Unbiased Reconstruction of a Mammalian Transcriptional Network Mediating Pathogen Responses

Ido Amit,[1,2,3,4] Manuel Garber,[1]* Nicolas Chevrier,[2,3]* Ana Paula Leite,[1,6]* Yoni Donner,[1]* Thomas Eisenhaure,[2,3] Mitchell Guttman,[1,4] Jennifer K. Grenier,[1] Weibo Li,[2,3] Or Zuk,[1] Lisa A. Schubert,[5] Brian Birditt,[5] Tal Shay,[1] Alon Goren,[1] Xiaolan Zhang,[1] Zachary Smith,[1] Raquel Deering,[2,3] Rebecca C. McDonald,[2,3] Moran Cabili,[1] Bradley E. Bernstein,[1] John L. Rinn,[1] Alex Meissner,[1] David E. Root,[1] Nir Hacohen,[1,2,3]†‡ Aviv Regev[1,4]‡

[1]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA. [2]Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, 149 13th St., Charlestown, MA 02129 USA. [3]Harvard Medical School, Boston, MA 02115, USA. [4]MIT Department of Biology, Cambridge, MA 02142, USA. [5]NanoString Technologies, 530 Fairview Ave. N, Suite 2000, Seattle. WA 98109, USA. [6]Computational and Systems Biology, MIT Cambridge, MA 02139, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: nhacohen@partners.org
‡These authors contributed equally to this work.

Models of mammalian regulatory networks controlling gene expression have been inferred from genomic data, yet have largely not been validated. We present an unbiased strategy to systematically perturb candidate regulators and monitor cellular transcriptional responses. We apply this approach to derive regulatory networks that control the transcriptional response of mouse primary dendritic cells (DCs) to pathogens. Our approach revealed the regulatory functions of 125 transcription factors, chromatin modifiers, and RNA binding proteins and constructed a network model consisting of two dozen core regulators and 76 fine-tuners that help explain how pathogen-sensing pathways achieve specificity. This study establishes a broadly-applicable, comprehensive and unbiased approach to reveal the wiring and functions of a regulatory network controlling a major transcriptional response in primary mammalian cells.

Regulatory networks controlling gene expression serve as decision-making circuits within cells. For example, when immune dendritic cells are exposed to viruses, bacteria or fungi, they respond with transcriptional programs that are specific to each pathogen (*1*)and are essential for establishing appropriate immunological outcomes (*2*). These responses are initiated through specific receptors, such as Toll-like receptors (TLRs), that distinguish broad pathogen classes, and are propagated through well-characterized signaling cascades (*2*). However, how the transcriptional network is wired to produce specific outputs remains largely unknown.

Two major observational strategies have associated regulators with their putative targets on a genome scale (*3*): *cis*-regulatory models rely on the presence of predicted transcription factor binding sites in the promoters of target genes (*3–5*), whereas *trans*-regulatory models are based on correlations between regulator and target expression (*3–6*). Since promoter binding sites and correlated expression are weak predictors of functional regulator-target linkages, such approaches are limited in their ability to produce reliable models of transcriptional networks (*3*). A complementary strategy is to systematically perturb every regulatory input and measure its effect on expression of gene targets. This strategy has been successfully employed in yeast (*7-9*) and sea urchin (*10*), but not in mammals.

A perturbation-based strategy for network reconstruction. We developed a perturbation strategy for reconstructing transcriptional networks in mammalian cells, and applied it to determine a network controlling the responses of DCs to pathogens (Fig. 1). First, we profiled gene expression at nine time points following stimulation with five pathogen-derived components and identified specific and shared genes that respond to each stimulus (fig. S1A). We used these profiles to identify 144 candidate regulators whose expression changed in response to at least one stimulus (SOM) (fig. S1B, top). We also identified a signature of 118 marker genes (fig. S1B, bottom) that captures the complexity of the response. We generated a validated lentiviral shRNA library for 125 (of the 144) candidate regulators (fig. S1C, top), used it to systematically perturb each of the regulators in DCs, stimulated the cells with a pathogen component, and profiled the expression of the 118 gene signature (*11*) (fig. S1C, bottom). Finally, we used the measurements from the perturbed cells to derive a validated model of the regulatory network (fig. S1D).

Gene expression programs in response to TLR agonists. We measured genome-wide expression profiles in

DCs exposed to PAM3CSK4 ('PAM'), a synthetic mimic of bacterial lipopeptides, polyI:C, a viral-like dsRNA, LPS, a purified component from Gram negative *E. coli,* gardiquimod, a small molecule agonist, and CpG, a synthetic ssDNA. These compounds are known agonists of TLR2, TLR3, TLR4, TLR7 and TLR9, respectively. PolyI:C also activates MDA-5, and LPS can also act through co-receptors such as CD14. We therefore refer to the ligands rather than their receptors for clarity. Based on pilot experiments (fig S2, SOM), we measured mRNA expression at 0.5, 1, 2, 4, 6, 8, 12, 16, and 24 hours following stimulation with these pathogen components.

The observed transcriptional responses were classified into a 'PAM-like' program and a 'polyI:C-like' program, as well as a shared response (24.5% shared by PAM/polyI:C/LPS). The LPS response (Fig. 2, A and B, and fig. S3) was largely the union of the 'PAM-like' and 'polyI:C-like' programs. This is partly explained by the known signaling pathways activated by these agonists. PAM binds TLR2 and signals through the MYD88 pathway; polyI:C binds TLR3 and MDA-5 and signals mostly through the TRIF and IPS-1 pathways, respectively; and LPS binds TLR4 and co-receptors and uses both pathways (*12*). It is also consistent with the known induction of an anti-viral response by polyI:C and LPS (*13*). The 'PAM-like' program is enriched for NFκB and inflammatory responsive genes ($P<6.1*10^{-8}$), whereas the 'polyI:C-like' program is enriched for IRFs, viral- and interferon-responsive genes (ISGs, $P<8.3*10^{-24}$). We thus term them the 'inflammatory-like' and 'anti-viral-like' programs. A small number of genes are specific to a single stimulus. For example, ~250 genes are polyI:C-specific (1250 are shared with LPS), including several Type I IFNs (*e.g.* IFNA2, IFNA4, Fig. 2A). Surprisingly, 82% of the gardiquimod (TLR7) and CpG (TLR9) response was shared with the LPS response, but with a weaker anti-viral component (fig. S4). This observation is unexpected given their different signaling mechanisms (*14*), but is highly reproducible and robust (fig. S4, SOM).

Identification of candidate regulators and a response signature. To select potential regulators that mediate the observed transcriptional response, we focused on regulator genes whose expression changes during pathogen sensing (a reasonable assumption for many mammalian responses (*15, 16*), including pathogen-sensing (*1, 4*). First, we reconstructed an observational *trans*-model of gene regulation (figs. S1B, top**,** S5A, SOM), that associated 80 modules of co-regulated genes with 608 predictive regulators (*4, 17, 18*) (SOM, fig. S5B), automatically chosen out of a curated list of 3287 candidate regulators (SOM). Filtering identified 117 regulators above a minimal expression signal in at least one experiment (fig. S5B)**.** These included known regulators from the NFKB, STAT and IRF families as well as unexpected

candidates such as the circadian regulator Timeless and the DNA methyltransferase Dnmt3a. Second, we added 5 constitutively expressed regulators whose *cis*-regulatory elements are enriched in the responsive genes (SOM). Third, to capture delayed responses or nonlinear relations, we incorporated 22 regulators with at least a 2-fold change in expression. This resulted in 144 candidate regulators, with a distribution of expression patterns similar to the general response (figs. S6–S8 and table S1), The regulators' expression under LPS was conserved between DCs and functionally similar macrophages, (Pearson correlation r~0.9 at 1h, fig. S9A) as well as between human macrophages and mouse DCs (r~0.6 at 2h, fig. S9B) supporting the functional relevance of the regulators' transcription.

To identify highly informative reporter genes for monitoring the effects of perturbing regulators, we devised *GeneSelector* (fig. S10A, table S2, SOM). GeneSelector incrementally chooses genes (from our full expression dataset) whose expression profile improves our discrimination of stimuli given the previously chosen genes. Using this approach, we identified the optimal time point (six hours post activation, fig. S10B) and a set of 81 genes that distinguishes the stimuli (SOM). We added 37 candidate regulators with detectable expression at the 6h time point, creating a signature of 118 genes. Finally we added 10 control genes whose expression levels were unchanged under all stimuli, but whose (constant) basal levels varied from very low to high.

Perturbations, profiling and modeling. We generated validated lentiviral shRNAs that knocked down expression of 125 of our 144 candidate regulators by at least 75% (fig. S11, table S3, SOM) and 32 shRNAs with no known gene targets as controls in bone marrow DCs (fig S12, fig S13, table S4, SOM). To carry out our perturbational study, we selected a single treatment, LPS, that activates the majority of both the 'inflammatory-like' and 'anti-viral-like' programs. Following stimulation of shRNA-perturbed DCs with LPS for six hours, we used nCounter (*11*) to count transcripts of the 118 reporter and 10 control genes.

The changes in signature gene expression resulting from infection with each shRNA were used to construct a model that associated regulators to their targets. We expect increases in the transcript levels of reporter genes whose repressors are targeted by knockdown, and decreases in reporters whose activators are targeted. Our False Discovery Based (FDR) model estimates the significance of a change in transcripts in DCs infected with a given shRNA (SOM). We control for gene-specific noise by comparing to changes in the expression of each gene following perturbation with the control shRNAs (Fig. 3A), and for shRNA-specific noise by comparing to changes in the expression of the control genes following a given shRNA perturbation (Fig. 3B). We

estimated the sensitivity of our calls from the 37 regulators, which are also included as target reporters (fig. S14, SOM).

On the basis of these results we identified a densely overlapping network with 2322 significant regulatory connections, including 1728 activations and 594 repressions (Fig. 3B, red and blue, respectively, at 95% confidence tables S5 to S7). Of the 125 tested regulators, we confidently identified 100 with at least four targets. Among those were 24 hub regulators that were predicted to regulate over 25% of the 118 genes measured, and 76 specific regulators each affecting the expression of 4 to 25 genes. On average ~14 (±8 stdv) regulators activate a target gene, and 5 (±5.8) regulators repress it. Indirect effects may account for the large number of regulators we observe for each target.

Our perturbational model captured known regulatory features of the response, but also identified novel regulators. The reporter genes partition into two main clusters based on their response to perturbations (Fig. 3B and fig. S15A): the 'anti-viral (polyI:C) like' program reporters (*e.g.*CXCL10, ISG15, IFIT1), and the 'inflammatory (PAM) like' program reporters (*e.g.* IL1b, CXCL2, IL6, IL12b), consistent with the expression data. We also found many known regulatory relations, for example the NFκB family of transcription factors (Rel, Rela, Relb, Nfkb1, Nfkb2 and Nfkbiz) regulating their known inflammatory gene targets. Our network provided evidence for the involvement of at least 68 additional regulators in the response to pathogens, of which 11 were hubs not previously associated with this system. Interestingly, 12 regulators identified (*e.g.* Hhex, Fus, Bat5, Pa2g4) are in linkage disequilibrium with SNPs associated with autoimmune and related diseases in genome-wide association studies (table S8).

The core inflammatory and anti-viral programs. We next addressed how each regulator contributes to the generation of specific cell states. We first automatically defined the two major states induced by the five pathogen components using non-negative matrix factorization (NMF) (*19*) and the original array data (SOM). This procedure identified two major expression components (termed 'metagenes'): one predominantly determined by genes from the 'inflammatory-like' program and the other by genes from the 'anti-viral like' programs (Fig 2A). Next, we quantified the effects of each regulator's knockdown on these two states (Fig. 3B, fig. S15A, table S9), by classifying the nCounter expression measurements following a regulator's perturbation (*19, 20*).

Finally, we used a regulator ranking score (SOM) to assign 33 (8 known) genes as regulators of the inflammatory state and 33 (15 known) genes as regulators of the anti-viral state. This accurately classified the known activators of the inflammatory response (*e.g.* the NFκB factors Rela, Nfkbiz, Nfkb1, Fig. 3B, yellow in the inflammatory metagene) and of

the anti-viral response (*e.g.* Stat1, Stat2, Stat4, Irf8, Irf9 Fig. 3B, yellow in the viral metagene). Although all perturbation experiments were conducted only under LPS stimulation (a bacterial component), we correctly classified factors known to mediate the response to other stimuli. 34 additional regulators were associated with both responses, suggesting that a single regulator can control genes in either state depending on the differential timing of regulator activation, its level, or combinatorial regulation. Notably, for 12 of the transcription factors examined, we found an enriched *cis*-regulatory element in the appropriate metagene (SOM).

On the basis of the NMF scores (table S9), we identified an inflammatory subnetwork (fig. S15B), an anti-viral subnetwork (Fig. 4A and fig. S15C), and several fine-tuning subnetworks that affect smaller numbers of genes from both responses (Figs. S15D and S16, SOM). The inflammatory subnetwork (fig. S15B) consisted of three regulatory modes: dominant activators (Cebpb, Bcl3, Cited2) which induce more inflammatory targets than anti-viral ones; cross-inhibitors (Nfkbiz, Nfkb1, Atf4, Pnrc2) which induce inflammatory genes while repressing anti-viral ones, and specific activators (Runx1, Plagl2), that only target inflammatory genes. We observed that dominant activators mostly regulate effectors, whereas regulators are primarily controlled by cross-inhibitors.

Focusing on the network architecture, we found multiple feed-forward circuits in this response, where an upstream regulator controls a target gene both directly and indirectly through a secondary regulator (*21*) (e.g. Fig. 4B, and tables S10 and S11). The majority (76%, 4892 of 6444) of these feed-forward circuits were found to be coherent (*21*); having the same direct and indirect effect on the regulated gene. The vast majority (80%) are type I loops (*22*) with all-positive regulation (*e.g.* NFKBIZ activates E2F5 and both activate IL6). Such feed-forward circuits respond to persistent rather than transient stimulation, protecting the system from responding to spurious signals, as was shown for one circuit in LPS-stimulated macrophages (*23*). Our finding suggests that coherent feed-forward loops, especially class I (*21*), are a general design principle in this system and may physiologically impact this response.

In the anti-viral sub-network, we identified a two-tiered regulatory circuit combining feed-forward and feed-back loops (Fig. 4A and table S11). This circuit has at the top the anti-viral regulators Stat1 and Stat2, which regulate a full complement of anti-viral reporters. The second-tier regulators Timeless, Rbl1 and Hhex are controlled by Stat1 and 2 and most likely form coherent feed-forward loops that target specific sub-sets of genes. Timeless, Rbl1 and Hhex also feed-back and promote the expression of the Stat regulators. This circuit is repressed through the cell cycle regulator and

RNA binding protein Fus (*24*), acting as a single dominant inhibitor of 43 viral genes.

Finally, we derived a core network incorporating the regulators with the most substantial impact on each response, on the basis of the number, magnitude, and logic of targets that each regulator affects (SOM). The core network (Fig. 4C) has 24 regulators, 13 of which have previously been identified as key factors regulating the inflammatory or anti-viral responses, while 11 have not been previously implicated in either response. Of these, 19 are transcription factors, three are chromatin modifiers, and two are RNA binding proteins. The regulators apparently distinguish the two programs through cross-inhibition (Fig. 4C, gray lines) or dominant activation (Fig. 4C). The core network also explains how differential expression of secreted factors is specified, leading to the activation and migration of appropriate cell types for different pathogens (*25*) (fig. S17, SOM).

Embedded within the many known regulators of the anti-viral response (Fig. 4C and fig. S15C), we found a large set of regulators not previously associated with this response. These included several known regulators of the cell cycle and the circadian rhythm, including Rbl1, Jun, RB, E2F5, E2F8, Nmi, Fus, and Timeless, several of which were placed in our core network. This suggests that a cell cycle regulatory circuit was co-opted to function in the anti-viral response in DCs (with no observable effect on cell cycle progression, fig S18). Since we identified these anti-viral regulatory relations in perturbation experiments using DCs stimulated with the bacterial component LPS, we silenced four regulators (TIMELESS, RBL1, JUN and NMI) following exposure to the viral component polyI:C. Each of the four regulators strongly impacted the anti-viral program, more than was observed under LPS stimulation (Fig. 4D), and affected genes (*e.g.* Type I IFNs) whose expression is polyI:C-specific. Nmi affected a smaller set of genes, consistent with the model's prediction. These results demonstrate our ability to correctly predict function in unobserved conditions.

Although most anti-viral genes are induced following stimulation with the bacterial component LPS, a few critical ones are expressed specifically in polyI:C stimulation, or follow distinct patterns in each stimulus. In response to viral infection cells induce the production of interferon beta1 (IFNB1), a crucial mediator of the antiviral response. Because high levels of IFNB1 may be deleterious to the host if infected by specific bacteria (*26*), we predicted that specific mechanisms insulate IFNB1's regulation from the response to LPS. Indeed, although IFNB1 expression was induced in the first two hours of stimulation with LPS, this expression declined at subsequent time points, in contrast to its sustained induction following polyI:C treatment (Fig. 5A). Our model suggested that three regulators known to affect chromatin remodeling (*24, 27, 28*) are IFNB1 repressors in LPS (Fig.

5B): the Polycomb complex subunit Cbx4 (*27*), Fus (*24*), and the DNA methyltransferase Dnmt3a (*28*). Cbx4 appeared to confer anti-viral specificity to IFNB1 induction as it is induced within the first two hours of PAM and LPS treatment but not by polyI:C (Fig. 5C), and Cbx4 knockdown caused induction of IFNB1 mRNA and protein during LPS treatment (Fig. 5D and fig. S19A), but had no effect on the induction of the chemokine Cxcl10, a polyI:C and LPS-induced gene (fig. S19B). Cbx4 knockdown did not affect IFNB1 during PAM activation (Fig. 5E), when the anti-viral response is not induced. Combined with evidence for chromatin changes around the Ifnb1 locus and its closest neighbor gene, Ptplad2 (fig. S20A), which has a similar dependence on Cbx4, these data are consistent with an effect by Cbx4 on local chromatin organization (fig. S20, B and C). Cbx4 knockdown affected few genes (~120 up-regulated and ~120 down-regulated genome-wide, table S12). Because most up-regulated genes show a precise temporal pattern in unperturbed cells akin to that of Cbx4– they are induced quickly and return to basal level by 2-4 hours (fig. S21 A–F), we conclude that a chromatin modifier can act like a transcription factor controlling the precise expression of specific genes in the regulatory program.

Taken together, our results suggest a model of a transcriptional negative feedback loop, controlling IFNB1 expression in LPS stimulation, wherein the induced pro-inflammatory regulator and chromatin modifier Cbx4 represses transcription by modifying the chromatin in the Ifnb1 locus, generating the specificity needed to drive inflammatory versus the anti-viral response (Fig. 5F). The Type I coherent feedforward loop formed by Cbx4 and Dnmt3a (Fig. 4B) is consistent with a delayed repression of IFNB1.Since neither regulator carries a sequence-specific DNA binding domain, the factors responsible for their guidance to the Ifnb1 locus remain unknown.

Discussion. A central goal of our study was to address the mechanistic basis for pathogen-specific responses. Consistent with previous studies (*13*), we distinguished two key programs, a PAM (TLR2)-like inflammatory response and a polyI:C (TLR3/MDA-5)-like anti-viral response, which are together induced by LPS, a gram-negative bacterial component and a TLR4-ligand. These programs reflect both qualitative and quantitative differences between the required functional responses, and are consistent with the cross-protection between certain bacteria and virus infections (*13*). The broad effect of LPS allowed us to focus on a single stimulus and timepoint, but screens with other stimuli may identify additional unique regulators.

We found that these two responses are controlled by two corresponding regulatory arms, uncovering a mechanistic basis for the observed transcriptional responses. These two arms are integrated into a core network of two dozen

regulators which balances specific and shared responses through dominant activation and cross-inhibition. In the inflammatory response, we found several feed-forward loops, which may ensure response to only persistent, and not sporadic, signals. In the anti-viral response, we discovered a two-tiered circuit involving feedback and feed-forward loops, implicating a module of cell cycle regulators (Jun, Rbl1, Timeless and Nmi), which we directly validated. Over 75 additional genes work to further fine-tune the regulation of gene targets. This perturbational model identifies many regulatory relations that would have been missed by non-systematic approaches.

While we have benefited from the specific features of the DC system, our work establishes an unbiased, straightforward and general framework for network reconstruction in mammalian cells (SOM). In particular, we develop several strategies to leverage shRNA for the study of gene regulation. This approach can be executed at substantial scale and reasonable cost, and is compatible with the challenge of deciphering the multiple regulatory systems that operate in mammals. It can be expanded to derive increasingly detailed models, and distinguish direct from indirect targets.

Our study will facilitate the development of new computational approaches to infer regulatory models. While many computational approaches have attempted to derive observational models, their quality has been difficult to evaluate ([3]). The data generated here includes both expression profiles for training a model, and a perturbational unbiased screen for testing its quality (Web portal; ftp://ftp.broadinstitute.org/pub/papers/dc_network/). When we compared the perturbational model to our observational model, we found that many candidate regulators were correctly identified in both fig. S5, S22). However, there were also numerous false positive relations in the observational model, attributable to the fact that both the correct regulator and many others have indistinguishable expression (figs. S22, S23).

The high-resolution map we constructed has important biomedical implications. By identifying regulators that mediate the differential control of specific gene pairs (e.g. IL-23 vs. IL-12, fig. S17) and entire regulatory arms (*e.g.* viral vs. inflammatory), it opens the way for therapeutic targeting of specific pathways to control disease or enhance vaccine efficacy. Furthermore, 12 of our regulators reside in genetic loci that were in linkage disequilibrium with SNPs associated with autoimmune and related diseases. The identified genes and their impact on DCs provide hypotheses to help explain how alleles of genes in a cascade may alter susceptibility to specific infections or immune disorders in humans.

References and Notes

1. Q. Huang *et al.*, *Science* **294**, 870 (2001).
2. T. Kawai, S. Akira, *Int. Immunol.* **21**, 317 (2009).
3. H. D. Kim, T. Shay, E. K. O'Shea, A. Regev, accepted to *Science* (2009).
4. S. A. Ramsey *et al.*, *PLoS Comput. Biol.* **4**, e1000021 (2008).
5. H. Suzuki *et al.*, *Nat. Genet.* **41**, 553 (2009).
6. E. Segal *et al.*, *Nat. Genet.* **34**, 166 (2003).
7. A. P. Capaldi *et al.*, *Nat. Genet.* **40**, 1300 (2008).
8. Z. Hu, P. J. Killion, V. R. Iyer, *Nat. Genet.* **39**, 683 (2007).
9. C. T. Workman *et al.*, *Science* **312**, 1054 (2006).
10. D. H. Erwin, E. H. Davidson, *Nat. Rev. Genet.* **10**, 141 (2009).
11. G. K. Geiss *et al.*, *Nat. Biotechnol*. **26**, 317 (2008).
12. J. C. Kagan *et al.*, *Nat. Immunol.* **9**, 361 (2008).
13. S. Doyle *et al.*, *Immunity* 17, **251** (2002).
14. T. Kawai *et al.*, *Nat. Immunol.* **5**, 1061 (2004).
15. I. Amit *et al.*, *Nat Genet* **39**, 503 (2007).
16. D. Pe'er, A. Regev, A. Tanay, *Bioinformatics* **18** Suppl. 1, S258 (2002).
17. S. I. Lee *et al.*, *PLoS Genet.* **5**, e1000358 (2009).
18. H. Zou, T. Hastie, *Journal of the Royal Statistical Society B* **67**, 301 (2005).
19. J. P. Brunet, P. Tamayo, T. R. Golub, J. P. Mesirov, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4164 (2004).
20. J. P. Daily *et al.*, *Nature* **450**, 1091 (2007).
21. S. Mangan, U. Alon, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11980 (2003).
22. U. Alon, *Nat. Rev. Genet.* **8**, 450 (2007).
23. V. Litvak *et al.*, *Nat. Immunol.* **10**, 437 (2009).
24. X. Wang *et al.*, *Nature* **454**, 126 (2008).
25. A. D. Luster, *Curr. Opin. Immunol.* **14**, 129 (2002).
26. T. Decker, M. Muller, S. Stockinger, *Nat. Rev. Immunol.* **5**, 675 (2005).
27. E. Bernstein *et al.*, *Mol. Cell Biol.* **26**, 2560 (2006).
28. B. Li *et al.*, *Biochem. J.* **405**, 369 (2007).
29. We thank E. Lander, I. Wapinski, D. Pe'er, N. Friedman, J. Kagan, A. Luster, and V. Kuchroo for discussions and comments, L. Gaffney for assistance with artwork**,** S. Gupta and the Broad Genetic Analysis Platform for microarray processing, and T. Mikkelsen and the Broad Sequencing Platform for help with the ChIP-seq experiments. Supported by the Human Frontier Science Program Organization and he APF Fellowship's Claire & Emanuel G. Rosenblatt Award (IA); NIH R21 AI71060 and the NIH New Innovator Award (NH); a Career Award at the Scientific Interface from the Burroughs Wellcome Fund, the NIH Pioneer Award, and the Sloan Foundation (AR). Complete microarray data sets available at the Gene Expression Omnibus (GSEXXX).

Supporting Online Material
www.sciencemag.org/cgi/content/full/1179050/DC1
Materials and Methods

Fig. 1. A systematic strategy for network reconstruction. The strategy consists of four steps, from left, (1) state measurement using arrays, (2) selection of regulators and response signatures, (3) network perturbation with shRNAs against each regulator followed by measurement of signature genes, and (4) network reconstruction from the perturbational data.

Fig. 2. Gene expression response to pathogen components. (**A**) mRNA profiles of the 1800 genes whose expression was induced to at least 1.7-fold from baseline level in both duplicates of at least one time point in CD11c+ dendritic cells stimulated with the indicated pathogen component across a time course of 0, 0.5, 1, 2, 4, 6, 8, 12, 16, 24 hours (tick marks; GRD – gardiquimod). Replicates were collapsed and genes hierarchically clustered. Rows – genes, columns – experiments, red – induced from baseline, blue – repressed from baseline, white – unchanged from baseline. (**B**) Model illustrating the differential networks controlling the anti-viral ('polyI:C-like') and inflammatory ('PAM-like') programs.

Fig. 3. Gene regulatory programs controlling the response to pathogen components. (**A** and **B**), A strategy to minimize false discovery (FDR) calls of significant changes in an output target gene resulting from knockdown of a regulator gene. (A) The first FDR procedure (top) compares the expression of the gene following a perturbation with a regulator shRNA (right) to its expression upon perturbation with 32 non-targeting shRNAs (left). The dashed lines identify the gene-specific FDR-based thresholds for induction (blue line) and repression (red line). A discrete vector of significant calls (bottom) is derived from the raw data (blue – regulator represses the target gene; red – regulator induces the target gene). (B) A second FDR procedure (top) compares the expression of the target gene to that of eight control (target) genes upon perturbation with the same shRNA. In the example shown, the gene's induction (left) was significant relative to the variation in expression among the control target genes resulting in a high score (bottom, dark blue), but its repression did not significantly differ from the controls, resulting in a lower score (bottom, weaker red). (**C**) All significant (95% confidence) relations between regulators (columns) and targets (rows), colored as in (B). The gray bar (right) represents the NMF-based calls for each target gene; black – anti-viral program; dark gray – inflammatory

program; light gray – control genes. The bottom bar shows the degree of effect by the regulator on each program as determined by the NMF projection of the regulator's perturbation profile. (Yellow – positive effect, Green – negative effect; NMF scores are mean-normalized).

Fig. 4. The core regulatory circuits controlling the inflammatory and anti-viral responses. (**A**) The antiviral sub-network shows regulatory relations between the core anti-viral regulators (blue nodes, top), their targets (boxes, bottom), each other, and inflammatory regulators (green node, top right). The two top regulators, Stat1 and 2, activate all anti-viral targets (dashed blue arrows). The second-tier regulators activate sub-sets of targets (dashed purple arrows). (**B**) Examples of feed forward loop classes identified in the network, with fraction of each class. (**C**) A core regulatory network of the inflammatory and anti-viral programs, consisting of the most distinct regulators, and their relation to ligands and receptors (top). Pointed arrows – induction; blunt arrows – repression; green ovals – inflammatory regulators; blue ovals – antiviral regulators. Example target genes are noted. (**D**) nCounter expression profiles for the target genes (rows) upon perturbation with shRNAs against a subset of viral regulators (columns) and followed by stimulation with LPS (left) or polyI:C (right). All values are normalized by expression in cells infected with a control shRNA and under the same stimulus (shCtl).

Fig 5. The polycomb component Cbx4 selectively restricts IFNB1 production under bacterial perturbations. (**A** and **C**) LPS (red), polyI:C (blue) and PAM (green)-induced expression of ifnb1 (A) and cbx4 (C) derived from data in Fig. 2A. (**B**) IFNB1 expression (by nCounter) in response to LPS in DCs perturbed by control shRNAs or shRNAs targeting each of 125 regulators (format as in Fig. 3B). (**D**) Ifnb1 mRNA levels (by qRT-PCT) following LPS treatment in unsorted mouse bone marrow DCs perturbed with an shRNA against Cbx4 (black) or a control shRNA (gray); signals are relative to t=0. (**E**) ifnb1 mRNA levels (by nCounter) at 6h post-LPS or PAM in bone marrow DCs perturbed with an shRNA against Cbx4 or one of three control shRNAs. (**F**) Model for bacterial-specific repression of ifnb1 by Cbx4: both polyI:C and LPS induce ifnb1 expression early (left), but only LPS induces Cbx4, which then represses the ifnb1 locus at a later time (right, top).

Measure states

- Genomewide mRNA expression

Gene selection

- Select regulators

- Select representative signature of output genes

Network perturbation

- RNAi perturbation

Stimulation

- Measure mRNA levels of signature genes
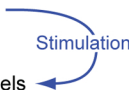
Construct network

- Develop network model based on data

A Complete response

Ctrl. LPS pIC PAM CpG GRD

Inflammatory

pIC specific

Anti-viral

Shared

1800 genes

Time (hours): 0.5 1 2 4 6 8 12 16 24

B

pIC — Tlr3
LPS — Tlr4
PAM — Tlr2

Plasma membrane

Nuclear membrane

GRN1 ← → GRN2

Anti-viral program   Inflammatory program

A

# of molecules

FDR #1
gene x

Control shRNA   Regulator shRNA

Discrete calling:
- Repressor
- Activator
- Not significant

Repressor   Activator

FDR #1

B

Z score

shRNA regulator #1

shRNA regulator #2

FDR #2

FDR #2

Control genes   Gene x

Control genes   Gene x

Discrete calling:
- Repressor
- Activator
- Not significant

Repressor   Activator

FDR #2

C

Regulators (shRNA perturbation)

Output genes (expression)

Cxcl1
Cxcl2
Il1a
Il1b
Il6
Ptgs2
Marco
Il12b

Il12a

Ifnb1
Cxcl11
Cxcl9
Cxcl10
Ifit1
Isg15
Il5
Il15ra
Mx2
Ifit3
Ifit2
Timeless
Stat2

Stat1

Inflammatory
Anti-viral

Stat1
Stat2
Stat4
Irf9
Irf8
Jun
Timeless
Irf4
Rbl1
Nmi

Fus
Atf4

Rela

Nfkbiz
Nfkb1
Rnnc2
Pum2
Cebpb

2   0

A (antiviral program network diagram with STAT1, STAT2, FUS, NFKBIZ, HHEX, RBL1, TIMELESS and gene sets including Ifnb1, Cxcl9, Ifit3, Tsc22d1, Tnfsf8, Isg20; Nmi, Iigp1, Irf7, Lhx2, Bbx; Fus, Tcf4, Pml, Usp12; Irf7; Ifit1, Isg15, Rbl1; Usp25, Daxx, Cd40, Atm, Irf1, Ligp2, Mertk; Cxcl11, Stat2, Trim12, Stat1, Trim21)

B Coherent feed-forward loops (75% of FFLs)

Type 1	Type 2	Type 3	Type 4
NFKBIZ → ATF4 → *il12b*	FUS ⊣ RBL1 ⊣ *Iigp1*	CBX4 → DNMT3A ⊣ *ifnb1*	NFKBIZ ⊣ FOS ⊣ *il12b*
82%	8%	7%	3%

Incoherent feed-forward loops (25% of FFLs)

Type 1	Type 2	Type 3	Type 4
JUN → FUS ⊣ *ifnb1*	FUS ⊣ NFKBIZ ⊣ *fos*	CBX4 → STAT1 ⊣ *ifnb1*	CEBPZ ⊣ HHEX → *iL12b*
57%	10%	7%	26%

C Anti-viral program (CXCL10, IFNB1, IL15) and Inflammatory program (IL12B, CXCL2, CXCL1, IL1B, CCL3) regulatory diagram with pIC/Tlr3, LPS/Tlr4, PAM/Tlr2 signaling through plasma membrane and nuclear membrane. Anti-viral TFs: STAT1, STAT2, IRF8, ETV6, JUN, STAT4, IRF9, ATF3, RBL1; HMGN3, FUS, SAP30. Inflammatory TFs: PLAGL2, NFKB1, NFKBIZ, RUNX1, CEBPB, ATF4, HAT1, RELA, PNRC2; NFE2L2, IRF4, FOS.

D Heatmap (LPS and pIC conditions; shCtl, shRbl1, shJun, shTimeless, shNmi) with labeled genes: Timeless, Nmi, Ifna2, Ifna4, Ifna5, Ifnb1, Rbl1, Cxcl10, Socs1, Il15ra, Il15, Isg15, Cxcl11, Daxx, Stat1, Irf7, Isg20