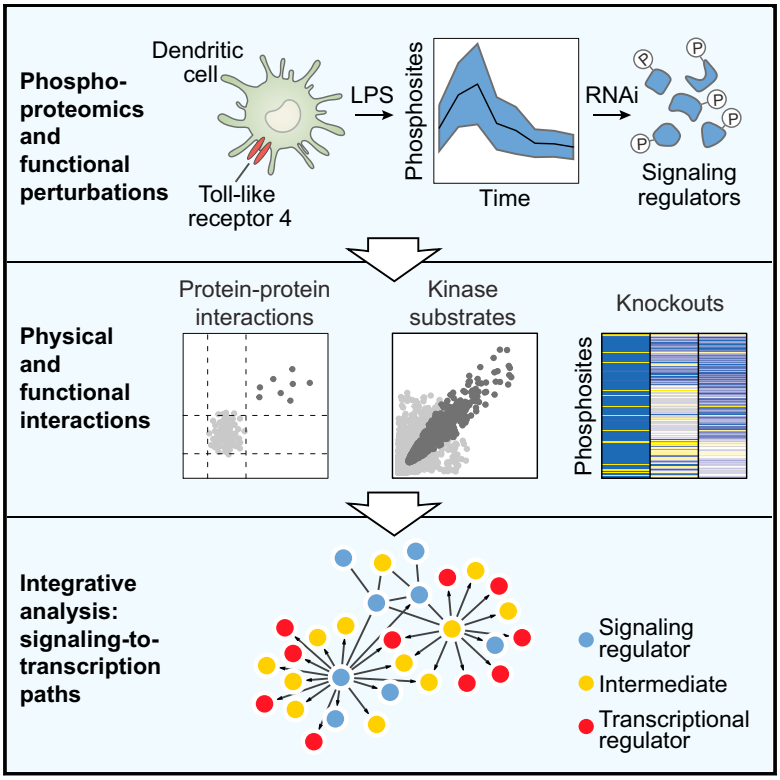
Resource

Cell Reports

An Integrative Framework Reveals Signaling-to-Transcription Events in Toll-like Receptor Signaling

Graphical Abstract



Highlights

- Temporal analysis of phosphorylation networks in LPS-treated dendritic cells (DCs)
- AP1AR and its binding partner, PICALM, as pro-inflammatory signaling regulators
- Biochemical framework to dissect physical, enzymatic, and functional interactions
- Integrative modeling of signaling-to-transcriptional paths downstream of TLR4

Authors

Philipp Mertins, Dariusz Przybylski, Nir Yosef, ..., Aviv Regev, Nir Hacohen, Nicolas Chevrier

Correspondence

aregev@broadinstitute.org (A.R.), nhacohen@mgh.harvard.edu (N.H.), chevrier@fas.harvard.edu (N.C.)

In Brief

Mertins et al. develop a systematic framework to unify the study of signaling and transcriptional regulatory networks. By integrating physical, enzymatic, and functional interaction data, they identify biochemical paths that connect Toll-like receptor (TLR) 4 signaling to transcription. Their analysis uncovers two dozen regulators in TLR signaling, including AP1AR and PICALM, which link vesicle transport to pro-inflammatory responses.

Mertins et al., 2017, Cell Reports *19*, 2853–2866
June 27, 2017
http://dx.doi.org/10.1016/j.celrep.2017.06.016

CellPress

Cell Reports
Resource

An Integrative Framework Reveals Signaling-to-Transcription Events in Toll-like Receptor Signaling

Philipp Mertins,[1,7] Dariusz Przybylski,[1,7] Nir Yosef,[1,2] Jana Qiao,[1] Karl Clauser,[1] Raktima Raychowdhury,[1] Thomas M. Eisenhaure,[1] Tanja Maritzen,[3] Volker Haucke,[3] Takashi Satoh,[4] Shizuo Akira,[4] Steven A. Carr,[1] Aviv Regev,[1,5,*] Nir Hacohen,[1,6,*] and Nicolas Chevrier[1,8,9,*]

[1]Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142, USA
[2]Department of Electrical Engineering and Computer Science and Center for Computational Biology, University of California, Berkeley, Berkeley, CA 94720, USA
[3]Molecular Physiology and Cell Biology Section, Leibniz-Institute for Molecular Pharmacology (FMP), 13125 Berlin, Germany
[4]WPI Immunology Frontier Research Center, Osaka University, 3-1 Yamada-oka, Suita, Osaka 565-0871, Japan
[5]Howard Hughes Medical Institute, Department of Biology, MIT, Cambridge, MA 02142, USA
[6]Center for Immunology and Inflammatory Diseases and Center for Cancer Research, Massachusetts General Hospital, Charlestown, MA 02129, USA
[7]These authors contributed equally
[8]Present address: FAS Center for Systems Biology, Harvard University, Cambridge, MA 02138, USA
[9]Lead Contact
*Correspondence: aregev@broadinstitute.org (A.R.), nhacohen@mgh.harvard.edu (N.H.), chevrier@fas.harvard.edu (N.C.)
http://dx.doi.org/10.1016/j.celrep.2017.06.016

SUMMARY

Building an integrated view of cellular responses to environmental cues remains a fundamental challenge due to the complexity of intracellular networks in mammalian cells. Here, we introduce an integrative biochemical and genetic framework to dissect signal transduction events using multiple data types and, in particular, to unify signaling and transcriptional networks. Using the Toll-like receptor (TLR) system as a model cellular response, we generate multifaceted datasets on physical, enzymatic, and functional interactions and integrate these data to reveal biochemical paths that connect TLR4 signaling to transcription. We define the roles of proximal TLR4 kinases, identify and functionally test two dozen candidate regulators, and demonstrate a role for *Ap1ar* (encoding the Gadkin protein) and its binding partner, *Picalm*, potentially linking vesicle transport with pro-inflammatory responses. Our study thus demonstrates how deciphering dynamic cellular responses by integrating datasets on various regulatory layers defines key components and higher-order logic underlying signaling-to-transcription pathways.

INTRODUCTION

Signaling networks must coordinate multiple layers of regulation throughout the cell to respond to environmental changes. For example, mammalian immune cells detect microbial molecules thanks to pathogen-sensing pathways such as Toll-like receptors (TLRs) (Takeuchi and Akira, 2010). Upon activation by their cognate ligands, TLRs follow general principles of signal transduction by recruiting cytosolic adaptors and downstream enzymes such as kinases, which triggers cascades of biochemical changes leading to cellular outputs such as gene expression changes (Figures 1A and 1B). A fundamental question in cellular response systems, such as TLRs, is how to generate and combine knowledge about signaling and transcription regulatory networks to build an integrated view of the flow of information in a cell. Answering this question will help close gaps in our knowledge of intracellular wiring and inform therapeutic targeting of cellular components that are central to disease.

Despite recent advances in measuring cellular processes and associated biochemical changes from many different angles (e.g., post-translational modifications, gene expression, and transcription factor binding), building integrated models of signaling pathways that take into account multiple regulatory layers remains an elusive task due to several challenges. First, using prior knowledge from databases alone, it is hard to compare and connect signaling nodes and processes that have been studied in disparate systems and with different readouts. Furthermore, existing databases are largely incomplete, as demonstrated by the fact that the vast majority of known phosphorylation sites remain orphans with respect to their matching kinases. Second, acquiring data within a single cellular context and across regulatory processes ranging from post-translational modifications (PTMs) to protein complexes to kinase substrates is difficult due to the various technical requirements of each assay, making them hard to adapt within a unique and relevant cellular context. Third, individual large-scale measurements are

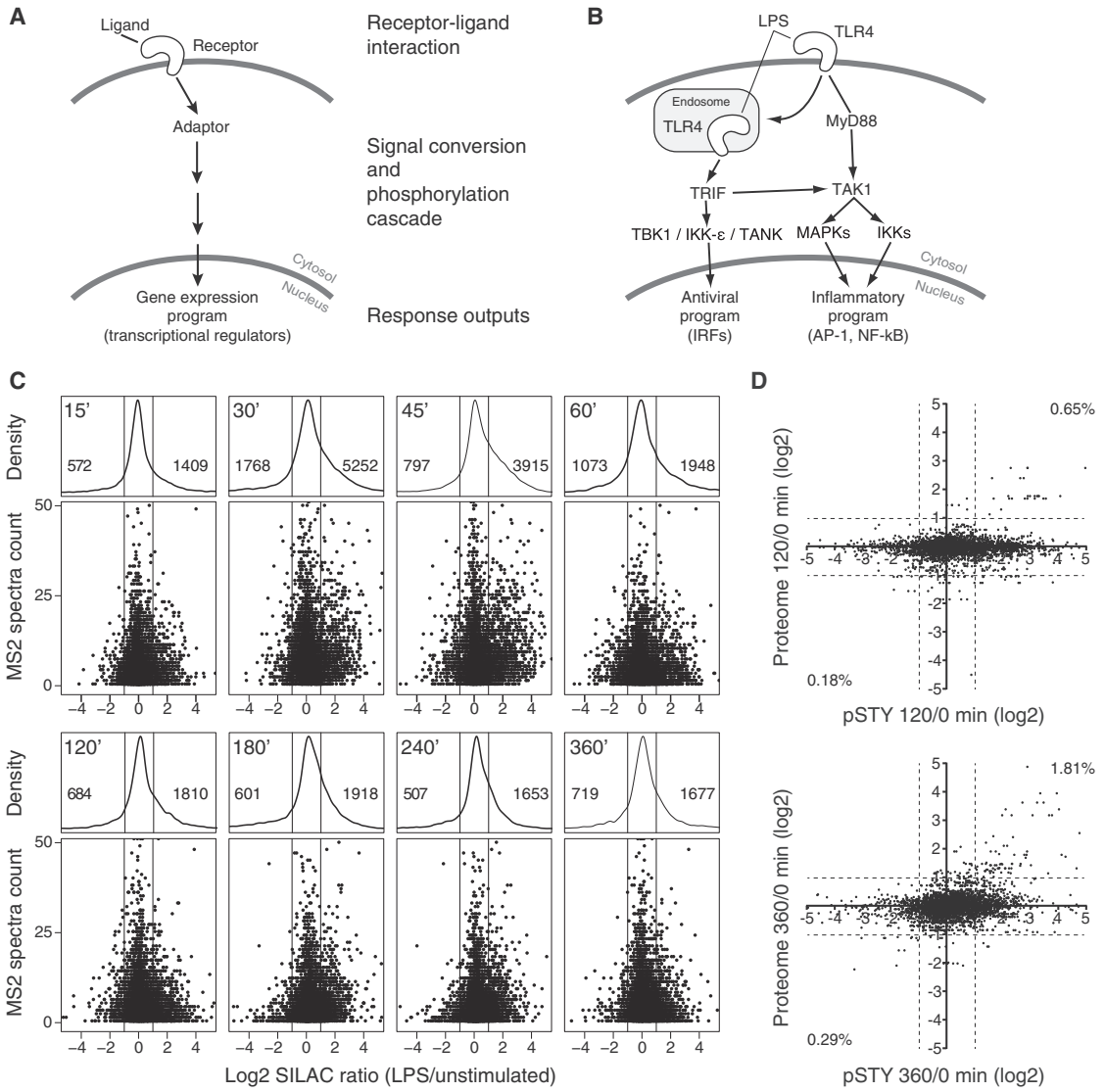


Figure 1. TLR4 Stimulation with LPS Leads to Global and Dynamic Changes in the Phosphoproteome of DCs

(A and B) Diagram highlighting general principles of cellular signaling-to-transcription events (A) and their transposition to the TLR4 pathway (B).

(C) Temporal changes in the phosphoproteome of LPS-stimulated DCs. Shown are the distributions of log2 fold changes of phosphosites (x axis) between LPS-treated and untreated cells at indicated times after LPS stimulation, as density (top of each panel) and dot plots (bottom of each panel, with MS2 spectra count in y axis and showing phosphosites measured in all eight time points).

(D) Comparison between the phosphoproteome and total proteome of LPS-stimulated DCs. Shown are distributions of log2 fold changes of phosphosites (x axis) and proteins (y axis) between LPS-treated and untreated cells at 120 (top) and 360 (bottom) min post-stimulation.

See also Figure S1 and Table S1.

inherently limited by their variability in sensitivity and specificity and are often used to capture static snapshots rather than the dynamic events of cellular responses. It is thus critical to address these challenges to help to dissect the connections that form the basis of multi-layered cellular responses (Bensimon et al., 2012; Santra et al., 2014; Yugi et al., 2016).

Here, we hypothesized that integrating measurements spanning, in the context of a single cellular response model, both signaling and transcriptional regulatory layers will help to reveal key network-wide properties that would otherwise not be observable. To test this, building upon prior work (Chevrier et al., 2011), we developed an experimental and computational framework to measure and integrate the information underlying signaling-to-transcription events in the TLR system, from the membrane to gene regulation. We measure dynamic changes in two types of interactions: physical (i.e., phosphorylation, kinase-substrate relationships, protein-protein and DNA-protein interactions) and functional (i.e., effects of genetic perturbations on gene expression or phosphorylation events) (Figure S1A). Using these datasets, we identify regulators of TLR4 responses in

dendritic cells (DCs), including AP1AR and its binding partner, PICALM, and introduce a network-based computational approach that takes advantage of these diverse measurements to decipher the higher-order logic governing TLR signaling-to-transcription events.

RESULTS

The Dynamic Phosphoproteome of LPS-Stimulated Dendritic Cells

We reasoned that large-scale, dynamic measurements of the changes in protein phosphorylation in lipopolysaccharide (LPS)-treated DCs would help to reconstruct signaling-to-transcription pathways, because TLR signaling functions through phosphorylation of its own constituents, from kinases such as mitogen-activated protein kinases (MAPKs), IRAKs, IKKs, or TBK1 to transcription factors such as nuclear factor κB (NF-κB) or IRFs (Figure 1B) (Takeuchi and Akira, 2010). Furthermore, work by others (Sharma et al., 2010; Sjoelund et al., 2014; Weintz et al., 2010) and us (Chevrier et al., 2011) showed that phosphoproteomics can identify regulators of the TLR system. We used stable isotope labeling with amino acids in cell culture (SILAC)-based phosphoproteomics to compare the levels of phospho-serine, phospho-threonine, and phospho-tyrosine sites between DCs left untreated as control or stimulated with LPS at eight time points (15, 30, 45, 60, 120, 180, 240, and 360 min) (Figures S1B and S1C). We identified and quantified a total of 20,975 phosphosites derived from 5,789 distinct proteins in at least two LPS-stimulated samples (false discovery rate [FDR] < 1%; Figure 1C; Table S1), of which 20.5% were present in all eight time points (4,310/20,975 phosphosites from 1,952 proteins; Figure S1D) due to undersampling of highly complex and low-signal-intensity phosphopeptide mixtures in individual SILAC experiments. The largest changes in the DC phosphoproteome were observed at 30 and 45 min after LPS stimulation, which covered 92.8% of all quantified phosphosites in this study (19,456/20,795) (Figure S1E). In addition, these changes in phosphorylation were not due to changes in protein amounts, as only 0.65 and 1.81% of proteins showed an increase in both phosphorylation and protein levels at 2 and 6 hr after LPS stimulation, respectively (Figure 1D). These results suggested that LPS stimulation modifies a large fraction of the DC phosphoproteome within an hour.

Temporal Analysis of Phosphorylation Changes Highlights Known and Candidate Regulators of TLR4 Signaling

Next, to study the dynamics of the LPS-regulated phosphoproteome, we focused on the 3,557 phosphosites mapping onto 1,606 proteins that were quantified in at least six out of eight time points and differentially regulated upon LPS stimulation in a single or two consecutive time points (2,071/3,557 phosphosites for the latter) (Table S2). Overall, 53.4% (3,557/6,659) of the phosphosites quantified in at least six independent time points were found to be differentially regulated by LPS, which corresponds to 61.4% (1,606/2,617) at the phosphoprotein level. We used *k*-means clustering to partition these 3,557 phosphosites into ten co-abundance clusters with distinct temporal profiles (Figures 2A and S2A). We found three general patterns of changes in phosphorylation levels: (1) early upregulation until 45 min (clusters I and II), (2) late upregulation after 120 min (cluster III), and (3) downregulation at various times (clusters IV–X) (Figure 2B). Each temporal cluster contained known TLR pathway proteins for a total of 43 out of 141 canonical TLR components, including 7.8% (11/141) and 10.6% (15/141) for clusters I and II, respectively (Figures 2B and S2B). Known TLR proteins identified in this data encompassed both positive (e.g., MAPK family, IRF3, and NF-κB) and negative (e.g., TANK and TNFAIP3) regulators, and were differentially phosphorylated at multiple sites in some cases (Figure S2C). The 1,606 phosphoproteins present in these ten temporal clusters were enriched for molecular functions, including kinases, transcriptional regulators, or protein binding (Figure S2D). Some of the enriched gene sets pointed to nascent areas of TLR biology, such as the organization and regulation of the TLR system within the framework of intracellular organelles and structures (e.g., activity and regulation of GTPases, cytoskeleton; Figure S2D). Cluster II, and to a lesser extent other clusters, showed a significant enrichment for other immune signaling pathways (e.g., B and T cell receptor signaling or DNA-sensing pathways), highlighting the existence of shared proteins between these immune response systems (Figure S2D). Taken together, these results reveal the dynamic changes imparted on the DC phosphoproteome by LPS, which include known and putative regulators of TLR4 signaling as well as processes linked to DC biology, such as changes in cell shape, motility, metabolism, and antigen processing.

Genetic Perturbations of Phosphorylated Proteins Identify Putative Regulators of TLR4 Signaling

To test if the phosphoproteins identified above play a role in the TLR system, we used our temporal and enrichment analyses to prioritize candidates for genetic perturbations (Figure 3A). We focused on 751 phosphoproteins from the 1,606 ones used for temporal clustering, which were upregulated at 30 and 45 min after LPS treatment (clusters I and II; Figure 2B). We reasoned that using early clusters would help to identify candidate regulators likely to be downstream of TLR4 by avoiding feedbacks from transcription or autocrine and paracrine signaling. Third, we selected 169 out of 751 phosphoproteins to test by retaining all enzymes (e.g., kinases and GTPases) and enzyme binders and regulators (e.g., GTPase regulators) (Figures S3A and S3B; Table S3). All selected phosphoproteins were also found to be expressed at the mRNA level in DCs (Garber et al., 2012). The two TLR4 adaptor proteins MYD88 and TRIF (encoded by *Ticam1*) were part of these candidate genes. MYD88 was added manually as a positive control, although it was not found to be differentially phosphorylated, whereas TRIF matched our selection criteria above.

We successfully perturbed 131 out of 168 candidate genes with an average knockdown efficiency of 81% ± 9% SD (Figure S3C). We stimulated DCs with LPS and measured the effect of gene silencing on the mRNA levels of 263 TLR response signature genes, representing the inflammatory and antiviral programs (Table S3). We determined statistically significant changes in the expression of signature transcripts upon

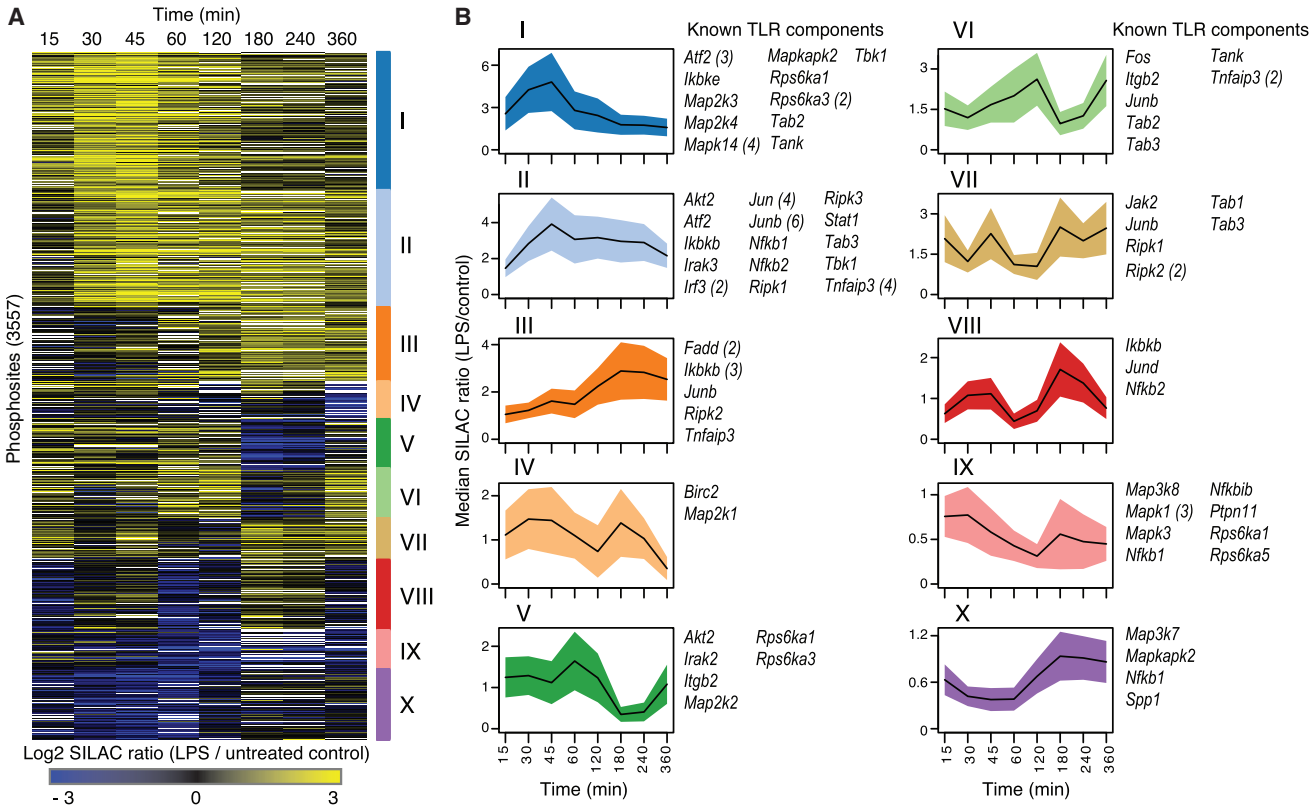


Figure 2. Temporal Analysis of the LPS-Induced Phosphoproteome Reveals Known and Candidate Regulators of TLR4 Signaling

(A) Temporal phosphorylation profiles during LPS stimulation in DCs. Log2 fold changes between LPS-treated and untreated cells for 3,557 phosphosites (rows) detected in at least six out of eight time points (columns). Phosphosites are partitioned into ten clusters using k-means (color bars, right). White indicates missing values.

(B) Median log2 fold changes between LPS-treated and untreated cells (y axis) and median absolute deviation (MAD; colored error bar) at each time point (x axis) for phosphosites in all ten k-means clusters from (A). Known TLR pathway proteins detected in each cluster are indicated on the right. Parentheses indicate the number of phosphosites per proteins (when >1).

See also Figure S2 and Table S2.

individual knockdowns based on comparisons to 16 control genes, whose expression remains unchanged upon TLR activation, and to 38 control short hairpin RNAs (shRNAs) that did not affect TLR signature genes. 27 out of the 131 genes tested significantly affected TLR signature gene expression, which included known TLR signaling components such as TICAM1, TBK1, MAPK9, RIPK3, and IRAK2 (Figure 3B). Furthermore, several phosphoproteins were reported to function in TLR signaling by independent studies: TRAFD1 (Sanada et al., 2008), STK3 (Geng et al., 2015), ULK1 (Eriksen et al., 2015), and CORO1A (Tanigawa et al., 2009). Interestingly, known and candidate components had similar effects on the TLR gene signature upon knockdown. By measuring the pairwise similarity among these 27 perturbation profiles (using Pearson's correlation), we observed three major modules of signaling regulators: MYD88 and a set of four proteins (SAMHD1, TBC1D17, AP1AR, and PDLIM7) affecting inflammatory gene expression (module I), TICAM1 and five proteins (module II), and 16 proteins displaying effects that overlap with MYD88 and/or TICAM1 (module III) (Figure 3B).

Validation of AP1AR and Other Candidate Regulators of the Myd88-Dependent Inflammatory Pathway

We next sought to validate the putative roles of the four phosphoproteins (AP1AR, PDLIM7, SAMHD1, and TBC1D17) whose perturbation profiles closely resembled that of MYD88 in controlling pro-inflammatory genes (Figure 4A). We measured the expression levels of inflammatory and antiviral cytokines in LPS-stimulated DCs infected by two independent, gene-specific lentiviral shRNAs per candidate phosphoprotein. We observed a decrease in inflammatory cytokine mRNA expression compared to eight control hairpins in all cases (*Il6*, *Cxcl1*, and, to a lesser extent, *Tnf*), whereas antiviral cytokines Ifit1 and Cxcl10 were mostly unaffected (Figure 4B). Similarly, using mouse *Ap1ar*$^{-/-}$ knockout DCs (Maritzen et al., 2012), we observed a strong decrease in inflammatory cytokines, especially *Il1b*, *Il12b*, and *Tnf*, whereas antiviral cytokines were not affected (*Ifnb1*) or slightly reduced (*Cxcl10*) (Figure 4C).

To generate mechanistic insights about the putative role of AP1AR in the TLR4 pathway, we sought to identify binding partners of AP1AR in LPS-stimulated DCs using affinity purification

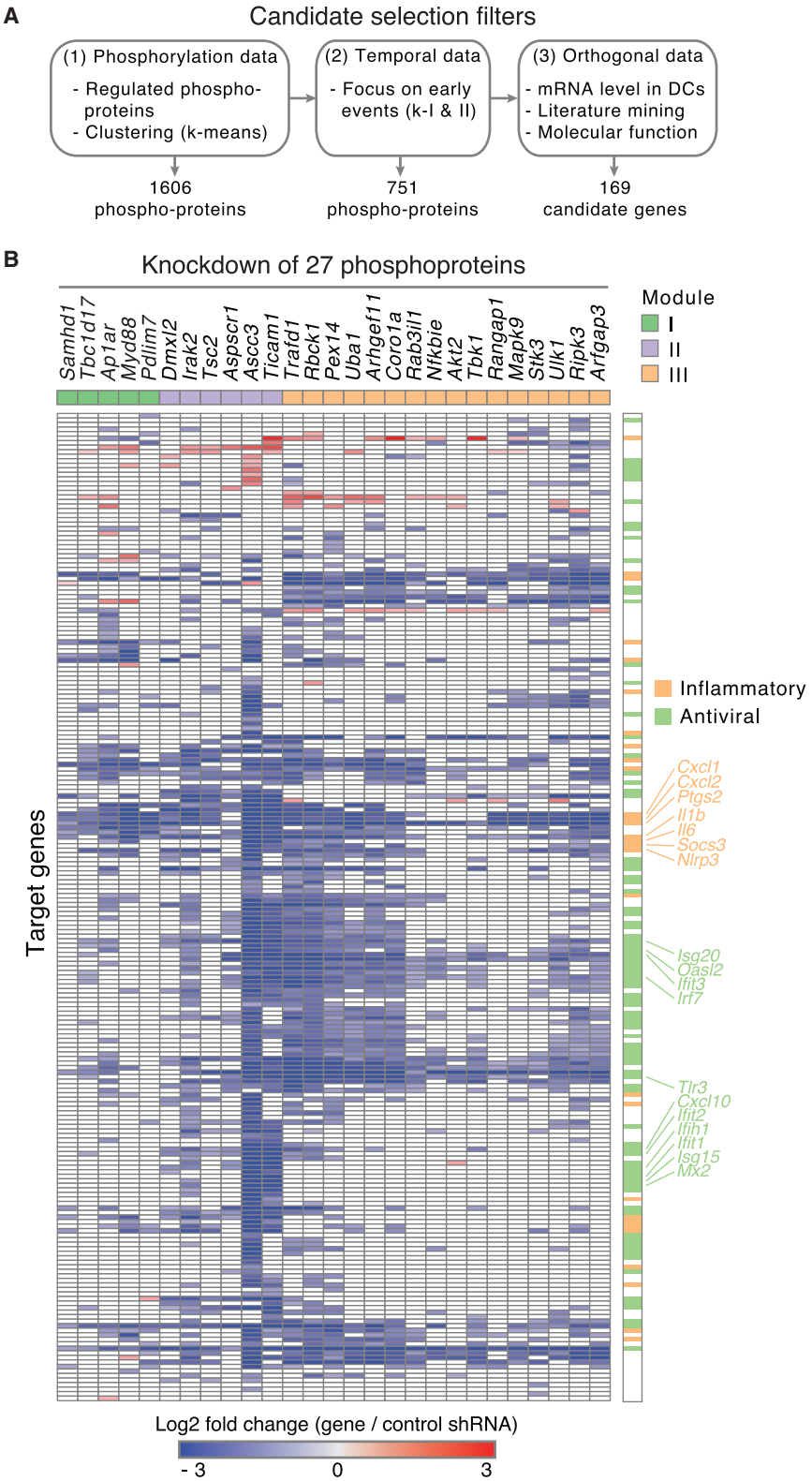


Figure 3. Genetic Perturbations of Phosphorylated Proteins Identify Putative Regulators of TLR4 Signaling

(A) Overview of phosphoprotein candidate selection for functional analysis.

(B) Perturbation profiles of the 27 phosphoproteins that significantly impacted TLR4 outputs. Shown are the perturbed candidates and control phosphoproteins (columns) and the log2 fold changes for each target gene (rows) between gene-specific and control shRNAs. The rightmost column categorizes target genes into antiviral (light green) and inflammatory (light orange) programs.

See also Figure S3 and Table S3.

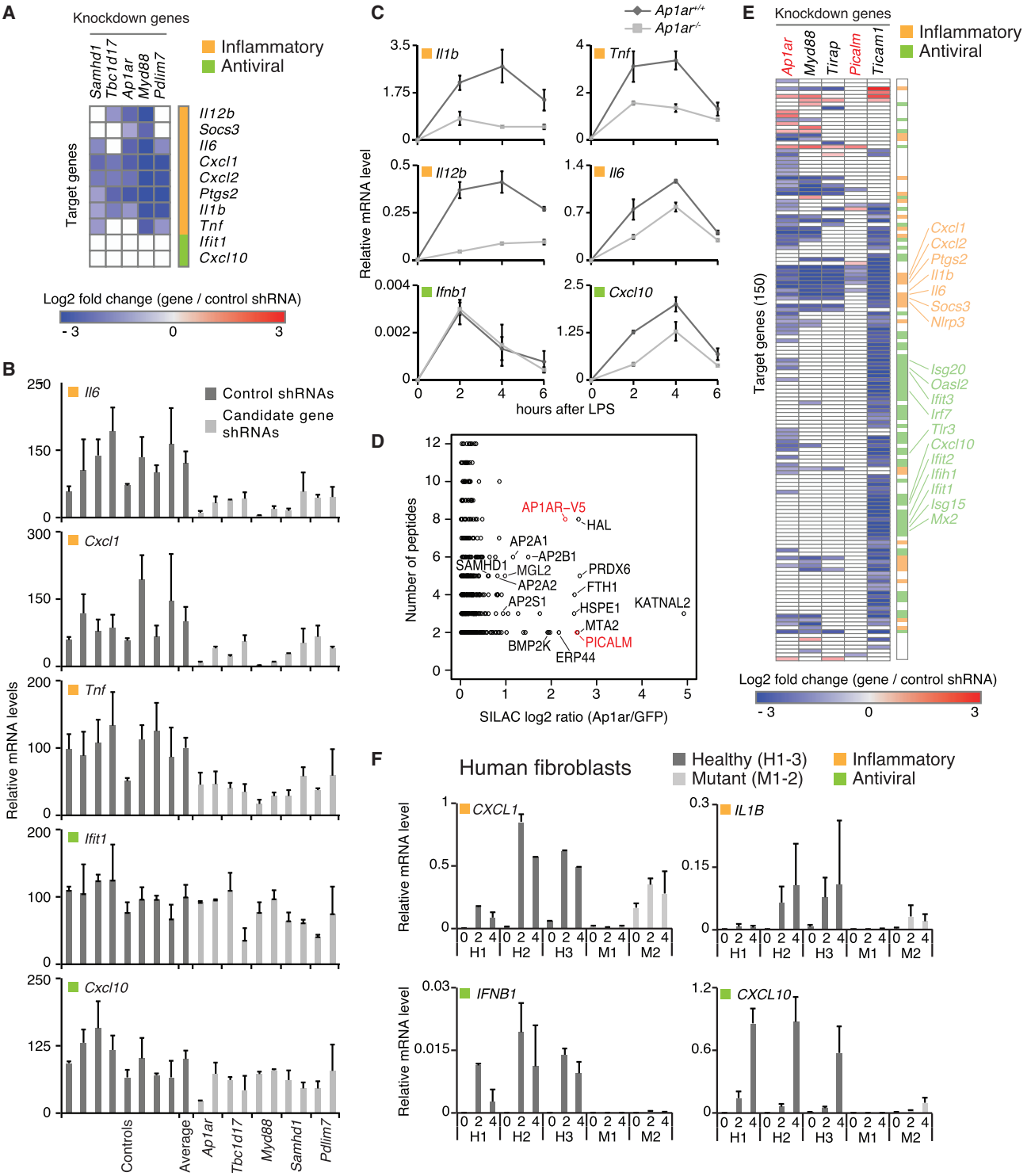


Figure 4. Identification of Candidate Regulators in the MYD88-Dependent Inflammatory Pathway

(A) Perturbation profiles of genes affecting the MYD88 pathway. Shown are four perturbed candidate genes and MYD88 (columns) and the log2 fold changes between gene-specific and control shRNAs (rows) of ten target genes. The rightmost column categorizes target genes into antiviral (light green) and inflammatory (light orange) programs.

(B) Expression levels (qPCR) relative to control shRNAs (left bars, dark gray) for two antiviral cytokines (*Ifit1* and *Cxcl10*) and three inflammatory cytokines (*Il6*, *Cxcl1*, and *Tnf*) following LPS stimulation in DCs using two independent shRNAs. Bottom tick marks separate shRNAs controls and each gene (average indicates the mean value for all eight control shRNAs). Two to three replicates for each experiment; error bars represent SD.

(legend continued on next page)

followed by mass spectrometry in primary mouse DCs (Figure S4A). Protein overexpression was effective in nearly all transduced cells as measured by GFP fluorescence and did not impact cell responsiveness to LPS, as shown by strong morphological changes (Figure S4B). We overexpressed V5-tagged AP1AR and GFP as control bait in SILAC-labeled DCs stimulated with LPS for 30 min (Figure S4C), which led to the identification of proteins that co-precipitated with AP1AR, but not GFP (Figure 4D; Table S4). Several known components of the assembly protein complex 2 (AP-2) were pulled down with AP1AR (AP2A1, AP2B1, and AP2S1), as well as the AP-2 binding partner PICALM, which is an important component of clathrin-mediated endocytosis (Miller et al., 2015). Next, to test if some of these AP1AR binders affect TLR4 signaling outputs, we turned back to genetic perturbations followed by gene signature measurements. Out of ten putative AP1AR binders (at least two peptides identified and >1.5 log2 SILAC ratio of AP1AR/GFP), six showed a knockdown efficiency >50%. We found that PICALM led to a decrease in the induction of LPS-induced inflammatory genes similarly to MYD88, TIRAP, and AP1AR (Figure 4E). Altogether, these results suggest a potential mechanism whereby AP1AR and PICALM act together in the regulation of MYD88-dependent inflammatory signaling.

For another candidate identified based on phosphorylation changes, SAMHD1, we further tested its potential involvement in TLR signaling using human skin fibroblasts derived from Aicardi-Goutières syndrome (AGS) patients that carry deleterious SAMHD1 mutations (Crow and Manel, 2015). We observed a decrease in both inflammatory and antiviral gene expression upon LPS stimulation in two independent patient cell lines compared to three healthy controls (Figure 4F), which differed from knockout mouse DC data (Figure S4D). The latter observation might be attributable to the difference in cellular context or to compensatory mechanisms in the mouse knockout cells. Interestingly, physical interactions between SAMHD1 and TLR pathway proteins have been reported previously, such as the TLR4 adaptor protein TIRAP (Li et al., 2011), and also with CCNA2 and CDK2, which can be activated by TLR4 signaling (Hasan et al., 2007; Huttlin et al., 2015). Altogether, we gathered evidence supporting that AP1AR, its binding partner PICALM, and SAMHD1 are likely to act as regulators of pro-inflammatory TLR4 signaling.

Signaling Regulator Perturbation Profiles Overlap with Transcription Factor Target Genes, Suggesting Potential Signaling-to-Transcription Paths

Having shown that phosphorylation dynamics can help identify potential regulators of TLR signaling-to-transcription events, we next sought to identify how signaling regulators are connected to downstream transcriptional regulators. The two targeted screens for candidate (1) phosphoproteins and (2) AP1AR binders led to 29 perturbation profiles showing significant changes in TLR signature genes upon LPS stimulation (Figures 3B and 4E). Based on the similarity of these perturbation-induced expression profiles (Pearson's correlation), we partitioned these 29 proteins into three modules (Figure 5A; similar to Figure 3B). Next, we asked what transcription factors (TFs) are likely to act downstream of these three modules of proteins by taking advantage of existing data on the binding sites across the genome of 23 TFs involved in TLR4 signaling (Garber et al., 2012). We reasoned that measuring the overlaps between genes whose promoters are bound by a TF, and genes whose mRNA levels are impacted by knockdown of a phosphoprotein, would help to infer some of the signaling regulator-TF relationships likely to be active upon TLR4 activation (Figure S5). For 20 out of 23 TFs tested, we identified significant overlaps (p value < 0.05; hypergeometric test) between gene sets whose promoters were bound by one or several TFs and those whose mRNA levels were impacted by knockdown of 25 out of 29 candidate and known regulators (Figure 5B; Table S5). Some of these overlaps recapitulated known signaling regulator-TF relationships in the TLR pathways, such as MYD88 and NF-κB family members REL and RELB, or TRIF and IRFs and STATs. Gene targets of AP1AR, MYD88, and PICALM overlapped significantly with genes bound by RUNX1 and REL. Taken together, these results further support a role for the 29 phosphoproteins identified downstream of TLR4 and suggest the existence of signaling regulator-TF relationships between 25 phosphoproteins and 20 TFs.

Physical and Functional Proteomics Pinpoint Binding and Phosphorylation Events Downstream of the Myd88 Adaptor and Associate Kinases

Next, to decipher the biochemical events linking the signaling to transcriptional regulator relationships identified above, we measured protein-protein and kinase-substrate interactions by focusing on MYD88-dependent signaling. First, in DCs stimulated with LPS for 30 min, we rediscovered most known MYD88 binding partners, including TIRAP, TRAF6, or IRAK family kinases, which support the validity of our affinity-purification coupled with mass spectrometry (AP-MS) assay in primary DCs (Figures 6A and S6A; Table S6). IRAK2 immunoprecipitation identified several interaction partners such as MYD88 and TRAF6 but with lower enrichment ratios compared to MYD88, which is likely due to the short-lived interaction dynamics of kinases (Figure 6B).

(C) Inhibition of transcription of inflammation cytokines in *Ap1ar*$^{-/-}$ DCs. mRNA levels (qPCR; relative to Gapdh) for indicated inflammatory (light orange) and antiviral (light green) cytokines in three replicates per time point. Error bars represent SD.
(D) Interaction proteomics identified putative binders for AP1AR in DCs. Log2 fold change (x axis) of proteins enriched differentially between DCs expressing V5-tagged-AP1AR and -GFP (control bait) plotted against the number of peptides identified per protein (y axis).
(E) Perturbation profiles of indicated genes (columns) and the log2 fold changes between gene-specific and control shRNAs (rows) of 150 target genes. The rightmost column categorizes target genes into antiviral (light green) and inflammatory (light orange) programs.
(F) Impact of SAMHD1 mutations on human fibroblast cell response to LPS. Human fibroblasts from healthy (H) or mutant-carrying patients (M; with homozygous c.445C > T p.Gln149* for M1 and c.1609-1G > C for M2) were stimulated with LPS or left untreated as control, and indicated inflammatory (light orange) and antiviral (light green) cytokine levels were measured by qPCR (relative to GAPDH). Error bars represent SD.
See also Figure S4 and Table S4.

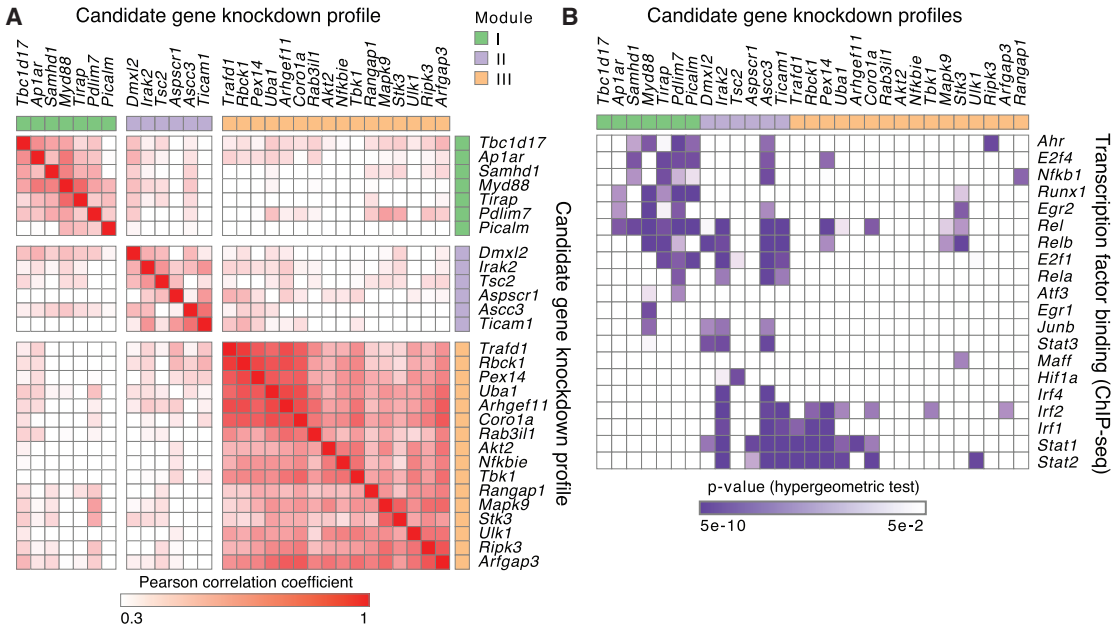


Figure 5. Similarities in Perturbation Profiles and Overlap with TF Target Genes Suggest Three Functional Modules for the 29 Candidate Phosphoproteins

(A) Functional classification based on similarity of perturbation profiles. Shown is a correlation matrix (Pearson correlation coefficient) of the perturbation profiles from Figures 3B and 4E combined.

(B) Intersection between genes affected by a phosphoprotein perturbation and genes whose promoters are bound by transcription factors (TFs). Shown are the overlaps between genes affected by 29 candidate signaling regulators knockdowns (columns, including positive control genes) and genes whose promoters are bound by 20 TFs (rows). P values, hypergeometric test (purple: significant correlation; white: no correlation).

See also Figure S5 and Table S5.

Second, we used two complementary approaches to identify the substrates downstream of MYD88-associated kinases, which remain poorly characterized. Perturbation approaches followed by phosphoproteomics have proven useful in determining functional pathway components downstream of a given network node (Bodenmiller et al., 2010; Chevrier et al., 2011). We measured the impact of four knockout (KO) models (*Myd88*$^{-/-}$, *Myd88*$^{-/-}$/*Ticam1*$^{-/-}$, which abrogates all TLR4 signals), *Irak2*$^{-/-}$ and *Irak4*$^{-/-}$) on the DC phosphoproteome upon LPS stimulation for 30 min (Figures 6C and 6D). To stringently evaluate KO effects on the LPS-dependent DC phosphoproteome, we focused on the 1,628 phosphosites mapping onto 990 unique proteins that were differentially regulated in both (1) LPS-treated wild-type DCs at 30–45 min (time course data; Table S1) and (2) *Myd88*$^{-/-}$/*Ticam1*$^{-/-}$ DCs compared to wild-type (Table S6). Out of these 1628 phosphosites, a third (38.1%, 621/1,628) were only affected by *Myd88*$^{-/-}$/*Ticam1*$^{-/-}$ double deletion, whereas the remaining sites were affected by both double and single mutants: 45.6% (742/1628) for *Irak4*$^{-/-}$, 31.1% (506/1628) for *Myd88*$^{-/-}$, and 8.1% (132/1,628) for *Irak2*$^{-/-}$. These numbers agree with the essential role of IRAK4 in TLR signaling (Picard et al., 2003; Suzuki et al., 2002) and the partially redundant function of IRAK2 with IRAK1 (Kawagoe et al., 2008). Furthermore, these 990 Myd88/Ticam1-dependent phosphoproteins captured 32.6% (46/141) of the canonical TLR proteins, including known phosphosites such as TBK1 S716 and JUN S63/S73 downregulated in *Myd88*$^{-/-}$ and *Irak4*$^{-/-}$ cells, IRF3 S379 impacted upon double KO only, or MAPK9 T183/Y185 by MYD88- and TRIF-dependent pathways (Figure S6B).

To complement this genetic approach, we developed a large-scale in vitro kinase (IVK) assay using recombinant kinases IRAK4, TBK1, and IRAK2 mixed with native protein lysates from SILAC-labeled DCs followed by phosphoproteomics (Figure 6E). We identified a total of 967 phosphosites upregulated by IRAK4, 325 by TBK1, and 201 by IRAK2, which included sites also upregulated in LPS-treated DCs: 55 out of 967 (5.7%) for IRAK4 and 62 out of 325 (19.1%) for TBK1 (Figures 6F, 6G, and S6C; Table S6). These results suggest that some of the phosphosites identified by IVK are likely to be physiologically relevant, although others might be due to off targets effects (e.g., activation of secondary kinases, or proximity with proteins in solution that would not exist in cells).

An Integrated Model Reveals Signaling-to-Transcription Paths across the TLR4 System

Lastly, we sought to combine our measurements on physical and functional interactions into an integrated model of signaling-to-transcription relationships in the TLR4 system (Figure S7A). We used a network-based approach that relies on three main steps (Figure 7A). First, we assembled a "background" network of 92,610 protein-protein and 5,533 kinase-substrate interactions from public repositories and 43 protein-protein and 230 kinase-substrate interactions identified from this study using DCs (Table S7). Second, we assigned weights to the edges

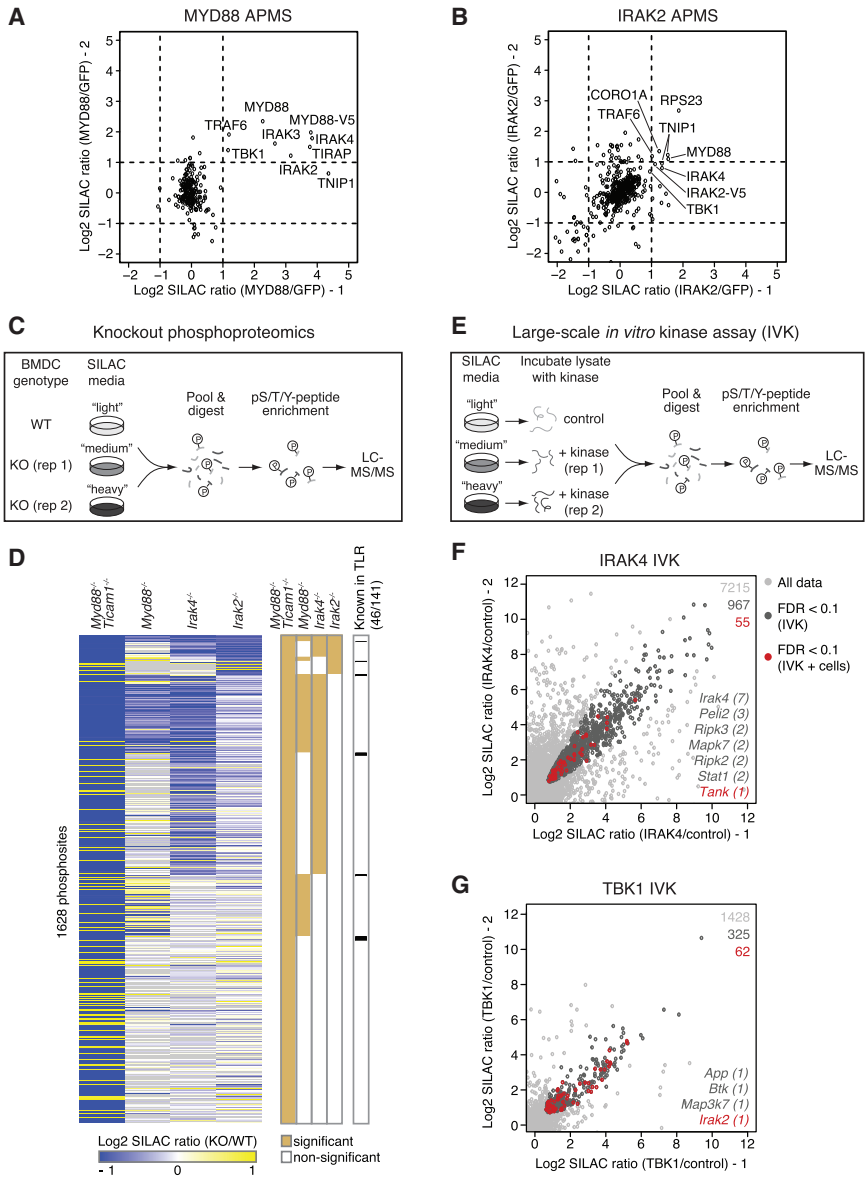


Figure 6. Physical and Functional Proteomics Assays Pinpoint Binding and Phosphorylation Events Downstream of the Myd88 Adaptor and Associated Kinases

(A and B) Affinity purification followed by proteomics. Shown are dot plots of SILAC ratios for proteins identified in DCs overexpressing V5-tagged MYD88 (A) or IRAK2 (B). Cells were stimulated with LPS for 30 min and protein complexes purified using anti-V5 antibodies coupled to magnetic beads. Each axis represents an independent experiment.

(C) Diagram depicting our experimental approach for measuring the impact of gene KO on the TLR4-regulated phosphoproteome of mouse BMDCs.

(D) Phosphoproteomics in KO cells. Left: heatmap for SILAC ratios of phosphosites (rows) in four KO models (columns) at 30 min after LPS stimulation compared to control wild-type cells, as indicated (gray, missing values). Middle: phosphosites with significant up- or downregulation in KO versus WT (light brown). Right: phosphosites belonging to known TLR proteins (black).

(E) Diagram depicting our experimental approach for large-scale in vitro kinase assays using native protein lysates from BMDCs and phosphoproteomics.

(F and G) In vitro kinase (IVK) assay followed by phosphoproteomics. Shown are scatterplots of SILAC ratios of phosphosites identified using the purified kinases IRAK4 (F) and TBK1 (G). Light gray, all data points; dark gray, phosphosites with FDR < 0.1 in IVK; red, phosphosites with FDR < 0.1 in both IVK and cells stimulated with LPS, which highlights the overlap between IVK and phosphoproteome measurements on stimulated cells (denoted as IVK + cells). Gene names at the bottom right of each plot indicate known TLR components with the number of phosphosites in parenthesis.

See also Figure S6 and Table S6.

(i.e., protein-protein and kinase-substrate interactions) and nodes (i.e., signaling or transcriptional regulators) of the background network to create a "weighted" interaction network based on the phosphorylation changes driven by LPS stimulation and specific kinases (based on KO and IVK data). Third, we searched the weighted network for biochemical paths linking the 29 phosphoproteins/signaling regulators or "seed nodes" to transcriptional regulators or "target nodes." To test the validity of this integrative algorithm, we quantified its performance in retrieving known seed-target relationships between canonical TLR pathway components using receiver-operator characteristic (ROC) curves. In the high precision regime, using a weighted network outperformed methods that used the background interaction network or phosphorylation data alone. For example, at a false positive rate (FPR) of 0.001, the weighted network method yielded a true positive rate (TPR) that was 3.9 and 10.4 times higher than background network and "phosphorylation only" approaches, respectively (Figure S7B). Thus, our network-based approach correctly identified known signaling-to-transcription relationships between canonical TLR pathway components thanks to the information collected using DCs in this study.

Next, we searched for biochemical paths connecting the 29 signaling regulators highlighted above as "seeds" (Figure 5A) and the 782 TFs detected by mass spectrometry in bone marrow-derived dendritic cells (BMDCs) as "targets" (Table S1). We identified 420 significant relationships between 27 out of 29 seed(except for seed DMXL2 and RAB3IL1) and 95 out of 782 target nodes ($p < 0.0005$, FDR < 0.05), whereas only 12 relationships linking 7 seeds to 11 targets can be found without integrating our DC-specific datasets with publicly available interactions within our algorithmic framework (Figure 7B; Table S7). Each signaling node reached between 51 (TBK1) and 3 (ARHGEF11) TFs, for an

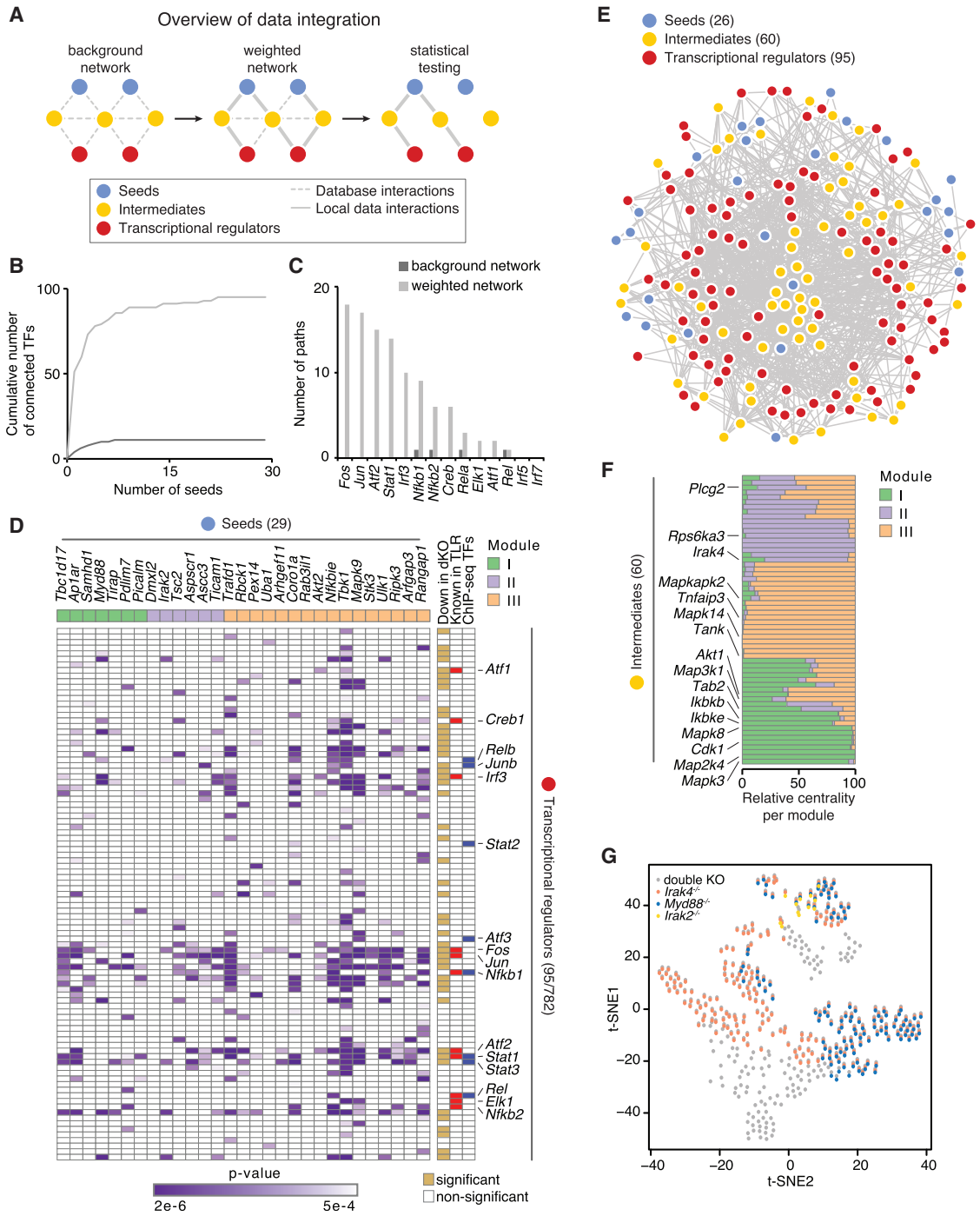


Figure 7. An Integrative Analysis Reveals Known and Candidate Signaling-to-Transcription Paths and Helps Parse the Effects of Myd88 and Associated Kinases in the TLR4 System

(A) A computational framework for integrative analysis of the functional and physical proteomics datasets collected in this study (from left to right). A background interaction network is assembled using database and local data, nodes and edges are scored based on experimental evidence from this work, and statistically significant relationships determined by bootstrap analysis.

(B) Cumulative number of significant relationships (bootstrap p value < 0.0005, FDR < 0.05) identified between seed nodes (29) and any of the transcriptional regulators detected in BMDCs (782 possible target nodes in total) using background network (dark gray) and weighted network (light gray) methods.

(C) Total number of relationships linking seeds (29) and known TLR transcription regulators (14) for background network (dark gray) and weighted network (light gray) methods.

(legend continued on next page)

average of 14.5 ± 11.6 SD. TFs via 1.5 ± 0.7 SD intermediate nodes (Figure S7C). Importantly, these signaling-to-transcription relationships captured 11 out of the 14 canonical TLR TFs and 8 out of 20 of the TFs whose binding sites were compared to knockdown effects (Figures 5B, 7C, and 7D). Furthermore, 49% (47/95) of the TFs were both upregulated at the phosphorylation level upon LPS stimulation and downregulated in *Myd88*$^{-/-}$/*Ticam1*$^{-/-}$ cells. Overall, each of the three modules identified based on co-phenotypes upon knockdown (Figure 5) appeared to be biochemically linked to similar downstream TFs (Figure 7D).

We asked which intermediate nodes were most central between seed and target nodes (i.e., most connected to target TF nodes). For the 420 significant relationships linking the 27 seed and 95 target nodes, we ranked the top 25 intermediates present across each of our three modules (Figure 7D), which lead to a total of 60 non-overlapping intermediate nodes that included 16 canonical TLR pathway components (Figures 7E and 7F). These 60 intermediate nodes displayed various levels of specificity across the three modules identified above, with for example IRAK4 being central to module II (i.e., connected to a relatively high number of nodes), whereas MAPK8 (JNK) and MAPK14 (P38) were more connected across modules I and III, respectively. Other nodes appeared shared between modules such as AKT1 for I and III or TAB2 for I, II, and III. Thus, intermediate nodes display both specific and shared roles across the regulatory modules of the TLR4 pathway, which likely reflects crosstalk within pathways leading to the regulation of overlapping sets of target genes.

To gain insights into how signal is distributed downstream of TLR4, we asked how the 420 seed-target relationships identified here were affected by the four KO strains used in this study (Figure 6D). We quantified how many of the nodes (seed, intermediate, and target) present in each of the 420 seed-target pairs were impacted at their phosphorylation level by KO. 391 out of the 420 pairs were significantly affected by *Myd88*$^{-/-}$/*Ticam1*$^{-/-}$, and 261 out of these 391 pairs were also impacted by *Myd88*$^{-/-}$, *Irak2*$^{-/-}$, and/or *Irak4*$^{-/-}$, leading to four clusters of effects: (1) double KO only or together with (2) IRAK4 alone, (3) IRAK4 and MYD88, or (4) IRAK4, MYD88, and IRAK2 (although to a lesser extent) (Figure 7G). Interestingly, a large fraction of TLR4 signals were impacted by MYD88 deletion, as expected, but IRAK4 was responsible for broader effects despite the presence of IRAK4 and MYD88 in the same complex. Seed-target pairs that were impacted by double-KO cells, but not MYD88 KO cells, are likely to be important for TRIF-dependent signaling (i.e., module II). Overall, this quantitative measurement of KO effects on signaling-to-transcription paths provides additional information on how signal is transmitted and partitioned from MYD88 and some of its kinase partners to downstream signaling and transcriptional regulatory layers.

DISCUSSION

We established an integrative framework to dissect signal propagation in the TLR system using data spanning both signaling and transcriptional regulatory events. Previous studies have connected paths within networks largely using protein-protein interaction or phosphorylation data alone or in conjunction with one to two different types of experimental data (Gitter et al., 2013; Huang and Fraenkel, 2009; Huang et al., 2013; Terfve et al., 2015). This study provides a proof-of-principle example of the power of integrative analyses that take into account regulatory layers not typically studied in conjunction, from phosphorylation dynamics to relationships between kinases substrates to proteins forming complexes or binding to DNA to gene regulation. In future work, it will be crucial to take into account additional regulatory layers such as the spatial distribution of proteins (Brubaker et al., 2015), other PTMs and their enzymes (e.g., ubiquitination, acetylation) (Mertins et al., 2013), and post-transcriptional modifications (RNA) or translational control events.

The observations that AP1AR, its binding partner PICALM, and SAMHD1 might play a role in pro-inflammatory signaling will require future mechanistic studies. Interestingly, both AP1AR and its binding partner, PICALM, interact with clathrin adaptor proteins (Maritzen and Haucke, 2010; Miller et al., 2015), suggesting a link between the TLR4-MYD88 pathway and intracellular vesicle transport regulation that is reminiscent of the TLR4-TRIF axis (Kagan et al., 2008). In addition, previous work linked the *Ap1ar* locus to TNF production by DCs triggering colitis (Ermann et al., 2011), which further support our results on the role of AP1AR in pro-inflammatory signaling. The other candidate regulator reported here, SAMHD1, is best characterized in viral restriction (Ballana and Esté, 2015), but it also plays a role in processes such as TNF-mediated pro-inflammatory signaling in fibroblasts (Liao et al., 2008), cell cycle (Pauls et al., 2014), or DNA damage (Clifford et al., 2014) and in disease such as AGS (Crow and Manel, 2015) and cancer (Schuh et al., 2012). LPS regulated both known and previously unrecognized phosphosites on SAMHD1 such as T634, mouse ortholog site for the known human T592 regulatory site targeted by CDK2 (Pauls et al., 2014), or T52 found in the poorly characterized SAM domain and that was regulated in a MYD88-dependent manner. Taken together, these observations provide valuable information for future mechanistic investigations.

The multi-layer datasets reported here will be useful for further analyses, mining and hypothesis-generating purposes

(D) Significant relationships (420 pairs) found between 29 seeds (columns) and 95 transcriptional regulators (rows). Modules from Figures 3B and 5A are shown (columns) in light green (I), purple (II), and orange (III). Transcriptional regulators with phosphosites with significant up- or downregulation in *Myd88*$^{-/-}$/*Ticam1*$^{-/-}$ versus WT and in time series are indicated on the right (light brown). P values, bootstrap (purple).

(E) An interaction network connects 27 seeds (blue) to 95 transcriptional regulators (red) through the top 60 intermediate (yellow) nodes that were ranked based on centrality measure (see Experimental Procedures).

(F) Centrality score of the top 60 intermediate nodes across the three modules from (D).

(G) t-distributed stochastic neighbor embedding (t-SNE) analysis of the effects of gene KO (data from Figure 6D) on the phosphorylation levels of nodes present in the paths mediating the seed-transcriptional regulator relationships identified in (D). Shown are all of the 391 out of 420 relationships affected by *Myd88*$^{-/-}$/*Ticam1*$^{-/-}$ (gray dots). The effects of Irak4, Myd88, and Irak2 on these paths are overlaid in orange, blue, and yellow, respectively.

See also Figure S7 and Table S7.

on additional candidate regulators, from the protein to the phosphosite level. First, many of the 131 phosphoproteins selected for genetic screening had little to no effect on gene expression. While poor knockdown efficiency and functional redundancy can likely explain some of these cases, measuring the effects of perturbing these proteins on other aspects of DC biology such as motility or antigen presentation might help uncover important mechanisms. Second, we focused our targeted screen for regulators of gene expression on enzymes and their regulators, but screening additional molecular functions is likely to uncover additional regulators. For example, 24 phosphoproteins downregulated between 180 and 240 min after LPS stimulation are involved in RNA binding and include the known pathogen-sensing regulators *Ddx21*, *Ddx3x*, or *Adar* and a host of potential candidates for this nascent area in TLR biology (Anderson, 2010).

Lastly, it will be critical to build upon this work to systematically identify functional phosphosites and their matching kinases. Our study correctly identified many phosphosites of canonical TLR components or other pathogen-sensing pathways such as NLRC4 S533, which is a key site for host immunity (Qu et al., 2012). Our large-scale IVK assay uncovered many known and candidate substrates that will be important to validate using in vivo chemical genetics approaches (Allen et al., 2007) and shorter timescales to increase confidence about substrate specificity as shown in bacteria and yeast (Kanshin et al., 2015; Skerker et al., 2008). Thus, future research on screening functional phosphosites using site-directed mutagenesis will help to reveal phosphorylated residues with functional significance and potential therapeutic value.

EXPERIMENTAL PROCEDURES

Cells

Bone-marrow-derived DCs were generated from 6- to 8-week-old female C57BL/6J (The Jackson Laboratory), *Ap1ar*$^{-/-}$ (Maritzen et al., 2012), *Samhd1*$^{-/-}$ (Rehwinkel et al., 2013), *Myd88*$^{-/-}$, *Myd88*$^{-/-}$/*Ticam*$^{-/-}$, *Irak2*$^{-/-}$, *Irak4*$^{-/-}$ mice. All stimulations were performed using ultra-pure *E. coli* K12 LPS (Invivogen) at 100 ng/mL. For shRNA knockdowns, high-titer lentiviruses expressing shRNAs were used to infect bone marrow cells as previously described (Chevrier et al., 2011).

mRNA Measurements

Total or poly(A)$^+$ RNA was extracted and reverse transcribed prior to qPCR analysis with SYBR green (Roche) in triplicate with *Gapdh* for normalization. For mRNA counting, 5 × 10^4 bone-marrow-derived DCs were lysed in RLT buffer (QIAGEN) with 1% β-ME (beta-mercaptoethanol). 10% of the lysate was used for mRNA counting using the nCounter Digital Analyzer (NanoString) and a custom CodeSet constructed to detect a total of 267 genes (including 16 control genes whose expression remain unaffected by TLR stimulation). To determine significantly affected signature genes, a fold-change ratio is computed for each pairwise comparison of a knockdown sample versus a set of control samples (i.e., non-targeting shRNA; at least ten per experimental batch).

Affinity Purification followed by Mass Spectrometry

Analysis of interaction partners of V5-tagged proteins (MYD88, IRAK2 and AP1AR) was performed using a single-step purification procedure as previously described (Hubner and Mann, 2011), with several modifications. Peptide samples were analyzed on a Q Exactive mass spectrometer (Thermo Fisher Scientific), and mass spectra were processed as described above.

Large-Scale Phosphoproteome and Proteome Analyses

For temporal phosphoproteome analysis, BMDCs grown in SILAC media were stimulated with LPS and lysed and processed for enrichment of phosphopeptides using strong cation exchange chromatography (SCX)/IMAC (immobilized metal affinity chromatography) as described previously (Chevrier et al., 2011). For IVK and KO phosphoproteome analysis, peptide samples were separated by basic reversed-phase (RP) prior to IMAC enrichment as described previously (Mertins et al., 2013). IVK reactions were performed with recombinant kinases for IRAK2, IRAK4, or TBK1 on SILAC-labeled native cell lysates from DCs. For proteome analysis, total peptides were separated into 12 fractions using an Agilent 3100 Offgel fractionator. Peptide samples were analyzed on LTQ Orbitrap, LTQ Orbitrap Velos, or Q Exactive mass spectrometer (Thermo Fisher Scientific). To identify and quantify peptides, mass spectra were processed with the Spectrum Mill (Agilent Technologies) and the MaxQuant (version 1.2.2.5) software packages (Cox and Mann, 2008). Details on differential expression, clustering, pathway enrichment, and network analyses are in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the original mass spectra data reported in this paper is MassIVE : MSV000081071. The data are directly accessible via ftp://massive.ucsd.edu/MSV000081071.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2017.06.016.

AUTHOR CONTRIBUTIONS

N.C. designed the study and performed experiments; P.M. and J.Q. performed mass spectrometry; R.R. and T.M.E. contributed to experiments; D.P., P.M., N.Y., and N.C. performed computational analyses; T.M., V.H., T.S., and S.A. contributed knockout bone marrow cells; and P.M., D.P., S.A.C., A.R., N.H., and N.C. wrote the manuscript with input from all authors.

ACKNOWLEDGMENTS

We thank Diane Mathis, Laurie Glimcher, Paul Anderson, Mikael Pittet, Jeroen Saeij, Jonathan Kagan, Ido Amit, D.R. Mani, and members of the Hacohen, Regev, and Carr laboratories for helpful discussions. We are grateful to Yanick Crow and Gillian Rice for human fibroblast cell lines from Aicardi-Goutières syndrome patients and to Jan Rehwinkel for *Samhd1*$^{-/-}$ mouse bone marrow cells. This work was supported by NIH grant U54 AI057159, the NIH New Innovator Award DP2 OD002230 (N.H.), NIH grant P50 HG006193 (A.R., N.H.), and the Bauer Fellows Program (N.C.).

Received: November 6, 2016
Revised: April 11, 2017
Accepted: June 1, 2017
Published: June 27, 2017

REFERENCES

Allen, J.J., Li, M., Brinkworth, C.S., Paulson, J.L., Wang, D., Hübner, A., Chou, W.-H., Davis, R.J., Burlingame, A.L., Messing, R.O., et al. (2007). A semisynthetic epitope for kinase substrates. Nat. Methods *4*, 511–516.

Anderson, P. (2010). Post-transcriptional regulons coordinate the initiation and resolution of inflammation. Nat. Rev. Immunol. *10*, 24–35.

Ballana, E., and Esté, J.A. (2015). SAMHD1: at the crossroads of cell proliferation, immune responses, and virus restriction. Trends Microbiol. *23*, 680–692.

Bensimon, A., Heck, A.J.R., and Aebersold, R. (2012). Mass spectrometry-based proteomics and network biology. Annu. Rev. Biochem. *81*, 379–405.

Bodenmiller, B., Wanka, S., Kraft, C., Urban, J., Campbell, D., Pedrioli, P.G., Gerrits, B., Picotti, P., Lam, H., Vitek, O., et al. (2010). Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. Sci. Signal. *3*, rs4.

Brubaker, S.W., Bonham, K.S., Zanoni, I., and Kagan, J.C. (2015). Innate immune pattern recognition: a cell biological perspective. Annu. Rev. Immunol. *33*, 257–290.

Chevrier, N., Mertins, P., Artyomov, M.N., Shalek, A.K., Iannacone, M., Ciaccio, M.F., Gat-Viks, I., Tonti, E., DeGrace, M.M., Clauser, K.R., et al. (2011). Systematic discovery of TLR signaling components delineates viral-sensing circuits. Cell *147*, 853–867.

Clifford, R., Louis, T., Robbe, P., Ackroyd, S., Burns, A., Timbs, A.T., Wright Colopy, G., Dreau, H., Sigaux, F., Judde, J.G., et al. (2014). SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage. Blood *123*, 1021–1031.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Crow, Y.J., and Manel, N. (2015). Aicardi-Goutières syndrome and the type I interferonopathies. Nat. Rev. Immunol. *15*, 429–440.

Eriksen, A.B., Torgersen, M.L., Holm, K.L., Abrahamsen, G., Spurkland, A., Moskaug, J.Ø., Simonsen, A., and Blomhoff, H.K. (2015). Retinoic acid-induced IgG production in TLR-activated human primary B cells involves ULK1-mediated autophagy. Autophagy *11*, 460–471.

Ermann, J., Garrett, W.S., Kuchroo, J., Rourida, K., Glickman, J.N., Bleich, A., and Glimcher, L.H. (2011). Severity of innate immune-mediated colitis is controlled by the cytokine deficiency-induced colitis susceptibility-1 (Cdcs1) locus. Proc. Natl. Acad. Sci. USA *108*, 7137–7141.

Garber, M., Yosef, N., Goren, A., Raychowdhury, R., Thielke, A., Guttman, M., Robinson, J., Minie, B., Chevrier, N., Itzhaki, Z., et al. (2012). A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. Mol. Cell *47*, 810–822.

Geng, J., Sun, X., Wang, P., Zhang, S., Wang, X., Wu, H., Hong, L., Xie, C., Li, X., Zhao, H., et al. (2015). Kinases Mst1 and Mst2 positively regulate phagocytic induction of reactive oxygen species and bactericidal activity. Nat. Immunol. *16*, 1142–1152.

Gitter, A., Carmi, M., Barkai, N., and Bar-Joseph, Z. (2013). Linking the signaling cascades and dynamic regulatory networks controlling stress responses. Genome Res. *23*, 365–376.

Hasan, U.A., Caux, C., Perrot, I., Doffin, A.-C., Menetrier-Caux, C., Trinchieri, G., Tommasino, M., and Vlach, J. (2007). Cell proliferation and survival induced by Toll-like receptors is antagonized by type I IFNs. Proc. Natl. Acad. Sci. USA *104*, 8047–8052.

Huang, S.-S.C., and Fraenkel, E. (2009). Integrating proteomic, transcriptional, and interactome data reveals hidden components of signaling and regulatory networks. Sci. Signal. *2*, ra40.

Huang, S.-S.C., Clarke, D.C., Gosline, S.J.C., Labadorf, A., Chouinard, C.R., Gordon, W., Lauffenburger, D.A., and Fraenkel, E. (2013). Linking proteomic and transcriptional data through the interactome and epigenome reveals a map of oncogene-induced signaling. PLoS Comput. Biol. *9*, e1002887.

Hubner, N.C., and Mann, M. (2011). Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). Methods *53*, 453–459.

Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K., et al. (2015). The BioPlex network: a systematic exploration of the human interactome. Cell *162*, 425–440.

Kagan, J.C., Su, T., Horng, T., Chow, A., Akira, S., and Medzhitov, R. (2008). TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon-beta. Nat. Immunol. *9*, 361–368.

Kanshin, E., Bergeron-Sandoval, L.-P., Isik, S.S., Thibault, P., and Michnick, S.W. (2015). A cell-signaling network temporally resolves specific versus promiscuous phosphorylation. Cell Rep. *10*, 1202–1214.

Kawagoe, T., Sato, S., Matsushita, K., Kato, H., Matsui, K., Kumagai, Y., Saitoh, T., Kawai, T., Takeuchi, O., and Akira, S. (2008). Sequential control of Toll-like receptor-dependent responses by IRAK1 and IRAK2. Nat. Immunol. *9*, 684–691.

Li, S., Wang, L., Berman, M., Kong, Y.-Y., and Dorf, M.E. (2011). Mapping a dynamic innate immunity protein interaction network regulating type I interferon production. Immunity *35*, 426–440.

Liao, W., Bao, Z., Cheng, C., Mok, Y.-K., and Wong, W.S.F. (2008). Dendritic cell-derived interferon-γ-induced protein mediates tumor necrosis factor-α stimulation of human lung fibroblasts. Proteomics *8*, 2640–2650.

Maritzen, T., and Haucke, V. (2010). Gadkin: A novel link between endosomal vesicles and microtubule tracks. Commun. Integr. Biol. *3*, 299–302.

Maritzen, T., Zech, T., Schmidt, M.R., Krause, E., Machesky, L.M., and Haucke, V. (2012). Gadkin negatively regulates cell spreading and motility via sequestration of the actin-nucleating ARP2/3 complex. Proc. Natl. Acad. Sci. USA *109*, 10382–10387.

Mertins, P., Qiao, J.W., Patel, J., Udeshi, N.D., Clauser, K.R., Mani, D.R., Burgess, M.W., Gillette, M.A., Jaffe, J.D., and Carr, S.A. (2013). Integrated proteomic analysis of post-translational modifications by serial enrichment. Nat. Methods *10*, 634–637.

Miller, S.E., Mathiasen, S., Bright, N.A., Pierre, F., Kelly, B.T., Kladt, N., Schauss, A., Merrifield, C.J., Stamou, D., Höning, S., and Owen, D.J. (2015). CALM regulates clathrin-coated vesicle size and maturation by directly sensing and driving membrane curvature. Dev. Cell *33*, 163–175.

Pauls, E., Ruiz, A., Badia, R., Permanyer, M., Gubern, A., Riveira-Muñoz, E., Torres-Torronteras, J., Álvarez, M., Mothe, B., Brander, C., et al. (2014). Cell cycle control and HIV-1 susceptibility are linked by CDK6-dependent CDK2 phosphorylation of SAMHD1 in myeloid and lymphoid cells. J. Immunol. *193*, 1988–1997.

Picard, C., Puel, A., Bonnet, M., Ku, C.-L., Bustamante, J., Yang, K., Soudais, C., Dupuis, S., Feinberg, J., Fieschi, C., et al. (2003). Pyogenic bacterial infections in humans with IRAK-4 deficiency. Science *299*, 2076–2079.

Qu, Y., Misaghi, S., Izrael-Tomasevic, A., Newton, K., Gilmour, L.L., Lamkanfi, M., Louie, S., Kayagaki, N., Liu, J., Kömüves, L., et al. (2012). Phosphorylation of NLRC4 is critical for inflammasome activation. Nature *490*, 539–542.

Rehwinkel, J., Maelfait, J., Bridgeman, A., Rigby, R., Hayward, B., Liberatore, R.A., Bieniasz, P.D., Towers, G.J., Moita, L.F., Crow, Y.J., et al. (2013). SAMHD1-dependent retroviral control and escape in mice. EMBO J. *32*, 2454–2462.

Sanada, T., Takaesu, G., Mashima, R., Yoshida, R., Kobayashi, T., and Yoshimura, A. (2008). FLN29 deficiency reveals its negative regulatory role in the Toll-like receptor (TLR) and retinoic acid-inducible gene I (RIG-I)-like helicase signaling pathway. J. Biol. Chem. *283*, 33858–33864.

Santra, T., Kolch, W., and Kholodenko, B.N. (2014). Navigating the multilayered organization of eukaryotic signaling: a new trend in data integration. PLoS Comput. Biol. *10*, e1003385.

Schuh, A., Becq, J., Humphray, S., Alexa, A., Burns, A., Clifford, R., Feller, S.M., Grocock, R., Henderson, S., Khrebtukova, I., et al. (2012). Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. Blood *120*, 4191–4196.

Sharma, K., Kumar, C., Kéri, G., Breitkopf, S.B., Oppermann, F.S., and Daub, H. (2010). Quantitative analysis of kinase-proximal signaling in lipopolysaccharide-induced innate immune response. J. Proteome Res. *9*, 2539–2549.

Sjoelund, V., Smelkinson, M., and Nita-Lazar, A. (2014). Phosphoproteome profiling of the macrophage response to different toll-like receptor ligands identifies differences in global phosphorylation dynamics. J. Proteome Res. *13*, 5185–5197.

Skerker, J.M., Perchuk, B.S., Siryaporn, A., Lubin, E.A., Ashenberg, O., Goulian, M., and Laub, M.T. (2008). Rewiring the specificity of two-component signal transduction systems. Cell *133*, 1043–1054.

Suzuki, N., Suzuki, S., Duncan, G.S., Millar, D.G., Wada, T., Mirtsos, C., Takada, H., Wakeham, A., Itie, A., Li, S., et al. (2002). Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4. Nature *416*, 750–756.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. Cell *140*, 805–820.

Tanigawa, K., Suzuki, K., Kimura, H., Takeshita, F., Wu, H., Akama, T., Kawashima, A., and Ishii, N. (2009). Tryptophan aspartate-containing coat protein (CORO1A) suppresses Toll-like receptor signalling in Mycobacterium leprae infection. Clin. Exp. Immunol. *156*, 495–501.

Terfve, C.D.A., Wilkes, E.H., Casado, P., Cutillas, P.R., and Saez-Rodriguez, J. (2015). Large-scale models of signal propagation in human cells derived from discovery phosphoproteomic data. Nat. Commun. *6*, 8033.

Weintz, G., Olsen, J.V., Frühauf, K., Niedzielska, M., Amit, I., Jantsch, J., Mages, J., Frech, C., Dölken, L., Mann, M., and Lang, R. (2010). The phosphoproteome of toll-like receptor-activated macrophages. Mol. Syst. Biol. *6*, 371.

Yugi, K., Kubota, H., Hatano, A., and Kuroda, S. (2016). Trans-Omics: how to reconstruct biochemical networks across multiple 'omic' layers. Trends Biotechnol. *34*, 276–290.

Cell Reports, Volume *19*

Supplemental Information

An Integrative Framework Reveals Signaling-to-Transcription Events in Toll-like Receptor Signaling

Philipp Mertins, Dariusz Przybylski, Nir Yosef, Jana Qiao, Karl Clauser, Raktima Raychowdhury, Thomas M. Eisenhaure, Tanja Maritzen, Volker Haucke, Takashi Satoh, Shizuo Akira, Steven A. Carr, Aviv Regev, Nir Hacohen, and Nicolas Chevrier

Figure S1

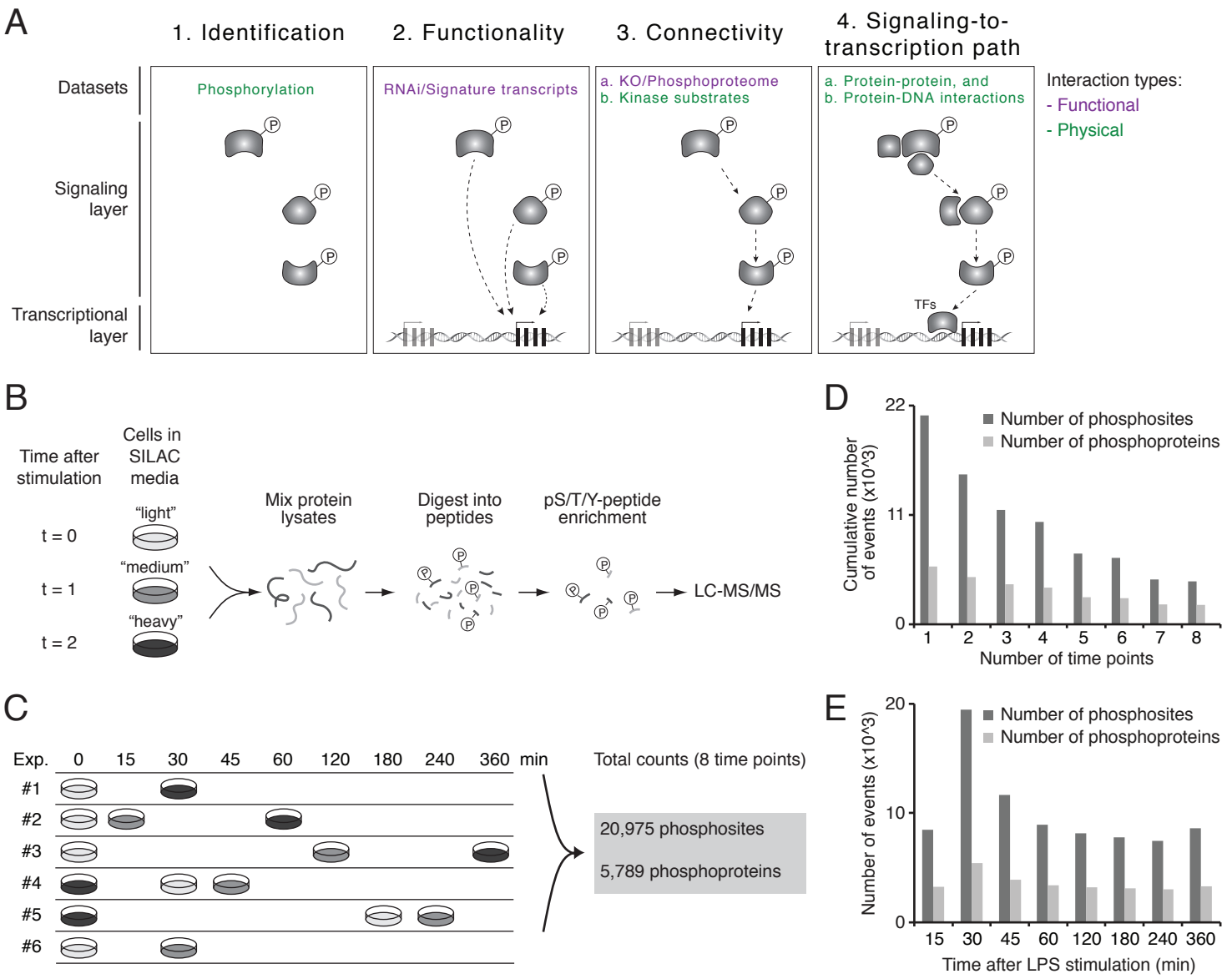


Figure S1. Overview of the experimental design and outcome of temporal phosphoproteomic profiling of TLR4-stimulated DCs, Related to Figure 1.

(A) Diagram highlighting the general steps of our integrative approach that includes the identification of temporal changes in phosphorylation (1), functional testing using genetic perturbations (2), connecting kinases and substrates (3), and inferring biochemical paths linking signaling to transcription events (4). All these steps rely on data capturing physical and functional interactions in mouse primary DCs stimulated with LPS.

(B) Schematic depiction of the experimental workflow for phosphoproteomics. From left: Protein lysates from unstimulated (t = 0) and LPS-treated cells (t = 1 and 2) grown in "light", "medium" or "heavy" SILAC media were mixed (1:1:1) and digested into peptides with trypsin before phospho-serine, -threonine, and -tyrosine (pS/T/Y) peptide enrichment using immobilized metal affinity chromatography (IMAC), and LC-MS/MS analysis (see Experimental Procedures).

(C) Schematic summary of the temporal phosphoproteomic profiling. Six experiments were conducted (from top to bottom) to cover control and stimulated cells (top, 8 time points post-LPS stimulation) with SILAC label switching as indicated with the colored-dishes matching the nomenclature from B. Total counts of phosphosites and phosphoproteins detected across all time points are indicated on the right.

(D-E) Distributions of phosphosites and phosphoproteins detected across multiple (cumulative counts; D) and individual time points (E).

Figure S2

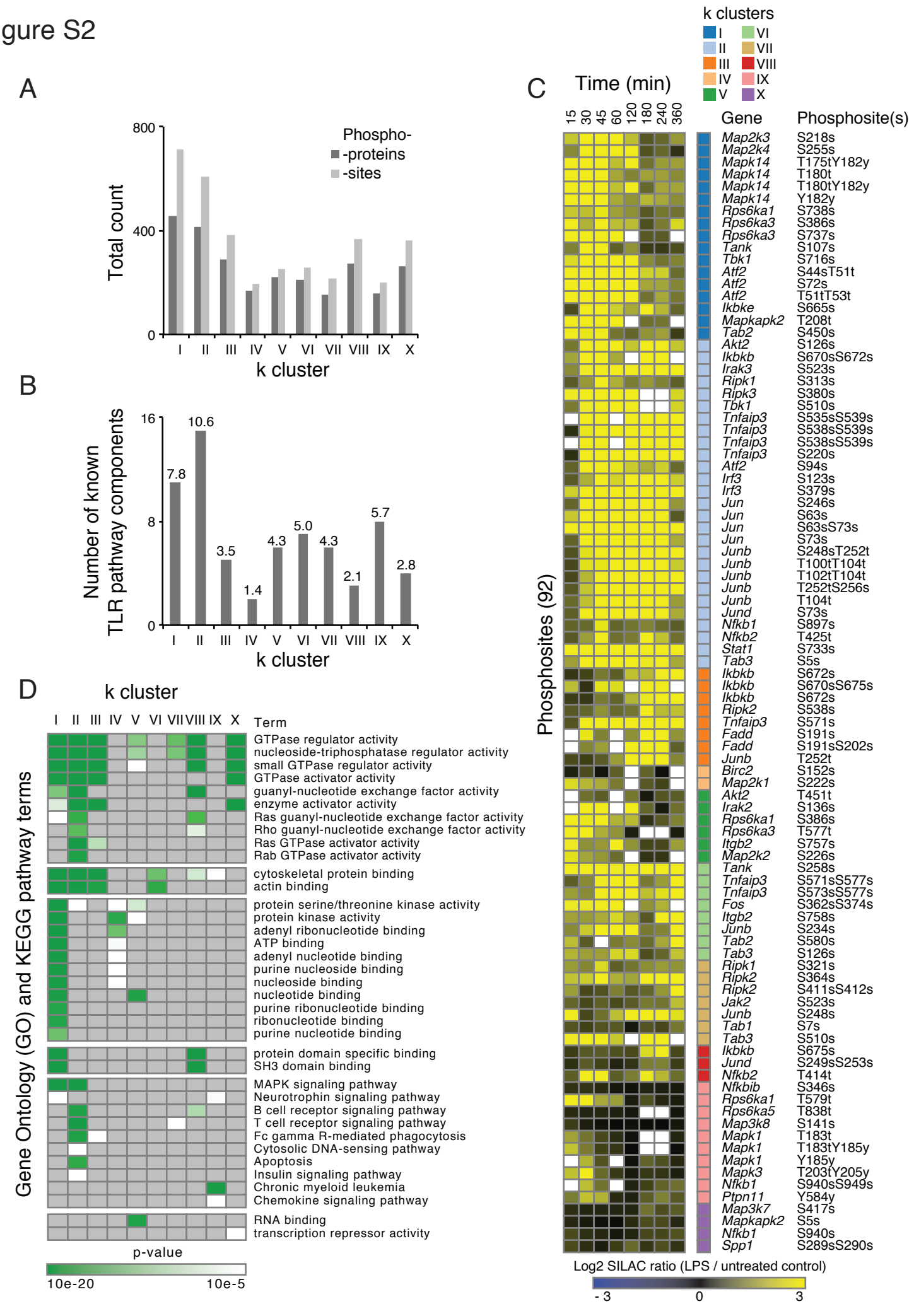

A
800
400
0
Total count
Phospho-
-proteins
-sites
I II III IV V VI VII VIII IX X
k cluster
B
16
8
0
Number of known
TLR pathway components
7.8
10.6
3.5
1.4
4.3
5.0
4.3
2.1
5.7
2.8
I II III IV V VI VII VIII IX X
k cluster
C
Time (min)
15 30 45 60 120 180 240 360
k clusters
I II III IV V VI VII VIII IX X
Gene
Phosphosite(s)
Phosphosites (92)
Map2k3 S218s
Map2k4 S255s
Mapk14 T175tY182y
Mapk14 T180t
Mapk14 T180tY182y
Mapk14 Y182y
Rps6ka1 S738s
Rps6ka3 S386s
Rps6ka3 S737s
Tank S107s
Tbk1 S716s
Atf2 S44sT51t
Atf2 S72s
Atf2 T51tT53t
Ikbke S665s
Mapkapk2 T208t
Tab2 S450s
Akt2 S126s
Ikbkb S670sS672s
Irak3 S523s
Ripk1 S313s
Ripk3 S380s
Tbk1 S510s
Tnfaip3 S535sS539s
Tnfaip3 S538sS539s
Tnfaip3 S538sS539s
Tnfaip3 S220s
Atf2 S94s
Irf3 S123s
Irf3 S379s
Jun S246s
Jun S63s
Jun S63sS73s
Jun S73s
Junb S248sT252t
Junb T100tT104t
Junb T102tT104t
Junb T252tS256s
Junb T104t
Jund S73s
Nfkb1 S897s
Nfkb2 T425t
Stat1 S733s
Tab3 S5s
Ikbkb S672s
Ikbkb S670sS675s
Ikbkb S672s
Ripk2 S538s
Tnfaip3 S571s
Fadd S191s
Fadd S191sS202s
Junb T252t
Birc2 S152s
Map2k1 S222s
Akt2 T451t
Irak2 S136s
Rps6ka1 S386s
Rps6ka3 T577t
Itgb2 S757s
Map2k2 S226s
Tank S258s
Tnfaip3 S571sS577s
Tnfaip3 S573sS577s
Fos S362sS374s
Itgb2 S758s
Junb S234s
Tab2 S580s
Tab3 S126s
Ripk1 S321s
Ripk2 S364s
Ripk2 S411sS412s
Jak2 S523s
Junb S248s
Tab1 S7s
Tab3 S510s
Ikbkb S675s
Jund S249sS253s
Nfkb2 T414t
Nfkbib S346s
Rps6ka1 T579t
Rps6ka5 T838t
Map3k8 S141s
Mapk1 T183t
Mapk1 T183tY185y
Mapk1 Y185y
Mapk3 T203tY205y
Nfkb1 S940sS949s
Ptpn11 Y584y
Map3k7 S417s
Mapkapk2 S5s
Nfkb1 S940s
Spp1 S289sS290s
Log2 SILAC ratio (LPS / untreated control)
- 3
0
3
D
k cluster
I II III IV V VI VII VIII IX X
Term
Gene Ontology (GO) and KEGG pathway terms
GTPase regulator activity
nucleoside-triphosphatase regulator activity
small GTPase regulator activity
GTPase activator activity
guanyl-nucleotide exchange factor activity
enzyme activator activity
Ras guanyl-nucleotide exchange factor activity
Rho guanyl-nucleotide exchange factor activity
Ras GTPase activator activity
Rab GTPase activator activity
cytoskeletal protein binding
actin binding
protein serine/threonine kinase activity
protein kinase activity
adenyl ribonucleotide binding
ATP binding
adenyl nucleotide binding
purine nucleoside binding
nucleoside binding
nucleotide binding
purine ribonucleotide binding
ribonucleotide binding
purine nucleotide binding
protein domain specific binding
SH3 domain binding
MAPK signaling pathway
Neurotrophin signaling pathway
B cell receptor signaling pathway
T cell receptor signaling pathway
Fc gamma R-mediated phagocytosis
Cytosolic DNA-sensing pathway
Apoptosis
Insulin signaling pathway
Chronic myeloid leukemia
Chemokine signaling pathway
RNA binding
transcription repressor activity
p-value
10e-20
10e-5

Figure S2. Dynamic phosphoproteomic profiles reveal phosphosites regulated on known TLR components and differential pathway enrichment over time, Related to Figure 2.

(A-B) Distributions of phosphosites and phosphoproteins (A; Y axis), and known TLR pathway proteins (B; Y axis) detected across the 10 k-means clusters from Figure 1A (X axis). Numbers on top of each bar indicate the percent of known TLR proteins within each cluster (B).

(C) Phosphorylation profiles of known TLR pathway proteins. Log2 fold changes between LPS-treated and untreated cells for 92 phosphosites on known TLR proteins (rows) detected in at least 6 out of 8 time points (columns). Phosphosites are partitioned into 10 clusters using k-means (legend, top; color bars, right). Right, gene names and phosphosites localization (S, serine; T, threonine; Y, tyrosine). White indicates missing values.

(D) Gene enrichment analysis of LPS-dependent phosphoproteins. Enrichment p-values (modified Fisher's exact test, showing values $< 10^{-5}$; grey boxes indicate values above this cutoff) for Gene Ontology (GO) and KEGG pathway terms (rows; term name indicated on the right) across all 10 k-means clusters (columns).

Figure S3

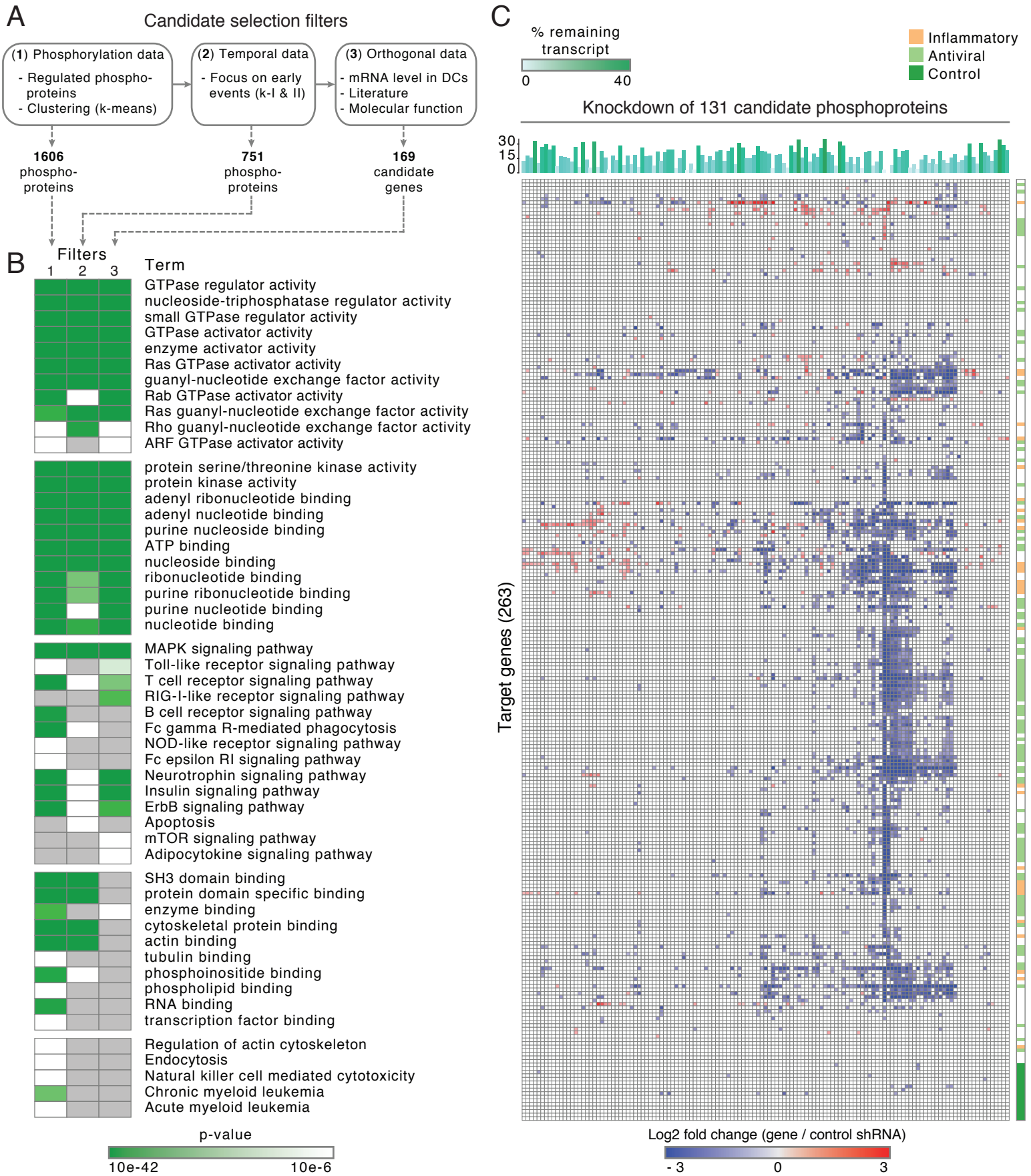

A
Candidate selection filters
(1) Phosphorylation data
- Regulated phospho-proteins
- Clustering (k-means)
(2) Temporal data
- Focus on early events (k-I & II)
(3) Orthogonal data
- mRNA level in DCs
- Literature
- Molecular function
1606 phospho-proteins
751 phospho-proteins
169 candidate genes
B
Filters
1 2 3
Term
GTPase regulator activity
nucleoside-triphosphatase regulator activity
small GTPase regulator activity
GTPase activator activity
enzyme activator activity
Ras GTPase activator activity
guanyl-nucleotide exchange factor activity
Rab GTPase activator activity
Ras guanyl-nucleotide exchange factor activity
Rho guanyl-nucleotide exchange factor activity
ARF GTPase activator activity
protein serine/threonine kinase activity
protein kinase activity
adenyl ribonucleotide binding
adenyl nucleotide binding
purine nucleoside binding
ATP binding
nucleoside binding
ribonucleotide binding
purine ribonucleotide binding
purine nucleotide binding
nucleotide binding
MAPK signaling pathway
Toll-like receptor signaling pathway
T cell receptor signaling pathway
RIG-I-like receptor signaling pathway
B cell receptor signaling pathway
Fc gamma R-mediated phagocytosis
NOD-like receptor signaling pathway
Fc epsilon RI signaling pathway
Neurotrophin signaling pathway
Insulin signaling pathway
ErbB signaling pathway
Apoptosis
mTOR signaling pathway
Adipocytokine signaling pathway
SH3 domain binding
protein domain specific binding
enzyme binding
cytoskeletal protein binding
actin binding
tubulin binding
phosphoinositide binding
phospholipid binding
RNA binding
transcription factor binding
Regulation of actin cytoskeleton
Endocytosis
Natural killer cell mediated cytotoxicity
Chronic myeloid leukemia
Acute myeloid leukemia
p-value
10e-42
10e-6
C
% remaining transcript
0
40
Inflammatory
Antiviral
Control
Knockdown of 131 candidate phosphoproteins
30
15
0
Target genes (263)
Log2 fold change (gene / control shRNA)
- 3
0
3

Figure S3. Genetic perturbation profiles of the 131 phosphoproteins selected from phosphoproteomic profiles, Related to Figure 3.

(A-B) Candidate filtering and associated gene enrichment analysis of LPS-dependent phosphoproteins selected for functional analysis. Enrichment p-values (modified Fisher's exact test, showing values < 10^{-5}; grey boxes indicate values above this cutoff) for Gene Ontology (GO) and KEGG pathway terms (rows; term name indicated on the right) across all 3 filters for candidate gene selection (columns).

(C) Perturbation profiles of 131 phosphoproteins. Shown are the perturbed candidates and control genes (columns) and the log2 fold changes between gene-specific and control shRNAs (rows) of 263 target genes (including control targets used as "housekeeping", unchanged genes for normalization). The right-most column categorizes target genes into controls (dark green), and antiviral (light green) and inflammatory (light orange) programs. Top, bar plot indicating knockdown efficiency for each perturbed gene (top left, percentage of remaining mRNA transcripts for indicated genes upon knockdown).

Figure S4

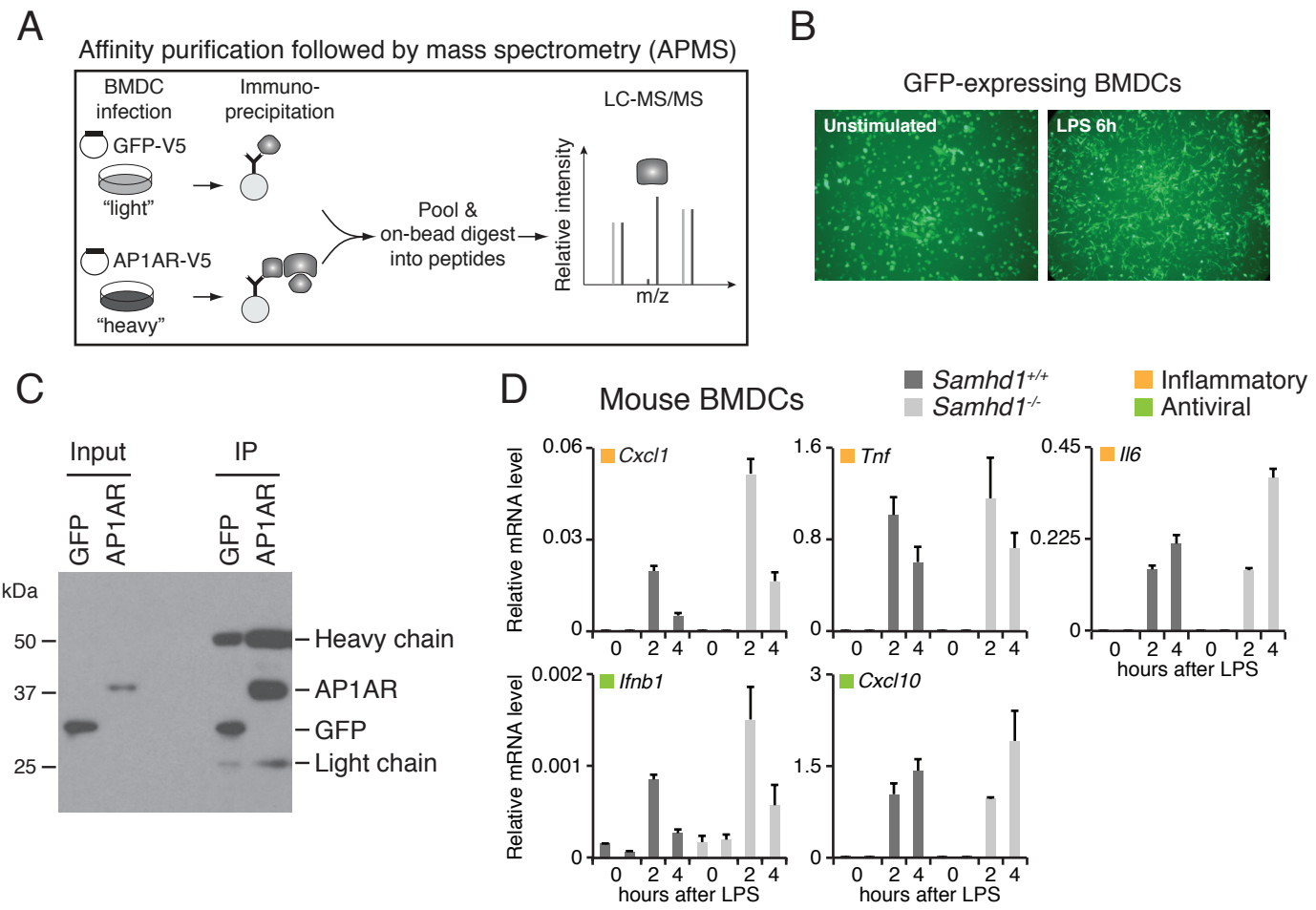


Figure S4. A method for affinity-purification followed by MS (AP-MS) in BMDCs, and analysis of Samhd1$^{-/-}$ mouse BMDCs, Related to Figure 4.

(A) Diagram depicting our experimental approach for AP-MS in mouse BMDCs.

(B) BMDCs overexpressing V5-tagged proteins through lentiviral infection respond normally to LPS. Micrographs of GFP-expressing DCs before and after LPS stimulation.

(C) Immunoblot analysis of input and immunoprecipitated (IP) samples from DCs expressing V5-GFP or V5-AP1AR.

(D) Expression levels (relative to Gapdh; qPCR) of indicated inflammatory (light orange) and antiviral (light green) cytokines in Samhd1$^{+/+}$ and Samhd1$^{-/-}$ BMDCs stimulated with LPS for 2 and 4 h or left untreated as control (duplicate wells are shown).

Figure S5

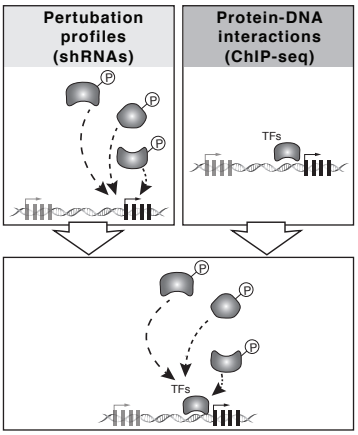


Figure S5. Measuring overlaps between knockdown profiles and TF binding sites, Related to Figure 5.

Diagram depicting overlaps between genes affected by knockdown and genes whose promoters are bound by a given transcription factor (TF), as a schematic example of the analysis performed in Figure 5B.

Figure S6

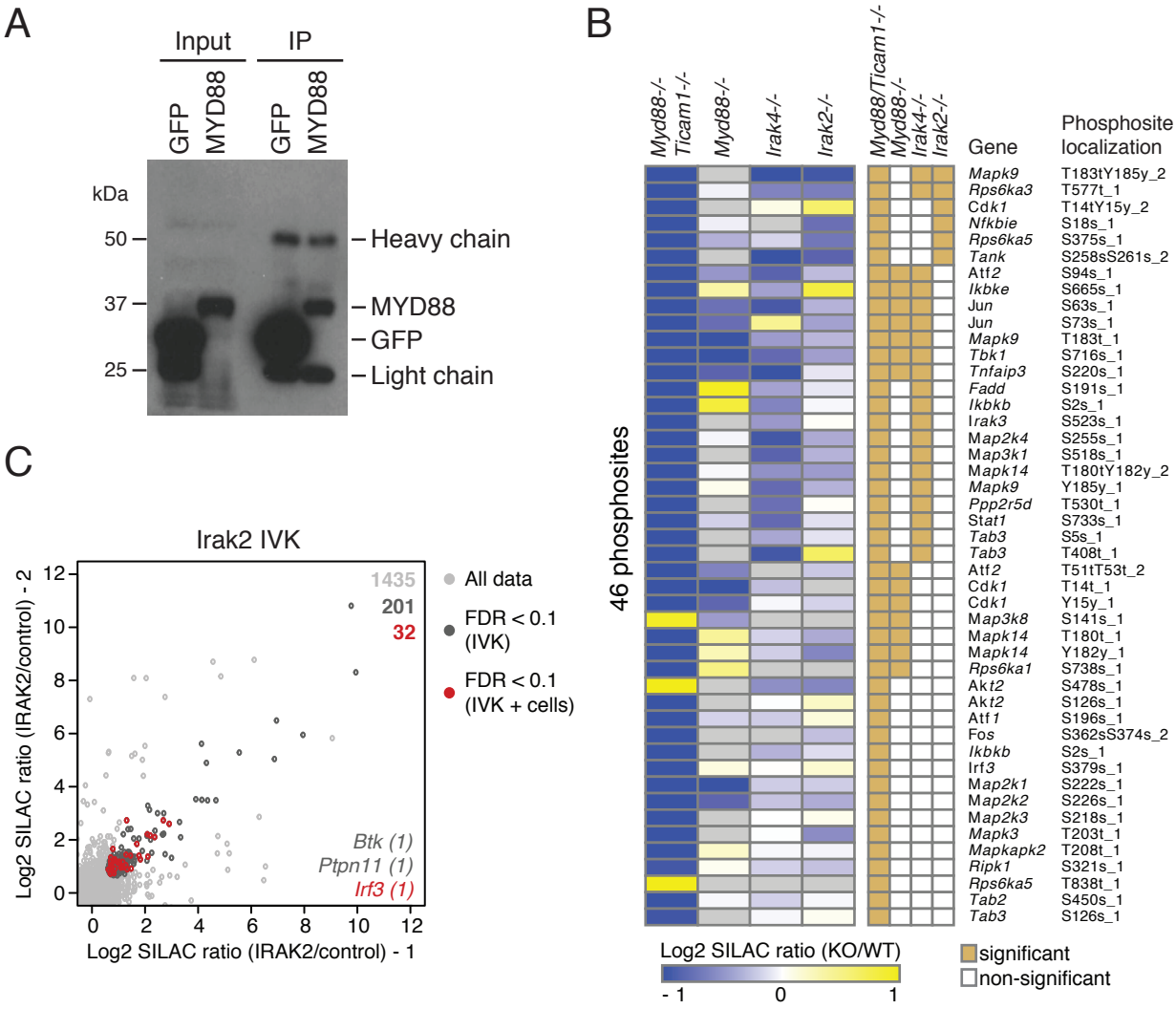


Figure S6. Physical and functional proteomics assays pinpoint how phosphorylation of known TLR pathway regulators is modulated by KO, Related to Figure 6.

(A) Immunoblot analysis of input and immunoprecipitated (IP) samples from DCs expressing V5-GFP or V5-MYD88.

(B) Phosphoproteomics in KO cells. Left, shown is a heatmap for SILAC ratios of phosphosites (rows) in 4 KO models (columns) at 30 min after LPS stimulation compared to control wild-type cells, as indicated (grey, missing values). Middle, shown in light brown are phosphosites with significant up- or down-regulation in KO vs WT. Right, shown are the 46 phosphosites mapping to known TLR signaling regulators (Letters on the right indicated the phosphorylated residue: S, serine; T, threonine; Y, tyrosine. Numbers indicated the amino-acid position in the protein).

(C) In vitro kinase (IVK) assay followed by phosphoproteomics. Shown are dot plots of SILAC ratios of phosphosites identified using purified IRAK2. Light grey, all data points; dark grey, phosphosites with FDR < 0.1 in IVK; red, phosphosites with FDR < 0.1 in both IVK and in cells stimulated with LPS, which highlights the overlap between IVK and phosphoproteome measurements on stimulated cells (denoted as IVK + cells). Gene names at the bottom right of the plot indicates known TLR components with the number of phosphosites in parenthesis.

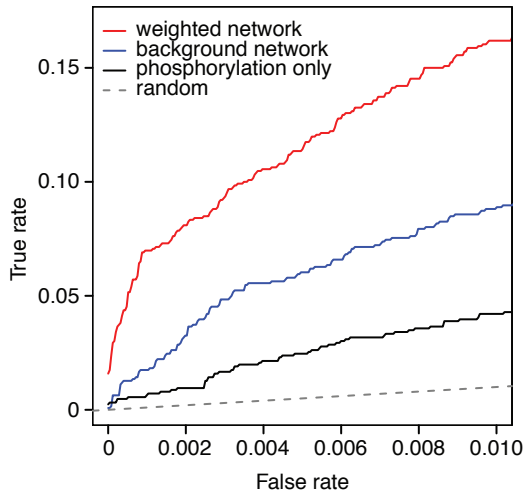
Figure S7

A

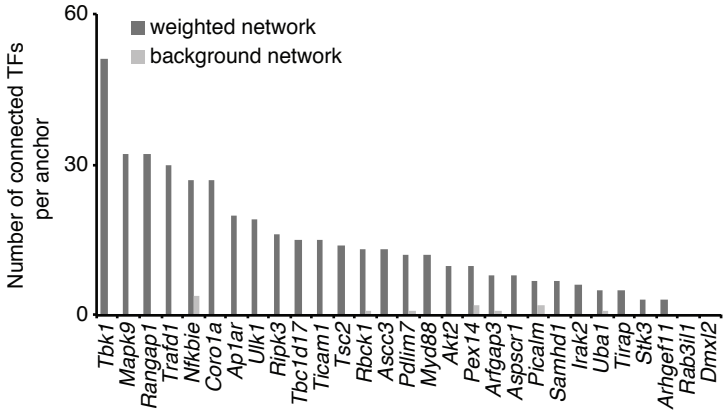
Summary of data collected in BMDCs on physical and functional interactions

	Datasets	Measurements	Matching figures
Physical interactions	Phosphorylation	Phosphoproteomics	Figure 1C
	Large-scale IVK	Phosphoproteomics	Figure 6F, 6G, and S6C
	Protein-protein interactions	AP-MS	Figure 4D, 6A, and 6B
	Protein-DNA interactions	TF ChIP-seq	Figure 5B
Functional interactions	shRNA / Signature transcripts	Multiplex mRNA counting	Figure 3B and S3C
	KO / Phosphoproteome	Phosphoproteomics	Figure 6D

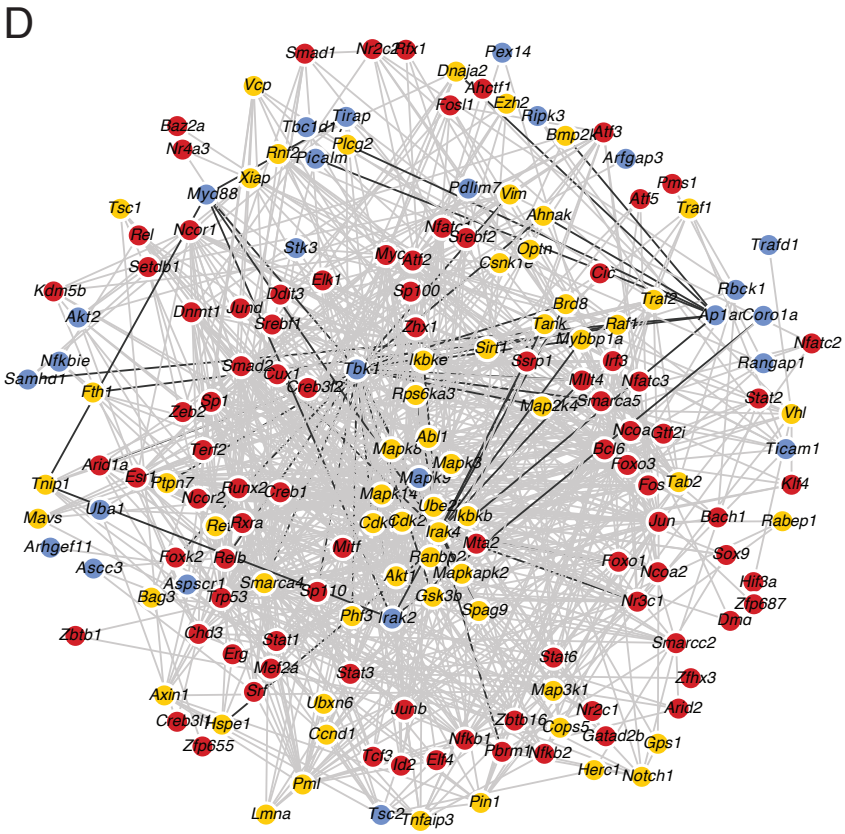
B



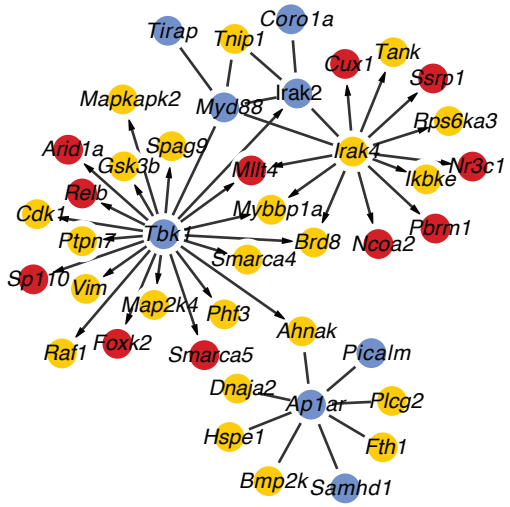
C



D



E



Seeds (26)
Intermediates (60)
Transcriptional regulators (95)
Public database interactions
Local interaction data:
(i) APMS
(ii) IVK

Figure S7. An integrative analysis reveals signaling-to-transcription paths in the TLR4 system, Related to Figure 7.
(A) List summarizing the physical and functional interaction data sets collected in the cellular context of BMDCs stimulated with LPS.
(B) Receiver operator characteristic (ROC) analysis demonstrates the ability of our integrative algorithm to retrieve known seed-target node relationships between canonical TLR pathway components.
(C) Total number of transcription regulators (TRs) for which significant relationships were found with the 29 'seed' nodes using 'weighted network' (dark grey) and 'background network' (light grey) methods.
(D) An interaction network connects 27 seeds (blue) to 95 transcriptional regulators (red) through the top 60 intermediate (yellow) nodes. Shown is a close-up view of the network from Figure 7E depicting protein names on nodes and edge types (database, AP-MS, and/or IVK).
(E) Sub-network extracted from panel D showing interactions from local data sets as dark edges (AP-MS, undirected; IVK, directed).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cells

Bone marrow-derived dendritic cells (BMDCs) were generated from 6-8 week old female C57BL/6J (Jackson Laboratories), *Ap1ar*$^{-/-}$ (Maritzen et al., 2012), *Samhd1*$^{-/-}$ (Rehwinkel et al., 2013), *Myd88*$^{-/-}$, *Myd88*$^{-/-}$*/Ticam*$^{-/-}$, *Irak2*$^{-/-}$, *Irak4*$^{-/-}$ mice. Bone marrow cells were collected from femora and tibiae and plated at 10^6 cells/mL on non-tissue culture treated petri dishes or 96-well plates in RPMI-1640 medium (Gibco), supplemented with 10% FBS, L-glutamine, penicillin/streptomycin, MEM non-essential amino acids, HEPES, sodium pyruvate, β-mercaptoethanol, and murine GM-CSF (15 ng/mL; Peprotech). GM-CSF-derived BMDCs were used directly for all RNAi experiments. For all other experiments, floating cells from GM-CSF cultures were sorted at day 5 by MACS using the CD11c (N418) MicroBeads kit (Miltenyi Biotec), or used directly at day 8. Sorted CD11c$^+$ or floating cells were used as GM-CSF-derived BMDCs, and plated at 10^6 cells/mL and stimulated at 16 h post sorting or collection. Human fibroblasts used in this study were AGS128 SI (c.445C>T p.Gln149* hom, referred to as M1), AGS495 SA (c.1609-1G>C hom, referred to as M2), and F10Y, F8Y, and PBX2 are controls (referred to as H1-3), and were maintained in DMEM supplemented with 10% FBS. All cell stimulations were performed using ultra-pure *E. coli* K12 LPS (lipopolysaccharide) from Invivogen at 100 ng/mL for indicated times.

mRNA isolation

Total RNA was extracted with QIAzol reagent following the miRNeasy kit's procedure (Qiagen), and reverse transcribed with the High Capacity cDNA Reverse Transcription kit (Applied

Biosystems). For experiments with more than 12 samples, we harvested PolyA+ RNA in 96- or 384-well plates with the Turbocapture mRNA kit (Qiagen) and reverse transcribed with the Sensiscript RT kit (Qiagen).

qPCR measurements

Real time quantitative PCR reactions were performed on the LightCycler 480 system (Roche) with FastStart Universal SYBR Green Master Mix (Roche). Every reaction was run in triplicate and GAPDH levels were used as an endogenous control for normalization.

shRNA knockdowns

High titer lentiviruses encoding shRNAs targeting genes of interest were obtained from The RNAi Consortium (TRC; Broad Institute, Cambridge, MA, USA). Bone marrow cells were infected with lentiviruses as described (Chevrier et al., 2011). For each gene of interest, we tested five shRNAs for knockdown efficiency using qPCR of the target gene and selected shRNAs with best knockdown efficacy (typically >75%).

mRNA counting and data analysis

5×10^4 bone marrow-derived DCs were lysed in RLT buffer (Qiagen) with 1% β-ME. 10% of the lysate was used for mRNA counting using the nCounter Digital Analyzer (NanoString) and a custom CodeSet constructed to detect a total of 267 genes (including 16 control genes whose expression remain unaffected by TLR stimulation). We normalize data by dividing the nCounter mRNA count values for each gene by the sum of counts obtained for the 16 control genes present in our custom CodeSet. To determine significantly affected signature genes, a fold-change ratio is

computed for each pairwise comparison of a knockdown sample versus a set of control samples (i.e., non-targeting shRNA; at least 10 per experimental batch). As a threshold, we require a substantial fold-change (above a threshold value t) in the same direction (up- or down-regulation) in more than half of the pairwise comparisons sample vs. control shRNA. The threshold value t is determined as max (q, d), d being the mean + 1.645 times the standard deviation in the fold change shown by the control genes (corresponding top = 0.05, under the assumption of normality). The threshold q is similar for all comparisons and is based on the noise level estimated from the control shRNA samples. Specifically, we compute gene expression fold changes in all possible pairs of control shRNA samples (which are supposed to be consistent). We set the threshold q such that 95% of the comparisons exhibit lower fold change than q. The resulting value of q is 1.961. Notably, we ignore all pairwise comparisons in which both control and knockdown samples had low counts before normalization (<50). All heatmaps and distance matrix analyses were generated using the software Gene-E (https://software.broadinstitute.org/GENE-E/index.html).

Metabolic labeling of cells

For stable isotope labeling of amino acids in cell culture (SILAC) experiments, GM-CSF-derived BMDCs were grown for seven days in media containing either normal L-arginine (Arg-0) and L-lysine (Lys-0) (Sigma), L-arginine 13C6 (Arg-6) and L-lysine D4 (Lys-4), or L-arginine 13C6-15N4 (Arg-10) and L-lysine 13C6-15N2 (Lys-8) (Sigma Isotec). Concentrations for L-arginine and L-lysine were 42 mg/L and 73 mg/L, respectively. To prevent metabolic conversion of L-arginine to L-proline we added 200 mg/L L-proline to the cell culture medium. The cell culture media, Roswell Park Memorial Institute-1640 (RPMI) deficient in L-arginine and L-lysine, was a custom media preparation from Caisson Laboratories (North Logan, UT) and dialyzed serum was

obtained from SAFC-Sigma. We followed all standard SILAC media preparation and labeling steps as previously described (Chevrier et al., 2011).

Global serine, threonine, and tyrosine phosphorylation analysis for LPS time course experiments

BMDCs grown in SILAC media were stimulated with LPS, and lysed and processed for enrichment of phosphopeptides as described previously (Chevrier et al., 2011; Mertins et al., 2013). Briefly, after LPS stimulation, cells grown in non-TC treated Petri dishes were placed on ice and scraped. Cell suspensions were washed in ice-cold PBS and sedimented by centrifugation at 4°C and 1,000 g for 5 minutes. The supernatant was removed and cell pellets were immediately frozen in liquid nitrogen. Cell pellets were lysed for 20 minutes in ice-cold lysis buffer containing 8 M Urea, 75 mM NaCl, 50 mM Tris pH 8.0, 1 mM EDTA, 2 µg/ml Aprotinin (Sigma, A6103), 10 µg/ml Leupeptin (Roche, #11017101001), 1 mM PMSF, 10 mM NaF, 2 mM Na3VO4, 50 ng/ml Calyculin A (Calbiochem, #208851), Phosphatase inhibitor cocktail 1 (1/100, Sigma, P2850) and Phosphatase inhibitor cocktail 2 (1/100, Sigma, P5726). Lysates were pre-cleared by centrifugation at 16,500 g for 10 min and protein concentrations were determined by BCA assay (Pierce). We obtained on average 1 mg of total protein per label out of 10 million cells. Cell lysates were mixed in equal protein amounts per label and proteins were reduced with 5 mM dithiothreitol and alkylated with 10 mM iodoacetamide. Samples were diluted 1:4 with HPLC water (Baker) and sequencing-grade modified trypsin (Promega, V5113) was added in an enzyme to substrate ratio of 1:150. After 16 h digest, samples were acidified with 0.5% trifluoroacetic acid (final concentration). Tryptic peptides were desalted on reverse phase tC18 SepPak columns (Waters, 500 mg, WAT036790) and dried in a vacuum concentrator centrifuge. Before phosphopeptide

enrichment peptides were separated using strong cation exchange (SCX) chromatography. Peptides were reconstituted in 500 µl strong cation exchange buffer A (7 mM KH2PO4, pH 2.65, 30% MeCN) and separated on a Polysulfoethyl A column from PolyLC (250 x 9.4 mm, 5 µm particle size, 200 A pore size) using an Akta Purifier 10 system (GE Healthcare). We used an 80-min gradient with a 20-min equilibration phase with buffer A, a linear increase to 30% buffer B (7 mM KH2PO4, pH 2.65, 350 mM KCL, 30% MeCN) within 33 min, 100% B for 7 min and a final equilibration with Buffer A for 20 min. The flow rate was 3 ml/min and the sample was injected after the initial 20 min equilibration phase. Upon injection, 3 ml fractions were collected with a P950 fraction collector throughout the run. 60 fractions were collected of which 3-4 adjacent fractions were combined to obtain 12 samples. The 12 fractions were desalted with reverse phase tC18 SepPak columns (Waters, 100 mg, WAT036820) and lyophilized to dryness. SCX-separated samples were enriched for phosphopeptdies by immobilized metal affinity chromatography (IMAC) as described previously (Chevrier et al., 2011). Peptides were reconstituted in 200 µl IMAC binding buffer (40% MeCN, 0.1% FA) and incubated for 1 h with 5 µl of packed Phos-Select beads (Sigma, P9740) in batch mode. After incubation, samples were loaded on C18 StageTips, washed twice with 50µl IMAC binding buffer and washed once with 50µl 1% formic acid. Phosphorylated peptides were eluted from the Phos-Select resin to the C18 material by loading 3 times 70 µl of 500 mM K2HPO4 (pH 7.0). StageTips were washed with 50 µl of 1% formic acid to remove phosphate salts and eluted with 80 µl of 50% MeCN / 0.1 % formic acid. Samples were dried down by vacuum centrifugation and reconstituted in 8 µl 3% MeCN / 0.1 % formic acid. Peptide samples were separated on an online nanoflow HPLC system (Agilent 1200) and analyzed on a LTQ Orbitrap and a LTQ Orbitrap Velos instrument, as described (Chevrier et al., 2011). Briefly, 50% of the enriched phosphopeptide samples were loaded onto a 14-cm reverse

phase fused-silica capillary column (New Objective, PicoFrit PF360-75-10-N-5 with 10 µm tip opening and 75 µm inner diameter) packed in-house with 3 µm ReproSil-Pur C18-AQ media (Dr. Maisch GmbH). The HPLC setup was connected via a custom-made electrospray ion source to the mass spectrometer. After sample injection, peptides were separated at an analytical flowrate of 200 nL/min with a 70-min linear gradient (~ 0.29 %B/min) from 10% solvent A (0.1% formic acid in water) to 30% solvent B (0.1% formic acid/90% acetonitrile). The run time was 130 min for a single sample, including sample loading and column reconditioning. Data-dependent acquisition was performed using the Xcalibur 2.1 software in positive ion mode. Survey spectra were acquired in the orbitrap with a resolution of 60,000 and a mass range from 350 to 1750 m/z. In parallel, up to 16 of the most intense ions per cycle were isolated, fragmented and analyzed in the LTQ part of the instrument. Ions selected for MS/MS were dynamically excluded for 20 s after fragmentation.

Analysis of relative total protein expression

BMDCs grown in SILAC media were left untreated or stimulated with LPS for 2 and 6 h. SILAC samples were lysed, digested and desalted as described for the global phosphoproteome analysis. To reduce sample complexity, 100 µg of total peptides were separated using an Agilent 3100 Offgel fractionator (Agilent, G3100A) as described in the manual. For separation into 12 fractions, we used Immobiline DryStrips, 13cm, pH 3-10 (GE Healtcare, 17-6001-14) that were rehydrated in a 1:50 dilution of IPG buffer, pH 3-10 (GE Healthcare, 17-6000-87) containing 5% glycerol. Peptides were reconstituted in IPG buffer (1:50 dilution) containing 5% glycerol and focused for 20kV*h with a maximum current of 50 µA and power of 200 mW. After separation, fractions were acidified by adding 1% formic acid and desalted using StageTips. For global proteome analysis 1 µg of peptide sample was separated on an online nanoflow HPLC system (Agilent 1200) and

analyzed on a LTQ Orbitrap and a LTQ Orbitrap Velos instrument, as described for the global phosphoproteome analysis of LPS timecourse samples.

Affinity purification followed by mass spectrometry (APMS) for V5-tagged MYD88, IRAK2 and AP1AR

Analysis of interaction partners of V5-tagged proteins was performed using a fast and low-stringency single-step purification procedure (to retain weak binders and potentially transient interactions) as previously decribed (Hubner and Mann, 2011), with several modifications to fit our experimental system. 2 x 10^6 bone marrow cells were plated in SILAC complete medium supplemented with 15 ng/mL GM-CSF in 10-cm Petri dishes, and infected two days later with lentiviruses (MOI ~10-20) containing V5-tagged ORFs (Yang et al., 2011) in 10-cm Petri dishes. 2-4 h after infection, cells were fed with GM-CSF-containing complete medium. Two days after infection, GM-CSF-containing complete medium supplemented with blasticidin (10 µg/mL) was added to cells, which were further incubated for 3 days. ORF expression and size was validated using standard Western blotting with anti-V5 antibody (Invitrogen). For immunoprecipitation (IP) of protein complexes, BMDCs expressing a V5-tagged ORF encoding human MYD88 (81.8% amino acid (AA) identity to mouse counterpart), IRAK2 (69.4% AA identity to mouse), or AP1AR (89.8% AA identity to mouse) were stimulated with LPS for 30 min, scraped on ice and washed in ice-cold PBS. Cell pellets were lysed for 30 min on ice in a lysis buffer containing 150 mM NaCl, 50 mM Tris pH 7.5, 5% Glycerol, 1% IGPAL-CA-630 (Sigma, #I8896), and freshly added protease and phosphatase inhibitors (Roche). After centrifugation at 4°C for 10 min at 14,000 g, protein concentration in supernatants was measured by BCA (Pierce), and equal amounts (~2.5-3 mg) of lysates from each SILAC sample were used for subsequent IP. Cell lysates were incubated

for ~16 h at 4°C on a roller with anti-V5 tag antibody covalently bound to magnetic beads (MBL). APMS experiments for MYD88 and IRAK2 were performed as on-bead digests with single-shot mass spectrometry runs, and the AP1AR interaction partners were analyzed by 8 slice in-gel digests with 8 LC-MS/MS runs. For MYD88 and IRAK2, beads were washed after anti-V5 immunoprecipitation twice with wash buffer (150 mM NaCl, 50 mM Tris pH 7.5, 5% Glycerol) containing 1% IGEPAL-CA-630, and twice with wash buffer alone. Beads from each SILAC state were combined after the first wash. Purified protein complexes were then eluted by direct on-bead digestion with trypsin using a buffer containing 2 M urea, 50 mM Tris pH 7.5, 1 mM DTT, and 5 µg/mL Trypsin. After elution, samples were reduced (4 mM DTT) and alkylated (10 mM iodoacetamide) following standard procedures, and further digested with trypsin overnight. Digestion were stopped by adding 1% TFA, and peptides were desalted purified on C18 StageTips before LC-MS/MS analysis. For AP1AR, anti-V5 enriched samples were washed three times with wash buffer and eluted from beads by heating to 100°C in SDS sample buffer (Life Technologies) for 5 min. Samples were separated on a 4-12% gradient gel (NuPAGE; Life Technologies) and cut into 8 slices that were subjected to in-gel trypsin digest and desalting on C18 StageTips as described previously (Lee et al., 2013). Desalted peptide samples for MYD88, IRAK2 and AP1AR APMS experiments were separated on an online nanoflow UHPLC system (Proxeon EASY-nLC 1000) and analyzed on a Q Exactive (Thermo Fisher Scientific) mass spectrometer. We used a 13-cm reversed phase fused-silica capillary column (New Objective, PicoFrit PF360-75-10-N-5 with 10 µm tip opening and 75 µm inner diameter) packed in-house with 3 µm ReproSil-Pur C18-AQ media (Dr. Maisch GmbH) and separated peptides at a flow rate of 200 nL/min in a 82 min linear gradient from 6 to 30% composition of solvent A (3% acetonitrile /0.1% formic acid) and solvent B (90% acetonitrile /0.1% formic acid). The Q Exactive was operated at a spray voltage of 2 kV,

a capillary temperature of 250 C and a S-lens RF level of 50. Data was acquired in positive ion mode, with MS1 scans at a resolution of 70,000 at m/z=200, a mass range of 300-1800, AGC target of 1e6 and 5 ms maximum ion time. Up to 12 of the most intense ions per duty cycle were isolated using an isolation window of 2.5 m/z and fragmented by HCD at a NCE of 25 with an underfill ratio set at 5%. For data-dependent MS2 scans we used a resolution of 17,500, an AGC target of 5e4 and a maximum ion time of 120ms. All ions selected for MS2 scans were dynamically excluded for 20 s after fragmentation.

Massively parallel *in vitro* kinase (IVK) assay

In vitro kinase reactions were performed with recombinant kinases on SILAC-labeled native cell lysates, as follows: 10 million cells were lysed in 1 ml of IVK lysis buffer (0.5% CHAPS, 50 mM Tris pH 7.5, 150 mM NaCl, 5 mM MgCl2, 5 mM MnCl2, 2 µg/ml Aprotinin (Sigma, A6103), 10 µg/ml Leupeptin (Roche, 11017101001) and 1 mM PMSF) for 20 minutes on ice to obtain a ~1 mg/ml protein lysate. Cell debris was removed by centrifugation for 15 minutes at 20,000 g and the protein concentration was measured using a Bradford assay. The buffer and low molecular weight components of the cell lysate were exchanged by size-exclusion chromatography using Zeba Spin Desalting Columns (Thermo Fisher Scientific, 89891, 5 ml column, 7K MWCO) at 4 C. The storage solution of the column was removed by centrifugation at 1,000 g for 2 minutes. The column was washed and equilibrated 4 times with 2.5 ml of ice-cold IVK reaction buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM MgCl2, 5 mM MnCl2, 0.5 mM DTT, 2 µg/ml Aprotinin (Sigma, A6103), 10 µg/ml Leupeptin (Roche, 11017101001) and 1 mM PMSF) by centrifugation at 1,000 g for 2 minutes. The column was then placed in a new collection tube, 1 ml of the IVK cell lysate was applied and the buffer exchanged by centrifugation at 1,000 g for 2 min. The column

flow-through contained cellular components of >7 kDa molecular weight in IVK reaction buffer. Before the in vitro kinase reaction phosphatase inhibitors were added as 1:100 dilutions of PIC2 (Sigma, P5726) and PIC3 (Sigma, P0044). For each IVK reaction 500 µg of SILAC-labeled total cellular proteins in IVK reaction buffer were used. Directly before the reaction, 1 mM of adenosine triphosphate (ATP) and 0.5 µg of recombinant kinase were added and the reaction was incubated for 1 h at 25 C. IVK reactions were performed for IRAK2 (SignalChem, I10-10BG), IRAK4 (EMD Millipore, 14-599), and TBK1 (EMD Millipore, 14-628). IVK reactions were stopped by adding 480 mg of urea per 500 µl of reaction buffer, resulting in a final concentration of 8M urea. Different SILAC samples for kinase and control (no kinase) reactions were combined, reduced with 5 mM DTT for 30 min, alkylated with 10 mM iodoacetamide for 30 min in the dark, diluted 1:4 with 50 mM Tris/HCl pH 7.5 and proteolytically digested with trypsin at a 1:50 enzyme to substrate ratio at 25 C for 16 h. The digests were acidified with 1% formic acid, precipitated urea was removed by centrifugation at 1,000 g for 10 min, and the samples were desalted using SepPak columns (Waters, 100 mg tC18, WAT036820). For single-shot IVK analysis samples were directly enriched for phosphopeptides by IMAC, whereas for deep coverage IVK analysis samples were separated into 6 basic reversed-phase (RP) fractions and then enriched by IMAC (see below).

Global phosphoproteome analysis of IVK samples and knock-out samples

For IVK and KO phosphoproteome analysis desalted peptide samples were separated by basic reversed-phase (RP) prior to IMAC enrichment as described previously (Mertins et al., 2013). Total peptide amounts were 0.5 mg per SILAC state for IVK and 1.5 to 2 mg per SILAC state for wt and KO samples. For basic RP separation, desalted peptides were reconstituted in 900 µL of 20 mM ammonium formate, pH 10. Basic reversed-phase chromatography was performed on 4.6 mm

× 250 mm Zorbax 300 Å Extend-C18 columns (Agilent, 3.5 µm bead size), using an Agilent 1100 Series HPLC instrument. Prior to each separation, columns were monitored for efficient separation with standard mixtures containing 6 peptides. Solvent A (2% acetonitrile, 5 mM ammonium formate, pH 10), and a nonlinear increasing concentration of solvent B (90% acetonitrile, 5 mM ammonium formate, pH 10) were used to separate peptides by their hydrophobicity at a high pH. The flow rate was 1 ml/min and the percentage of solvent B was increased in a nonlinear gradient with 4 different slopes (0% for 9 min; 0% to 6% in 4 min; 6% to 28.5% in 50 min; 28.5% to 34% in 5.5 min; 34% to 60% in 13 min; 60% for 8.5 min). Eluted peptides were collected in 96 × 2 mL deepwell plates (Whatman, #7701-5200) with 1 min (= 1 ml) fractions. Early eluting peptides were collected in fraction "A", which is a combined sample of all fractions collected before any major UV-214 signals were detected. Samples were combined into 6 or 12 subfractions, in a serpentine, concatenated pattern, combining every 12th fraction (1,13,25,37,..; 2,14,26,38,...; ...), or every 6th fraction (1,7,13,19,25,…; 2,8,14,20,26,…; ...). Subfractions were acidified to a final concentration of 1% formic acid and dried in a vacuum concentrator. For IMAC enrichment, iron-chelated IMAC beads were prepared from Ni-NTA superflow agarose beads (Qiagen, #1018611) that were stripped of nickel with 100 mM EDTA and incubated in an aqueous solution of 10 mM FeCl3 (Sigma, 451649). Dried phosphopeptide fractions were reconstituted in 50% acetonitrile/0.1% trifluoroacetic acid and then diluted 1:1 with 100% acetonitrile/0.1% trifluoroacetic acid to obtain a final 80% acetonitrile/0.1% TFA peptide solution at a concentration of 0.5 µg/µl. Peptide mixtures were enriched for phosphorylated peptides with 10 µL IMAC beads for each sample for 30 min. Enriched IMAC beads were loaded on Empore C18 silica-packed Stage tips (3M, 2315). Stage tips were equilibrated with 2 × 100 µL washes of methanol, 2 × 50 µL washes of 50% acetonitrile/0.1% formic acid, and 2 × 100 µL washes of 1% formic acid. Samples were then loaded

onto stage tips and washed twice with 50 µL of 80% acetonitrile/0.1% trifluoroacetic acid and 100 µL of 1% formic acid. Phosphorylated peptides were eluted from IMAC beads with 3 × 70 µL washes of 500 mM dibasic sodium phosphate, pH 7.0, (Sigma, S9763) and washed twice with 100 µL of 1% formic acid before being eluted from stage tips with 60 µL 50% acetonitrile/0.1% formic acid. All washes were performed on a tabletop centrifuge at a maximum speed of 3,500g. Prior to LC-MS/MS analysis IMAC enriched samples were dried in a vacuum concentrator and reconstituted in 9 ul of 3% ACN / 0.1% FA.

IMAC samples were analyzed on an online nanoflow EASY-nLC 1000 UHPLC system (Thermo Fisher Scientific) coupled to a benchtop Orbitrap Q Exactive mass spectrometer (Thermo Fisher Scientific). Fifty percent of each phosphopeptide sample were injected onto a Picofrit column (10 µm tip opening / 75 µm diameter, New Objective, PF360-75-10-N-5) packed in-house with 20 cm C18 silica material (1.9 µm ReproSil-Pur C18-AQ medium, Dr. Maisch GmbH, r119.aq). The UHPLC setup was connected with a custom-fit microadapting tee (360 µm, IDEX Health & Science, UH-753), and capillary columns were heated to 50 °C in column heater sleeves (Phoenix-ST) to reduce backpressure during UHPLC separation. Injected peptides were separated at a flow rate of 200 nL/min with a linear 80 min gradient from 100% solvent A (3% acetonitrile, 0.1% formic acid) to 30% solvent B (90% acetonitrile, 0.1% formic acid), followed by a linear 6 min gradient from 30% solvent B to 90% solvent B. Each sample was run for 150 min, including sample loading and column equilibration times. Data-dependent acquisition was obtained using Xcalibur 2.2 software in positive ion mode at a spray voltage of 2.00 kV. MS1 Spectra were measured with a resolution of 70,000, an AGC target of 3e6 and a mass range from 300 to 1800 m/z. Up to 12 MS2 spectra per duty cycle were triggered at a resolution of 17,500, an AGC target

of 5e4, an isolation window of 2.5 m/z and a normalized collision energy of 25. Peptides that triggered MS2 scans were dynamically excluded from further MS2 scans for 20 s.

Identification and quantification of phosphopeptides and proteins

Mass spectra were processed within the Spectrum Mill (Agilent Technologies) and the MaxQuant (version 1.2.2.5) software packages (Cox and Mann, 2008) using a Uniprot mouse database containing 59,348 entries. The mass tolerance for precursor ions and for fragment ions was set to 20 ppm and 0.7 Da for LTQ-Orbitrap data and 20 ppm and 20 ppm for Q Exactive data, respectively. Cysteine carbamidomethylation was searched as a fixed modification, whereas variable oxidation on methionine and N-acetylation (Protein) was used for all analyzed datasets and phosphorylation on serine, threonine or tyrosine residues were considered as variable modifications for all phosphoproteome analyses. The enzyme specificity was set to trypsin and cleavage N-terminal of proline was allowed. The maximum of missed cleavages was set to 3. For peptide identification the maximum peptide FDR was set to 1%. SILAC ratios for phosphosites were obtained from the proteinPeptideComparisonColumnsExport table in Spectrum Mill and the Phospho (STY)Sites table in MaxQuant. The median ratios of all non-phosphorylated peptides derived from separate unmodified peptide exports in Spectrum Mill and MaxQuant were used to normalize the M/L and H/L ratios of all phosphorylated peptides and corresponding phosphosites. To allow better phosphosite grouping, Spectrum Mill and MaxQuant phosophosite annotations were converted to a unique identifier containing the uniprot accession number, the modified amino acid location and the number of phosphorylated residues on a peptide for each phosphosite quantification event. Median SILAC ratios of phosphopeptides for each experiment were calculated over all versions of the same peptide including different charge states and methionine

oxidation states. The highest scoring versions of each distinct peptide were reported per experiment. Data derived from both software packages was combined and Spectrum Mill data was reported when the same phosphopeptide was identified and quantified by both programs. Site-specific phosphosite localization scores were provided for both Spectrum Mill and MaxQuant. Lastly, only phosphosites that were observed in at least two independent SILAC experiments are reported.

Differential expression (DE) analysis of phosphoproteomic data

To identify differentially regulated phosphosites in the time series, knockout, and IVK data sets, we used sets of 2 replicates of SILAC ratios for each experimental condition and filtered them to retain reproducible data. We deemed two ratios as reproducible if found within the 95% agreement limits of a Bland-Altman plot (Bland and Altman, 1986; Krönke et al., 2015). Then, we assessed statistical significance of differential phosphorylation using a moderated T-test (Smyth, 2004) and by correcting for multiple hypothesis testing using false discovery rate (FDR). When 2 replicates were not available, we proceeded as follows: (1) for the time series data, we paired 15- and 45-min samples with the corresponding, most correlated 30-min replicates (Pearson's correlation of 0.62 and 0.71 respectively); and (2) for the double knockout *Ticam*$^{-/-}$*Myd88*$^{-/-}$ data, we used an absolute log2 fold-change threshold of 0.79 to filter phosphosites that were also asymmetrically differentially phosphorylated at 30 and 45 min in the time series data set. Lastly, for single KO datasets (*i.e.*, *Irak2*$^{-/-}$, *Irak4*$^{-/-}$, *Myd88*$^{-/-}$), we focused our analyses on differentially regulated phosphosites that were also found to be (1) regulated in the time series data, and (2) affected by the double knockout *Ticam*$^{-/-}$*Myd88*$^{-/-}$.

Clustering of phosphoproteomic time series and candidate selection

We performed supervised *k*-means clustering to partition the differentially expressed phosphopeptides from the time series experiment. We focused this analysis on phosphosites that were independently measured in 6 out of 8 time points, and showed differential regulation – based on the statistical criteria defined above – in at least 2 consecutive time points following LPS stimulation. We set the parameter *k* (number of clusters) as the minimal number of clusters that provided a sufficient level of within-cluster similarity. For every cluster, we define the within-cluster similarity as the average r^2 between the members of the cluster and the centroid of the cluster. We used the following cutoffs for: the minimum within-cluster similarity (across all clusters) to be >0.7, and the average (across all clusters) to be >0.75. Using clustering and DE results, we selected 168 candidate phosphoproteins for genetic perturbations (**Figure 3**), including: (i) 121 that were both present in k clusters and differentially regulated in at least 2 consecutive, early time points (15 and 30 or 30 and 45 min); and (ii) 47 candidates manually added that were regulated in the same consecutive time points, but not measured in enough independent time points to pass the filters set prior to k-means clustering. Because several well-known TLR components (*e.g.*, TICAM1) were found to be strongly differentially regulated at early time points but were absent from the data at later time points, we reasoned that adding back several phosphoproteins showing a similar trend would alleviate some of the false negative problem associated with phosphoproteomic measurements.

Pathway enrichment analysis

We measured enriched pathways in our data sets using DAVID (http://david.abcc.ncifcrf.gov), focusing on KEGG pathways and GO terms, and heatmaps were generated with Gene-E

(https://software.broadinstitute.org/GENE-E/index.html). Furthermore, to compile a gene set capturing most of the well described signaling and transcriptional regulators that have been implicated in the TLR pathways, we compiled a list of 141 genes (**Table S2**) by merging information from multiple databases: KEGG, InnateDB, Panther, Reactome, and Uniprot.

Integrative network analysis

Overview of the computational framework to identify signaling-to-transcription relationships. To help uncover signaling-to-transcription relationships in the TLR4 system, we developed a computational framework to integrate biochemical datasets from this study and publicly available databases. First, we assembled a background interaction network combining protein-protein and kinase-substrate interactions (referred to as PPIs and KSIs, respectively) from several public repositories and augmented by adding APMS and IVK data from BMDCs. Second, we assigned weights to the edges (*i.e.*, PPIs, KSIs) and nodes (*i.e.*, signaling or transcriptional regulators) of this network to reflect prior knowledge and the biochemical changes affecting them based on our data sets (*i.e.*, differential phosphorylation upon LPS stimulation, kinase KO and IVK). Third, we searched for potential paths within the network that linked our 29 phosphoproteins ('seed nodes') to selected transcriptional regulators ('target nodes') and determined their significance using network randomization whereby weights were shuffled iteratively to compute statistical significance. We describe each step below.

Step 1: Assembling the background interaction network. We assembled an input set of interactions: 92,610 PPIs and 5,533 KSIs from public databases and AP-MS and IVK data sets from this study (**Table S7**). We collected the union of interactions found in mouse and/or human model systems in the following databases: (i) BioGRID (Stark et al., 2006), by keeping interactions

supported by at least one of the following experiments: Affinity Capture-Luminescence, Affinity Capture-MS, Affinity Capture-Western, Biochemical Activity, Co-crystal Structure, FRET, Far Western, PCA, Proximity Label-MS, Two-hybrid; (ii) PhosphoSitePlus (2014-09-03) (Hornbeck et al., 2015); and (iii) the Human Protein Reference Database (HPRD release 9) (Keshava Prasad et al., 2009), by focusing on interactions identified from two-hybrid experiments. Next, we filtered the resulting network by removing interactions with extracellular proteins (according to Swiss-Prot annotations) (Bairoch et al., 2004) and ubiquitously interacting protein modifiers (*i.e.*, *Ubc*, *Ubd*, *Sumo1*, *Sumo2*, *Sumo3*, *Nedd8*).

Step 2: Defining seed-target relationship score. First, we computed 'relationship scores' for each seed-target pair linked in the interaction network through at least one path (whereby a path represents a set of nodes and edges linking a seed to a target). The interaction network can be represented as a graph $G=(V,E,w_n,w_e)$, where V is a set of nodes (proteins), E is a set of edges (interactions), and w_n and w_e represent corresponding sets of weights assigned to nodes and edges respectively. The weight values were designed to be bigger than or equal to 0 but smaller than 1. Given a set of targets $T \subset V$ and a set of candidate seeds A, we compute a 'relationship score' $S(A,T)=R$ (where R is a set of non-negative, real numbers). To do so, we take each simple path k from $a \subset A$ to $t \subset T$, and assign to it a path score $S_k = \prod_i w_{ni} * \prod_j w_{ej}$ where i and j go over all nodes and edges in the path. This score reflects the evidence that a interacts with t through the path k. Given that every path between a and t is increasing the total evidence that a influences t we compute the 'relationship score' as $S(a,t) = \sum_k S_k$ where k enumerates all simple paths from a to t having at most 2 intermediate nodes. We note that in general, the longer the path, the smaller its score S_k.

Step 3: Assigning weights to edges and nodes from the background interaction network. First, to assign edge weights, we used two components: one to reflect the experimental evidence supporting a given PPI or KSI, and another to capture the local network topology as measured by the number of edges arriving to and leaving from a given node. The first component w_{e1} depended on the number n_{exper} of different experiment types supporting the interaction underlying a given edge (*e.g.*, AP-MS, Y2H, etc.):

$$w_{e1} = {}^{1}/_{(1 + e^{-2n_{exper}})}$$

Therefore, the larger the number of experiment types n_{exper}, the higher the interaction weight. The second component w_{e2} depended on the local network topology: for an edge originating at node *1* and ending at node *2*, nd_{out}, nu_{out} are the numbers of all directed and undirected edges, respectively, originating at node 1, and nd_{in}, nu_{in} the numbers of all edges ending at node *2*. Then:

$$w_{e2} = \left[\frac{1}{(0.5 * nd_{out} + nu_{out} + 1) * (0.5 * nd_{in} + nu_{in} + 1)}\right]^{c}$$

where the value of $c = 0.01$, was derived in the optimization step described below. The composite edge weight value w_e was set to $0.9 * \sqrt[2]{w_{e1} * w_{e2}}$ for directed edges and $0.45 * \sqrt[2]{w_{e1} * w_{e2}}$ for undirected ones. In addition, when an interaction was supported by AP-MS and/or IVK data from this study, we set the corresponding values of w_e to 0.999 and 0.99 for directed and undirected edges, respectively.

Second, to assign node weights, we first considered the background case, where differential phosphorylation information was not used. In this case, all network nodes were given

the same weight of 0.1, which maximized the performance on the following task: take KEGG pathway database (Kanehisa et al., 2016) (without 'Toll-like receptor signaling pathway'), and use the transcription regulators as targets. Next, compute the seed-target relationship score for each possible seed-target pair. The pairs found to belong to the same KEGG pathway were then considered as true, and the pairs crossing the pathways as false (the intermediate nodes in network paths joining seed-target pairs were not restricted to KEGG proteins). The 'relationship score' depended on both the uniform node weight w_{n0} and the edge weights. Therefore, we concurrently searched for the optimal values of w_{n0} and parameter c (from the edge weight equation above) that maximized the area under the ROC curve. The resultant values for w_{n0} and c were 0.10 and 0.01, and they were used to define the 'background network' that is the input set of interactions with w_e and w_{n0}. Lastly, we reasoned that network nodes corresponding to differentially phosphorylated proteins in our experiments were more likely to transduce the signal from TLR4 to downstream gene regulation events. Therefore, for a given node i its weight w_{ni} increased with the number of experiments that indicated i as differentially phosphorylated:

Data set	$\mathbf{w_{ni}}$	**Number of nodes**
Non-phosphorylated	0.10	8710
Time series (TS; 15, 30, 45 min)	0.20	1784
K-means	0.30	223
TS and double KO (dKO)	0.60	335
TS, dKO and 1 KO	0.63	302
TS, dKO and 2 Kos	0.67	217
TS, dKO and 3 Kos	0.70	31

(where TS stands for time series, KO stands for Irak2$^{-/-}$, Irak4$^{-/-}$, or *Myd88*$^{-/-}$, and dKO for *Myd88*$^{-/-}$/*Ticam1*$^{-/-}$ experiments). Resulting values w_{ni} were normalized such that their average was equal to 0.1 (*i.e.*, the same value as the uniform node weights w_{n0}).

Resulting weights on network nodes defined the 'weighted network' that was used for ROC analysis (**Figure S7B**), as follows: (1) we used the set of known TLR components defined above (**Table S2**) and split it into targets (*i.e.*, transcriptional regulators) and seeds (*i.e.*, all other nodes). Furthermore, we focused on the 11 out of 14 TLR targets that were found as differentially phosphorylated in our time series phosphoproteomic data. (2) Seed-target pairs among the known TLR components were set as 'true', whereas pairs between TLR 'targets' and non-TLR 'seeds' present in the KEGG database set as 'false'. (3) Lastly, we ordered seed-target pairs using only the sum of seed and target node weights as their relationship scores (*i.e.*, without using network paths scores). The corresponding ROC curve ('phosphorylation only') served as a reference benchmark for methods that used network information (**Figure S7B**).

Step 4: Identify significant seed-target relationships. We used bootstrapping to assign statistical significance of the 'relationship scores' computed above. To do so, we created 1000 randomized networks by swapping edges while keeping constant the node degrees (*i.e.*, the number of edges linking a given node) to maintain the network topology (Maslov and Sneppen, 2002). We randomized directed and undirected edges separately. The phosphorylation dependent node weights were randomly permuted among all nodes in the randomized networks. Next, 'relationship scores' between seeds and targets were computed in each random network. The resulting scores were used as a background distribution for empirical *p*-value computations. Specifically, we used marginal distributions, parametrized by the minimal path length. For example, if in the original non-randomized interaction network, a given seed was connected to its closest target by a path of

length *l,* then we used a distribution of relationship scores computed with paths of lengths bigger or equal *l*.

Centrality score analysis. After identifying significant seed-target pairs, we aimed to assign and compare the signaling centralities of the intermediate nodes (*i.e.*, excluding seed and target nodes) in each module of seeds (**Figures 3B and 5A**). For each module and each intermediate node, we summed scores of paths linking it with the corresponding set of significant seed-target pairs. This sum was then divided by the number of seeds in the module to arrive at the centrality score of an intermediate node in a module.

Visualization of knockout effects seed-target relationships. We quantified the effects of KO (*i.e.* $Myd88^{-/-}$, $Irak2^{-/-}$, $Irak4^{-/-}$) on the 'relationship scores' of the significant (bootstrap $p < 5 \times 10^{-4}$, Benjamini-Hochberg FDR < 0.05) seed-target pairs that were also affected by $Myd88^{-/-}/Ticam1^{-/-}$. For each simple seed-target path, we computed the percentage of the nodes affected by each KO. The total KO effect on a given seed-target pair was computed as the weighted average percentage from all paths, where weights were proportional to path scores S_k. The resulting KO effects were visualized using t-Distributed Stochastic Neighbor Embedding (t-SNE) (**Figure 7G**) and KOs with the weighted percentage of affected nodes bigger or equal to 50 were colored.

SUPPLEMENTAL REFERENCES

Bairoch, A., Boeckmann, B., Ferro, S., and Gasteiger, E. (2004). Swiss-Prot: juggling between evolution and stability. Brief. Bioinformatics *5*, 39–55.

Bland, J.M., and Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. Lancet *1*, 307–310.

Chevrier, N., Mertins, P., Artyomov, M.N., Shalek, A.K., Iannacone, M., Ciaccio, M.F., Gat-Viks, I., Tonti, E., DeGrace, M.M., Clauser, K.R., et al. (2011). Systematic discovery of TLR signaling components delineates viral-sensing circuits. Cell *147*, 853–867.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology *26*, 1367–1372.

Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Research *43*, D512–D520.

Hubner, N.C., and Mann, M. (2011). Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). Methods *53*, 453–459.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Research *44*, D457–D462.

Keshava Prasad, T.S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A., et al. (2009). Human Protein Reference Database--2009 update. Nucleic Acids Research *37*, D767–D772.

Krönke, J., Fink, E.C., Hollenbach, P.W., MacBeth, K.J., Hurst, S.N., Udeshi, N.D., Chamberlain, P.P., Mani, D.R., Man, H.W., Gandhi, A.K., et al. (2015). Lenalidomide induces ubiquitination and degradation of CK1α in del(5q) MDS. Nature *523*, 183–188.

Lee, M.N., Roy, M., Ong, S.-E., Mertins, P., Villani, A.-C., Li, W., Dotiwala, F., Sen, J., Doench, J.G., Orzalli, M.H., et al. (2013). Identification of regulators of the innate immune response to cytosolic DNA and retroviral infection by an integrative approach. Nat Immunol *14*, 179–185.

Maritzen, T., Zech, T., Schmidt, M.R., Krause, E., Machesky, L.M., and Haucke, V. (2012). Gadkin negatively regulates cell spreading and motility via sequestration of the actin-nucleating ARP2/3 complex. Proc. Natl. Acad. Sci. U.S.A. *109*, 10382–10387.

Maslov, S., and Sneppen, K. (2002). Specificity and stability in topology of protein networks. Science *296*, 910–913.

Mertins, P., Qiao, J.W., Patel, J., Udeshi, N.D., Clauser, K.R., Mani, D.R., Burgess, M.W., Gillette, M.A., Jaffe, J.D., and Carr, S.A. (2013). Integrated proteomic analysis of post-translational modifications by serial enrichment. Nat Meth *10*, 634–637.

Rehwinkel, J., Maelfait, J., Bridgeman, A., Rigby, R., Hayward, B., Liberatore, R.A., Bieniasz, P.D., Towers, G.J., Moita, L.F., Crow, Y.J., et al. (2013). SAMHD1-dependent retroviral control and escape in mice. The EMBO Journal *32*, 2454–2462.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol *3*, Article3–Article25.

Stark, C., Breitkreutz, B.-J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. Nucleic Acids Research *34*, D535–D539.

Yang, X., Boehm, J.S., Yang, X., Salehi-Ashtiani, K., Hao, T., Shen, Y., Lubonja, R., Thomas, S.R., Alkan, O., Bhimdi, T., et al. (2011). A public genome-scale lentiviral expression library of human ORFs. Nat Meth *8*, 659–661.