



## Antiviral Response Dictated by Choreographed Cascade of Transcription Factors

Elena Zaslavsky, Uri Hershberg, Jeremy Seto, Alissa M. Pham, Susanna Marquez, Jamie L. Duke, James G. Wetmur, Benjamin R. tenOever, Stuart C. Sealfon and Steven H. Kleinstein

This information is current as of May 8, 2012

*J Immunol* 2010;184;2908-2917; Prepublished online 17 February 2010;  
doi:10.4049/jimmunol.0903453  
<http://www.jimmunol.org/content/184/6/2908>

- 
- Supplementary Data** <http://www.jimmunol.org/content/suppl/2010/02/15/jimmunol.0903453.DC1.html>
- References** This article **cites 40 articles**, 19 of which can be accessed free at: <http://www.jimmunol.org/content/184/6/2908.full.html#ref-list-1>
- Subscriptions** Information about subscribing to *The Journal of Immunology* is online at <http://www.jimmunol.org/subscriptions>
- Permissions** Submit copyright permission requests at <http://www.aai.org/ji/copyright.html>
- Email Alerts** Receive free email-alerts when new articles cite this article. Sign up at <http://www.jimmunol.org/etoc/subscriptions.shtml/>



# Antiviral Response Dictated by Choreographed Cascade of Transcription Factors

Elena Zaslavsky,<sup>\*,†,1</sup> Uri Hershberg,<sup>‡,1</sup> Jeremy Seto,<sup>\*,†</sup> Alissa M. Pham,<sup>§</sup>  
 Susanna Marquez,<sup>‡</sup> Jamie L. Duke,<sup>¶</sup> James G. Wetmur,<sup>§</sup> Benjamin R. tenOever,<sup>§</sup>  
 Stuart C. Sealfon,<sup>\*,†</sup> and Steven H. Kleinstein<sup>‡,¶</sup>

The dendritic cell (DC) is a master regulator of immune responses. Pathogenic viruses subvert normal immune function in DCs through the expression of immune antagonists. Understanding how these antagonists interact with the host immune system requires knowledge of the underlying genetic regulatory network that operates during an uninhibited antiviral response. To isolate and identify this network, we studied DCs infected with Newcastle disease virus, which is able to stimulate innate immunity and DC maturation through activation of RIG-I signaling, but lacks the ability to evade the human IFN response. To analyze this experimental model, we developed a new approach integrating genome-wide expression kinetics and time-dependent promoter analysis. We found that the genetic program underlying the antiviral cell-state transition during the first 18 h postinfection could be explained by a single convergent regulatory network. Gene expression changes were driven by a stepwise multifactor cascading control mechanism, where the specific transcription factors controlling expression changed over time. Within this network, most individual genes were regulated by multiple factors, indicating robustness against virus-encoded immune evasion genes. In addition to effectively recapitulating current biological knowledge, we predicted, and validated experimentally, antiviral roles for several novel transcription factors. More generally, our results show how a genetic program can be temporally controlled through a single regulatory network to achieve the large-scale genetic reprogramming characteristic of cell-state transitions. *The Journal of Immunology*, 2010, 184: 2908–2917.

Dendritic cells (DCs) play a key role in the early immune response to viral infections. Activation of viral pattern sensors, such as RIG-I and MDA5, triggers a transcriptional program that includes the production of type I IFNs. These antiviral cytokines signal in both autocrine and paracrine fashion through the JAK-STAT pathway leading to additional transcription events involving the differential expression of many hundreds of genes. The antiviral state produced by this extensive genetic reprogramming involves a core set of genes as well as pathogen-specific components (1).

The DC response to individual pathogens involves multiple signals that must be integrated to initiate an appropriate immune

response. Pathogenic viruses attempt to subvert normal immune function through the expression of IFN antagonists (2, 3). For example, IFN regulatory factor (IRF) 3 activation and IFN- $\beta$  expression are blocked by the NS1 protein of influenza (4). Unraveling the impact of these immune antagonists would be aided by a detailed understanding of the genetic regulatory network that operates during an uninhibited antiviral response. This knowledge is lacking because previous human studies have used viruses that interfere with the immune response (4–6). One fundamental unresolved question is to what extent the antiviral response is a single interconnected transcriptional cascade (convergent architecture) or a combination of transcriptional events operating independently in reaction to the multiple signals that arise following viral insult (parallel architecture). Newcastle disease virus (NDV) infection of human DCs provides an ideal system to define the uninhibited regulatory network (7, 8). NDV is an avian virus that is able to stimulate innate immunity and DC maturation, but lacks the ability to evade the human interferon response (9). By focusing on NDV, we can accurately depict the baseline network of transcription factor (TF) interactions that underlie a broad range of immune responses. Through comparative studies, this network will enable detailed analysis of other infections and greatly improve our understanding of the control mechanisms in antiviral immunity and the myriad ways through which pathogenic viruses subvert normal immune function.

Systems biology methods combined with high-throughput experimental technologies are providing new insights into virus-host interactions (10, 11). Genome-wide transcriptional profiling has suggested that the DC antiviral response is characterized by temporal waves of gene activation, which may be controlled by different combinations of transcriptional regulators (1). Potential regulators can be implicated using direct approaches, such as differential expression of the TF mRNA (1) or indirectly by identification of

\*Center for Translational Systems Biology, <sup>†</sup>Department of Neurology, and <sup>§</sup>Department of Microbiology, Mount Sinai School of Medicine, New York, NY 10029; and <sup>‡</sup>Department of Pathology, Yale University School of Medicine, and <sup>¶</sup>Interdepartmental Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT 06511

<sup>1</sup>E.Z. and U.H. contributed equally.

Received for publication October 30, 2009. Accepted for publication December 31, 2009.

This work was supported by National Institutes of Health, National Institute of Allergy and Infectious Diseases Contract No. HHSN266200500021C. J.L.D. was supported in part by National Institutes of Health Grant T15 LM07056 from the National Library of Medicine.

The microarray data presented in this article have been submitted to the Gene Expression Omnibus under accession number GSE18791.

Address correspondence and reprint requests to Dr. Steven H. Kleinstein, Pathology Informatics, Yale University School of Medicine, 300 George Street, Suite 505, New Haven, CT 06511. E-mail address: steven.kleinstein@yale.edu

The online version of this article contains supplemental material.

Abbreviations used in this paper: DC, dendritic cell; GO, gene ontology; IRF, IFN regulatory factor; NDV, Newcastle disease virus; TF, transcription factor; TSS, transcription start site.

Copyright © 2010 by The American Association of Immunologists, Inc. 0022-1767/10/\$16.00

common *cis*-regulatory motifs (12, 13). These methods typically provide a static view of the network. Other computational methods have been proposed to identify TFs driving time-dependent changes in expression, but these do not explicitly account for the regulation of the TF itself (14, 15). The most common approaches are based on the hypothesis that genes sharing a similar temporal profile are regulated by common TFs (16). In mammals, a variety of posttranscriptional regulatory mechanisms can impact mRNA kinetics following upregulation (17), so that requiring similarity across the entire time series may be unnecessarily restrictive. As an alternative, we propose to focus on the initial upregulation time as a criterion to identify genes that are likely to share common regulatory control logic.

We developed and validated by experiment a new approach integrating genome-wide expression kinetics and time-dependent promoter analysis. Our method infers the TFs driving initial gene expression changes, determines the timing of their activity and identifies a causal chain of regulation. We have applied this approach to the anti-NDV response in human DCs to deduce the causality and coherence of the transcriptional events responsible for the complex gene regulation elicited by virus infection. Along with many expected factors (e.g., IRFs and STATs), key roles were predicted and experimentally validated for several regulators with no previously known role in antiviral responses. Our approach identified a single transcriptional cascade that spanned the experimental time series and could account for observed expression changes in a majority of all upregulated genes.

## Materials and Methods

### Multiple sample gene array time course following NDV infection of isolated DCs

**Differentiation of DCs.** Monocyte-derived conventional DCs were obtained from human blood donors following a standard protocol (18, 19). Briefly, human monocytes from buffy coats were isolated by Ficoll density gradient centrifugation (Histopaque; Sigma-Aldrich, St. Louis, MO), and CD14<sup>+</sup> monocytes were immunomagnetically purified by using a MACS CD14 isolation kit (Miltenyi Biotec, Auburn, CA). CD14<sup>+</sup> monocytes ( $0.7 \times 10^6$  cells/ml) were later differentiated into immature conventional DCs after 5 to 6 d incubation in DC growth media (RPMI 1640; Life Technologies, Carlsbad, CA), 10% FCS (Hyclone; Thermo Fisher Scientific, Waltham, MA), 2 mM L-glutamine, 100 U/ml penicillin, 100  $\mu$ g/ml streptomycin (Pen/Strep; Invitrogen, Carlsbad, CA), 500 U/ml hGM-CSF (PeproTech, Rocky Hill, NJ) and 1000 U/ml hIL-4 (PeproTech) at 37°C.

**Virus preparation and viral infection.** The recombinant Hitcher strain of NDV (rNDV/B1) was prepared, and aliquots of allantoic fluid were harvested as previously described (20). NDV virus stock was titered by infection of Vero cell plates and identification of viral growth by the addition of mAbs specific for NDV HN protein (Mount Sinai Hybridoma Core Facility, New York, NY) followed by addition of anti-mouse IgG-FITC and visualization using fluorescent microscopy. Titered NDV stock was diluted in DMEM and added directly into pelleted DCs at a multiplicity of infection of 0.5 prepared as previously described (21). Postincubation for 30 min at 37°C, fresh DC growth medium was added back to the infected DCs ( $1.0 \times 10^6$  cells/ml). Virus-free allantoic fluid was added to additional tubes of cells to serve as a negative control.

**Microarray experiments.** Total RNA was purified from  $2.5 \times 10^6$  DCs per time sample using RNeasyMicro (Qiagen, Valencia, CA) with DNase treatment. RNA was eluted from columns using water and quantified by spectrophotometry. Quality control of RNAs was performed using the Agilent Bioanalyzer (Agilent, Santa Clara, CA). Total RNA from cells of two different donors were infected with NDV or control and harvested at 0, 1, 2, 6, 10, and 18 h for control and 1, 2, 4, 6, 8, 10, 12, 14, 16, and 18 h for NDV infection. Biological replicates were performed for both of these donors. A third donor was not included in the current analysis because no replicate measurements were made. Good-quality RNA was reverse transcribed using T7-oligo(dT)24 to yield double-stranded cDNA. cRNA was transcribed and biotinylated from cDNA templates. Fragmented cRNA was hybridized to Affymetrix HU133plus2 Gene Chip Arrays (Affymetrix, Santa Clara, CA) by the Mount Sinai Microarray Shared Resource Facility. All microarrays studied in this paper have been deposited in the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) with the series accession number GSE18791.

**Microarray analysis.** Raw Affymetrix microarray data were normalized using GCRMA. Differential expression was defined for each probeset using four criteria. For at least one time point in each of the experiments, we required: 1) a minimum log<sub>2</sub> intensity of 5, which was indicative of nonexpressed genes by visual inspection of the expression intensity histogram at time-zero; 2) an absolute fold-change of at least 2 relative to time-zero; and 3) a significant change in expression by LIMMA (BioConductor implementation) after correction for multiple hypothesis testing ( $q < 0.005$ ) (22). The fourth criterion for defining differential expression required that the probeset did not meet criteria 2 and 3 above in the control time series. In analyses where a background set of genes was required, it was defined as the set of genes from all time points that met criterion 3 (significant change by LIMMA), without regard to either of the other criteria. All of this analysis was performed using the BioConductor software package (23) in R. Gene ontology (GO) analysis was performed using the conditional test in the GStats package (24).

### Comparing the timing of response progression

**Estimating the upregulation time for genes.** The initial upregulation time for each gene was estimated by fitting the mRNA expression time series following NDV infection to a logistic function (Equation 1). Parameters ( $A$ ,  $K$ ,  $I$ , and  $C$ ) were estimated by minimizing the sum-squared error using differential evolution with modified sampling (25). To neutralize the undue influence of high-intensity gene array results, we divided the error at each time point by its intensity. The upregulation time for each gene was defined as the time of maximal change in the acceleration of the logistic function (star in Fig. 1). This point can be calculated by finding the maximum of the third derivative (i.e., the first time at which the fourth derivative of the logistic function equals zero). Estimates of the upregulation time were considered acceptable if they fell between 0 and 18 h (i.e., the time points included in the fit). Note that the value for time point zero in the fitting was taken from the appropriate control gene arrays because measurements of time zero were done only for the control experiments.

$$A + \frac{K}{1 + e^{-I(t-C)}} \quad (1)$$

### TF binding site analysis

Using the University of California Santa Cruz Genome Bioinformatics site (<http://genome.ucsc.edu/>), we downloaded the transcription start site (TSS) for all human RefSeq genes, defined by the March 2006 refGene table (26). The region 2 kb upstream of each TSS was identified within a genome-wide multiple alignment of 27 vertebrate species to the human genome (27), also available through University of California Santa Cruz Genome Bioinformatics site. To identify putative TF binding sites, the human sequences, along with aligned regions from chimp and mouse, were analyzed using the TRANSFAC MATCH (28) algorithm with a cutoff chosen to minimize the sum of false positives and false negatives. The analysis was performed for all vertebrate TF matrices in the 2008.3 release of TRANSFAC (29), and putative binding sites were considered to be evolutionarily conserved if matches were also found at the aligned positions in both chimp and mouse sequences and had no gaps present in the multiple alignment. Each TRANSFAC matrix was linked to a set of gene symbols describing potential binding factors using annotations present in the "Binding Factor" field of the database. Only vertebrate TRANSFAC matrices that could be linked to a human (HGNC) gene symbol (30), either directly or through an alias listed in National Center for Biotechnology Information gene, were included. To build the regulatory network, we further refined the list of TFs to include only those that were upregulated in our microarray time series.

### Time-dependent promoter analysis and regulatory network construction

We constructed a regulatory network by inferring transcriptional regulators of each time slice. The TF inference was based on statistical enrichment of putative TF targets. More precisely, let  $G$ , called the foreground set, be the set of differentially expressed genes at a particular time point, and let  $T$  be the set of all genes in the dataset with binding sites for a given TRANSFAC matrix  $M$ . The set  $T$  is determined by choosing an appropriate sequence match cutoff for matrix  $M$  and further filtering the genes with matches to  $M$  to include only those with a conserved binding site upstream of an orthologous mouse and chimp gene (as described above). The background set  $B$ , which serves to catalog the expected distribution of binding sites for  $M$  and dictate whether the observations in the foreground set  $G$  are unusual, is computed as described above (see *Microarray analysis*). Next, we find how many of the genes in the foreground and

background contain binding sites for matrix  $M$ , defining subsets in the foreground and the background as intersections of  $G$  and  $B$ , respectively, with  $T$ . With  $N_g = |G|$ ,  $N_b = |B|$ ,  $H_g = |G \cap T|$  and  $H_b = |B \cap T|$ , we computed the hypergeometric  $p$  value  $HGP(N_g, N_b, H_g, H_b)$  (Equation 2), which is the probability that, if binding sites for  $M$  were assigned randomly to genes in the foreground and the background, one might observe at least  $H_g$  binding sites in the foreground set.

$$HGP = \sum_{j=H_g}^{\min(H_g+H_b, N_g)} \frac{\binom{N_g}{j} \binom{N_b}{H_g+H_b-j}}{\binom{N_g+N_b}{H_g+H_b}} \quad (2)$$

At each time point, we computed the  $p$  value for every TRANSFAC matrix mapped to a gene found to be differentially expressed somewhere in the time course and retained those passing a false discovery rate-corrected threshold of 0.05 (22). To connect the individual TFs into a regulatory network, we placed each TF gene represented by one of the enriched matrices at the time it is first differentially expressed. An activity window was defined for every TF in the network. The temporal boundaries of the activity window of each TF were defined to include the earliest and latest times of differential expression and enrichment of associated matrices. Once the activity window has been identified, the TF is linked to all genes with predicted binding sites for which the upregulation time falls within the window. Regulatory links appear explicitly in the network diagram when both the target and regulator appear as nodes in the network and are implicit otherwise. Feedforward links connect regulators to targets that are differentially expressed at later time points, whereas feedback links are predicted when the target is differentially expressed before the regulator.

### EMSA

Whole cell extracts were obtained from NDV-infected DCs at 0, 4, 6, and 8 h postinfection using Nonidet P-40 lysis buffer containing 50 mM Tris, 150 mM NaCl, 5 mM EDTA, 10% glycerol, 40 mM  $\beta$ -glycerophosphate, 1% Nonidet P-40, and 1 mM PMSF. In the analysis of ISRE target sites, the 0 and 4 h time points were from the same experiment, whereas the 6 and 8 h time points were not. In all other cases, 0 and 6 h were from the same experiment, whereas 4 and 8 h came from different infections. Cells were rinsed in PBS and allowed to lyse on ice for 30 min. Samples were then spun at 14,000 rpm for 10 min, and supernatant was collected. Protein concentrations were quantified using Bradford reagent (Bio-Rad, Hercules, CA) per the manufacturer's instructions. To begin the protein-DNA binding reaction, 10  $\mu$ g whole cell extracts were incubated in a total volume of 15  $\mu$ L containing 50 mM HEPES (pH 7.9), 10% glycerol, 200 mM KCl, 5 mM EDTA (pH 8), 1 mM MgCl<sub>2</sub>, 5 mM DTT, and 1  $\mu$ g poly (deoxyinosine-deoxycytidylic acid; Sigma-Aldrich). Samples were incubated on ice for 10 min, followed by the addition of 150,000 CPU of <sup>32</sup>P-labeled DNA probe and 20 min of incubation at room temperature. Oligonucleotide probes were ordered from Sigma-Aldrich and annealed to their complementary oligonucleotides using annealing buffer containing 100 mM NaCl and 50 mM HEPES (pH 7.6). Probe sequences are shown in Table II. The reaction was placed at 95°C and allowed to cool to room temperature. Annealed probes were end labeled with [<sup>32</sup>P] using T4 polynucleotide kinase (NEB). Samples were electrophoresed at 180 V in 0.5% Trisborate-EDTA buffer on an 8% native polyacrylamide gel composed of 49:1 acrylamide to bis-acrylamide. Gels were dried on Whatman paper at 80°C for 1 h and exposed by autoradiogram.

## Results

To define the transcriptional regulatory network underlying an uninhibited antiviral response, we performed genome-wide transcriptional profiling of human monocyte-derived DCs infected with NDV. Differentiated DCs from two human donors, each at two different instances, were infected in vitro with NDV, and microarray analysis was performed over the first 18 h postinfection.

### Timing of gene expression is highly conserved

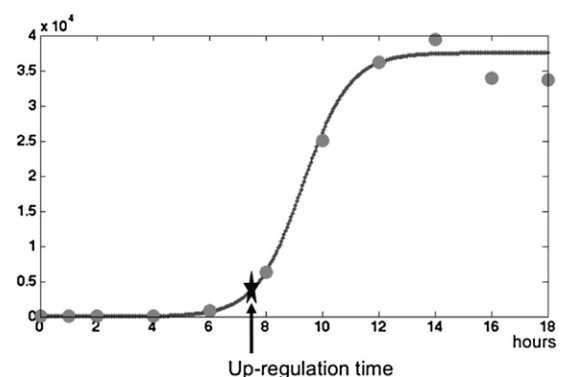
To determine the degree of variability present during the initial phases of the anti-NDV response, we compared the timing of gene expression changes between individuals and across experiments. The initial upregulation time for each gene was estimated by fitting a logistic function to the mRNA expression time series (Fig. 1 and

*Materials and Methods*). The logistic function is a commonly used model for transcriptional regulation, and visual inspection of the data suggested this was appropriate for many of the genes known to participate in the antiviral response (e.g., IFNB1; Fig. 1). This model-based approach allowed us to estimate the upregulation time at a finer resolution than the microarray sampling. Another advantage is that the estimate is less sensitive to variation in individual expression values because the entire time course is fit by the model. Focusing on a representative set of genes known to participate in pathogen responses (31), we identified 94 upregulated genes that displayed expression patterns with an acceptable fit to the logistic function. These were used to measure the degree of conservation in the anti-NDV response. The selected genes spanned a wide range of upregulation times, although none were found to be upregulated at later time points (after 12 h postinfection; Fig. 2), indicating either the relatively early initiation of the response or the difficulty of fitting expression profiles that do not reach saturation.

To quantify the overall conservation of the regulatory progression between samples and subjects, we assessed the relative degree of similarity in the upregulation times of the selected pathogen response genes. As can be readily seen in Fig. 2, the estimated upregulation times were very highly correlated both between donors and across samples, resulting in correlation coefficients ranging between 0.8 and 0.9. This indicated that the initial steps of the antiviral response are ordered in a highly conserved progression of gene expression. Such a high degree of conservation in the time progression, as our results indicated, is most likely the result of a tightly controlled set of regulatory mechanisms that operate as the system progresses through time. Surprisingly, the variance in upregulation times appeared relatively constant over time (Fig. 2), suggesting the existence of feedback mechanisms to dampen any noise that might accumulate as the signal propagates through the network (32, 33). To identify the regulatory cascade underlying this tightly controlled system, we first focused on the events occurring at distinct time-points and then connected this information across time into a coherent higher-level transcriptional network. We focused specifically on upregulated genes because downregulation times were not correlated across donors or samples (Supplemental Fig. 1).

### Multifactor cascading control of the antiviral genetic program

In this section, we implemented our time-centric view by identifying and analyzing sets of genes that were first upregulated at each time point (i.e., 1, 2, 4, 6, 8, 10, 12, 14, 16, and 18 h postinfection; Fig. 3A). This approach contrasts with the more standard analyses



**FIGURE 1.** Model-based estimation of the upregulation time for individual genes. The upregulation time for each gene was determined by fitting a logistic function (solid line) to the experimentally measured mRNA expression values from a single experiment (circles). The estimated time of upregulation was defined as the maximal change in the acceleration of the logistic function and is indicated by a star. This example shows IFNB1 expression from donor 2, sample 1.



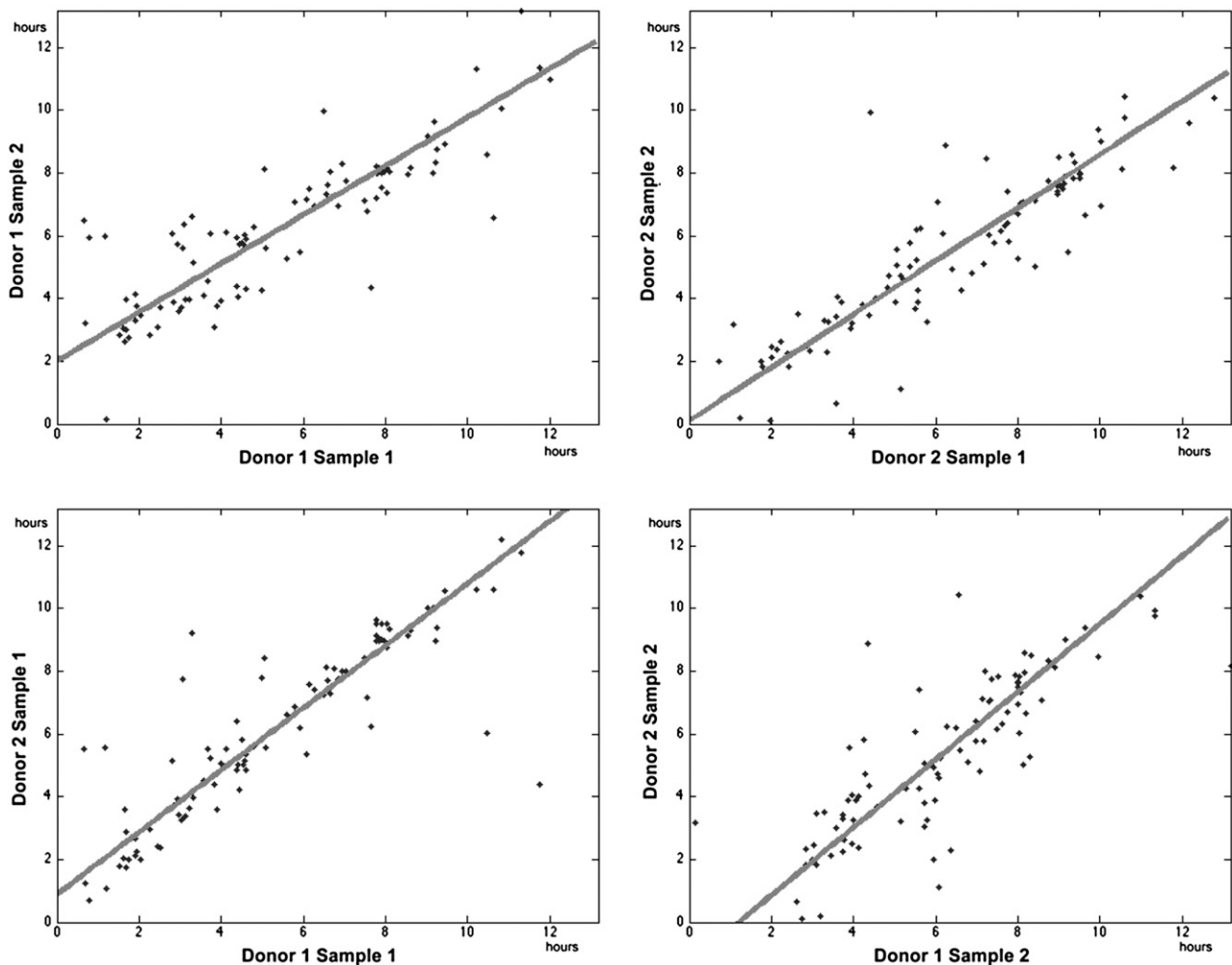
that first cluster genes across a set of experiments and then analyze these refined subsets for functional enrichment, TF activity, etc. (16, 34, 35). It also differs from the straightforward approach of treating each time point independently, because the entire time series is used to determine the point of initial upregulation.

**Transcriptional regulators.** The activation of the antiviral response is complex and involves many viral and cellular determinants. Due to the high level of correlation we found between samples and the careful temporal progression it implied, we analyzed the underlying regulatory controls in a temporal succession. Starting from the 1351 upregulated genes, we considered sets of genes that were first differentially expressed at each time point separately and attempted to infer the TFs involved in regulating them. Transcriptional regulators for each time slice were identified by testing for statistical enrichment of their putative targets (as described in *Materials and Methods*). That is, if among the set of upregulated genes an unusually high number contain binding sites for a particular TF, that factor is likely involved in regulating the genes under consideration. Statistical significance was assessed using the hypergeometric distribution to measure the degree to which a given regulator's set of targets is overrepresented in the sample of genes under consideration compared with a representative background set.

Using information from the TRANSFAC database (29), we identified 80 TF binding matrices (motifs) annotated to human TF

genes that were upregulated during at least one point in our microarray time series. We restricted the TRANSFAC database in this way because we were only interested in those TFs most likely to participate in the regulatory program as conduits of the transcriptional signals (16, 34, 35), implying differential-expression at the mRNA level. Although it would be conceptually simple to extend the analysis to all TFs included in TRANSFAC, computing the  $p$  value for all TRANSFAC matrices would also greatly reduce statistical power as a result of the multiple hypothesis correction. For each of these 80 matrices, we determined the set of putative target genes by scanning the promoter regions and requiring that potential binding sites be conserved in chimp and mouse (*Materials and Methods*). This kind of phylogenetic footprinting has been shown to decrease the number of false-positive binding sites (36).

Our enrichment analysis found a total of 30 matrices whose targets were significantly over-represented during at least one time-point (false discovery rate,  $q < 0.05$ ). We generated a temporal enrichment profile for each of these matrices by identifying the different time points during which their targets were over-represented. As seen in Fig. 3B, the activity of most of the matrices, and by implication the TFs they represent, spans several hours. Moreover, we observed multiple temporal phases in the response, each driven by distinct groups of TFs. In agreement with many previous studies of the innate antiviral response (6), IRF-based



**FIGURE 2.** The anti-NDV response is highly correlated across samples and donors. The initial upregulation times were estimated for a set of genes known to be involved in pathogen responses using a model-based analysis. Upregulation times for individual genes (points) were compared across samples from the same donor (*upper row*) and between donors for the same sample (*bottom row*). The extent of conservation was determined using orthogonal least-squares linear regression (solid lines). The correlations for each comparison were, starting at the top left corner and moving clockwise: 0.85, 0.88, 0.81, and 0.87.

activation was evident in the initial wave of transcriptional upregulation. In our analysis, this initial wave begins 4 h postinfection, somewhat later than expected. This is likely because genes are assigned to time points somewhat later than their actual upregulation due to the stringent differential-expression criteria used in this study. Similarly, STAT activity was apparent by 4–6 h postinfection and peaked at 14 h postinfection. The middle phase of the response was driven by a variety of TFs, many of which have not been previously implicated in antiviral responses (see *DC antiviral transcriptional network*). Overall, the inferred enrichment profiles strongly suggested a cascade of TFs that drive different temporal phases of the response (Fig. 3B).

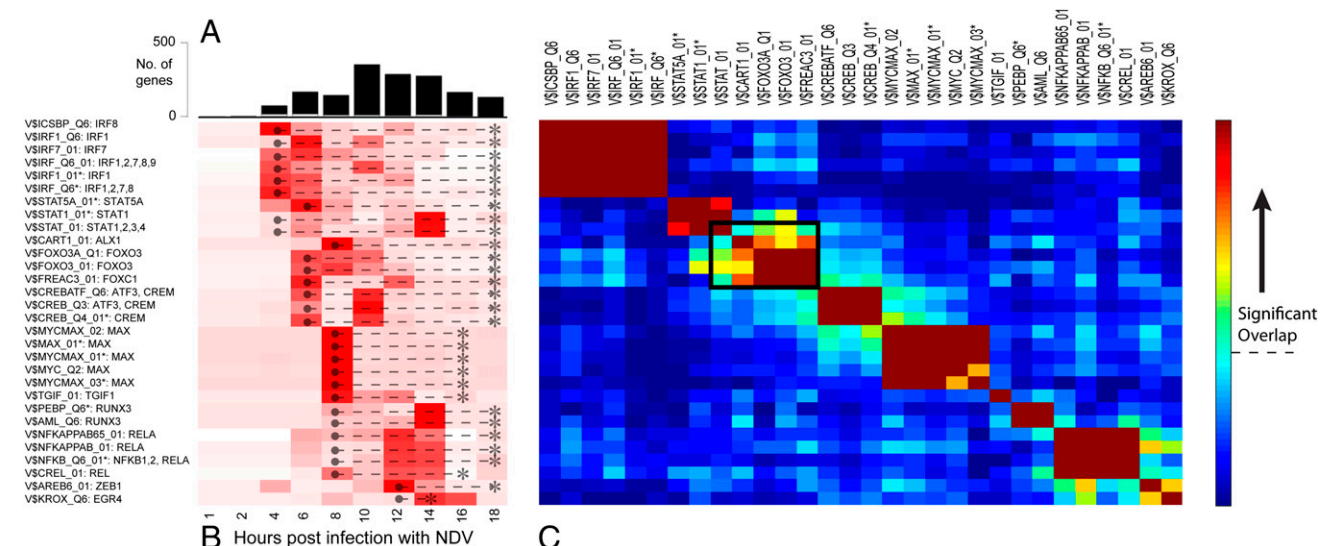
**Distributed overlap and redundancy of transcription target genes.** Having determined the individual regulators of the response, we focused on their target genes to determine the extent of cooperative regulation and redundancy. It is quite clear that there is a great amount of redundancy between the targets of the different matrices. Of the 30 TRANSFAC matrices implicated, none has  $> \sim 10\%$  unique targets (and many have less). Moreover, most individual genes are targeted by  $> 1$  matrix. This can arise when a single TRANSFAC matrix is associated with multiple genes (e.g., the association of VSIRF\_Q6 with several IRFs) or when the TFs are members of a complex that is represented by multiple matrices (e.g., NFKB1, NFKB2, REL, and RELA). In other cases, a high overlap of target genes could indicate cooperative regulation. To identify these potential interactions, we quantified the extent of overlap in the target gene sets for all of the TRANSFAC matrices implicated in the anti-NDV response (Fig. 3C). As expected, the most significant overlap of target genes occurred for matrices that describe members of the same TF family or complex. These are readily apparent in the dark red blocks along the diagonal in Fig. 3C. However, in a few cases, significant overlap was observed between different TF families. The most striking example involves the matrices encoding the STATs, ALX1, FOXO3, and FOXC1, whose targets overlap  $\sim 30\%$  on average (Fig. 3C, black box). Within this group of core TFs, the highest pairwise overlap of target

genes occurs for ALX1 and FOXO3, and we hypothesize that these two TFs may be part of a single *cis*-regulatory module. In general, we predicted that most genes would be regulated by multiple TFs, with only 150 genes containing binding sites corresponding to a single matrix.

#### Causal linkage generates unified regulatory cascade

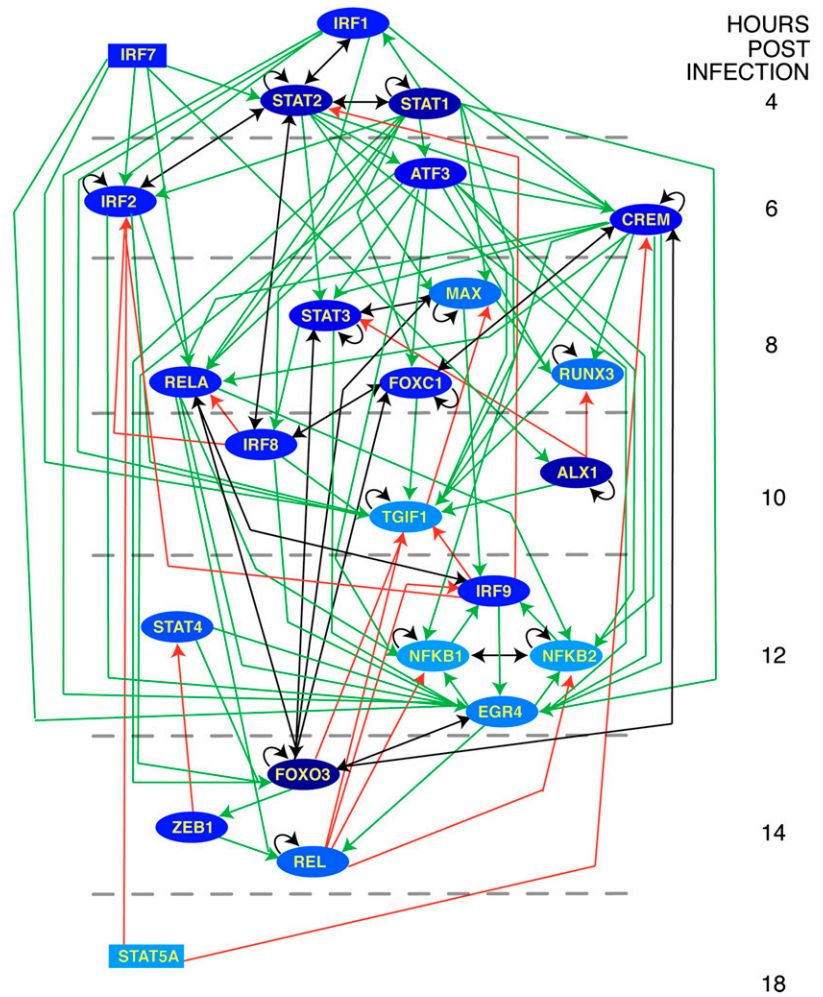
We sought to generate a transcriptional regulatory network that could connect the TFs associated with each phase of the response into a coherent linked cascade. This kind of representation required us to shift our focus from TRANSFAC matrices, which can each be annotated to multiple TFs, to a framework that considers individual TF activity over time.

**DC antiviral transcriptional network.** We generated an initial set of network links by identifying all pairs of TFs in which one was a putative regulator of the other. To reduce the potential for false-positive regulatory relationships, we limited the time frame during which a TF could be implicated as having an important role in regulating a potential target. This was accomplished by defining an activity window for each TF (*Materials and Methods*). We removed all network links where the initial upregulation of the target gene occurred outside the activity window of the regulator. This pruning removed 37 links, producing a final network that consisted of 24 TFs and 133 regulatory links. To visualize the resulting network (Fig. 4), each TF was placed at the time its mRNA was first upregulated in the response. We also considered the placement of TFs based on their activity, but ruled this out for two main reasons. First, many of the TRANSFAC matrices were enriched at multiple time points, making the choice of placement somewhat arbitrary. Second, this scheme does not allow for differentiation between TFs that are annotated to the same TRANSFAC matrix. Placing genes at the time of their first differential expression provides TF-specific information and, in combination with the activity heatmap (Fig. 3B), leads to a clearer view of the temporal progression of the network. This placement scheme is also consistent with our view of the biological process



**FIGURE 3.** Dynamic activity profiles and target overlap for TFs in the NDV response network. **A**, The number of genes first upregulated at each time point, split by TF (white bar) and non-TF (black bar) genes. **B**, Heat map showing the overrepresentation of targets associated with each of the TRANSFAC matrices in the network (rows) over time (columns). The colors are row normalized  $-\log(p)$  values. Darker red indicates greater inferred activity of the TF. The temporal activity window of each TF matrix (●—\*) was inferred from the union of the activity of all of the individual TFs represented by that matrix. **C**, Heat map quantifying the extent of overlap between predicted targets of each TRANSFAC matrix (rows and columns). Colors indicate  $-\log(p)$  values of the hypergeometric test. The threshold for statistical significance assuming a Bonferroni-corrected  $p$  value of 0.05 is indicated by the dashed line on the color bar. Matrices that could be individually removed without any loss of coverage are marked with an asterisk (\*). The black box indicates a region of high overlap between TFs from different families, suggesting possible cooperativity.

**FIGURE 4.** Transcription regulatory network for the uninhibited antiviral response. Each node represents a TF with inferred activity in the anti-NDV response. Edges connect regulators to targets so that arrowtails indicate upregulation of the regulator, whereas arrowheads indicate activity of the regulator on the target. Regulatory relationships can be feedforward (green links), feedback (red links), or reciprocal (black links). Time in the figure progresses vertically down, with nodes placed in the time slice during which the gene is first differentially expressed. Within each time slice, TFs are ordered using the model-based analysis wherever possible. Node color reflects importance measured by number of outgoing links to all gene targets (i.e., total number of genes, not just TFs, targeted by the node in question), with darker colors corresponding to more highly connected nodes. Rectangular nodes indicate TFs with no predicted regulators.

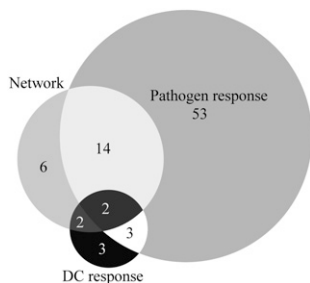


as a stepwise cascade where the upregulation of the TF is tied to its activity. To provide additional resolution, genes appearing within the same time slice were ordered based on the upregulation time estimated by our model-based analysis (see *Timing of gene expression is highly conserved*). To interpret the network visualization in Fig. 4, it is helpful to keep in mind that links connect regulators to targets and arrowtails indicate upregulation of the regulator, whereas arrowheads indicate activity of the regulator.

The resulting network of 24 TFs spans virtually the entire time period analyzed. The TFs in this network are predicted to regulate 779 of 1351 (58%) upregulated genes. As mentioned above, the

transcriptional cascade we have discovered includes many TFs already associated with the interferon response (e.g., IRFs, STATs, NF-κB) (6). It additionally includes *ATF3* and *TGF1*, which were previously implicated by other systems biology analyses (16, 35). In all, 18 of 24 TFs implicated in the transcriptional cascade controlling the antiviral response appear in the known general pathogen response signature (31) or the core DC response signature (1) (Fig. 5). Clearly, our approach is effective at capturing the known biology of the system.

The network contains both feedforward links, which propagate the transcriptional signal through time, as well as feedback links, where TFs may influence the activity of targets that have previously been upregulated (Fig. 4). Feedback links, although serving as potential conduits of regulatory activity, do not neatly fit into our idealized view of a regulatory cascade where upregulation of TFs at the



**FIGURE 5.** Comparison of the anti-NDV transcriptional response with other pathogen responses. TFs defined by the inferred network were compared with TFs from two published gene signatures: the general pathogen response defined in Ref. 31 and the core DC response defined in Ref. 1. TFs in these signatures were identified by their annotation in Ref. 37 or linkage to a vertebrate TRANSFAC matrix (*Materials and Methods*).

Table I. Novel TFs predicted by the network analysis

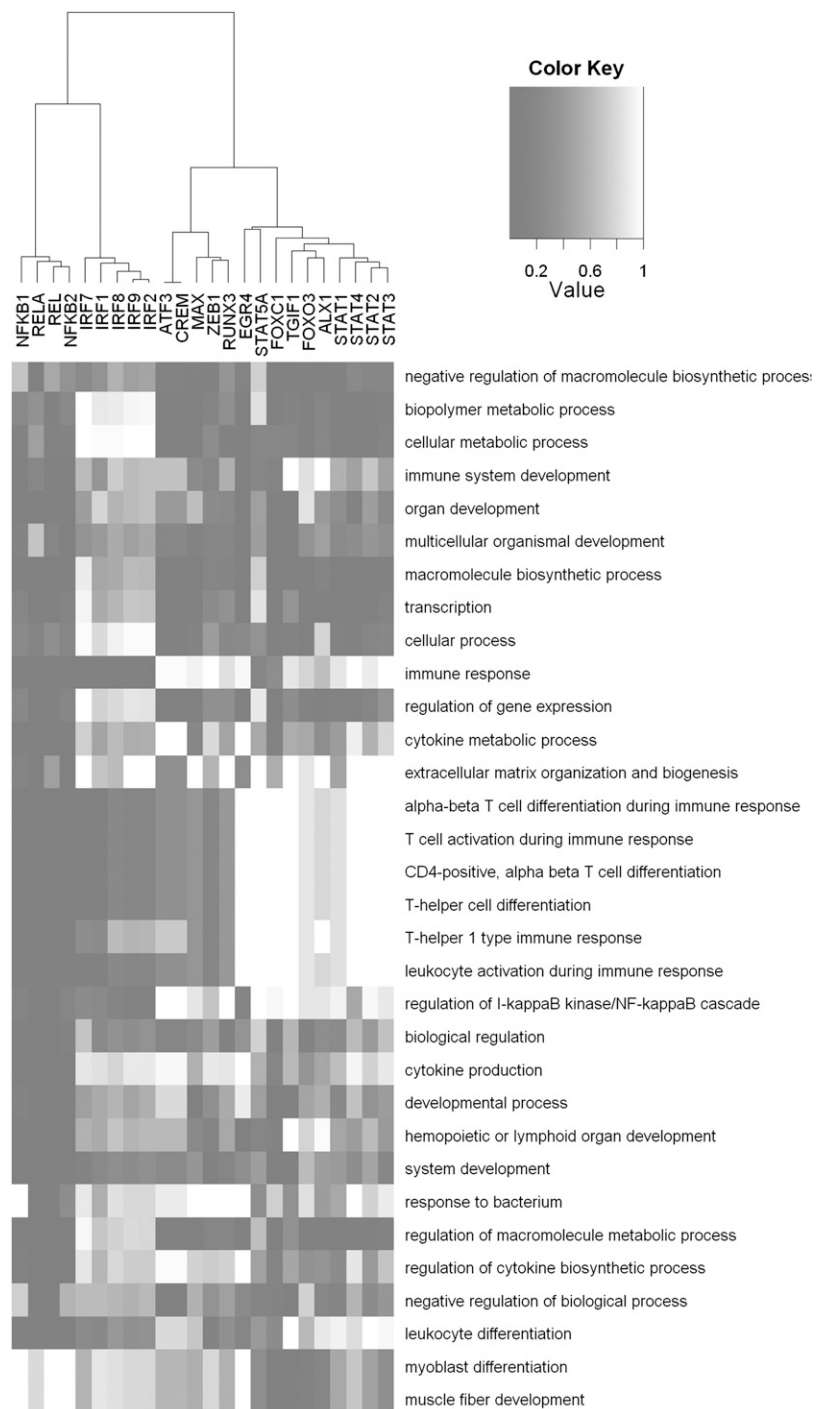
Gene symbol	TRANSFAC Matrix	TRANSFAC Motif	First Expressed (h)
<b>ALX1</b>	V\$CART1_01	...TAAT...ATTA...	10
<b>FOXC1</b>	V\$FREAC3_01	...GTAAA:AAACA...	8
<b>FOXO3</b>	V\$FOXO3_01 <sup>a</sup>	...TTTAC...	14
<b>MAX</b>	V\$MYC MAX_02 <sup>a</sup>	...CACATG...	8
<b>RUNX3</b>	V\$AML_Q6 <sup>a</sup>	...TTTGT...	8
<b>ZEB1</b>	V\$AREB6_01	...ACCT...	14

Factors listed in boldface were experimentally tested using EMSA. <sup>a</sup>Multiple TRANSFAC matrices are significant; only the one with the most significant *p* value is indicated.

mRNA level drives the next wave of transcriptional changes. Thus, it is important to note that excluding these links only disconnects a single TF (STAT5A) from our network. We chose to visualize these links to present potentially important information on candidate feedback circuits. To understand whether the inferred network architecture was effective in capturing the underlying transcriptional cascade, we reasoned that true regulatory interactions would link genes that appeared close together in time. To test this hypothesis, we compared the average link length (i.e., the absolute difference in upregulation time between the regulator and target) in the inferred network with that found in a set of random networks. To generate each random network, all upregulated genes were randomly assigned an initial differential expression time, and TFs were randomly assigned an activity window from the set of observed

profiles. Networks were inferred from these randomized data using the same method as the observed data, including the determination of enriched TRANSFAC matrices, etc. We found that the average link length in the real network (3.2 h) was significantly shorter than those of the random networks ( $p < 0.002$  for 1000 random networks). A similar result was obtained when the link lengths were computed over all the target genes and not just the TFs appearing in Fig. 4 ( $p < 0.005$  for 1000 random networks). We conclude that our network is effective in capturing the underlying biology and produces a pattern that is consistent with a stepwise transcriptional signal propagation.

*Experimental validation of regulatory links.* The network implicated six novel TFs in the antiviral response (Fig. 5). We applied multiple criteria for selecting the most promising of these factors



**FIGURE 6.** GO analysis of TF targets. Each row corresponds to a GO term that was significantly enriched ( $p < 0.001$ ) for the targets of at least one TF. The colors indicate absolute  $p$  values (darker for increased significance) computed using the conditional test in the GOstats package in Bioconductor. Hierarchical clustering of the columns was carried out using the Euclidean distance metric and Ward's linkage.



along with a set of their predicted regulators and targets for experimental validation by EMSA. First, we looked for a robust upregulation pattern in the TF mRNA expression and focused particularly on those upregulated in the middle phase of the response. Second, we eliminated TFs with nonspecific binding signatures so that the true sites would be easier to identify (Table I). Finally, we selected TFs with targets either in the constructed network or in the known pathogen response gene set, additionally selecting targets with very good matches to the TF's binding signature. To ensure the reliability of the EMSA validation, we verified that such targets did not overlap with predicted binding sites for any other matrixes annotated to TFs associated with pathogen responses (1, 31) or appearing in our network. Our set of chosen TFs included ALX1, FOXC1, and RUNX3. Of these, ALX1 and FOXC1 were of particular interest, because they belonged to the group of TFs with evidence for a significant overlap among their target sets (see *Distributed overlap and redundancy of transcription target genes*; Fig. 3C). As a group, these TFs had at least one binding site in 70% of all the network targets, suggesting they are critical factors driving the observed upregulation pattern. We sought to place these TFs in a functional context by performing a GO analysis of their predicted target genes. FOXC1, RUNX3, and ALX1 targets share many of the same functions; however, they have distinct profiles (Fig. 6). The targets of RUNX3 include genes related to the T cell response, and the factor appears functionally similar to known players *ATF3* and *CREM*. In contrast, ALX1 and FOXC1 both cluster together with several STATs, although FOXC1 targets more genes related to development.

To validate the chosen target sites associated with our three novel TFs, we performed EMSAs with NDV-infected whole cell extract using <sup>32</sup>P-labeled oligonucleotides. EMSA probes were designed to include core target sites flanked with the endogenous genomic sequence (Table II). In comparing uninfected whole cell extract to extract derived from 4, 6, and 8 h postinfection, we were able to verify four uncharacterized virus-inducible elements (Fig. 7). TF binding assays using a characterized ISRE element of ISG15 and a putative IRF7 site upstream of the ALX1 TSS both demonstrated virus-inducible binding that was evident at 4 h postinfection and was sustained for the duration of the sampling (Fig. 7A, 7B). Furthermore, a putative ALX1 target element also demonstrated virus-inducible occupancy during infection (Fig. 7C). This element upstream of the *PLA1A* gene mediated the assembly of a single DNA-binding complex estimated to be 75–150 kDa. Moreover, binding of the ALX1 promoter element in *PLA1A* appeared to be dynamic, as a second faster migrating complex was

visible by 8 h postinfection, suggesting a time-dependent change in the composition of this complex (Fig. 7C). These results place ALX1 within a well-known network motif (38) (Fig. 8). Two additional non-ISRE binding sites were also verified by EMSA, including putative binding sites for FOXC1 and RUNX3 (Fig. 7D, 7E). Taken together, these data corroborate both microarray data and computational analyses, suggesting a complex transcriptional regulatory network that may depend on the cooperative activities of multiple TFs.

## Discussion

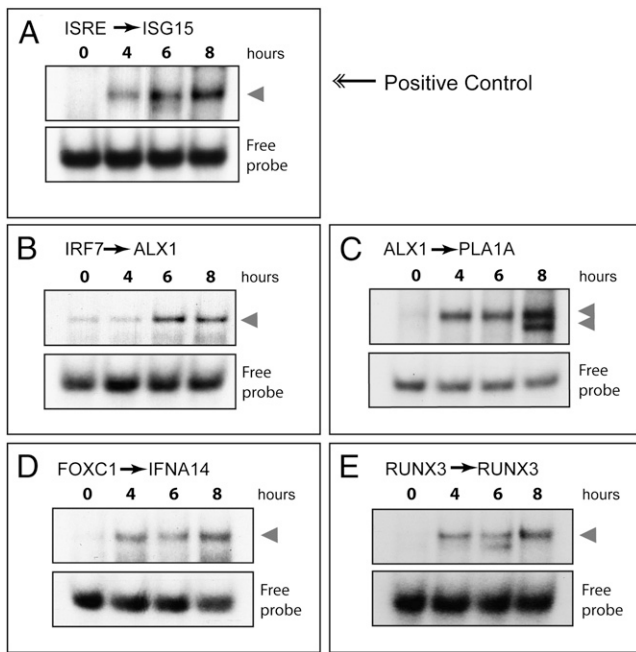
The activation of viral pattern sensors combined with cytokine signaling initiates a genetic program in DCs that is critical to controlling infection. We have found that the timing of this program is highly conserved. To deduce the underlying regulatory architecture, we developed an integrative method to reverse engineer the transcriptional network and its progression through time. Using genome-wide transcriptional profiling data covering the first 18 h postinfection, our approach identified a single regulatory network composed of 24 TFs that accounts for the upregulation of a majority of genes (Fig. 4). This network captures and connects through time many known elements of the antiviral response, including IRFs, STATs, and NF-κB. The cascading pattern of TF activity (Fig. 3B) and the interconnectedness of the network (Fig. 4) combined with the tight conservation of target upregulation times (Fig. 2), all indicate an integrated control as opposed to an underlying parallel architecture. It is important to carefully interpret the network structure and ordering of events defined by the predicted regulatory network because the overall genetic program results from the integration of multiple signals that may arrive at different times. The NDV infection was performed using a multiplicity of infection of 0.5 so that only approximately half the cells are infected and thus many cells are simply responding to the IFN and other cytokines secreted by infected cells, rather than directly to the virus. Although this implies that the gene expression measurements reflect the operation of two distinct signaling pathways, the inferred network structure suggests that both rapidly converge on an overlapping set of downstream transcriptional regulators. Nevertheless, this could explain why NF-κB activity follows STAT activity in our network. We conclude that the uninhibited antiviral response is the result of a single transcriptional cascade with a convergent architecture and is not implemented as a series of independent transcriptional events.

Six TFs whose involvement in the network was predicted by our method are not part of the general pathogen response signature (31) or the DC core response (1). These represent potentially novel antiviral

Table II. Probe sequences used in EMSA experiments

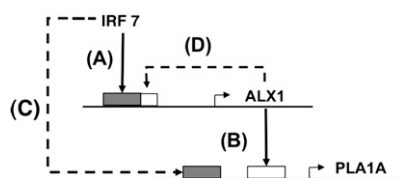
Gene symbol	Specific TRANSFAC Matrix	TF/Target Site	Probe Sequence (5'–3')
ISG15 (control) <sup>a</sup>		ISRE	GATCGGAAAGGGAAACCGAAACTGAAGCC
ALX1 <sup>a</sup>	V\$IRF7_01	IRF7	AGTTC TAATGAATATGAAAATAGGCGGGCG
PLA1A <sup>a</sup>	V\$CART1_01	ALX1	GAGGCTAGTTAATGATTAGTGAAAATCCACG
IFNA5	V\$CART1_01	ALX1	ATTAAAATTTAATGGGATTTTTAGTTAGAA
IFNA7	V\$CART1_01	ALX1	TGAACATACTAATTTCCATTTTCTAAATGC
FOXC1	V\$STAT1_01	STAT1	GGGCTCCGCTGCCCGGAAAAAGTGTAAGT
IFNA14 <sup>a</sup>	V\$FREAC3_01	FOXC1	GTGCATAGGCTCTAAAATAAGGAACATAC
CREM	V\$FREAC3_01	FOXC1	CAAATGGCTCAGCAAATAAAAATGTTAA
CXCL9	V\$FREAC3_01	FOXC1	GAAATGAATATCCTAAAATAAATATGATCCC
PEL1	V\$FREAC3_01	FOXC1	CCTTATTAACCTGCAAATAAAAATGCTGTGA
RUNX3	V\$CREB_Q4_01	CREM	CACCTGGGCCGTGATGTACAGCCCTTTTA
RUNX3 <sup>a</sup>	V\$AML_Q6	RUNX3	CACCTGGGCCGTGATGTACAGCCCTTTTA
CCL3	V\$AML_Q6	RUNX3	TATCCTGAGCCCTGTGGTACCAGGGACC

<sup>a</sup>Validated by EMSA.



**FIGURE 7.** Experimental validation of TF network connections. EMSAs were performed at 0, 4, 6, and 8 h postinfection to test for virus-inducible binding at predicted binding sites. *A*, Positive control demonstrating virus-inducible binding to an ISRE element in the promoter of ISG15. As validation of the TF network connections, virus-inducible binding was demonstrated for predicted binding sites of IRF7 (*B*), ALX1 (*C*), FOXC1 (*D*), and RUNX3 (*E*). The probe sequences (detailed in Table II) encompass predicted binding sites found in the promoters of ALX1 (*B*), PLA1A (*C*), IFNA14 (*D*), and RUNX3 (*E*) as specified in *Materials and Methods*. TF-probe complexes are indicated with arrows.

TFs: ALX1, FOXC1, FOXO3, MAX, RUNX3, and ZEB1. To validate the transcriptional network, we attempted to experimentally validate the regulatory connections for three of these novel TFs (ALX1, FOXC1, and RUNX3). We considered both incoming links, testing regulation of the novel TF, and outgoing links, testing the ability of the novel TF to regulate other genes. Using EMSA, we validated 4 out of the 12 regulatory relationships tested, indicating an antiviral role for all of the TFs tested. Moreover, this success rate likely represents a lower bound for the specificity of our method in determining regulatory relationships for two main reasons. First, we did not include the many well-known players in our experimental validation. Second, a negative EMSA result may simply reflect suboptimal binding conditions or be indicative of a relatively unstable complex. Furthermore, although the temporal activation of PLA1A, IFNA14, and RUNX3 does not necessarily correlate to the binding kinetics of ALX1, FOXC1, and RUNX3 putative binding sites, respectively, this is likely a reflection of their individual



**FIGURE 8.** Network motif involving the putative antiviral TF ALX1. *A*, The ALX1 gene is activated by IRF7. *B*, ALX1 activates other genes such as PLA1A. The solid lines (*A*, *B*) represent EMSA-validated network connections. The dashed lines, completing a feedforward loop (*C*) and indicating autoregulation (*D*), were predicted computationally.

transcriptional potentials or ability to recruit DNA modifying enzymes. In the case of the putative FOXC1 binding site, this element represents the first documented example of a non-IRF controlling the transcription of this genetic cluster. Although this site was found to overlap the TATA-box, it has also been observed that FOX proteins can regulate genes through direct binding to TATA-boxes (39). Further experiments will be needed to verify the true impact of this site and the cause for the FOXC1 binding site we have found.

ALX1 had the highest number of experimentally validated links. Based on the occupancy of the IRF7 binding site upstream of the ALX1 TSS, it would appear that transcription is, in part, mediated directly by type I IFN signaling (Fig. 7*B*). In the network, ALX1 is predicted to cooperate with IRFs in the regulation of downstream genes as part of a feed-forward loop (Fig. 8). ALX1 encodes CART1, a paired-class homeodomain protein that is necessary for survival of the forebrain mesenchyme in rodents (40). Although the binding site preferences for CART1 have been characterized (41), the function of this protein in humans is not known. A GO analysis of ALX1-predicted targets suggests that this factor is functionally similar to STAT1 and FOXO3 (Fig. 6). ALX1 target genes are involved in a variety of functions, but the most enriched GO category predicts that this factor is most likely involved in the negative regulation of macromolecule biosynthetic processes (GO:0010558; Fig. 6). Our network analysis places ALX1 within a well-known network motif (38), supporting a critical role for this TF in coordinating the antiviral response (Fig. 8). Our experimental results also appear to indicate the workings of such a feedforward control element of PLA1A by ALX1 and another virus-activated TF. Looking at our EMSA results, we see that following the assembly of a single DNA-binding complex at 4 h, a second faster migrating complex was visible by 8 h postinfection (Fig. 7*B*). These results suggest the exciting possibility that ALX1, and other novel TFs identified in this study, may cooperate both with each other, or with known virus activated factors such as IRF3, IRF7, STAT1, and/or NF- $\kappa$ B.

Combinatorial regulation can play an important role in mammalian gene regulation but is not specifically incorporated in our approach. However, multiple regulators are predicted for most target genes, and the potential for combinatorial control can be investigated indirectly by analysis of TF target overlap (Fig. 3*C*). Consistent with our idea of cascading control, we found that most of the implicated TFs do not exhibit statistically significant pairwise overlap of target genes. However, a cluster of high overlap, including ALX1 (mentioned above) and also FOXO3, FOXC1, and STAT, was identified and may also reflect cooperative activity (Fig. 3*C*, black box). Together, these TFs account for  $\sim 70\%$  of the genes controlled by the network. The cooperative nature of the network is further evidenced by the fact that only  $\sim 19\%$  of genes controlled by the network are targeted by a single matrix. Thus, although TFs pass along the baton in an orderly fashion, individual genes may be controlled by multiple TFs at various points in the cascade, and the network as a whole is the minimal unit of control.

The network-building approach developed in this study is generally applicable to transcriptional profiling time series. It differs from most current analyses in several important ways. First, we grouped genes according to their initial time of upregulation. Second, instead of simply identifying the controlling TFs, we connected these together mechanistically to generate a complete transcriptional cascade. In cases where the time resolution of microarray sampling may not be sufficiently dense to isolate genes with common *cis*-regulatory logic, we suggest that the model-based analysis might be generalized to estimate upregulation times on a finer scale. Several assumptions underlying our network reconstruction method are important to keep in mind. The analysis is limited to TFs whose

binding preferences are known and are included in the TRANSFAC database. Nevertheless, as we have demonstrated, it is possible to implicate factors with previously unknown functions such as ALX1. Our method is also restricted to TFs that are differentially expressed at the mRNA level. Consequently, it will miss factors that are posttranscriptionally regulated, which may be particularly important in the earliest stages of the response. Furthermore, because only upregulated genes are included, the analysis will also not identify transcriptional repressors. Although it is conceptually possible to extend the network reconstruction procedure to down-regulated genes, the lack of correlation among downregulation times (Supplemental Fig. 1) suggested that a transcriptional cascade may not be the best model for these data.

To study the highly dynamic processes involved in immune interaction with pathogens requires a methodology that incorporates time directly into the analysis rather than subsuming it. The method presented in this study is a first step in the construction of such a methodology. Using our time-centric promoter analysis methods, we have elucidated the transcriptional network underlying an uninhibited antiviral response in human DCs. Our results indicate a robust convergent design and stepwise execution of the antiviral program. Although inherently limited to discovering regulation by TFs whose binding preferences are known (and annotated in TRANSFAC), the inferred network nonetheless accounts for a majority of upregulated genes and time points. The identification of key transcriptional players in the antiviral response, along with knowledge of their timing and regulatory architecture, provides a framework to identify the specific mechanisms used by human pathogens to subvert normal immune function.

## Disclosures

The authors have no financial conflicts of interest.

## References

- Huang, Q., D. Liu, P. Majewski, L. C. Schulte, J. M. Korn, R. A. Young, E. S. Lander, and N. Hacohen. 2001. The plasticity of dendritic cell responses to pathogens and their components. *Science* 294: 870–875.
- Bonjardim, C. A., P. C. Ferreira, and E. G. Kroon. 2009. Interferons: signaling, antiviral and viral evasion. *Immunol. Lett.* 122: 1–11.
- Geiss, G. K., V. S. Carter, Y. He, B. K. Kwiciszewski, T. Holzman, M. J. Korh, C. A. Lazaro, N. Fausto, R. E. Bumgarner, and M. G. Katze. 2003. Gene expression profiling of the cellular transcriptional network regulated by alpha/beta interferon and its partial attenuation by the hepatitis C virus nonstructural 5A protein. *J. Virol.* 77: 6367–6375.
- Mibayashi, M., L. Martínez-Sobrido, Y. M. Loo, W. B. Cárdenas, M. Gale, Jr., and A. García-Sastre. 2007. Inhibition of retinoic acid-inducible gene I-mediated induction of beta interferon by the NS1 protein of influenza A virus. *J. Virol.* 81: 514–524.
- Zilliox, M., G. Parmigiani, and D. Griffin. 2006. Gene expression patterns in dendritic cells infected with measles virus compared with other pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 103: 3363–3368.
- Haller, O., G. Kochs, and F. Weber. 2006. The interferon response circuit: induction and suppression by pathogenic viruses. *Virology* 344: 119–130.
- Park, M.-S., M. L. Shaw, J. Muñoz-Jordan, J. F. Cros, T. Nakaya, N. Bouvier, P. Palese, A. García-Sastre, and C. F. Basler. 2003. Newcastle disease virus (NDV)-based assay demonstrates interferon-antagonist activity for the NDV V protein and the Nipah virus V, W, and C proteins. *J. Virol.* 77: 1501–1511.
- Honda, K., S. Sakaguchi, C. Nakajima, A. Watanabe, H. Yanai, M. Matsumoto, T. Ohteki, T. Kaisho, A. Takaoka, S. Akira, et al. 2003. Selective contribution of IFN- $\alpha/\beta$  signaling to the maturation of dendritic cells induced by double-stranded RNA or viral infection. *Proc. Natl. Acad. Sci. U.S.A.* 100: 10872–10877.
- Martinez-Sobrido, L., N. Gitiban, A. Fernandez-Sesma, J. Cros, S. E. Mertz, N. A. Jewell, S. Hammond, E. Flano, R. K. Durbin, A. García-Sastre, and J. E. Durbin. 2006. Protection against respiratory syncytial virus by a recombinant Newcastle disease virus vector. *J. Virol.* 80: 1130–1139.
- Peng, X., E. Y. Chan, Y. Li, D. L. Diamond, M. J. Korh, and M. G. Katze. 2009. Virus-host interactions: from systems biology to translational research. *Curr. Opin. Microbiol.* 12: 432–438.
- Tan, K., J. Tegner, and T. Ravasi. 2008. Integrated approaches to uncovering transcription regulatory networks in mammalian cells. *Genomics* 91: 219–231.
- Segal, E., M. Shapira, A. Regev, D. Pe'er, D. Botstein, D. Koller, and N. Friedman. 2003. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat. Genet.* 34: 166–176.
- Bar-Joseph, Z., G. K. Gerber, T. I. Lee, N. J. Rinaldi, J. Y. Yoo, F. Robert, D. B. Gordon, E. Fraenkel, T. S. Jaakkola, R. A. Young, and D. K. Gifford. 2003. Computational discovery of gene modules and regulatory networks. *Nat. Biotechnol.* 21: 1337–1342.
- Ernst, J., O. Vainas, C. Harbison, I. Simon, and Z. Bar-Joseph. 2007. Reconstructing dynamic regulatory maps. *Mol. Syst. Biol.* 3: 74.
- Nilsson, R., V. B. Bajic, H. Suzuki, D. di Bernardo, J. Björkregren, S. Katayama, J. F. Reid, M. J. Sweet, M. Gariboldi, P. Carninci, et al. 2006. Transcriptional network dynamics in macrophage activation. *Genomics* 88: 133–142.
- Gilchrist, M., V. Thorsson, B. Li, A. G. Rust, M. Korb, J. C. Roach, K. Kennedy, T. Hai, H. Bolouri, and A. Aderem. 2006. Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. *Nature* 441: 173–178.
- Petrocca, F., and J. Lieberman. 2009. Micromanagers of immune cell fate and function. *Adv. Immunol.* 102: 227–244.
- Sallusto, F., C. Nicolò, R. De Maria, S. Corinti, and R. Testi. 1996. Ceramide inhibits antigen uptake and presentation by dendritic cells. *J. Exp. Med.* 184: 2411–2416.
- Chang, W. L. W., N. Baumgarth, D. Yu, and P. A. Barry. 2004. Human cytomegalovirus-encoded interleukin-10 homolog inhibits maturation of dendritic cells and alters their functionality. *J. Virol.* 78: 8720–8731.
- Park, M. S., A. García-Sastre, J. F. Cros, C. F. Basler, and P. Palese. 2003. Newcastle disease virus V protein is a determinant of host range restriction. *J. Virol.* 77: 9522–9532.
- Fernandez, N. C., A. Lozier, C. Flament, P. Ricciardi-Castagnoli, D. Bellet, M. Suter, M. Perricaudet, T. Tursz, E. Maraskovsky, and L. Zitvogel. 1999. Dendritic cells directly trigger NK cell functions: cross-talk relevant in innate anti-tumor immune responses in vivo. *Nat. Med.* 5: 405–411.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* 57: 289–300.
- Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, et al. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5: R80.
- Falcon, S., and R. Gentleman. 2007. Using GOstats to test gene lists for GO term association. *Bioinformatics* 23: 257–258.
- Kleinstein, S. H., D. Bottino, A. Georgieva, R. Sarangapani, and G. S. Lett. 2006. Nonuniform sampling for global optimization of kinetic rate constants in biological pathways. In *Winter Simulation Conference*. F. P. W. L. F. Perrone, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, ed. Winter Simulation Conference, p. 1611–1616.
- Pruitt, K. D., T. Tatusova, and D. R. Maglott. 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 33(Database issue): D501–D504.
- Blanchette, M., W. J. Kent, C. Riemer, L. Elnitski, A. F. Smit, K. M. Roskin, R. Baertsch, K. Rosenbloom, H. Clawson, E. D. Green, et al. 2004. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* 14: 708–715.
- Kel, A. E., E. Gössling, I. Reuter, E. Cheremushkin, O. V. Kel-Margoulis, and E. Wingender. 2003. MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* 31: 3576–3579.
- Matys, V., E. Fricke, R. Geffers, E. Gössling, M. Haubrock, R. Hehl, K. Hornischer, D. Karas, A. E. Kel, O. V. Kel-Margoulis, et al. 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* 31: 374–378.
- Bruford, E. A., M. J. Lush, M. W. Wright, T. P. Sneddon, S. Povey, and E. Birney. 2008. The HGNC Database in 2008: a resource for the human genome. *Nucleic Acids Res.* 36(Database issue): D445–D448.
- Jenner, R. G., and R. A. Young. 2005. Insights into host responses against pathogens from transcriptional profiling. *Nat. Rev. Microbiol.* 3: 281–294.
- Dublanché, Y., K. Michalodimitrakis, N. Kummerer, M. Foglierini, and L. Serrano. 2006. Noise in transcription negative feedback loops: simulation and experimental analysis. *Mol. Syst. Biol.* 2: 41.
- Skotheim, J. M., S. Di Talia, E. D. Siggia, and F. R. Cross. 2008. Positive feedback of G1 cyclins ensures coherent cell cycle entry. *Nature* 454: 291–296.
- Ravasi, T., C. A. Wells, and D. A. Hume. 2007. Systems biology of transcription control in macrophages. *Bioessays* 29: 1215–1226.
- Ramsey, S. A., S. L. Klemm, D. E. Zak, K. A. Kennedy, V. Thorsson, B. Li, M. Gilchrist, E. S. Gold, C. D. Johnson, V. Litvak, et al. 2008. Uncovering a macrophage transcriptional program by integrating evidence from motif scanning and expression dynamics. *PLoS Comput. Biol.* 4: e1000021.
- Sauer, T., E. Shelest, and E. Wingender. 2006. Evaluating phylogenetic footprinting for human-rodent comparisons. *Bioinformatics* 22: 430–437.
- Roach, J. C., K. D. Smith, K. L. Strobe, S. M. Nissen, C. D. Haudenschild, D. Zhou, T. J. Vasicek, G. A. Held, G. A. Stolovitzky, L. E. Hood, and A. Aderem. 2007. Transcription factor expression in lipopolysaccharide-activated peripheral-blood-derived mononuclear cells. *Proc. Natl. Acad. Sci. U.S.A.* 104: 16245–16250.
- Milo, R., S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. 2002. Network motifs: simple building blocks of complex networks. *Science* 298: 824–827.
- Wierstra, I., and J. Alves. 2006. FOXM1c transactivates the human *c-myc* promoter directly via the two TATA boxes P1 and P2. *FEBS J.* 273: 4645–4667.
- Zhao, Q., R. R. Behringer, and B. de Crombrughe. 1996. Prenatal folic acid treatment suppresses acrania and meroanencephaly in mice mutant for the *Cart1* homeobox gene. *Nat. Genet.* 13: 275–283.
- Cai, R. L. 1998. Human CALT1, a paired-class homeodomain protein, activates transcription through palindromic binding sites. *Biochem. Biophys. Res. Commun.* 250: 305–311.