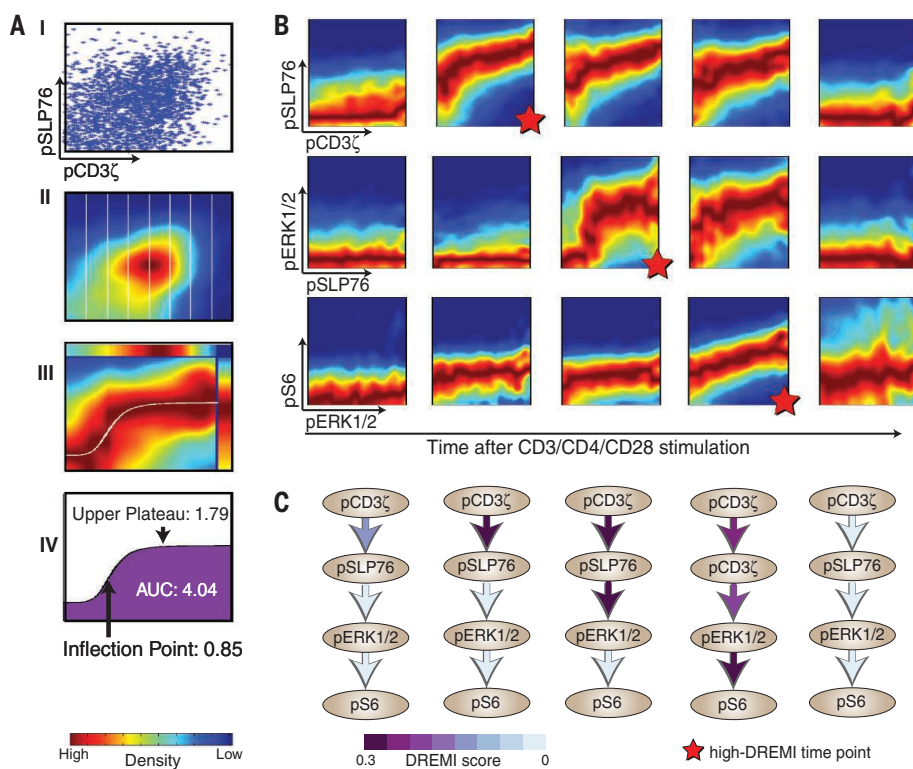
Quantitative characterization of T cell signaling. (**A**) The pCD3ζ-pSLP76 signaling interaction shown as (I) a scatterplot, (II) a kernel density estimate, and (III) by using a conditional DREVI method. (IV) Shape features are extracted and quantified. (**B**) DREVI plots of a signaling cascade downstream of TCR show the time-varying nature of edge shapes and strengths. (**C**) Edge strengths are quantified by using conditional DREMI.

RESEARCH ARTICLE

Conditional density-based analysis of T cell signaling in single-cell data

Smita Krishnaswamy,[1] Matthew H. Spitzer,[2] Michael Mingueneau,[3] Sean C. Bendall,[2] Oren Litvin,[1] Erica Stone,[4] Dana Pe'er,[1]*† Garry P. Nolan[2]†

Cellular circuits sense the environment, process signals, and compute decisions using networks of interacting proteins. To model such a system, the abundance of each activated protein species can be described as a stochastic function of the abundance of other proteins. High-dimensional single-cell technologies, such as mass cytometry, offer an opportunity to characterize signaling circuit-wide. However, the challenge of developing and applying computational approaches to interpret such complex data remains. Here, we developed computational methods, based on established statistical concepts, to characterize signaling network relationships by quantifying the strengths of network edges and deriving signaling response functions. In comparing signaling between naïve and antigen-exposed CD4[+] T lymphocytes, we find that although these two cell subtypes had similarly wired networks, naïve cells transmitted more information along a key signaling cascade than did antigen-exposed cells. We validated our characterization on mice lacking the extracellular-regulated mitogen-activated protein kinase (MAPK) ERK2, which showed stronger influence of pERK on pS6 (phosphorylated-ribosomal protein S6), in naïve cells as compared with antigen-exposed cells, as predicted. We demonstrate that by using cell-to-cell variation inherent in single-cell data, we can derive response functions underlying molecular circuits and drive the understanding of how cells process signals.

Cells process external cues through the biological circuitry of signaling networks in which each protein species processes information pertaining to other proteins, whose activities themselves are determined by biochemical modifications (such as phosphorylation) or other allosteric interactions. Signaling networks can be remarkably attuned to distinguishing subtle features of stimuli to enable key decisions regarding cellular response or fate. For example, naïve CD4[+] T cells take into account both dose and duration of T cell receptor (TCR) engagement, the strength of peptide binding in the major histocompatibility complex (MHC) cleft, and coreceptor cues in making a decision to differentiate into either regulatory or T helper cells (*1–4*). With this example as one among many, it follows then that to properly understand normal cellular responses and how these are dysregulated in disease, robust quantitative characterizations of signaling relationships will be required to enable more accurate models of signaling.

Despite progress in the quest to understand and represent the complexities of signaling biology, graph diagrams typically used as depictions of signaling relationships only offer qualitative abstractions. In such graphs, the vertices correspond to proteins, and a directional edge indicates the influence of one protein or molecular species on another, and as such, these graphs fail to capture many of the more complex ways through which signaling networks process information. Further, such representations are not designed to readily enable predictions of response to stimuli or therapeutic intervention. Although quantitative models have been proposed to describe signaling networks (*3, 5, 6*), these are specific to each system and require measurements of biochemical rates and many additional parameters. To scale to a large number of signaling networks and cell types, a robust data-driven approach that can quantify signaling interactions in molecular circuits is required. A data-driven approach would take advantage of statistically relevant differences in complex cell populations so as to better inform the function that is encoded by an inferred circuit diagram.

To this end, single-cell measurement technologies can offer quantitatively precise, even absolute (given appropriate probes and experimental design), measures of dozens of cellular components representing important biochemical functions. Variation in a complex cell population can be discerned in a functionally relevant context and enable insights into the underlying relationships between signaling molecules. Mass cytometry, for example, can assay the abundance of dozens of internal and surface protein epitopes simultaneously in millions of individual cells (*7, 8*), offering an opportunity to quantitatively characterize signaling at circuit-wide scales. Modeling a signaling network as a computational system, in which each signaling protein computes a stochastic function of other proteins, and treating each single cell as an example of possible input-out enables the recovery of how a signaling network functions. With many thousands of individual cells, each providing a point of data about relationships between proteins, we can infer the network function.

However, a major challenge in deciphering single-cell signaling data are developing computational methods that can handle the complexity, noise (which can be either natural stochasticity or actual instrument noise), and bias in the measurements. First, because cell populations are rarely homogeneous, different cell subpopulations can manifest distinct behaviors—and therefore, the relationships between signaling proteins may be obscured beneath a mixture of multiple network states. For example, naïve primary B cells can have weak and stochastic responses to stimuli so that only a small fraction of the population responds (via activation of signaling pathways), whereas memory B cells are considered primed and even evolved toward a more avid binding of antigen. Similarly, naïve T cells manifest different kinetics of response to T cell receptor engagement than do effector T cells. Second, technical noise in the measurements can further confound the quantification of molecular interactions. Third, marker abundance (which often correlates with cell size) can lead to biased correlations and hence be misinterpreted as an influence between the assayed signaling proteins.

We addressed these challenges by developing an algorithm, based on the statistical concepts of conditional probability (*9*) and density estimation (*10*), termed "conditional-Density Resampled Estimate of Mutual Information" (DREMI) to quantify the strength of molecular interactions. Given a relationship between two proteins, where X influences Y, DREMI considers the abundance or activity of protein Y as a stochastic function of the abundance or activity of protein X. DREMI uses the variation in a population of individually measured single cells to quantify the amount of information transmitted from protein X to protein Y in the signaling network.

A key conceptual shift compared with previous approaches to single-cell analysis is that DREMI computes mutual information on the estimated "conditional probability" of $Y \mid X$ rather than the "joint probability" of X and Y (the latter being the preferred approached in most other mutual information-based metrics). Joint probability describes the density of cell states (such as in a traditional scatter or density plot), whereas conditional probability describes how the state of Y varies with different states of X. To explore such relationships, we couple DREMI to an algorithm we term "conditional-Density Rescaled

[1]Department of Biological Sciences, Department of Systems Biology, Columbia University, New York, NY, USA. [2]Baxter Laboratory in Stem Cell Biology, Department of Microbiology and Immunology, Stanford University, Stanford, CA, USA. [3]Division of Immunology, Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA, USA. [4]Molecular Biology Section, Division of Biological Sciences, Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, USA.
*Corresponding author. E-mail: dpeer@biology.columbia.edu
†These authors contributed equally to this work.

Visualization" (DREVI). We use DREVI to visually render the estimated conditional density function as a rescaled heat map, which allows us to visualize the function underlying the molecular interaction in the cell population.

T cells offer an opportune system within which to characterize signaling relationships because several well-characterized but subtly distinct T cell subsets exist (such as regulatory, effector, and memory)—each of which have a distinct function, and yet the signaling distinctions between them are not well appreciated. For example, in naïve CD4+ T cells, which have not been exposed to antigen, the activation of the TCR (dependent on the engagement of coreceptors and availability of certain cytokines) leads to differentiation into functionally distinct effector cell types (11). In response to the same TCR activation, antigen-experienced effector or memory T cells tend to proliferate and mount responses more rapidly and in greater magnitude than naïve T cells (12–14). However, although the wiring diagrams that broadly describe the basic pathways are partly understood, the molecular mechanisms that lead to these differences are far from fully resolved.

Using a combination of DREMI and DREVI, we modeled signaling in pathways downstream of the TCR. We find that the functional form of interactions between signaling proteins demonstrates sigmoidal behaviors that are dynamically and systematically altered upon activation and with increases in stimulus strength, leading to complex but definable outcomes. We applied this to a physiological case using mass cytometry to track the abundance of 11 phosphoproteins after stimulation, ranging from 30 s to several hours, of mouse TCR-activation in more than 2 million individual cells. A comparison of T cell subtypes reveals subtle reconfigurations in the strength and functional shape of signal transfer between previously known pairs of interacting signaling molecules, all of which affects the response to TCR activation. Our model enabled predicting the outcome of a perturbation that quantitatively differed between T cell subtypes that we subsequently verified experimentally. The computational framework developed here can be generally applied to model any signaling system for which appropriate antibodies, binding agents, or surrogate readouts are available. The approach discerns signaling behaviors that exist in complex cell populations and, in a departure from our previous work (15), uses the inherent stochasticity to quantify and characterize the underlying signaling behaviors, demonstrating that computation can help reveal mechanistic insight from multidimensional single-cell data sets.

Results

The dynamics of T cell signaling was characterized and compared across mouse T cell subsets. Time-series data was collected for phospho-signaling proteins in T-lymphocyte populations of B6 mice (16) after stimulation of the TCR (by cross-linking TCR and the CD28 costimulator with biotinylated antibodies). We assayed the abundance of nine surface markers and 11 phospho-epitopes,

among which were key nodes in TCR signaling and related pathways (fig. S1 and table S1). Samples were collected at 13 time points after TCR activation ranging from 30 s to 80 min with two different types of stimuli (CD3/CD28 and CD3/CD4/CD28). We measured ~10,000 cells in each sample, cumulatively resulting in deep profiling of signaling data for more than 2 million individual cells.

The nine surface markers in our panel distinguish between T cell subsets, including naïve CD44low and antigen-experienced CD44high T cells (fig. S2) (17, 18). Overall, with 20 markers per individual cell, these studies simultaneously resolve six T cell subsets in the lymph nodes along with their signaling responses, thus reducing technical limitations inherent in the assay of post-sorted T cell subsets.

Conditional density-rescaled visualization of single-cell data

To analyze these data, we considered each individual cell in our mass cytometry data as an instance of co-occurring protein states and used this information to quantify the signaling relationship between the assayed protein parameters. We demonstrate our approach using a signaling relationship between two measured protein epitopes in our data. A primary locus for intracellular signaling initiation of the TCR activation cascade is a protein scaffold known as the cluster of differentiation 3 zeta-chain (CD3ζ). Upon engagement with antigen, CD3ζ is phosphorylated (to pCD3ζ), which creates a scaffold target for the zeta-chain–associated protein kinase 70 (ZAP-70), which in turn phosphorylates SLP76 (SH2 domain containing leukocyte protein of 76 kD) and
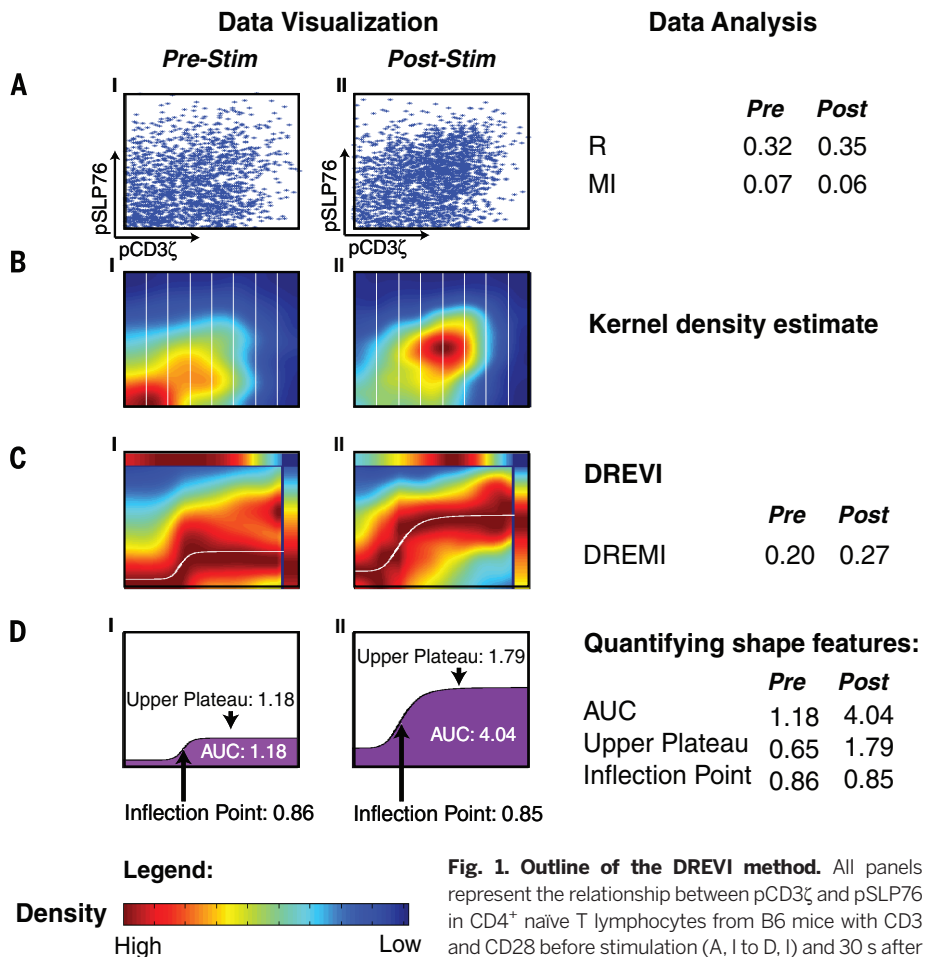

Fig. 1. Outline of the DREVI method. All panels represent the relationship between pCD3ζ and pSLP76 in CD4+ naïve T lymphocytes from B6 mice with CD3 and CD28 before stimulation (A, I to D, I) and 30 s after stimulation (A, II to D, II). (**A**) Each dot in the scatterplot represents the value of a single cell, where the x axis represents the amount of pCD3ζ levels, and the y axis represents the amount of pSLP76 levels. (**B**) The same data are shown using kernel density estimation, with red corresponding to dense regions and blue corresponding to sparse regions. The x axis is partitioned into slices (in practice, we use 256 such slices). (**C**) A DREVI plot of the same data, renormalizing the density estimate to obtain a conditional-density estimate of the abundance of pSLP76 levels given the abundance of pCD3ζ levels. Within each column, dark red (maximal color) represents the highest density in that slice. A response function (white curve) is fit to the region of highest conditional density. The top bar plots the marginal distribution of pSLP76 levels, and the right bar plots the marginal distribution of pS6 levels. (**D**) Following curve fitting, the inflection point, upper plateau, and AUC can be computed and compared between conditions.

LAT [the adaptor protein SH2 domain containing leukocyte protein and Linker of activated T cells, respectively (*19*)].

When the relation between pCD3ζ and pSLP76 prestimulation and 30 s after stimulation (Fig. 1A) is visualized as a scatterplot (a traditional means used to analyze flow cytometry data), it is difficult to identify any clear characteristics of the data other than the range of expression values. Visualizing the data by using a density plot (Fig. 1B) (*10*), one observes that the abundance of pCD3ζ and pSLP76 is concentrated at a particular value, but there is no clear statistical dependency between the two molecules, nor does this visualization method illuminate the nature of the well-researched influence that pCD3ζ has on pSLP76 (*20*).

Why do traditional methods of visualizing data, such as scatterplots or density estimations, fail to reveal clear relationships between molecules? In Fig. 1, the majority of cells can be found in a narrow range of values for each marker. Therefore, the joint density of pCD3ζ and pSLP76 is dominated by a narrow set of values and does not reveal the overall "functional response"—how protein Y changes as a function of the activity of protein X. To understand Y as a function of X, we have to explore X's full dynamic range and how values of Y are dependent on values of X.

To visualize and characterize signal transfer between proteins X and Y, we developed DREVI (Fig. 1, Box 1, and materials and methods), which represents in a visual form the stochastic function of how X influences Y. By using the natural variation in the amount of X and Y from cell to cell, we empirically learn the probability density P(X,Y) from the data. To characterize the influence that X extends on Y, we shift from the joint density and empirically estimate the conditional density P(Y|X) (materials and methods) by means of a kernel density estimation method based on heat diffusion (*10*, *21*). The conditional density enables us to ascertain how Y's values are changing with respect to the values of X, regardless of where the majority of cells are concentrated in the joint distribution. Although statisticians have previously developed sophisticated methods to directly compute the conditional density (*22*), these perform poorly on our data (fig. S3).

DREVI plots are shown in Fig. 1C that depict pSLP76 abundance (Y) when conditioned on pCD3ζ (X) abundance. This view of the data reveals a distinct difference in the relation between pCD3ζ and pSLP76 before and after TCR activation that was not apparent in the scatterplots or joint distribution. Given an equal amount of pCD3ζ, we observe that pSLP76 has a stronger response to increasing pCD3ζ levels after TCR activation (in Fig. 1C, the increased values of the median response are represented by the white line). This suggests that after TCR activation, additional factors (perhaps ZAP70 or LAT) modulate the relationship between pCD3ζ and pSLP76. Similarly, DREVI can help clarify the relationship between additional protein pairs (fig. S4) and visualize how the relationships change through time (fig. S5).

Conditional density-resampled estimate of mutual information

To systematically compare between conditions, time points, and cell types, a measure that quantifies the relationship between two proteins is needed. Often, mutual information (MI)–based metrics are used to evaluate relationships between gene or protein pairs. However, MI is difficult to compute on continuous data. To resolve this, a first step is to discretize the data. Currently, adaptive partitioning (*23*) is one of the most widely used approaches for such discretization (*24*, *25*). However, adaptive partitioning assumes that denser regions of the data are more important than sparser regions, and therefore, the dense regions are partitioned more finely than are sparse regions and dominate the resultant mutual-information metric. In contrast, we developed a measure that allows for sparser populations to be accounted for, thus preventing bias against small cell populations that could have distinct and interesting biology that might inform understanding of the signaling relations in question.

To quantify the strength of the influence protein X has on protein Y, we developed DREMI (Box 2, Fig. 2, and materials and methods). Like MI (*26*), DREMI is a shape-agnostic measure that scores how predictive X is of Y, but—unlike MI—it is not symmetric (and therefore might inform directionality) and also captures the strength of this relationship over all populated regions of the dynamic range, regardless of the (often peaked) distribution of X in the data. This is achieved by computing mutual information on the conditional density estimate of the data rather than the raw data itself. DREMI begins with the conditional density estimate, computed for DREVI (Box 1). The data are then resampled evenly through the range of the conditional probability density, and MI is computed on the resampled data spanning the entire range.

To understand the differences between MI and DREMI, we use the data from Fig. 1. MI works well for data that is well distributed across the range of X and Y (Fig. 2A, for example); by equipartitioning slices of X, the range of Y drastically drops within each slice, and therefore, knowing the value of X provides substantial

---

Box 1. Conditional density-rescaled visualization.

DREVI reveals the influence of protein X on protein Y. Most methods for visualizing pairwise data (such as scatterplots and density contours) depict the joint probability of X and Y. In contrast, we reveal signaling behavior along its full dynamic range by visualizing conditional probability. The main steps of DREVI are as follows:

(i) Compute the joint kernel density estimate $\hat{f}(x,y)$ (*10*).

(ii) Compute the marginal density of X, $\hat{f}(x)$, and the conditional density estimate of Y given X as $\hat{f}(y|x) = \frac{\hat{f}(y,x)}{\hat{f}(x)}$ on a fine grid of points $G = \{(x_i, y_j), 1 < i < n, 1 < j < m\}$ that span the range of X and Y (*2*, *12*, *15*).

(iii) Rescale each value of the conditional density estimate by its column-maximum to obtain $\hat{f}^*(y_j|x_i) = \hat{f}(y_j|x_i)/\max_k[\hat{f}(y_k|x_i)]$

(iv) Visualize $\hat{f}^*(G)$ as a heatmap, adding side-bars depicting the marginal densities of X and Y.

Box 2. Conditional density-resampled estimate of mutual information.

DREMI provides a score for the strength of the influence protein X has on protein Y. In many physiological conditions, only a small fraction of the cells have activated protein X in response to stimuli, and these active populations have little influence on the mutual information metric. DREMI explicitly factors these populations by computing a score based on the conditional distribution of Y | X rather than joint distribution. DREMI estimates the computation I^c, given by $I^c(x,y) = \sum_x \sum_y p(y|x) \log\left[\frac{p(y,x)}{p(x)p(y)}\right]$

DREMI is computed as follows:

(i) Begin with the rescaled conditional-density estimate $\hat{f}^*(G)$ as in DREVI (Box 1), computed on the grid of points $G = \{(x_i, y_j), 1 < i < n, 1 < j < m\}$.

(ii) Round values $\hat{f}^*(x_i, y_j) < \varepsilon$ to 0, for user-defined threshold ε, to eliminate technical noise from the measurements.

(iii) Resample from G according to the conditional density estimate $\hat{f}(G)$.

(iv) Equi-partition the full range of the data and calculate mutual information, using this partition as the discretization.

Corrected 5 March 2015; see full text.

information on the likely value of Y. However, in distributions typical of single-cell data (Fig. 2B), MI is dominated by a peak in density (a narrow range of X) in which only minor changes in Y are observed; thus, the relationship between X and Y is obscured. In contrast, DREMI (Fig. 2C) resamples from the conditional density estimate and equally weighs all regions along the entire range of X, as long as there are enough cells to form a robust conditional density estimate. Thus, DREMI takes into account sparse populations along the x axis and factors for the full range.

DREMI takes advantage of mass cytometry's ability to collect large sample sizes (data for millions of cells), facilitating the estimation of the conditional distribution in relatively sparse regions. For example, with 50,000 cells, an X range with only 1% of the data (typically treated as an outlier) still contains 500 cells, a sufficient number to robustly estimate $P(Y \mid X = x)$. To further ensure a robust estimation of $P(Y \mid X = x)$, DREMI incorporates an automated noise-filtering step in which points are eliminated if they have a low rescaled conditional density.

We evaluated the robustness of DREMI to noise (fig. S6) and subsampling (fig. S7). As shown in fig. S6, DREMI values generally decrease linearly as the SD of the noise increases, and the noise-elimination step can be used to adjust signal detection in noisy experimental conditions. A key strength of DREMI is that it is similarly resilient to noise, whether the underlying function is linear or sigmoid, making it well suited to capture the range of behaviors observed in biology. Further, DREMI is a robust measure that remains consistent under multiple subsamples of the data (fig. S7).

Characterization of signaling relations by curve-fitting

In the case of a strong relationship (high DREMI score), we can derive a "response function" from the conditional density using curve-fitting. Conditional density is particularly suited for the derivation of an edge response function because it allows for identification of the consensus Y response (densest region in DREVI; the deep red in the heat maps) for each value of X. We compare curve-fitting of the conditional density with directly fitting the raw data in order to demonstrate the superiority of the former—both in terms of fit [root mean squared error (RMSE)] and in terms of the interpretability of the parameters (fig. S8). For a well-fit curve, we expect most of the data points to fall in close proximity to this curve, as measured with RMSE. A fit to the conditional mean results in a sigmoidal curve that closely follows the data points (indicated by low RMSE), whereas the raw data best fits a line—but the fit is of lower quality (indicated by high RMSE). An optimal sigmoidal fit on the raw data results in degenerate curves, in which most data points reside at a large distance.

Fitting a curve in this manner allows for a parametric description of the relationship between the two proteins X and Y. Such a functional description of the relationship between X

and Y enables us to potentially predict the value of Y, if the value of X is altered by an intervention or drug. To derive the edge-response function, we fit points sampled from the conditional density to one of three models using regularized regression: linear, sigmoidal, or double sigmoidal (27). If none of the models result in a good fit, the model is then fit to a free-form curve.

In Fig. 1, we fit DREVI plots of the pCD3ζ-pSLP76 edge before and after TCR activation to a sigmoid curve (Fig. 1D, I and II). Among other characteristics, the parameters of the fitted sig-

moid specify the lower and upper asymptote that correspond to a digital "low" and "high" mode of activation for the protein Y, and the inflection point (the activation threshold of X at which the protein Y transitions from low to high state). Comparing the edge before TCR activation with after, there is a similar inflection point activation threshold in the pCD3ζ axis but a considerably higher response on the SLP76 axis, and a corresponding increase in total activation of SLP76 (Fig. 1D). The postactivation state results in higher "area under the curve" (AUC),

A Mutual information measures the decrease in uncertainty of Y given X
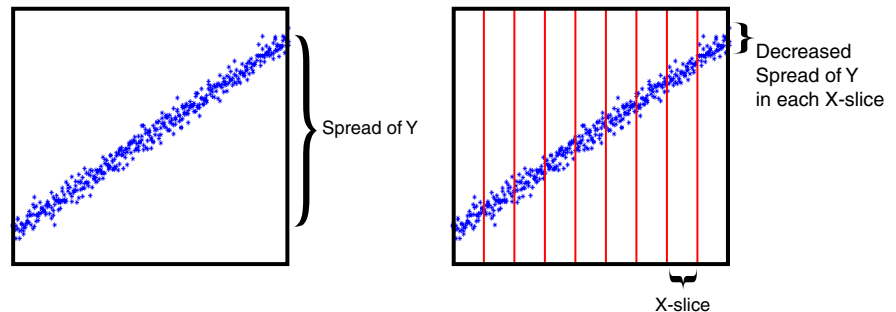
B The high local joint density biases mutual information assessment
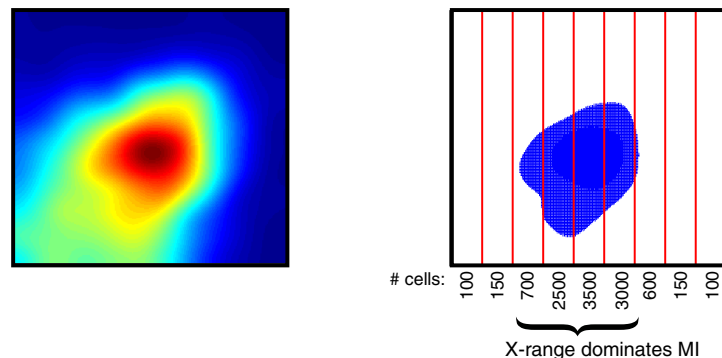
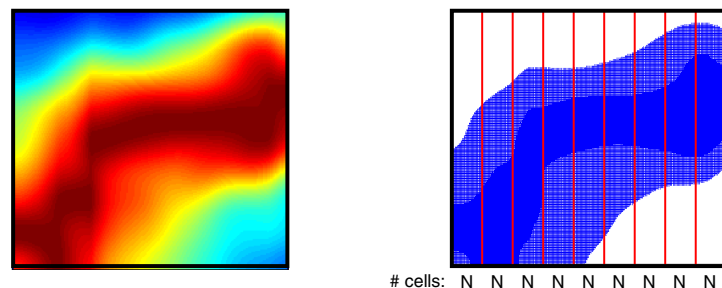**C DREMI resamples Y from conditional density in each X-slice to reveal relationship between X and Y**

Fig. 2. Outline of the DREMI method. (**A**) Simulated data depicting a relationship with high MI. Across the entire range of X, Y has a wide range of values (left). After conditioning on X (red lines), the range of Y within each X-slice is substantially smaller (right). Thus, knowledge of X's value provides information on Y's likely value. (**B**) The data represent the relationship between pCD3ζ and pSLP76 30 s after stimulation of the TCR, as in Fig. 1B, II. Most of the cells are concentrated around a single peak, as shown in the 2D density estimate (left). If we resample from this density estimate, samples will fall within a narrow range in which there is little change in Y's range. This narrow range dominates the MI metric. (**C**) Instead, DREMI samples from the renormalized conditional density estimate, which covers entire X-range. So there is the same number of cells in each X-slice, as long as there are sufficient cells in the original data. MI information is calculated over this resampled data.

which is proportional to the average amount of pSLP76 generated per quantity of pCD3ζ. A larger AUC implies more signal transfer activity so that after stimulation, less pCD3ζ is required to induce a response in pSLP76. Thus, when the TCR is activated, not only do pCD3ζ amounts increase, but this change is also more impactful than a similar increase would have been before stimulation. This increase in edge strength indicates that the pairwise relation between the two proteins is boosted by changes in the recruitment and localization of additional and potentially unknown proteins (the effects of which are successfully captured with DREVI and DREMI).

Not only do the DREMI and AUC scores increase after T cell stimulation, but they also enable distinguishing between different strengths and forms of stimulus (Fig. 3). Mouse T cells were stimulated with either CD3/CD28 or CD3/CD28/CD4, and mass cytometry analysis was undertaken as per above. CD4-costimulus, in addition to CD3/CD28, is known to boost signaling responses in CD4$^+$ T cells by engaging additional pathways that reinforce the signal transmission

(*28*). When comparing these two forms of stimuli, we derive a higher DREMI score under the stronger stimulus in all three of the edges shown in Fig. 3. The CD3/CD28/CD4 stimulus clearly leads to a higher activation (magnitude) and/or a lower activation threshold in the edge-response functions. This translates to a higher AUC in both the sigmoidal (pCD3ζ-pSLP76) (Fig. 3A, IV) and linear (pERK-pS6) (Fig. 3C, IV) cases. The pSLP76-pERK edge (Fig. 3B) can be interpreted as a stochastic "digital" response, with a larger percentage of cells responding upon CD3/CD28/CD4 stimulus (Fig. 3B, IV) (*5*). Together, DREMI, DREVI, and the edge-response function form a powerful toolbox for the analysis of signaling interactions.

Canonically ordered signaling transfer in the TCR pathway

We used DREMI to analyze the dynamics of the pCD3ζ-pSLP76-pERK-pS6 signaling cascade downstream of the TCR (Fig. 4). The edge strengths and shapes were highly dynamic within just a few minutes after TCR activation. The DREMI scores

pinpoint the "peak timing" of signal transduction, meaning the timing when protein activation of *Y* is most sensitive to changes in the activity of protein *X*. The peak timings of all edges in this signaling cascade followed the expected canonical order (Fig. 4). For example, a pCD3ζ and pSLP76 relationship, or edge, was detectable in a small population of cells when pCD3ζ is at its highest value even before stimulation (Fig. 4, top left). However, after TCR activation, pCD3ζ reached maximum signal transfer within 30 s and decreased soon after. The pSLP76-pERK edge scored highest at 1 min after activation, and the pERK-pS6 edge peaked at 2 min after TCR activation. Thus, the peak response follows the canonical ordering of the pathway in a manner analogous to the "just-in-time" transcriptional networks of metabolic pathways in Zaslaver *et al.* (*29*).

The observed dynamics highlight the importance of time series data and the impact of signaling dynamics on the relationships that can be derived from each individual time point. For instance, the pSLP76-pERK edge (Fig. 4A) is only active in the 1- to 2-min time frame with a strong

Fig. 3. Comparison between stimulation conditions. The cells are CD4$^+$ naïve T lymphocytes of B6 mice. (**A**) (I and II) The pCD3ζ-pSLP76 edge at 30s after stimulation under both stimulation conditions. (III) As stimulation strength increases, DREMI increases. (IV) The inflection threshold moves to the left, and the upper plateau of the response curve also increases, resulting in a higher AUC. (**B**) The pSLP76-pERK edge 2 min after stimulation. The blue lines indicate a partitioning between "low" and "high" states of pERK. The pie charts in IV show an increase in the percentage of cells that reach the high state at the higher dose. (**C**) The pERK-pS6 edge has a linear edge-response function. Here, increasing stimulation strength results in an increase in DREMI, shown in III, and AUC, shown in IV.
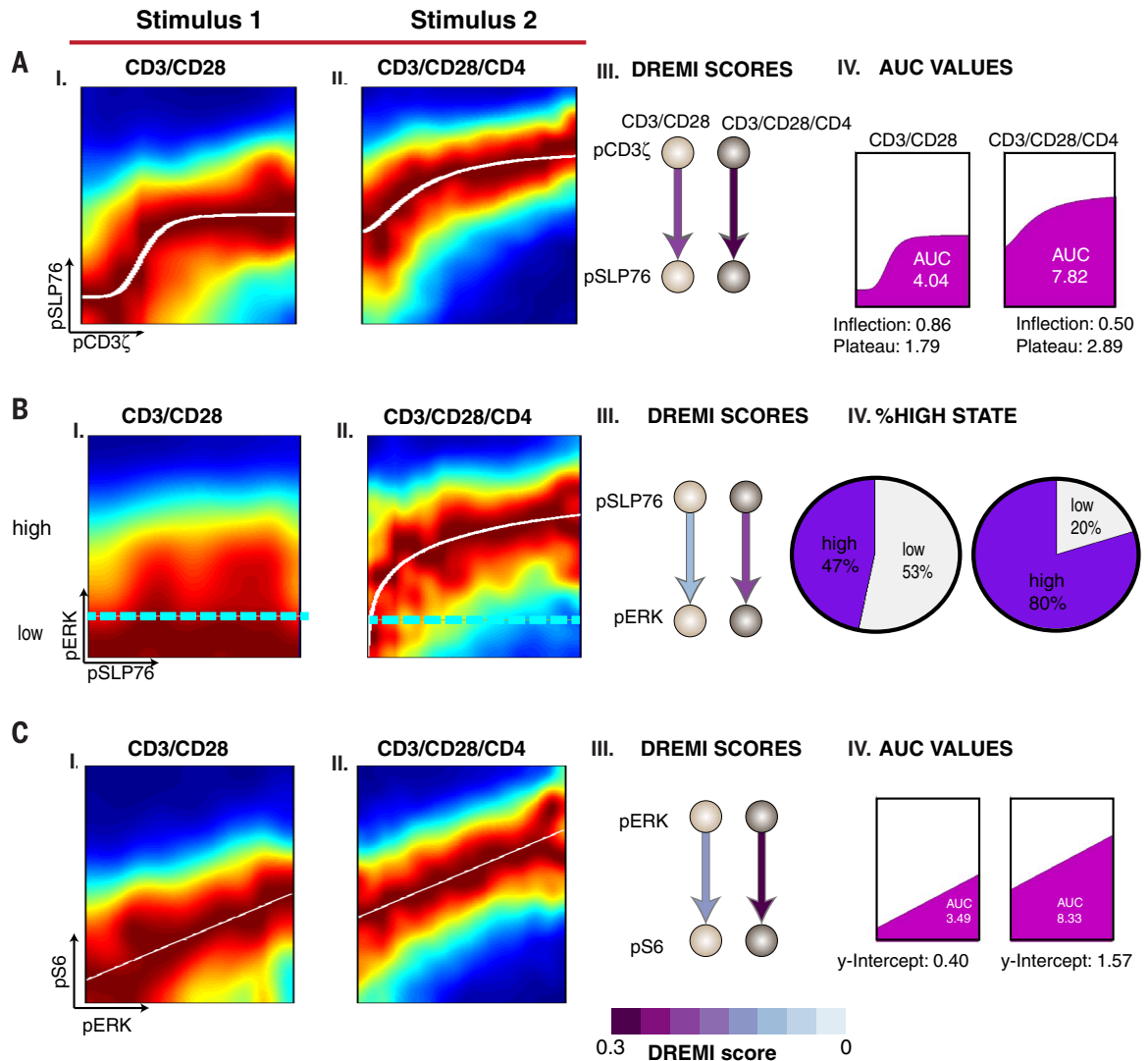
sigmoidal shape. The pSLP76-pERK edge is flat at other times. Such signaling dynamics are suggestive of the observed negative feedback loop, resulting in dephosphorylation of ERK perhaps by PTPN22 or SHP-1 (*30, 31*) because the influence of pSLP76 on pERK drastically decreased at 4 min, although pSLP76 remained activated at the 4-min time point (Fig. 4A, top row, fifth from left). The DREMI scores quantify the intuitive visual interpretation of the DREVI plots (Fig. 4B).

This canonical pathway ordering of peak timing is demonstrated again under the CD3/CD28 stimulation (fig. S9). However, we see that under the CD3/CD28/CD4 stimulation, signals are transferred downstream faster, and the signal transfer is sustained for longer periods. For example, the pCD3ζ-pSLP76 edge is only activated up until 2 min under CD3/CD28 stimulation, whereas it remains active until 4 min under a CD3/CD28/CD4 stimulatory environment. Additionally, the pSLP76-pERK edge activates only after 2 min under the CD3/CD28 stimulation but is activated within 1 min under CD3/CD28/CD4 stimulation. This is in accordance with upstream influences that change the nature and speed of the signaling behavior downstream as reflected in the edges more distal to the TCR. This increased activation possibly is due to additional signaling pathways that converge through other signaling network components downstream of the TCR.

DREMI quantified the strength of each known edge in a given network, and the dynamics of how the molecular interactions underlying this edge changed over time. Together, they enabled an elucidation of a classically understood pathway in a single experiment while providing subtle insights into the timing of signaling relationships within the network.

Performance evaluation of DREMI

To compare the performance of DREMI with other approaches that quantify relations typically used for this data type, we used the canonical timing of signaling edges described in the previous section as a basis for comparison. We compared these metrics on six edges to evaluate the ability of each method to rank edge strengths at various time points after TCR stimulation. DREMI outperformed Pearson correlation and recently developed dependency measures such as maximal information coefficient (MIC) (*32*) and adaptively partitioned MI (Fig. 5I and fig. S10) (*23*). DREMI identified the peak timing of signal transfer in all six test cases. Adaptive partitioning missed the peak timing in four of the six cases, including two cases in which adaptive partitioning picked the prestimulation time as the peak, and typically does not perform well when the joint density is concentrated in a narrow range. In such cases, adaptive partitioning places many partitions in the dense region (so that the partitions are often finer than the sensitivity of the experimental technology) and only a few in the sparse regions. MIC tends to pick peaks similarly to adaptive partitioning, identifying the correct peak in only one case. MIC has a further limitation in that there is very

little range in the scores, possibly because it aims for equitability (*33*), which is a condition that does not hold in these data. Pearson correlation missed the peak timing in two cases and typically did not perform well for noisy and nonlinear relationships. We also compared the performance of DREMI in an additional biological system using published mass-cytometry data focused on human B cells after B cell receptor stimulation, as described in (*8*). Similar to T cell signaling, DREMI was able to quantify and rank relationships better than metrics typically used for such data (fig. S11).

Last, we compared DREMI with other approaches on synthetic data that was designed to exemplify features of realistic data (Fig. 5II). We

compared two weak, low-noise linear relationships against two strong, nonlinear noisy relationships. Specifically, we included a sigmoidal relationship and a nonmonotonic quadratic function typical of systems that simultaneously induce both protagonist and antagonist pathways (*34*). The weak relationships were characterized by only slight changes in the distribution of *Y* compared with its dynamic range, as *X* increases (small slope), whereas the stronger nonlinear relationships have more pronounced change in the distribution of *Y* with changes in *X*. In both examples, the other methods scored the weaker linear relationship higher (Fig. 5II).

A key advantage of single-cell technologies such as mass cytometry and flow cytometry are
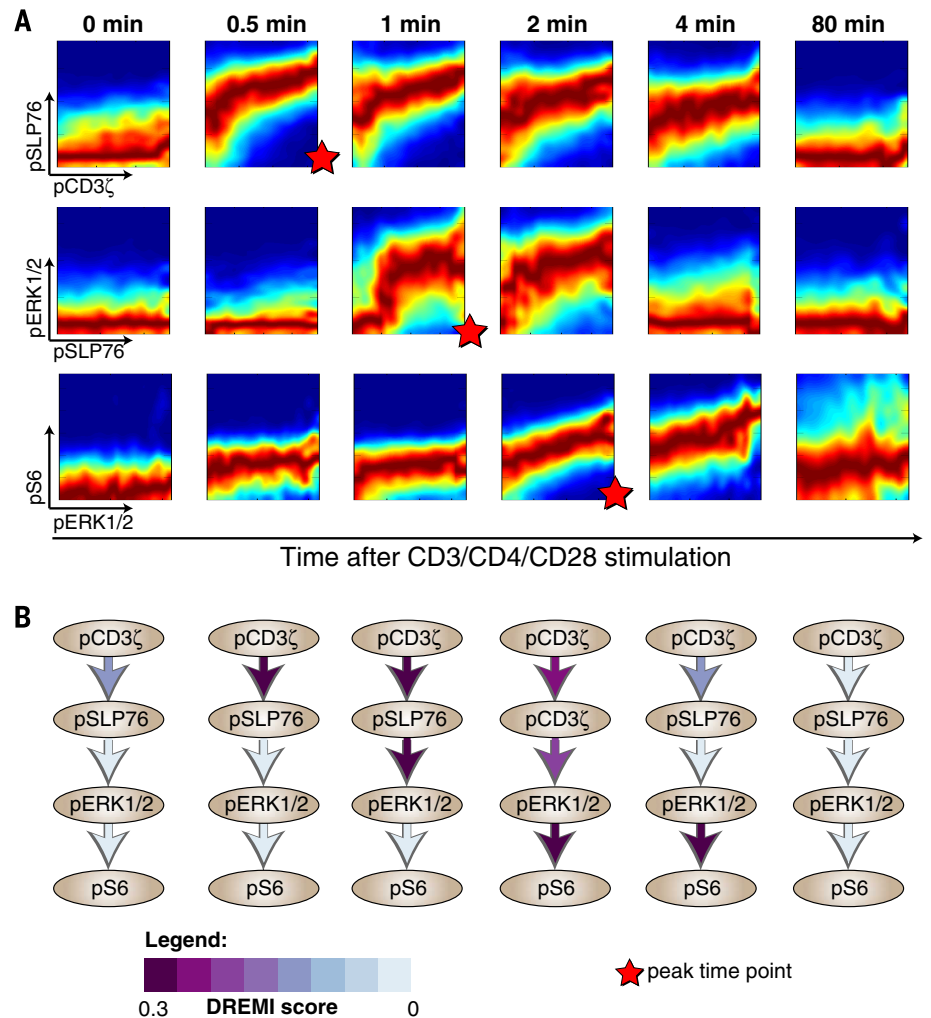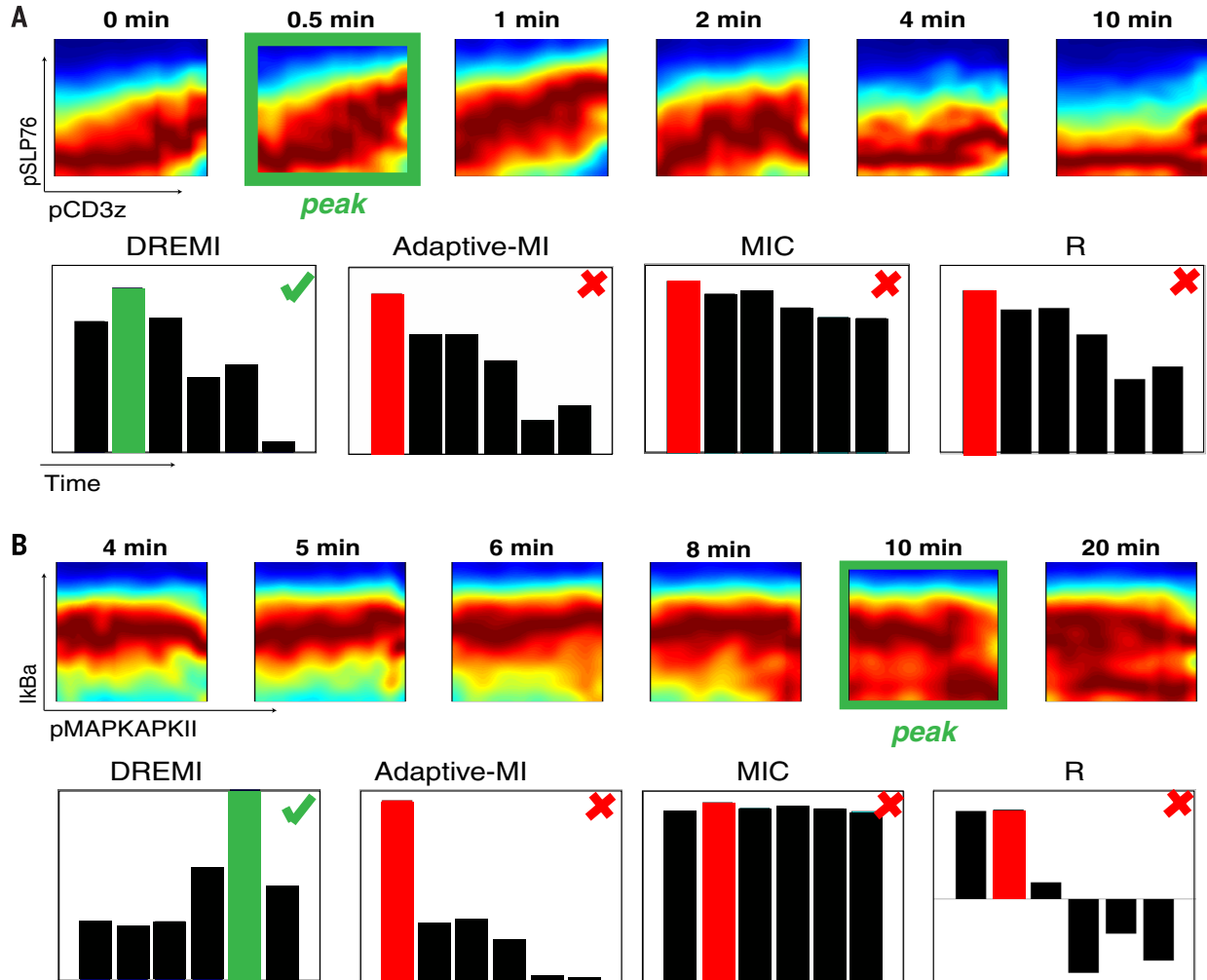

Fig. 4. DREVI and DREMI reveal the dynamics of TCR signaling. The pCD3ζ -pSLP76-pERK-pS6 cascade at various time points after TCR activation. All panels represent data from CD4⁺ naïve T lymphocytes from B6 mice with CD3, CD28, and CD4 cross-linking. (**A**) Each row represents an edge in the cascade, and each column represents a time point after TCR activation. DREVI plots show the dynamics of TCR signaling; edges change their strength and shape as signaling proceeds. The scale used is the same as in Fig. 1, with dark red representing the behavior of the majority of cells, whereas the other colors depict the heterogeneity and spread of the response. (**B**) Network diagrams with DREMI values indicated on each edge, columns matching the same time points as in (A). For each edge, the color indicates the computed DREMI values, with darker purple indicating a higher DREMI score. DREMI captures peak signal transfer timings in each of the three edges.

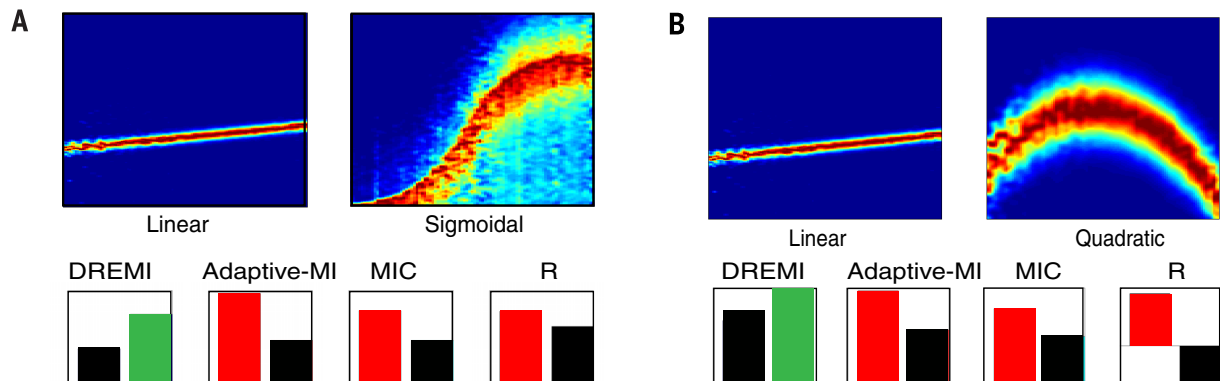## I. Comparison on mass cytometry Tcell data

Fig. 5. Comparison of DREMI with three other metrics. The metrics compared are mutual information using adaptive partitioning (*23*), MIC (*32*), and Pearson correlation. (**I**) Comparisons are performed on edges in naïve CD4[+] T cells under CD3 and CD28 stimulation. In each case, the green outline indicates the point with the strongest relationship upon visual inspection of the DREVI plot. (**A**) The pCD3ζ-pSLP76 edge. Only DREMI identifies the poststimulation 0.5 m time point as having a stronger signaling relationship than the prestimulation data. (**B**) The pMAPKAPKII-IkBa relationship as an example of a negative rela-tionship. DREMI successfully identifies this negative relationship, whereas other methods are unable to detect it. Four additional edges are compared in fig S10. (**II**) Comparison of dependency metrics on synthetic data. Both parts (**A**) and (**B**) show comparisons between a noiseless but weak dependence of *Y* on *X*, compared with a noisy but strong dependence of *Y* on *X*, (A) sigmoidal and (B) quadratic relationships. The other three metrics score the linear, weaker depen-dence higher than the noisier but stronger dependence. Because single-cell signaling data are often noisy and nonlinear, DREMI is better suited.

the provision of millions of data points in each experiment. Therefore, the scalability of the analysis is imperative to effectively analyze such data. DREMI is faster and more scalable than either MIC or adaptive partitioning (fig. S12). On a data set with 500,000 cells, DREMI analysis can be computed within 10 s. MI with adaptive partitioning takes over 90 min to perform the same computation, and MIC is limited to data

sets of less than 10,000 cells. Thus, DREMI is a versatile metric that can handle a variety of characteristics, including large sample size, non-linearity, and different degrees of stochasticity.

Unlike the other metrics, DREMI is a directional metric; DREMI(X, Y) is not necessarily equal to DREMI(Y, X) (fig. S13). We based our analysis on the known network diagram and scored DREMI in the causal direction of influ-

ence. Evaluating DREMI in the reverse direction shows a slight degradation of performance compared with DREMI in the correct direction (figs. S10, S11, and S14). However, reverse DREMI outperforms the three symmetric measures used in our comparison and has the advantage that it can handle nonlinearity and noise (fig. S14). The correct causal direction scores approximately equal or higher in three of the four edges explored at
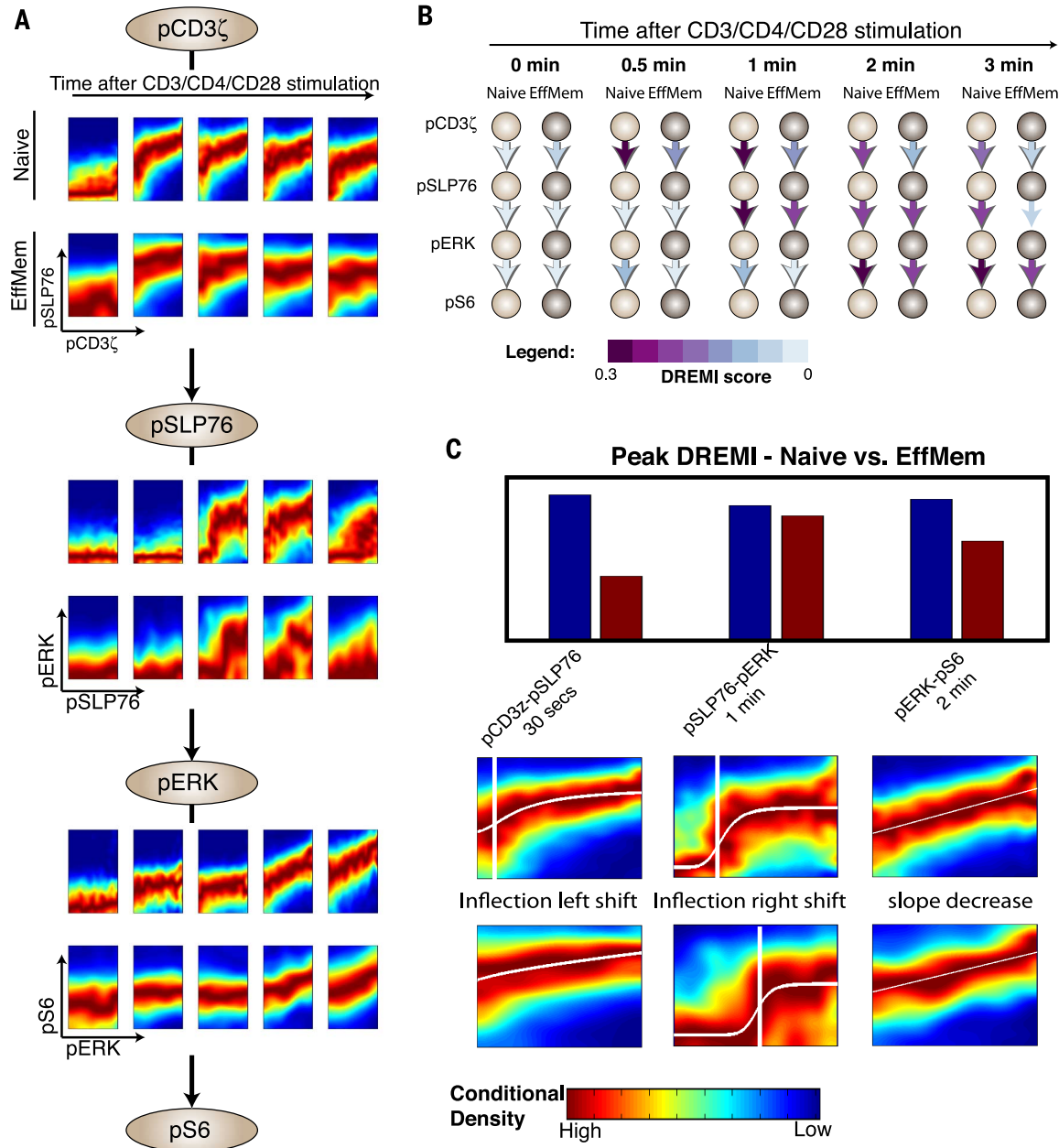
Fig. 6. Comparison of naïve with antigen-exposed T cells. Data from CD4[+] naïve (CD44[low]) and effector or memory (antigen-exposed, CD44[high]) T lymphocytes from B6 mice with CD3, CD28, and CD4 cross-linking. (**A** and **B**) We compare the pCD3ζ-pSLP76-pERK-pS6 cascade between the naïve and effector or memory CD4[+] T cells using (A) DREVI and (B) DREMI. The time points in (A) match the indicated time points in (B). Both depictions show that signal transmission is sharper and more sustained in naïve cells. (**C**) Bar graph depicts the higher peak-DREMI detected in naïve cells; DREVI plots with edge

response functions show alterations in edge behavior during peak signaling. The pCD3ζ-pSLP76 edge response function shows an earlier inflection point in effector or memory cells and is relatively flat after induction. This suggests only a small amount of pCD3ζ is needed to initiate a complete response. In contrast, pSLP76 needs more pCD3ζ for a similar response in naïve cells. The pSLP76-pERK edge indicates that the pERK response is weaker and requires more pSLP76 in these cells. Last, the pERK-pS6 edge shows a steeper slope in naïve cells, indicating that pS6 responds more strongly to pERK levels in naïve cells.

peak signaling times (fig. S13). The exception was pERK-pS6, where pS6 is known to have many additional factors influencing its activity, a recurring situation in cases in which DREMI in the reverse-causal direction scores higher. We show that increasing the number of parents decreases the DREMI in the causal direction more than in the reverse direction both in linear and non-linear functions (fig. S15).

Comparison of signaling interactions between T cell subtypes

The high dimensionality of mass cytometry enabled us to simultaneously treat and measure multiple T cell subsets, including both naïve and antigen-exposed CD4⁺ T cells. It is well documented that naïve T cells behave differently than cells that have previously seen and "remember" antigen. In naïve CD4⁺ T cells, the activation of the TCR and appropriate coreceptors under the right cytokine environment leads to differentiation toward different effector cell fates (11). In response to the same TCR activation, antigen-experienced effector or memory T cells proliferate and respond to the infection (12–14). Moreover, the activation response time is known to be faster and of greater magnitude in effector T cells than in naïve T cells. However, the signaling mechanisms that create these differ-

ences are not yet clear. The data set collected here enabled direct comparison of signaling responses in T cells before and after exposure to antigen because the cell subsets are in the same experimental tube and are resolved after measurement by means of gating techniques (such as the differential expression of the surface marker CD44).

We used DREMI and DREVI to analyze differences along the signaling cascade that starts at pCD3ζ and continues down through pSLP76, pERK, and pS6. Naïve and antigen-exposed T cells had three main differences in their responses to TCR cross-linking. First, we found that along each edge in this pathway, naïve cells had higher peak-DREMI, indicating that signals were transferred more faithfully from the upstream molecule to the downstream molecule (Fig. 6, A to C). Second, we found that naïve cells had more sustained periods of high-DREMI (Fig. 6B), which indicated that signal transduction events transpired longer. Third, despite the fact that there was a more faithful signal transfer in naïve T cells, the levels of most of the measured phosphoproteins along this cascade, except pCD3ζ, were observed to be higher (constitutively starting at time 0) in antigen-exposed cells (Fig. 6D).

The peak DREMI score (Fig. 6C) in either the naïve or the effector cascades occurs at 30 s for

the pCD3ζ -pSLP76 edge, at 1 min for the pSLP76-pERK edge, and at 2 min for the pERK-pS6 edge. All three of these were substantially higher in naïve cells than in effector-memory cells. Further, the pCD3ζ -pSLP76 relationship sustained signal transmission (high DREMI) for 2 min in naïve cells, compared with only 30 s in antigen-exposed cells. Similarly, the pSLP76-pERK edge had high DREMI between 1 and 3 min in naïve cells, whereas DREMI was only high between 1 and 2 min for antigen-exposed cells. Thus, at each stage in this cascade, Y transmitted more information about X, over a longer time period, in naïve cells than in antigen-exposed cells, suggesting tighter control in naïve cells. This difference between naïve and antigen-exposed cells was consistently observed across multiple replicates (fig. S16).

Although edge strength was higher in naïve cells, pSLP76 levels were higher in antigen-exposed cells independent of pCD3ζ levels (Fig. 7A). The near-saturation of the pSLP76 levels can also be seen in the high and flat slope of pCD3ζ-pSLP76 edge at 30 s (Fig. 6C). Similarly, pERK amounts were more faithfully transmitted to pS6 in naïve cells, whereas pS6 amounts were basally higher—independent of pERK—in antigen-exposed cells (Fig. 7A). In the case of TCR activation with CD4 costimulation, the analysis suggests that the
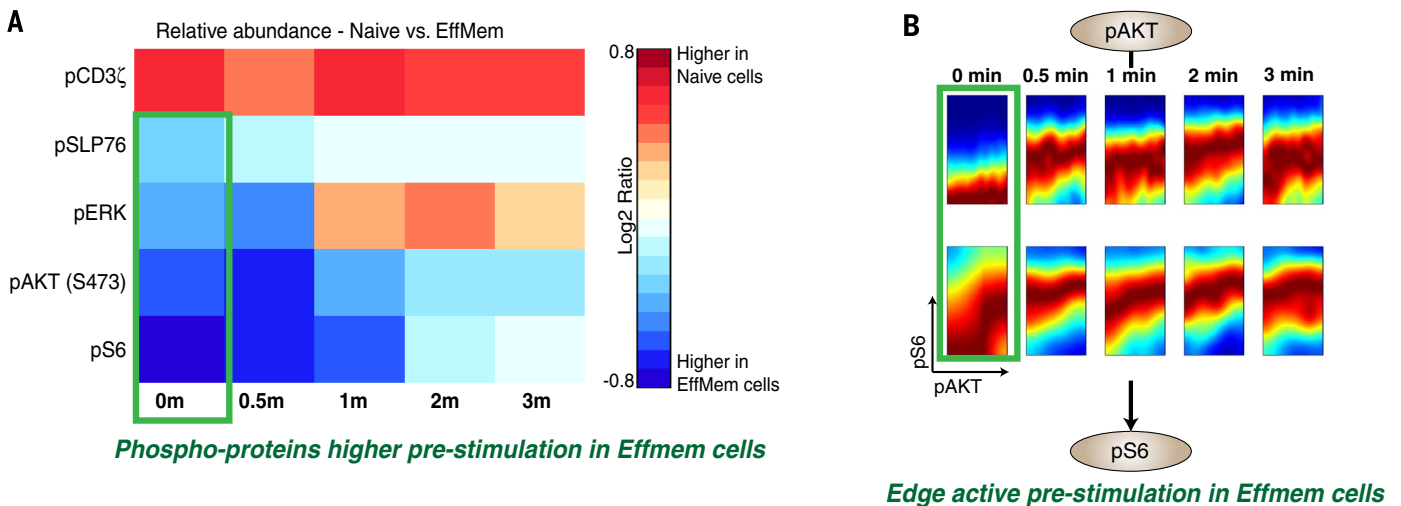
A Relative abundance - Naive vs. EffMem

Phospho-proteins higher pre-stimulation in Effmem cells

B

pAKT

0 min   0.5 min   1 min   2 min   3 min

pS6 / pAKT

pS6

Edge active pre-stimulation in Effmem cells

C

0 min   2 min

Naive   Effmem   Naive   Effmem

pCD3ζ
pSLP76
pERK
pS6

0.16 DREMI score 0

Fig. 7. Comparison of naïve with antigen-exposed T cells on the pAKT pathway. Data from CD4⁺ naïve (CD44^low) and effector or memory (antigen-exposed, CD44^high) T lymphocytes from B6 mice with CD3, CD28, and CD4 cross-linking. (**A**) Comparison of phophoprotein abundances between of naïve CD4⁺ T cells, with effector or memory CD4⁺ T cells. Red indicates the phosphoprotein levels amounts are higher in naïve CD4⁺ T cells, and blue indicates the phosphoprotein amounts are higher in effector or memory CD4⁺ T cells, demonstrating that differences in the average behavior of nodes (protein species) are not necessarily telling of the differences in edges (relationships between proteins). (**B**) DREVI plots comparing naïve and effector or memory cells on the pAKT-pS6 edge from 0 to 3 min after stimulation. (**C**) Network diagrams depicting DREMI scores in naïve and effector or memory cells, showing the combined pCD3ζ-pSLP76-pERK-pS6 and pCD3ζ-pAKT-pS6 pathways. The prestimulation pAKT-pS6 edge is stronger in effector-memory cells at 0 and 2 min, whereas responses downstream of pCD3ζ are stronger in naïve cells at 2 min.

Corrected 5 March 2015; see full text.

AKT pathway might explain the higher level of pS6 in antigen-exposed cells. The AKT-pS6 edge had higher DREMI in antigen-exposed cells than in naïve cells (Fig. 7, B and C, and fig. S17) in early time points. Moreover, the AKT-pS6 relation was active before stimulation in antigen-exposed cells, providing a constitutive boost in signaling, perhaps through positive feedback between pS6 and pAKT (35).

Together, these differences suggest that antigen-exposed cells are poised toward a more easily triggered response (12–14, 36). Because the basal levels of the phospho-proteins are higher, a shorter period of signal integration (as indicated by less sustained periods of high-DREMI) would be needed to elicit a cellular response. On the other hand, naïve cells, which are experiencing antigen for the first time, may need to mount a more cautious response, with the signaling cascade transmitting information carefully for longer periods of time to avoid spurious activation (12–14, 36).

ERK-knockout validation

Because DREMI analysis determined that the influence of pERK on pS6 is stronger in naïve cells, this predicts that knockout of ERK2 should have a larger impact on pS6 levels in naive cells (Fig. 8A), and this difference should be detected at the time the DREMI of this edge increases, 2 min after stimulation. We used the edge-response function and the measured difference in pERK abundance to approximate the expected change in pS6 levels at 2 min after knockout with

Fig. 8. Validation of differences in edge-strength using an inducible Erk2-knockout mouse. Data are from CD4+ naïve T lymphocytes from both B6 and Erk2-knockout mice with CD3 and CD28 cross-linking. (**A**) DREMI is used to compare the strength of the pERK-pS6 relationship between naïve and effector or memory T cells in B6 mice. Higher DREMI in naïve T cells predicts a bigger impact of Erk2 knockout on pS6 in naïve cells. (**B**) Heatmaps show the amounts of pERK and pS6 in the control and knockout mice cells. (**C**) Edge-response functions for B6 mice at the 2-min time point predict the outcome of a knockout experiment in (I) naïve cells and (II) effector or memory cells. The x axes of both figures are annotated with the measured change in mean pERK levels upon knockout, and the y axis shows the predicted decrease in pS6 levels based on the edge-response function. (**D**) I and II show two replicates demonstrating the change in pS6 amounts after ERK2 knockout. The blue bars represent the average difference between pS6 in normal and ERK2-knockout mice cells in naïve T cells. The maroon bars show the same difference in effector or memory cells. The blue bars are significantly higher than the maroon bars, especially after the 2-min time point (as predicted); therefore, our edge strength assessment is validated.
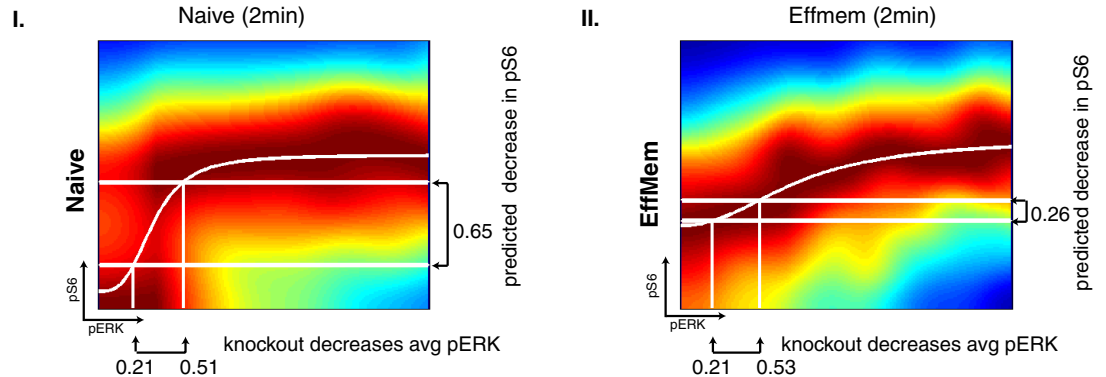

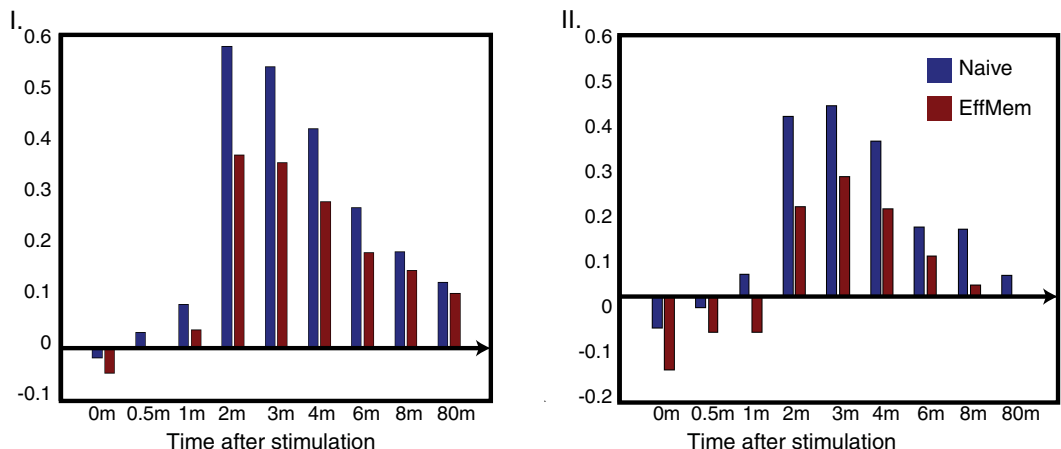**A pS6 depends on pERK more in naive cells**
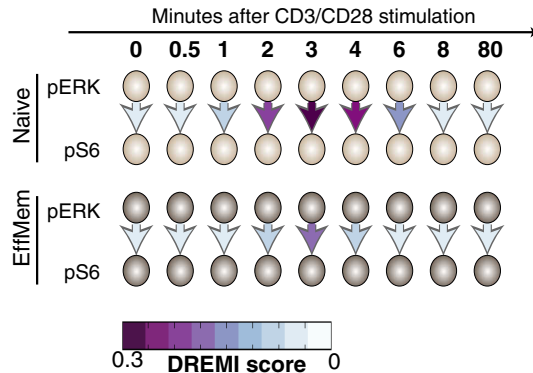
B Genetic deletion effects ERK abundance
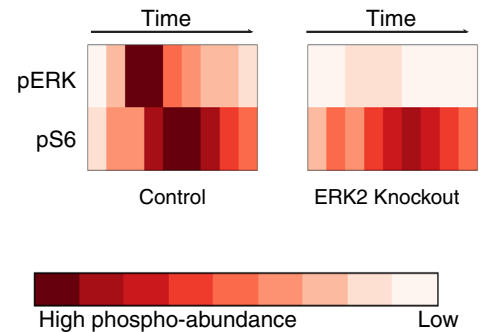
C Prediction of differential knockdown impact on pS6 using DREVI and edge response functions

D Change in pS6 levels under ERK knockdown

an expectation to observe a change that is ~50 to 60% larger in naïve cells relative to effector/memory cells at 2 min (Fig. 8, B and C). To validate these predictions, we collected data from a mouse model (*37*), which enables tamoxifen-induced knockout of ERK2 within a normally differentiated T cell population. To allow for direct comparison, immune cells from both the B6 and ERK knockout mice were stimulated and measured in the same tube and subsequently identified through congenic markers CD45.1 and CD45.2.

To test the hypothesis that naïve T cells showed a stronger relationship between pERK and pS6, based on the relatively higher DREMI in naïve cells, we assessed the impact of the Erk2 depletion on pS6 levels in both naïve and antigen-exposed cells. As predicted, changes in pS6 abundance are larger in naïve cells than in antigen-exposed cells. Approximately 50% greater response in naïve cells began at 2 min, as predicted (Fig. 8D). DREMI also predicts that the influence of pERK on pS6 is stronger in CD4$^+$ T cells than CD8$^+$ T cells (fig. S18), which was also confirmed experimentally (fig. S19), demonstrating that DREMI can correctly predict differences in signaling edge strengths between related T cell subsets.

Discussion

DREVI and DREMI combined with single-cell data provide solutions to a challenging problem that has obstructed the analysis of single-cell biology: that of quantifying the strength of the underlying complex relationships between proteins from data. Other approaches such as correlation, mutual information, and MIC—which are typically used to quantify the strengths of interactions—are not always reliable in the face of technical noise, natural stochasticity, bimodality, spurious correlations, and nonlinearity inherent in single-cell data. DREMI, on the other hand, consistently quantifies the relative strengths of relationships and how they change after stimulation through time and between different cell types.

We used DREVI and DREMI to functionally and quantitatively analyze edges within intracellular protein networks. DREVI and DREMI directly regard the single-cell population as instantiations of an underlying signal processing circuit and infer the form and strength of each edge in the circuit. Many DREVI plots have a sigmoid (or partial sigmoid) response curve, suggestive of the signaling kinetics. With absolute quantification of cytometry data, it would be possible to derive biochemical rate constants for multiple enzymatic pairs and multiple subsets in a single experiment. This would provide an enormous advantage over laborious experiments currently needed to compute such rate constants. Therefore, a versatile, data-driven approach such as that represented here enables the systematic and quantitative comparison of signaling relationships between different stimuli, time points, and cell types.

The analysis in this manuscript focused on mass cytometry data collected from mouse T cells measuring 11 phosphoproteins and nine surface markers. The analysis determined that the TCR signaling network is dynamic, with edges activating and deactivating quickly as the signals are transduced downstream. We used DREMI and DREVI to gain insights into how signaling is fine-tuned between related T cell populations. The differences we identified in TCR signaling between naïve and antigen-exposed T cells suggest that naïve cells more sensitively transmit upstream signaling inputs along a central signaling cascade. This more fine-tuned signaling response might help regulate the cell's decision to differentiate into either regulatory or helper T cells. In contrast, entrained effector or memory cells seem poised for fast responses upon secondary exposures. To explain this observation, we noted that several alternative pathways had higher activity in antigen-exposed cells, including AKT under CD4 costimulation, which likely bolsters phosphoprotein levels and shortens input integration time in antigen-exposed cells.

Cumulatively, these techniques can allow the tracking of subtle reconfigurations in signaling that occurs between related cell subtypes. These reconfigurations do not result in an essentially rewired network, but rather a network that is behaviorally reconfigured to produce changes in response. Although we have presented a specific application of our techniques, focused on TCR activation in T cell subsets, we believe that DREVI and DREMI are broadly applicable across biological systems and single-cell technologies (*38*, *39*). Our ability to characterize signaling edges and their strengths is of particular value for disease studies because it has been shown that many disease alleles affect the edge strengths of the signaling network (*40*). In (*16*), we report the direct application of these methods to identify and characterize the impact of complex genetic variation associated with type-1 diabetes on TCR signaling networks. As single-cell data become more abundant, our methods will allow construction of more quantitative models of cellular signaling. Such models can then elucidate how these signaling circuits differ between cell types as well as characterize essential and therapy-targetable differences between health and disease.

Materials and methods

Mice

Mice were maintained in specific pathogen-free facilities at Harvard Medical School (Institutional Animal Care and Use Committee 99–20, 02954). Age-matched, 4- to 6-week-old male B6^{g7} mice were used for all experiments. For TCR signaling responses in peripheral T cells from TCR transgenic mice, BDC2.5 and BDC2.5/B6^{g7}.Rag$^{-/-}$ mice were used. This mouse strain was previously described (*41*).

Cell stimulation

For analysis of phosphorylation events by means of mass cytometry, total lymph node cell suspensions from B6^{g7} were used. Cells were stimulated in medium containing 2 or 6 μg/ml of biotinylated stimulatory antibodies to CD3e and CD28 and incubated for 2 min at 37°C before addition of 8 or 24 μg/ml streptavidin. At various times

after cross-linking, the stimulation was stopped by addition of paraformaldehyde (final concentration 2%) and incubation at room temperature for 20 min. In some experiments, biotinylated antibodies to CD4 were also included in the stimulation mix. The cells were measured before stimulation and at various time points after stimulation: 30 s, 1, 2, 3, 4, 6, 8, 10, 20, and 80 min.

Erk knockout experiment

To study the role of ERK in peripheral T cells, we used an inducible model that was developed in (*37*) with the Erk2 floxed-gene and ER-CRE, which is a tamoxifen-inducible adenosine 3′,5′-monophosphate (cAMP) response element (CRE)–recombinase. Adult mice are injected with tamoxifen, every day (Sigma-Aldrich, St. Louis, MO) for 6 consecutive days, to induce the deletion of the floxed gene and create ERk2-deficient cells. The mice were analyzed 2 days later. There are normal numbers of peripheral T cells in these mice because they are analyzed only 2 days after the tamoxifen-induced deletion. Therefore, these peripheral cells are a result of normal differentiation and transportation from the thymus weeks before the tamoxifen treatment injections (*37*) but have ERK2 knocked-down for the TCR activation experiment.

Normal B6 and Erk knockout cells express different congenic CD45 surface antigens (Erk-knockout mice express CD45.1, and normal B6 express CD45.2), allowing for the stimulation of both cell types in the same tube. B6 and Erk knockout cells are therefore mixed before stimulation and identified after analysis with antibodies against CD45.1 and CD45.2. This strategy has the considerable advantage of eliminating any contribution of experimental variations during the stimulation, staining, and CyTOF processing of the samples (for each time point analyzed).

Measurement of phospho-signaling events with mass cytometry

In our experiments, we used mass cytometry to measure surface and internal proteins. Primary conjugates of the mass cytometry antibodies (complete list is provided in table S1) were prepared and titrated as described (*8*). Paraformaldehyde (PFA) fixed and frozen lymphocyte suspensions were thawed at room temperature and then transferred to ice.

Single samples of 1 million cells were stained in a 50 μl final volume. All samples from a given time series of a stimulation-condition were run in a single barcoded tube. Each barcoded tube contained 20 individual samples and was stained in 300 μl. Cells were stained with a cocktail of metal isotope conjugated antibodies (table S1) by using a two-step procedure (surface stains included CD4, CD8, TCRb, CD5, CD19, CD25, CD44, CD45.1, and CD45.2; intracellular stains included pCD3ζ, pSLP76, pERK1/2, pS6, pCREB, pMAP-KAPKII, IkBa, pNFkB, pRB, pFAK, and pAKT). The signaling network consisting of all markers is shown in fig. S1. Cell fixation, permeation, and staining were performed as previously described (*8*). To make all samples maximally comparable, all data was acquired by using internal metal

isotope bead standards as previously described in (*42*). Approximately 100,000 events were acquired per sample on a CyTOF I (DVS Sciences, Sunnyvale, CA) by using instrument settings previously described in (*42*).

Barcoding of samples

To reduce sample-to-sample variability, cells corresponding to each stimulation condition and time point were barcoded by using all six stable palladium isotopes with a method similar to that described by (*43*). In the barcoding scheme, each isotope is used in binary fashion, high for "1" and low for "0." Hence, the barcode itself is a binary string that can be decoded for each cell to determine the treatment condition for that cell. For this purpose, we developed an error-correcting 3-one barcoding scheme that both facilitates ease of debarcoding, and eliminates the majority of doublets. All legal barcodes in our scheme have exactly 3 ones and 3 zeroes, resulting in 20 possible barcodes. This barcoding scheme also enables the detection of most cell doublets because in a well-mixed tube, it is unlikely to have a doublet created by two cells from the same condition. Doublets with cells from two different conditions would have an illegal combined barcode (with more than 3 "1"s) that is represented by the logical OR of the individual barcodes. Therefore, this scheme eliminates almost all doublet events. Further, each pair of barcodes has a distance of at least 2, and therefore, this scheme is also robust to errors during the debarcoding.

Post-processing of mass cytometry data

Individual time series were normalized to the internal bead standards before analysis by using the method described in (*42*). In addition, as described in (*8*), for all analysis of the mass cytometry data, abundance values reported by the mass cytometer were transformed by using a scaled arcsinh, with scaling factor 5 to minimize the noise in measurements where the data are close to zero (the limit of detection on the mass cytometry scale). Events recorded by the mass cytometer were gated based on cell length and DNA content (to avoid debris and doublets), as described in (*8*).

Gating of cell subsets

Cells were then filtered through a series of gates, as depicted in fig. S2, for the purpose of isolating particular subpopulations of T cells such as CD4$^+$ naïve T cells, CD8$^+$ naïve T cells, and CD4$^+$ effector/memory T cells. These gates include (gate 1) a gate that separates T cells (TCRb$^+$ CD19$^-$) from B cells (TCRb$^-$ CD19$^+$), (gate 2) a gate that further separates CD4$^+$ T cells (CD4$^+$) from CD8$^+$ T cells (CD8$^+$), (gate 3) a gate that further separates CD4$^+$ T cells into conventional T cells (T$_{conv}$ cells) (CD25$^-$) and regulatory T cell (T$_{reg}$ cell) (CD25$^+$) populations, and (gate 4) a gate that further separates CD4$^+$ T$_{conv}$ cells and CD8$^+$ T cells into naïve (CD44$^-$) and antigen-exposed effector/memory populations (CD44$^+$). Because there is no known CD8$^+$ equivalent of the CD4$^+$CD25$^+$ T$_{reg}$ population, CD8$^+$ cells are not passed through gate 3. The final populations that we isolate and

use in our comparisons include (i) naïve CD4$^+$ T cells (CD44$^-$ CD25$^-$ CD4$^+$ CD8$^-$ TCRb$^+$ CD19$^-$), (ii) naïve CD8$^+$ T cells (CD44$^-$ CD25$^-$ CD8$^+$ CD4$^-$ TCRb$^+$CD19$^-$), and (iii) Effector/memory CD4$^+$ T cells (CD44$^+$ CD25$^-$ CD4$^+$ CD8$^-$ TCRb$^+$ CD19$^-$). The nodes, cell types, stimulation, and time points used in each figure are shown in table S2.

Data availability

All data are available for download at www.c2b2. columbia.edu/danapeerlab/html/dremi.html.

Description of computational methods

In the following sections, we provide a detailed explanation of how DREVI, DREMI, and the edge-response functions are computed,

Kernel density and conditional density estimation

A kernel density estimate $\hat{f}(x)$ is a nonparametric approximation of the probability density function $p(x)$. The standard method for computation of a kernel density estimate is to introduce a narrow Gaussian (or alternative) kernel at each point in the raw data and calculate the integral of the individual kernel values over a large set of points (*21*, *44*). If (x_1, x_2, x_3, ... x_n) are independent, identically distributed random samples from an unknown probability density function f, then the kernel density estimator is

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) \quad (1)$$

where K is the kernel function. The commonly used Gaussian kernel is given by

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x^2}{2}\right)} \quad (2)$$

K_h refers to the kernel density function scaled by $1/h$ (where h denotes the bandwidth). Intuitively, one can think of kernel density estimation as placing a small smooth kernel K_h at each sampled point and integrating over all such kernels to determine a smooth probability density function. The kernel density estimate in two dimensions is given by

$$\hat{f}_{h1h2}(x, y) = \sum_{i=1}^{n} K_{h1}(x - x_i) K_{h2}(y - y_i) \quad (3)$$

This standard method of kernel density estimation suffers from several deficiencies, including (i) sensitivity to selected bandwidth and the assumption of normalcy in many bandwidth estimation methods, (ii) boundary bias, and (iii) over- or under-smoothing.

Instead, we use the heat-diffusion method from (*10*) for two-dimensional (2D) kernel density estimation. This method views the kernel density estimate as a solution to the heat equation, which it evolves for a time period proportional to the bandwidth. This interpretation of a kernel density estimate derives from the notion of a "Weiner" process. A Weiner process W is a stochastic Markov process (the next state can be directly computed by the previous state) with

(i) Initial probability uniformly spread through the d-dimensional data points (x_1, x_2, x_3, ... x_n).

(ii) The transition probability of going from point x_i to x_j is given by the Gaussian kernel $p_{trans}(x_i, x_j) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x_j - x_i)^2}{2t}}$

The Gaussian kernel density estimation (KDE) can be interpreted as the probability distribution function (PDF) for this process at time t, which is the same as Eq. 2 with bandwidth h: $f_t(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x - x_i)^2}{2t}}$

But because this is an iterative process, the transition satisfies the following differential equation, which is also known as the "heat equation": $\frac{\partial \hat{f}(x;t)}{\partial t} = \frac{1}{2}\frac{\partial^2 \hat{f}(x;t)}{\partial x^2}, t > 0$

The initial condition is given by $\hat{f}(x; 0) = \frac{1}{n}\sum_{i=1}^{N} \delta_{x_i}(x)$, where δ is the dirac delta function (which gives point-masses to all data points). The advantage of writing this in PDF form is that if the PDF is known to be 0 outside the range, then boundary conditions enforcing this can be added, which results in a θ kernel rather than a Gaussian kernel:

$$\hat{f}(x; t) = \frac{1}{n}\sum_{i=1}^{n} \theta(x, x_i; t)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{k=-\infty}^{\infty} e^{-k^2\pi^2 t/2}\cos(k\pi x)\cos(k\pi x_i)$$
$$(4)$$

The θ kernel is locally adaptive and has the property that it behaves like a Gaussian kernel for small bandwidths and like a uniform kernel for large bandwidths. The evolution of the heat equation decides the shape of the kernel rather than having a fixed Gaussian estimate. This often results in a smoother density function (fig S3). Additionally, (i) it can be solved quickly by using fast Fourier transform methods, (ii) it can be extended to a more general diffusion equation easily, and last (iii) the bandwidth t can be solved as a fixed point solution to a recursion without a normal assumption on the distribution. The superior performance and speed of computation of the heat-equation–based KDE is shown in fig. S3.

To obtain the standard conditional density estimate, the joint density estimate $\hat{f}(x, y)$ is divided by the marginal density estimate $\hat{f}(x)$ as follows (*45*):

$$\hat{f}(y|x) = \frac{\hat{f}(y, x)}{\hat{f}(x)} \quad (5)$$

We use the conditional density estimate as a starting point for DREVI and DREMI as explained in the following sections.

Conditional density-rescaled visualization

DREVI is aimed at visualizing the stochastic function representing how X influences Y by depicting the distribution of Y for each value of X. To visualize 2D relationships in single-cell data, we begin by computing a 2D kernel density estimate as described in Eq. 4 (*10*) on a square mesh grid. The mesh grid of points is denoted $G = \{(x_i, y_j), 1 < i < n, 1 < j < m\}$, and the 2-D kernel density estimate is a matrix $D = \hat{f}(G)$ where $D(i, j)$ corresponds to $\hat{f}(x_i, y_j)$. The mesh grid of points G is selected so that the X-Y ranges are restricted to regions that are well-populated by

cells. Most markers (such as phosphoproteins) are distributed among the measured cells as heavy-tailed multimodal distributions (fig. S7C). For any pair of markers X and Y, we remove 1% of both tails of these distributions as outliers, removing ~200 cells from a population of ~10,000 cells (per condition and time point). The remaining range is generally well populated by cells and is amenable for a robust density estimate, which results in robust visualization and scoring, as shown in fig. S7.

After the 2D kernel density estimate is computed, each column of D (corresponding to a fixed X-value, $X = x_i$) is renormalized by the marginal density estimate of $X = x_i$ as shown in Eq. 5, $\hat{f}(x)$ for that column to obtain the conditional density matrix

$$\Delta(x_i, y_j) = \hat{f}(y_j|x_i) \quad (6)$$

This is an estimate of $P(Y|X)$ on the same mesh grid. This matrix is rescaled for the purposes of visualization as follows:

$$\hat{f}^*(y_j|x_i) = \hat{f}(y_j|x_i)/\max_k[\hat{f}(y_k|x_i)] \quad (7)$$

We denote the matrix of rescaled kernel density estimates as

$$\Delta^*(x_i, y_j) = \hat{f}^*(x_i|y_j) \quad (8)$$

The matrix of Eq. 8 is visualized as a heatmap using the "jet" color scale, where densest regions, corresponding to the highest rescaled conditional density, are maroon/red and the sparsest regions are blue. This way, the X-Y relationship is revealed even in regions populated by fewer cells because each X-slice is renormalized individually. The computation of DREVI is demonstrated Fig. 1.

Conditional Density-Resampled Mutual Information (DREMI)

We developed a method for scoring relationships termed DREMI. Like MI, this will depend on how well the value of X can predict the value of Y. However, unlike MI, we are not interested in how the entropy of Y decreases with X in a given sample on average. Instead, we are interested in looking at the underlying stochastic function $F(X) = Y$, with the aim of learning the strength of this relationship over the entire dynamic range of X, regardless of the distribution in the specific sample.

The key mathematical insight in DREMI is that we provide a score for the conditional distribution of $P(Y|X)$ rather than joint distribution $P(X, Y)$. DREMI resamples from the conditional density estimate as described above, taking care to generate a uniform number of samples from each column of the matrix Δ^*, so that the relationship between X-and-Y is scored through the full dynamic range of X that is well sampled by the data.

DREMI begins with the rescaled conditional density matrix Δ^* described in Eq. 8. To filter out noise, we round values of $\Delta^*(x_i, y_j) < \varepsilon$ to zero and eliminate the corresponding points from the grid G. We generate n new points $S_i = \{s_i^1, s_i^2, s_i^3, ..., s_i^n\}$ from the density distribution in each column i of Δ^*, denoted Δ_i^*. We generate a total of

$n \times m$ samples $S = \{S_1, S_2, ... S_m\}$. Because points in each column of Δ_i^* have a similar value $X = x_i$, we are generating an equal number of samples from the entire range of X (up to the fine-mesh grid level of granularity).

Last, we define a discrete partition of the X and Y axes and compute mutual information $I(X:Y)$ as in Eq. 10 using the samples in S. We generally use eight equi-partitions of the X and Y axes to compute the mutual information on these samples, but our software can use any number of equi-partitions based on a user-defined parameter. Our partitions are formed by pooling all data from conditions being compared for a particular variable, taking the maximal range of the X and Y and partitioning that space into eight equal units along each axis. We normalize DREMI by the log base 2 of the number of partitions to provide a score that lies between 0 and 1.

To derive the mathematical expression for DREMI, we start with standard mutual information, computed as the differential entropy between Y and $Y | X$:

$$\begin{aligned} I(Y;X) &= H(Y) - H(Y|X) = \\ &-\sum_j p(y_j)\log[p(y_j)] + \\ &\sum_i \sum_j p(x_i, y_j)\log[p(y_j|x_i)] = \\ &\sum_i \sum_j p(x_i, y_j)\log[p(y_j)] + \\ &\sum_i \sum_j p(x_i, y_j)\log[p(y_j|x_i)] \end{aligned} \quad (9)$$

Alternatively, MI can be regarded as a measure of the distance between the joint distribution $P(X, Y)$ and the marginal distributions $P(X)$ and $P(Y)$. An equivalent formula for computing mutual information is given by ([26])

$$I(x, y) = \sum_i \sum_j p(x_i, y_j)\log\left[\frac{p(x_i, y_j)}{p(x_i)p(y_j)}\right] \quad (10)$$

The first three steps of the DREMI algorithm involve estimating the conditional density by normalizing the joint kernel density estimate and filtering out noise. These steps result in a smoothed and de-noised conditional density estimate that will fill gaps in the raw data. The final step involves estimating MI from this computed conditional density. In fig. S20, we show intuitively how sampling from the conditional density estimate is equivalent to taking weighted samples from the joint distribution. Each point (x_i, y_j) from the joint distribution has sample weight proportional to $\frac{1}{p(x_i)}$.

Given this, DREMI can be written in terms of the differential entropy formulation of mutual information (Eq. 8) as follows:

$$I^c(Y|X) = H^c(Y) - H^c(Y|X) \quad (11)$$

where $H^c(Y)$ and $H^c(Y|X)$ are reweighted entropies so that each point (x_i, y_j) in the joint distribution is weighted by $\frac{1}{p(x_i)}$ given as follows:

$$H^c(Y) = -\sum_i \sum_j \frac{p(x_i, y_j)}{p(x_i)}\log[p(y_j)] \quad (12)$$

$$H^c(Y|X) = -\sum_i \sum_j \frac{p(x_i, y_j)}{p(x_i)}\log[p(y_j|x_i)] \quad (13)$$

Substituting Eq. 13 and Eq. 12 into Eq. 11 gives

$$\begin{aligned} I^c(Y|X) &= -\sum_i \sum_j \frac{p(x_i, y_j)}{p(x_i)}\log[p(y_j)] - \\ &\left\{-\sum_i \sum_j \frac{p(x_i, y_j)}{p(x_i)}\log[p(y_j|x_i)]\right\} \\ &= \sum_i \sum_j \frac{p(x_i, y_j)}{p(x_i)}\log[p(y_j|x_i)] - \log[p(y_j)] \\ &= \sum_i \sum_j \frac{p(x_i, y_j)}{p(x_i)}\log\left[\frac{p(y_i|x_j)}{p(y_j)}\right] \\ &= \sum_j \sum_i \frac{p(x_i, y_j)}{p(x_i)}\log\left[\frac{p(x_i, y_j)}{p(x_i)p(y_j)}\right] \\ &= \sum_j \sum_i p(y_i|x_j)\log\left[\frac{p(x_j, y_i)}{p(x_j)p(y_i)}\right] \end{aligned}$$

$$(14)$$

The final form in Eq. 14 is equivalent to Eq. 10, with each point reweighted by $\frac{1}{p(x_i)}$. Hence, I^c maintains many properties of mutual information, including the fact that $I^c(Y|X) = 0$ if $p(x, y) = p(x)p(y)$.

Robustness analysis

We demonstrate that DREVI and DREMI are robust methods. As shown in fig. S7, the downsampling of the data by 20% results in virtually identical DREVI plots and very similar DREMI scores (SD of 0.02 or less) for several edges. The reason underlying this robustness is our use of density estimation with a locally adaptive smoothing that results from the evolution of a differential equation whose solution is equivalent to the kernel density estimate ([10]). The density estimation method interpolates and fills gaps within the data by integrating using a smooth kernel.

Fitting edge-response functions

For strong relationships, $P(Y|X)$ is highly peaked for any value of X. Moreover, the peak of $P(Y|X)$ for adjacent X-slices tend to follow along a smooth and continuous curve, the "edge-response function," which we fit to the data. We derive the edge-response function by fitting points sampled from the noise-filtered conditional density Δ^* to one of three models using regularized regression: linear, sigmoidal, or double sigmoidal ([27]). If none of the models obtain significant fits (based on P values), we fit to a free-form curve (examples are provided in Figs. 1C, 3, and 5C).

We define the edge-response function as the function $f(X)$ that best fits the data. We derive the edge-response function by performing nonlinear, mixed-model regression ([46]) of the high-valued regions of the conditional-density estimate. The models we use are

(i) linear: $f_1(x) = ax + b$
(ii) sigmoidal: $f(x) = [(L - U)/1 + (x^n/K_d)] + U$
(iii) double-sigmoidal: $f(x) = \{[(L_1 - U_1)/1 + (x^n/K_{d1})] + U_1\} + \{[(L_2 - U_2)/1 + (x^n/K_{d2})] + U_2\}$
(iv) free-form curve

We minimize over both the models and parameters in (i) to (iii). If no model results in

sufficiently high significance of fit, we fit to a nonparametric curve.

Method availability

The methods described, along with an interactive graphic user inteface (GUI) are available for download at www.c2b2.columbia.edu/danapeerlab/html/dremi.html.

Performance of DREMI on synthetic data

We used synthetic data to evaluate how DREMI behaves under increasingly noisy conditions (fig. S6). We generated the synthetic databases on characteristics of real pairwise relations observed in single-cell cytometry data:

(i) X is concentrated in a narrow band within the entire X-range.

(ii) The high-density region of the conditional distribution reveals a consensus of response that is smooth and can be fit to a curve.

(iii) The noise in Y is variable.

To fulfill these conditions, we generated data as follows:

For condition (i), the X is a random variable that follows the distribution normal distribution with SD 0.25: $N(2, 0.25)$.

For condition (ii), the mean Y is functionally related to X: $E(Y \mid X = x) = f(x)$. We use linear, sigmoidal, and quadratic functional relationships in the synthetic data.

For condition (iii), Y has varying degrees of Gaussian noise added on top of its mean value: $Y = f(X) + N(0, s)$, for parameter s.

Each synthetic data set contained ~20,000 points. We generated synthetic data following the above conditions for five different values of noise: $s = \{0.25, 0.5, 0.75, 1, 1.25\}$. As shown in fig. S6, the results of computing DREMI on this synthetic data show that DREMI scores decrease roughly linearly with the SD of the Gaussian noise, regardless of the underlying functional shape of the relationship.

In addition to the analysis described in the main text, we also evaluate how varying the noise-filtering threshold influences the results. Also shown in fig. S6 is how the noise-filtering threshold ε can be used to eliminate experimental noise. Depicted in fig. S6B is a horizontal line that equalizes the scores of all of the noise levels. Thus, the value of ε can be tuned to compensate for experimental or measurement noise.

There are certain settings that provide naturally noisier signals than others. For instance, certain antibodies can have weak binding affinities with their target proteins, others can nonspecifically bind to a variety of proteins. In such cases, a higher level of the outlier-elimination threshold may be useful. In other situations, such as the pSLP76-pERK edge, the response of the protein itself is stochastic, and therefore, the stochasticity itself has meaning. In such cases, the outlier-elimination threshold can be set higher to capture the natural stochasticity in the data. Other global contexts, such as measurements on primary cells as opposed to cell lines, lead to higher noise and weaker response. In these cases, it may be useful to set the noise-filtering threshold

to a particular global value. We use a globally fixed noise-filtering threshold of 0.85 for all figures and DREMI values computed for this manuscript. As shown in fig. S6, within any set noise-filtering threshold, the DREMI score decreases linearly with the SD of noise, and as such is able to relatively rank DREMI scores for comparison purposes.

Comparing DREMI to other methods

We compared the performance of DREMI against other commonly used dependency measures: (i) MIC (*32*), (ii) adaptively partitioned MI (*47*), and (iii) Pearson correlation (Fig. 5). Because no gold standard exists to compare strength of signal transfer along an edge, we created a framework to conduct the comparisons based on the dynamics of signaling. We used DREVI to visually determine the peak timing of the dynamic signaling response, and in cases in which peak timing is known, our visual inspection matched published literature (*2*). We then assess whether each of the four metrics can recover the peak timing and also order the time points correctly according to strength of signal transfer. We systematically computed all of the metrics along the time series in each of the six edges (Fig. 5I and fig S10). DREMI outperformed all other metrics.

To show that DREMI is able to quantify and rank relationships in settings other than TCR-stimulated mice cells, we also compared DREMI on three known signaling edges on previously published B cell data (*8*) from the bone marrow of healthy human donors after BCR (B cell receptor) stimulation (fig. S11). These data only have a single time point after BCR stimulation, therefore instead of ranking edge strength along time points, we rank edge strength across increasingly mature B cell populations. The five populations, in order of maturity, are (i) Pre-BI, (ii) Pre-BII, (iii) Immature B cells, (iv) Mature CD38+ B cells, and (v) Mature CD38⁻ B cells. The cells are gated into these populations based on phenotypic B cell markers such as CD20, CD38, CD45RA, CD19, and CD123, as described in (*8*). We used prior biological knowledge to determine the cell population expected to have strongest signaling, determined before computing the metrics. DREMI outperformed all other metrics on this data.

An important goal of MIC is "equitability." Intuitively, this means that different types of edge shapes are given similar scores if they are similarly "fuzzy" (*33*). Evaluating MICs' performance, this does not seem to be a suitable assumption for molecular relations. Adaptive partitioning puts more weight into the densest part of the distribution. Although this might make sense, in essence, it makes a large number of microscopic partitions over the small range where most of the density lays. Thus, the partitions often represent random measurement noise rather than biological signal. Pearson correlation prefers linear dependencies to other types of dependencies, whereas nonlinear dependencies are frequently present in signaling relationships.

In Fig. 5, II, we compare DREMI on synthetic data, where the generating function is known. Here, we generate a weak linear relationship with low slope compared with the range of the data, $Y = aX + N(0, 0.1)$, to compare against strong nonlinear relationships:

(i) A quadratic relationship with a high degree of noise, $Y = -AX^2 + N(0, 0.75)$.

(ii) A sigmoidal relationship X-Y with both Gaussian noise and uniform noise in the range $[0, Y]$, for any value of Y: $Y = \text{sigmoid}(X) + N(0, 0.75) + U[0,\text{sigmoid}(X)]$, where U is a uniform random function.

In both cases, the noisy-yet-strong relationships scored lower than the weak linear relationship, with low noise for all metrics other than DREMI.

Evaluating edge-response functions on raw verses resampled conditional data

In fig. S8, we show that the data resampled from the conditional density is more suitable for curve-fitting than the raw data. In the first two columns, we use linear, sigmoidal, or double sigmoidal least-squares regression on the raw mass cytometry data. In the third column, we resample from the dense regions of the conditional density and perform the same types of regression on this resampled data. Last, in the fourth column we perform regression on the conditional mean (computed from the conditional density).

In each case, the best fit on the raw data has higher RMSE than the best-fit data that is resampled from the conditional density (measured on the raw data). In addition, the linear fit always results in a better fit on the raw data, whereas the conditional data result in a variety of shapes linear, sigmoidal, and double sigmoidal. The raw data tend to be concentrated in a small region of the X-Y plane, resulting in a weak penalty for deviations in most regions of the X-Y plane, and a heavy penalty for deviation only in the region of data concentration. Therefore, a line that passes through the concentrated region of the raw data generally results in the best fit.

REFERENCES AND NOTES

1. G. Iezzi, K. Karjalainen, A. Lanzavecchia, The duration of antigenic stimulation determines the fate of naive and effector T cells. *Immunity* **8**, 89–95 (1998). doi: 10.1016/S1074-7613(00)80461-6; pmid: 9462514
2. A. Lanzavecchia, G. Iezzi, A. Viola, From TCR engagement to T cell activation: A kinetic view of T cell behavior. *Cell* **96**, 1–4 (1999). doi: 10.1016/S0092-8674(00)80952-6; pmid: 9989490
3. N. Miskov-Zivanov, M. S. Turner, L. P. Kane, P. A. Morel, J. R. Faeder, The duration of T cell stimulation is a critical determinant of cell fate and plasticity. *Sci. Signal.* **6**, ra97 (2013). doi: 10.1126/scisignal.2004217; pmid: 24194584
4. M. S. Turner, L. P. Kane, P. A. Morel, Dominant role of antigen dose in CD4⁺Foxp3⁺ regulatory T cell induction and expansion. *J. Immunol.* **183**, 4895–4903 (2009). doi: 10.4049/jimmunol.0901459; pmid: 19801514
5. G. Altan-Bonnet, R. N. Germain, Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLOS Biol.* **3**, e356 (2005). doi: 10.1371/journal.pbio.0030356; pmid: 16231973
6. Y. E. Antebi *et al.*, Mapping differentiation under mixed culture conditions reveals a tunable continuum of T cell fates. *PLOS Biol.* **11**, e1001616 (2013). doi: 10.1371/journal.pbio.1001616; pmid: 23935451
7. D. R. Bandura *et al.*, Mass cytometry: Technique for real time single cell multitarget immunoassay based on inductively

coupled plasma time-of-flight mass spectrometry. *Anal. Chem.* **81**, 6813–6822 (2009). doi: 10.1021/ac901049w; pmid: 19601617

8. S. C. Bendall *et al.*, Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696 (2011). doi: 10.1126/science.1198704; pmid: 21551058

9. M. H. DeGroot, M. J. Schervish, *Probability and Statistics* (Addison-Wesley, Boston, ed. 4, 2012).

10. Z. I. Botev, J. F. Grotowski, D. P. Kroese, Kernel density estimation via diffusion. *Ann. Stat.* **38**, 2916–2957 (2010). doi: 10.1214/10-AOS799

11. J. Zhu, H. Yamane, W. E. Paul, Differentiation of effector CD4 T cell populations. *Annu. Rev. Immunol.* **28**, 445–489 (2010). doi: 10.1146/annurev-immunol-030409-101212; pmid: 20192806

12. M. Berard, D. F. Tough, Qualitative differences between naïve and memory T cells. *Immunology* **106**, 127–138 (2002). doi: 10.1046/j.1365-2567.2002.01447.x; pmid: 12047742

13. M. R. Chandok, D. L. Farber, Signaling control of memory T cell generation and function. *Semin. Immunol.* **16**, 285–293 (2004). doi: 10.1016/j.smim.2004.08.009; pmid: 15528073

14. M. K. MacLeod, J. W. Kappler, P. Marrack, Memory CD4 T cells: Generation, reactivation and re-assignment. *Immunology* **130**, 10–15 (2010). doi: 10.1111/j.1365-2567.2010.03260.x; pmid: 20331469

15. K. Sachs, O. Perez, D. Pe'er, D. A. Lauffenburger, G. P. Nolan, Causal protein-signaling networks derived from multiparameter single-cell data. *Science* **308**, 523–529 (2005). doi: 10.1126/science.1105809; pmid: 15845847

16. M. Mingueneau *et al.*, Single-cell mass cytometry of TCR signaling: Amplification of small initial differences results in low ERK activation in NOD mice. *Proc. Natl. Acad. Sci. U.S.A.* (2014). doi: 10.1073/pnas.1419337111

17. B. J. Baaten, C. R. Li, L. M. Bradley, Multifaceted regulation of T cells by CD44. *Commun. Integr. Biol.* **3**, 508–512 (2010). doi: 10.4161/cib.3.6.13495; pmid: 21331226

18. B. J. Baaten *et al.*, CD44 regulates survival and memory development in Th1 cells. *Immunity* **32**, 104–115 (2010). doi: 10.1016/j.immuni.2009.10.011; pmid: 20079666

19. J. E. Smith-Garvin, G. A. Koretzky, M. S. Jordan, T cell activation. *Annu. Rev. Immunol.* **27**, 591–619 (2009). doi: 10.1146/annurev.immunol.021908.132706; pmid: 19132916

20. R. J. Brownlie, R. Zamoyska, T cell receptor signalling networks: Branched, diversified and bounded. *Nat. Rev. Immunol.* **13**, 257–269 (2013). doi: 10.1038/nri3403; pmid: 23524462

21. M. Rosenblatt, Remarks on some nonparametric estimates of a density-function. *Ann. Math. Stat.* **27**, 832–837 (1956). doi: 10.1214/aoms/1177728190

22. P. Hall, J. Racine, Q. Li, Cross-validation and the estimation of conditional probability densities. *J. Am. Stat. Assoc.* **99**, 1015–1026 (2004). doi: 10.1198/016214504000000548

23. G. A. Darbellay, I. Vajda, Estimation of the information by an adaptive partitioning of the observation space. *IEEE Trans. Inf. Theory* **45**, 1315–1321 (1999). doi: 10.1109/18.761290

24. A. A. Margolin *et al.*, Reverse engineering cellular networks. *Nat. Protoc.* **1**, 662–671 (2006). doi: 10.1038/nprot.2006.106; pmid: 17406294

25. I. S. Jang, A. Margolin, A. Califano, hARACNe: Improving the accuracy of regulatory model reverse engineering via higher-order data processing inequality tests. *Interface Focus* **3**, 20130011 (2013). doi: 10.1098/rsfs.2013.0011; pmid: 24511376

26. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, Hoboken, NJ, ed. 2, 2006).

27. T. Hastie, R. Tibshirani, J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction: With 200 Full-Color Illustrations* (Springer, New York, 2001).

28. M. N. Artyomov, M. Lis, S. Devadas, M. M. Davis, A. K. Chakraborty, CD4 and CD8 binding to MHC molecules primarily acts to enhance Lck delivery. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16916–16921 (2010). doi: 10.1073/pnas.1010568107; pmid: 20837541

29. A. Zaslaver *et al.*, Just-in-time transcription program in metabolic pathways. *Nat. Genet.* **36**, 486–491 (2004). doi: 10.1038/ng1348; pmid: 15107854

30. C. J. Caunt, S. M. Keyse, Dual-specificity MAP kinase phosphatases (MKPs): Shaping the outcome of MAP kinase signalling. *FEBS J.* **280**, 489–504 (2013). doi: 10.1111/j.1742-4658.2012.08716.x; pmid: 22812510

31. G. Li *et al.*, Decline in miR-181a expression with age impairs T cell receptor sensitivity by increasing DUSP6 activity. *Nat. Med.* **18**, 1518–1524 (2012). doi: 10.1038/nm.2963; pmid: 23023500

32. D. N. Reshef *et al.*, Detecting novel associations in large data sets. *Science* **334**, 1518–1524 (2011). doi: 10.1126/science.1205438; pmid: 22174245

33. J. B. Kinney, G. S. Atwal, Equitability, mutual information, and the maximal information coefficient. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3354–3359 (2014). doi: 10.1073/pnas.1309933111; pmid: 24550517

34. K. A. Janes *et al.*, A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. *Science* **310**, 1646–1653 (2005). doi: 10.1126/science.1116598; pmid: 16339494

35. K. Okkenhaug, Signaling by the phosphoinositide 3-kinase family in immune cells. *Annu. Rev. Immunol.* **31**, 675–704 (2013). doi: 10.1146/annurev-immunol-032712-095946; pmid: 23330955

36. K. Adachi, M. M. Davis, T-cell receptor ligation induces distinct signaling pathways in naive vs. antigen-experienced T cells. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1549–1554 (2011). doi: 10.1073/pnas.1017340108; pmid: 21205892

37. C. F. Chang *et al.*, Polar opposites: Erk direction of CD4 T cell subsets. *J. Immunol.* **189**, 721–731 (2012). doi: 10.4049/jimmunol.1103015; pmid: 22675204

38. C. Giesen *et al.*, Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014). doi: 10.1038/nmeth.2869; pmid: 24584193

39. M. Angelo *et al.*, Quantitative imaging of ten markers in human breast tumors using multiplexed ion beam imaging (MIBI) and metal-tagged antibodies. *Lab. Invest.* **94**, 518A (2014).

40. Q. Zhong *et al.*, Edgetic perturbation models of human inherited disorders. *Mol. Syst. Biol.* **5**, 321 (2009). doi: 10.1038/msb.2009.80; pmid: 19888216

41. M. Mingueneau, W. Jiang, M. Feuerer, D. Mathis, C. Benoist, Thymic negative selection is functional in NOD mice. *J. Exp. Med.* **209**, 623–637 (2012). pmid: 22329992

42. R. Finck *et al.*, Normalization of mass cytometry data with bead standards. *Cytometry A* **83**, 483–494 (2013). doi: 10.1002/cyto.a.22271; pmid: 23512433

43. B. Bodenmiller *et al.*, Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat. Biotechnol.* **30**, 858–867 (2012). doi: 10.1038/nbt.2317; pmid: 22902532

44. E. Parzen, Estimation of a probability density-function and mode. *Ann. Math. Stat.* **33**, 1065–1076 (1962). doi: 10.1214/aoms/1177704472

45. M. Rosenblatt, in *Multivariate Analysis II*, K. P.R., Ed. (Academic Press, New York, 1969), pp. 25–31.

46. D. Koller, N. Friedman, *Probabilistic Graphical Models: Principles and Techniques.* (MIT Press, Cambridge, MA, 2009).

47. R. Steuer, J. Kurths, C. O. Daub, J. Weise, J. Selbig, The mutual information: Detecting and evaluating dependencies between variables. *Bioinformatics* **18** (suppl. 2), S231–S240 (2002). doi: 10.1093/bioinformatics/18.suppl_2.S231 pmid: 12386007

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6213/1250689/suppl/DC1
Figs. S1 to S20
Tables S1 and S2