Transcriptional Correlates of Disease Outcome in Anticoagulant-Treated Non-Human Primates Infected with Ebolavirus

Sara Garamszegi[1], Judy Y. Yen[2,3], Anna N. Honko[4], Joan B. Geisbert[5], Kathleen H. Rubins[6], Thomas W. Geisbert[5], Yu Xia[1,7], Lisa E. Hensley[8], John H. Connor[1,2,3]*

1 Bioinformatics Program, Boston University, Boston, Massachusetts, University of America, **2** Department of Microbiology, School of Medicine, Boston University, Boston, Massachusetts, University of America, **3** National Emerging Infectious Diseases Laboratories, Boston University, Boston, Massachusetts, University of America, **4** U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, Frederick, Maryland, United States of America, **5** Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, Texas, United States of America, **6** National Aeronautics and Space Administration, Houston, Texas, United States of America, **7** Department of Bioengineering, McGill University, Montreal, Canada, **8** Integrated Research Facility at Fort Detrick, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Fort Detrick, Frederick, Maryland, United States of America

Abstract

Ebola virus (EBOV) infection in humans and non-human primates (NHPs) is highly lethal, and there is limited understanding of the mechanisms associated with pathogenesis and survival. Here, we describe a transcriptomic analysis of NHPs that survived lethal EBOV infection, compared to NHPs that did not survive. It has been previously demonstrated that anticoagulant therapeutics increase the survival rate in EBOV-infected NHPs, and that the characteristic transcriptional profile of immune response changes in anticoagulant-treated NHPs. In order to identify transcriptional signatures that correlate with survival following EBOV infection, we compared the mRNA expression profile in peripheral blood mononuclear cells from EBOV-infected NHPs that received anticoagulant treatment, to those that did not receive treatment. We identified a small set of 20 genes that are highly confident predictors and can accurately distinguish between surviving and non-surviving animals. In addition, we identified a larger predictive signature of 238 genes that correlated with disease outcome and treatment; this latter signature was associated with a variety of host responses, such as the inflammatory response, T cell death, and inhibition of viral replication. Notably, among survival-associated genes were subsets of genes that are transcriptionally regulated by (1) CCAAT/enhancer-binding protein alpha, (2) tumor protein 53, and (3) megakaryoblastic leukemia 1 and myocardin-like protein 2. These pathways merit further investigation as potential transcriptional signatures of host immune response to EBOV infection.

Citation: Garamszegi S, Yen JY, Honko AN, Geisbert JB, Rubins KH, et al. (2014) Transcriptional Correlates of Disease Outcome in Anticoagulant-Treated Non-Human Primates Infected with Ebolavirus. PLoS Negl Trop Dis 8(7): e3061. doi:10.1371/journal.pntd.0003061

Editor: Anita K. McElroy, Centers for Disease Control and Prevention, United States of America

Received March 20, 2014; **Accepted** June 18, 2014; **Published** July 31, 2014

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. The raw microarray dataset was deposited in NCBI's Gene Expression Omnibus database (Accession number: GSE24943).

Funding: SG was supported by a fellowship from the National Science Foundation (NSF) Integrative Graduate Education and Research Traineeship (DGE-0654108) and an NSF Graduate Research Fellowship (DGE-0741448). YX was supported by a Research Starter Grant in Informatics from the Pharmaceutical Research and Manufacturers of America Foundation. Funding for experimental animal studies was provided by the Joint Science and Technology Office for Chemical and Biological Defense and the Defense Threat Reduction Agency (JSTO-CBD 4.0021.08.RD.B). This work was conducted at USAMRIID under project numbers 19595 and 188261 supported by JSTO-CBD. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: jhconnor@bu.edu

Introduction

Ebola virus (EBOV; *Filoviridae* [1]) infection of humans and non-human primates (NHPs) can cause viral hemorrhagic fever, an acute systemic illness characterized by fever, bleeding diathesis, fulminant shock, and death [2]. Although several studies have identified candidate therapeutics that may mitigate the effects of EBOV infection [3–12], there are currently no FDA-approved post-exposure treatments for human use. Additionally, despite extensive research on EBOV pathogenesis [13–15], lifecycle [16], and interactions with the host [17–19], there are no standard biomarkers to predict host immune response to EBOV infection, nor are there biomarkers of drug efficacy or survival following treatment. Here, we investigated the hypothesis that there are signatures of gene expression associated with survival in EBOV-infected, anticoagulant-treated NHPs.

Immune response pathways play a key role in EBOV pathogenesis. Infection is characterized by up-regulation of inflammatory mediators such as cytokines and chemokines, interleukins, interferon-inducible proteins, and tumor necrosis factor alpha (TNFα) [20–25]. In addition, EBOV infection is associated with an early loss of lymphocytes [26,27] and the dysregulation of coagulopathy. This dysregulation of coagulation and subsequent hemorrhage are characteristic of EBOV infection [14,15], and may be due to the fact that immune mediators are over-expressed by monocytes and macrophages, which, along with

Author Summary

Infection of humans and non-human primates (NHPs) with Ebola virus (EBOV) can cause viral hemorrhagic fever, an acute systemic illness which can lead to death. The high case fatality rates (25%–90%) make EBOV a virus of significant concern from a biodefense perspective. To date, there are no FDA-approved post-exposure treatments for human use, and there are no standard assays to predict how infected individuals will fare after becoming infected. We have analyzed how circulating immune cells respond to EBOV infection under conditions where NHPs either survive viral infection, or succumb to it. This analysis identified genes that are correlated with, and predictive of, survival following lethal EBOV infection in NHPs. Our results demonstrate that small gene sets and transcriptional regulatory networks can be used to identify individual markers associated with survival following EBOV infection.

dendritic cells, are primary targets for infection [13,28,29]. High-throughput microarray studies of the immune response of NHPs to EBOV infection have identified dramatic and early changes in host transcription of genes related to interferon response, cytokine signaling, and apoptosis [22]. Similarly, studies of endothelial cells suggest that accumulation of cytokines and other pro-inflammatory factors can contribute to the observed pathologies of EBOV infection [14]. Previous studies of clinical samples from humans infected with Sudan virus during the 2000–2001 outbreak, suggest that hemorrhagic symptoms and death may be associated with acute phase proteins and coagulation factors [25].

It has been demonstrated that anticoagulant therapeutics have a positive effect on the outcome of EBOV disease [7,9]. Notably, approximately 33% of EBOV-infected NHPs treated with anticoagulants, such as recombinant nematode anticoagulant protein c2 (rNAPc2) [7], and recombinant human activated protein C (rhAPC) [9] survived a 100% fatal EBOV infection model (Table 1). Animals responding to anticoagulant treatment had lower plasma viremia levels and attenuation of the pro-inflammatory and pro-coagulant responses in both studies [7,9], suggesting that these indicators could be markers of increased survival. However, these late-stage markers could not identify whether there were early transcriptional changes that were associated with survival. In addition, these results are limited in scope to individual gene and protein assays, which cannot assess the host immune response to infection from a global, transcriptional viewpoint [7,9]. To identify early-stage transcriptional changes, we analyzed an existing microarray dataset that examined the host gene expression in anticoagulant-treated NHPs. We used this dataset to identify critical transcriptional changes that differentiate between surviving and non-surviving NHPs following EBOV infection. These results provide critically distinct assessments of the host immune response to EBOV infection by identifying not only global changes in transcription, but also transcription factor activities associated with survival.

We analyzed the transcriptional profiles of peripheral blood mononuclear cell samples taken from EBOV-infected NHPs treated with either rNAPc2 or rhAPC, as described previously [7,9,24]. We investigated the hypothesis that gene expression patterns are associated with survival of EBOV challenge following anticoagulant treatment. A previous study assessed the global transcriptional response of NHPs to EBOV infection, but was unable to identify survival-associated profiles, due to a lack of anticoagulant treatment [22]. Previously, we assessed the global transcriptional response in the context of anticoagulant treatment, but did not seek to identify upstream transcriptional regulators associated with survival [24]. Here, we used a new method of analysis to identify gene sets which are associated with, and predictive of, survival following post-infection treatment. This approach identified two sets of statistically significant genes that distinguish between surviving and non-surviving NHPs. One, a minimal set of 20 genes, showed good discrimination between survivors and non-survivors, but provided little insight into signaling pathways that might be correlated with survival. The second, a larger set of 238 genes, identified a number of genes that were functionally controlled by common transcription factors that have not been previously associated with EBOV infection in NHPs.

Materials and Methods

Animal experiments and ethics statement

The datasets for this study were collected from previously published results studying the affect of anticoagulant therapeutics (rhAPC and rNAPc2) on the immune response of non-human primates (NHPs) to lethal Ebola virus (H.sapiens-tc/COD/1995/Kikwit-9510621; EBOV). Virus was passaged twice through VeroE6, and once through Vero cells, prior to use. Additional experimental methods and results of the rhAPC and rNAPc2 studies have previously been described in Geisbert *et al.* [7] and Hensley *et al.* [9], respectively (Table 1). The cumulative dataset used in this study includes 23 rhesus macaques: 4 untreated controls, 8 rNAPc2-treated NHPs, and 11 rhAPC-treated NHPs (Table 2). A total of 4 NHPs survived lethal challenge with EBOV virus (2 rNAPc2-treated and 2 rhAPC-treated).

Animal research for these previously published studies was conducted at the United States Army Medical Research Institute for Infectious Diseases, in compliance with the Animal Welfare Act and other federal statutes and regulations relating to animals and experiments involving animals, and adheres to the principles stated in the *Guide for the Care and Use of Laboratory Animals*, National Research Council, 1996. The facility is fully accredited by the Association for Assessment and Accreditation of Laboratory Animal Care, International. These experiments and procedures were approved by the USAMRIID Institutional Animal Care and Use Committee (IACUC).

Sample preparation and microarray processing

Experimental methods for RNA processing and DNA microarray preparation have been previously described in Yen *et al.* [24]. Microarrays were analyzed in R, using the LIMMA package in Bioconductor [30–32] and processed as follows: (i) background correction was done using the *subtract* method [33]; (ii) within-array normalization was done using the *loess* method [34]; (iii) log ratio and log intensity values were calculated; (iv) array control probes were removed from the dataset; and, (v) the data were zero-transformed within each animal using baseline (pre-infection) sample; in the case of multiple pre-infection samples, Day 0 was used (Figure S1). The raw microarray dataset was deposited in NCBI's Gene Expression Omnibus (GEO; [35]) database (Accession: GSE24943; [24]).

We organized the microarrays into three groups based on NHP response to anticoagulant drug treatment: (i) EBOV-infected NHPs that did not receive anticoagulant treatment ("EBOV Only"; EO); (ii) anticoagulant-treated NHPs that survived EBOV infection ("EBOV infected, Treated Survivors"; ETS); and, (iii) anticoagulant-treated NHPs that did not respond to treatment and did not survive EBOV infection ("EBOV infected, Treated Non-

Table 1. Summary of methods and results from previous publications.

	Dataset 1		Dataset 2	
Study	EBOV infected NHPs, treated with rNAPc2 [7]		EBOV infected NHPs, treated with rhAPC (Xigris) [9]	
Survival	33% (3/9 NHPs receiving rNAPc2). Mean time to death: 8.3 days (EBOV only); 11.7 days (EBOV+rNAPc2).		18% (2/11 NHPs receiving rhAPC). Mean time to death: 8.3 days (EBOV only); 12.6 days (EBOV+rhAPC).	
Number of Samples Used in This Study*				
	Survivor−rNAPc2	–	Survivor−rhAPC	–
	Survivor+rNAPc2	2	Survivor+rhAPC	2
	Non-Survivor−rNAPc2	2	Non-Survivor−rhAPC	2
	Non-Survivor+rNAPc2	6	Non-Survivor+rhAPC	9

This table briefly describes the infectious agent, treatment, methods, and results previously been described in Geisbert *et al.* [7] and Hensley *et al.* [9].
EBOV: Zaire Ebolavirus. NHP: non-human primate. rNAPc2: recombinant nematode anticoagulant protein c2. rhAPC: recombinant human activated protein C.
* Remaining animals did not have samples available for cDNA microarray analysis.
doi:10.1371/journal.pntd.0003061.t001

Survivors"; ETNS), which were characterized by a mean time to death indistinguishable from untreated NHPs, *i.e.* animals died prior to Day 10 post-infection. A fourth group, characterized by treated, non-surviving NHPs with a mean time to death greater than the untreated controls, was excluded because any results would have been uninformative with regard to survival or treatment-specific transcriptional signatures. We limited our microarrays to Days 3 and 6 post-infection, because these timepoints were available for all treatment groups. A total of 23 arrays were included in the comparison: 4 arrays for each treatment group on both Day 3 and Day 6, except for the EO group on Day 3, which only had 3 samples (Table 2).

Identification of a minimal survival-associated gene set

To identify the minimal number of genes which distinguish survivors from non-survivors, we grouped the "EBOV Only" and "EBOV infected, Treated Non-Survivors" groups together into one cumulative "Non-Survivor" (NS) group. We compared survivors against non-survivors on Day 3, Day 6, and Days 3 and 6 together (Figure S1). Gene expression was averaged within each treatment group for individual probes, and the difference in mean expression (ΔM) for individual probes was calculated as follows:

$$\Delta M = \left|\overline{ETS} - \overline{NS}\right|$$

where $\overline{ETS}$ is the mean expression in survivors (ETS), and $\overline{NS}$ is the mean expression in non-survivors (NS). The probe (and its corresponding gene) is considered biologically relevant if, in at least one of the three comparisons, it meets the following criteria: (i) statistical significance (Student's t-test, unequal variance; $P \leq 0.05$); (ii) large magnitude ($\Delta M \geq 2$); and (iii) the differences in individual values across a treatment group are in agreement with the difference in means (within-group agreement).

Identification of general survival-associated transcriptional profile

To determine whether there is a general transcriptional profile associated with survival, we analyzed Day 3 and 6 arrays in EBOV-infected survivors and compared them against NHPs that did not survive infection using six comparisons: the ETS group to the ETNS group, and the ETS group to the EO group, at three time points: Day 3, Day 6, and Days 3 and 6 together (Figure S1). Gene expression was averaged and the difference in mean expression (ΔM) for individual probes is calculated as follows:

$$\Delta M = \left|\overline{ETS} - \bar{X}\right|$$

where $\overline{ETS}$ is the mean expression in survivors (ETS), and $\bar{X}$ is the mean expression in either of the two non-survivor groups (ETNS or EO). The probe (and its corresponding gene) was considered biologically relevant if, in at least one of the six comparisons, it met the previously described criteria.

Statistical analysis and validation

The correlation between identified genes and survival was evaluated for statistical accuracy using a permutation test of hierarchical clustering. Expression values for each gene were randomly permutated among arrays (1000 trials); for each trial, complete linkage hierarchical clustering was used to evaluate whether the gene list separated NHPs which survived infection from NHPs that did not survive infection (Pearson correlation coefficient). Clustering was completed using PyCluster [36] and Cluster 3.0 [37]. Heatmaps and dendrograms were generated using Java TreeView [38], and networks were rendered in Cytoscape [39]. The classification ability of each gene set was evaluated using leave-one out cross-validation and receiver operating characteristic (ROC) curves. Classification was evaluated by comparing the sum of normalized distances of the left-out sample to the mean of the survivor and non-survivor groups, for each gene in the gene set. The area under the curve (AUC) was used to evaluate performance of each gene set.

Biologically and statistically relevant genes were probed for functional annotation, pathways, and upstream transcriptional regulators using the Ingenuity Pathway Analysis suite of software (IPA; Ingenuity Systems). We confirmed the expression profile of a subset of our gene set using two comparisons: (1) an independently derived microarray dataset which examined changes in gene expression in EBOV-infected NHPs without anticoagulant treatment; and, (2) reverse transcription-PCR. We examined expression over the course of infection using the following equation:

$$\Delta M_{MA,VDS} = \left|\left(\overline{MA_X}\right) - \left(\overline{VDS_X}\right)\right|$$

Where $\overline{MA_X}$ is the mean expression in our dataset on Day 3 or Day 6, and $\overline{VDS_X}$ is the mean expression in the "Validation Dataset", either the second microarray dataset or the RT-PCR, on Day 3 or Days 5/6. We consider "complete agreement" to be a case where the direction of expression in our dataset and the

validation dataset are the same (*e.g.* both up-regulated); "minor disagreement" is a case where the direction of expression is opposing in the microarray and validation datasets, but is within 1 $\log_2$ fold change of difference, and therefore not significantly different; and, "major disagreement" is a case where the direction of expression in the microarray and validation datasets are opposing and significant in magnitude (>1). We evaluated whether the RT-PCR results reflect similar trends in expression, compared to the microarray data, by calculating the percentage of "complete agreement" or "minor disagreement" cases.

We compared the 245 probes from our gene set to the second microarray, for EBOV-infected NHPs that did not receive anticoagulant treatment. Of the 245 probes in our gene set, only 182 were available on the second microarray for confirmation (74.3%). When comparing the two microarray datasets, an average of 132.5 probes (72.8%) were in complete agreement or minor disagreement regarding the changes in expression from baseline to Day 3 or Days 5/6. Of the remaining probes, the majority were cases in which one microarray dataset was not differentially expressed but the other was, or vice versa; there were only 8 cases of significant and opposing disagreement between the two microarray datasets (4.4%).

We also confirmed a subset of our gene set using RT-PCR. Of the 245 probes of interest, we tested 56 genes that were associated with the upstream transcriptional regulators we identified previously; of these 56 genes, 45 passed the quality testing. For EBOV-infected NHPs that did not receive anticoagulant treatment ("EBOV Only"), an average of 34.5 probes (76.7%) were either in complete agreement or minor disagreement with respect to expression on Days 3 and 6. For EBOV-infected, anticoagulant-treated NHPs that did not survive ("EBOV infected, Treated Non-Survivors"), an average of 34 (75.6%) probes were either in complete agreement or minor disagreement. Finally, for EBOV-infected, anticoagulant-treated NHPs that survived infection ("EBOV infected, Treated Survivors"), an average of 37.8 (83.9%) probes were either in complete agreement or minor disagreement between the microarray dataset and RT-PCR.

Results

Identification of a minimal survival-associated gene set in EBOV-infected, anticoagulant-treated NHPs

Previous studies have reported that NHPs exhibit strong transcriptional changes in response to EBOV infection [22,24]. We were interested in determining if these transcriptional changes were altered by anticoagulant treatment; specifically, we were interested to determine if there were any transcriptional correlates associated with survival of EBOV infection.

To determine the minimal number of genes which were associated with survival, we analyzed changes in mRNA expression in EBOV-infected NHPs with and without anticoagulant treatment. Previously published data was used to build a cumulative group of 23 EBOV-infected NHPs ([7,9]; Table 1): 4 EBOV-infected controls that did not receive anticoagulant treatment, 8 EBOV-infected NHPs that were treated with rNAPC2, and 11 EBOV-infected NHPs that were treated with rhAPC (Table 2). We organized the samples into three groups: (i) EBOV-infected, untreated non-survivors ("EBOV Only"; EO); (ii) EBOV-infected, anticoagulant-treated survivors ("EBOV-infected, Treated Survivors"; ETS); and, (iii) EBOV-infected, anticoagulant-treated, non-survivors ("EBOV-infected, Treated Non-Survivors"; ETNS), which were characterized by a mean time to death indistinguishable from untreated NHPs. In order to identify a set of probes which were associated with survival and not just

Table 2. Summary of samples used in this study.

Group	EBOV Only				EBOV-infected, Treated Non-Survivor					EBOV-infected, Treated Survivor			
Day/Sample	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**
−8			•	•		•		•	•	•			•
0	•	•		•	•	•	•			•	•	•	
3	•	•	•		•		•	•	•	•	•	•	•
6	•	•	•	•	•	•	•	•		•	•	•	•
rNAPc2	–	–	–	–	+	–	–	–	–	+	+	–	–
rhAPC	–	–	–	–	–	+	+	+	+	–	–	+	+
Designation	CH78	CH52	DB7U	GTX	AXX	L201-2	CH73	CL4R	94E136	AA016	CH04	DB1D	DB1L
Day of Death:	9	8	8	8	8	9	7	7	8	–	–	–	–

This table depicts the samples (dots) used in this study. Samples are grouped vertically based on treatment, then by individual animal. Samples are also grouped horizontally by day pre- and post-infection. Summary information includes applied treatment (if any), and date of death for each animal. Designations are from the original publications, Geisbert *et al.* [7] and Hensley *et al.* [9].
EBOV: Zaire Ebolavirus. rNAPc2: recombinant nematode anticoagulant protein c2. rhAPC: recombinant human activated protein C.
doi:10.1371/journal.pntd.0003061.t002

treatment responses, we compared the gene expression of NHPs that survived EBOV infection (ETS) to those that did not survive ("Non-Survivors", NS; a combination of the EO and ETNS groups) on Days 3 and 6 post-infection. Probes were considered of interest if the difference in expression between groups was statistically significant (Student's t-test, unequal variance; $P \leq 0.05$), with a large change in transcriptional magnitude ($\Delta M \geq 2$), and had within-group coherence (see *Materials and Methods*).

We identified a total of 20 unique probes which differentiated between the NHPs that survived EBOV infection (ETS) and those that did not (NS). These probes corresponded to 16 annotated genes, 3 genetic loci, and 1 microRNA (Figure 1). There are two obvious patterns of differentially expressed genes which exhibit significant and opposing regulation in the two groups (Figure 1A): (i) 6 genes have higher expression values in survivors compared to non-survivors (CLDN3, ILF2, ILF3, NDUFA12, RUVBL2, and SLC38A5); and, (ii) 10 genes have lower expression values in survivors compared to non-survivors (ACCN1, CEBPE, CRHR2, FAM63A, HMP19, IL2RA, LTF, PSMA1, RCHY1, and SLC9A7). Finally, the genetic loci (AC009283, LOC100289371, and LOC440871) and microRNA (miR-122) also appear to be down-regulated in survivors compared to non-survivors.

Due to the minimal size of the gene set, it is difficult to perform functional enrichment or pathway analysis; however, identification of individual gene traits and annotation is useful for elucidating how individual genes may be associated with viral infection. For example, over half of these genes have been previously associated with viral infection or replication, or have been found to physically interact with viral proteins (*e.g.* CLDN3 [40]; CRHR2 [41]; ILF2 [42,43]; ILF3 [44,45]; LTF [46–50]; miR-122 [51]; RCHY1 [52]; and, RUVBL2 [53]). Importantly, none of these genes have been linked to EBOV infection, with the exception of ILF3, of which an isoform called DRBP76 is known to bind EBOV protein VP35 [54]. Although miR-122 has not been associated with EBOV infection, it is a known positive regulator of the replication of hepatitis C virus, another RNA virus [55,56]. Additionally, several of the genes associated with survival encode for transport and membrane proteins (*e.g.* ACCN1, CLDN3, HMP19, NDUFA12, SLC9A7, and SLC38A5), suggesting that small-molecule transport proteins may play a role in survival of EBOV challenge. Several gene products are also associated with immune response or inflammatory response (*e.g.* CRHR2, IL2RA, LTF, and PSMA1), which is consistent with previously published studies investigating the effects of EBOV on gene expression [21,22,24].

Having identified a minimal survival-associated gene set, we were interested in determining whether it could accurately distinguish between survivors and non-survivors. Hierarchical clustering is a common way to use the characteristics of a gene set to determine if datasets are similar or dissimilar. We used hierarchical clustering to demonstrate that we could separate survivors (green) from non-survivors, using only this small gene set, with a high degree of significance (permutation test, one-tailed $P < 0.0001$; Figure 1B). In addition, we evaluated the cumulative ability of the gene set to distinguish survivors from non-survivors, based on individual gene traits, using leave-one-out cross-validation. This approach tests the robustness of a classifier, by using one sample as a testing set and the remaining samples as a training set, then classifying the left-out sample based on the training set. This procedure is repeated for all samples, and the robustness of a classifier is evaluated by a receiving operator characteristic (ROC) curve [57–59]. Leave-one-out cross-validation showed that our classifier correctly classified samples with 100% accuracy, i.e. no survivors were mistakenly classified as non-survivors, or vice versa (AUC = 1.00; Figure 1C).

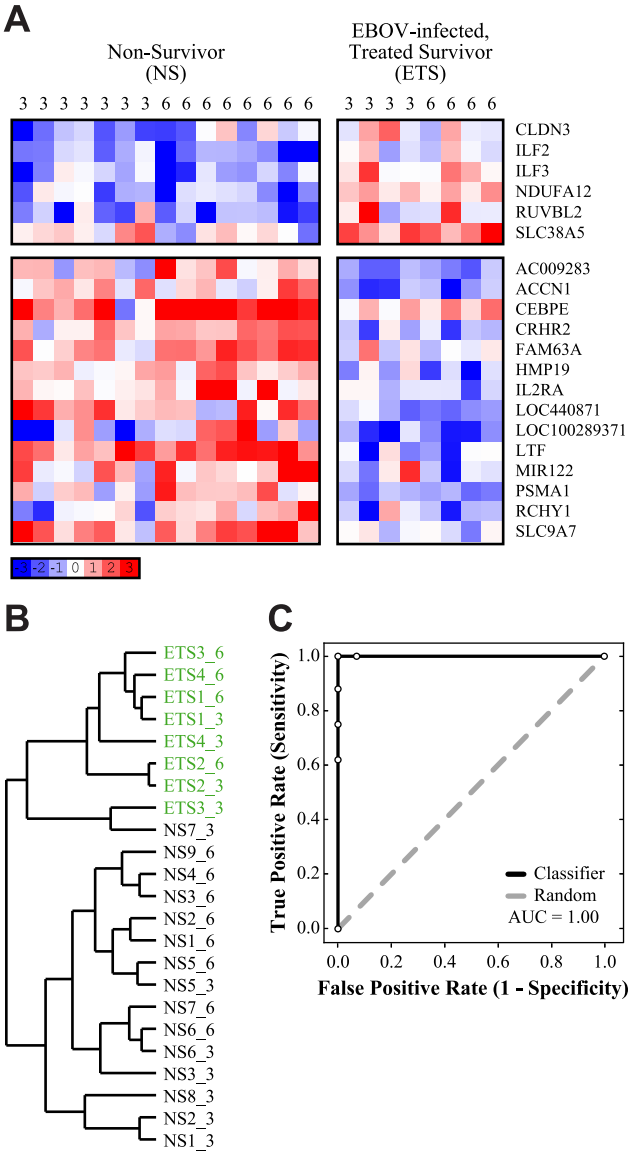

Figure 1. Identification of a small gene set is sufficient for classification of survival in EBOV-infected, anticoagulant-treated NHPs. (**A**) A heatmap depicting the gene expression values of 20 genes which were found to distinguish between "EBOV-infected, Treated Survivors" (ETS) and Non-Survivors ("EBOV Only" and "EBOV-infect, Treated Non-Survivors"; see *Materials and Methods*). The days post-infection are labeled at the top of each column. Blue indicates down-regulation, and red indicates up-regulation, of genes compared to pre-infection baseline, white indicates no change in expression (scale indicated). Gene names are listed to the right. (**B**) A dendrogram depicting hierarchical clustering of individual samples, using the set of 20 genes (permutation test, one-tailed $P<0.0001$). ETS samples are indicated in green; samples are labeled "groupX_Y", where X is the sample number and Y is the days post-infection. (**C**) A ROC curve depicting how accurately this gene set can classify survivors and non-survivors, based on leave-one-out cross-validation (AUC = 1.00). This gene set is indicated in black, and a random ROC curve is indicated in grey.
doi:10.1371/journal.pntd.0003061.g001

Identification of general survival-associated transcriptional profile

Although a minimal gene set can be useful for identifying individual genes associated with survival, we were also interested to know whether there was a general transcriptional profile

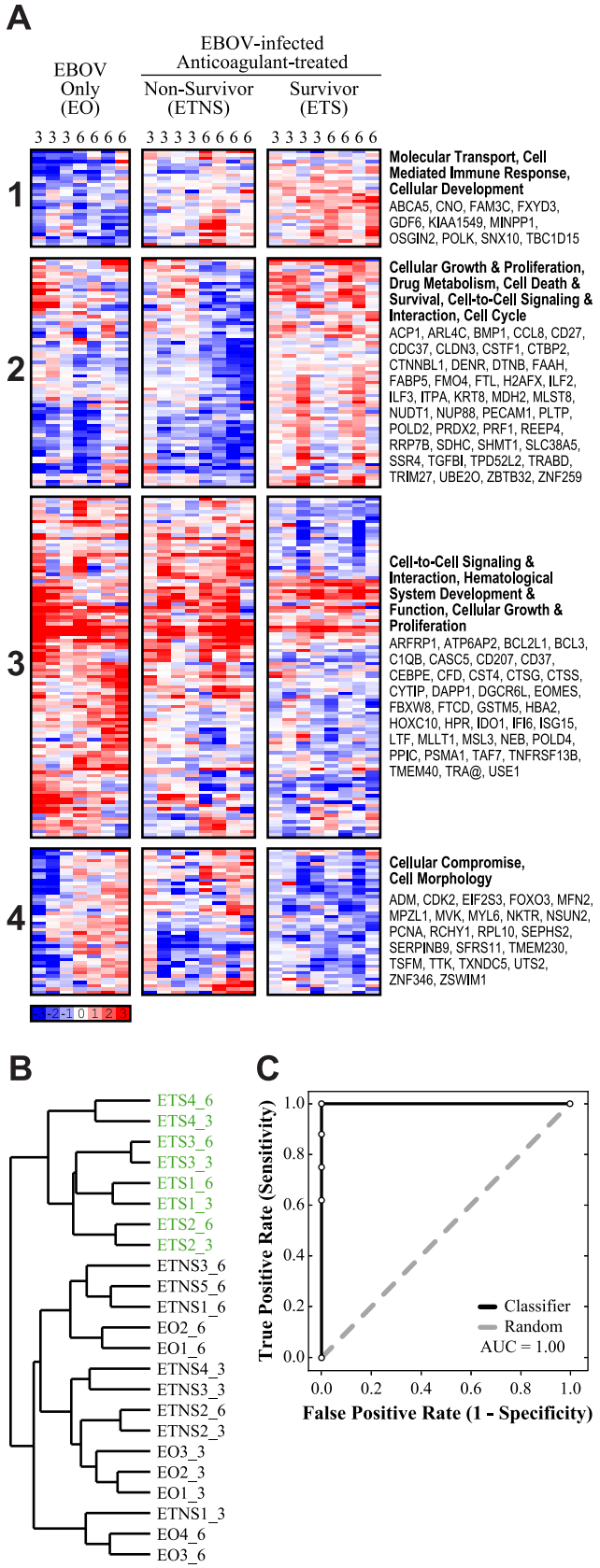

Figure 2. Survival following anticoagulant treatment is associated with a broad transcriptional profile. (**A**) A heatmap depicting the gene expression values of 245 probes (238 genes) which were found to distinguish between "EBOV-infected, Treated Survivors" (ETS), "EBOV-infect, Treated Non-Survivors", and NHPs that did not receive anticoagulant treatment ("EBOV Only"; see *Materials and Methods*). The days post-infection are labeled at the top of each column. Blue indicates down-regulation, and red indicates up-regulation, of genes compared to pre-infection baseline; white indicates no change in expression (scale indicated). Probes are grouped into 4 clusters, denoted by numbers to the left of the heatmap. Functional annotation of each cluster is listed in bold to the right of the heatmap, along with a list of representative genes associated with the annotation. (**B**) A dendrogram depicting hierarchical clustering of individual samples, using the gene set of 245 probes (permutation test, one-tailed $P<0.0001$). ETS samples are indicated in green; samples are labeled "groupX_Y", where X is the sample number and Y is the days post-infection. (**C**) A ROC curve depicting how accurately this gene set can classify survivors and non-survivors, based on leave-one-out cross-validation (AUC = 1.00). This gene set is indicated in black, and a random ROC curve is indicated in grey.
doi:10.1371/journal.pntd.0003061.g002

associated with survival. In particular, we were interested to know if NHPs that survived EBOV infection displayed coordinated enrichment of specific signaling pathways or transcriptional responses. To identify a broad, survival-associated transcriptional profile that was also associated with anticoagulant treatment, we compared the expression profiles of anticoagulant-treated NHPs that survived EBOV infection (ETS) to NHPs that did not receive anticoagulant treatment (EO) and to NHPs that received anticoagulant treatment but did not survive EBOV infection (ETNS). We compared survivors to the two non-survivor groups (EO or ETNS) on Day 3, Day 6, and Days 3 and 6 together; probes were chosen according to the previously described criteria (see *Materials and Methods*).

Using this approach, we identified a total of 245 unique probes, corresponding to 238 annotated genes, which accurately differentiated the three groups (Figure 2). Hierarchical clustering of the 245 probes showed that the probes clustered into four gene expression patterns: (1) down-regulation in the EO group on both days, compared with up-regulation in survivors on both days; (2) down-regulation in the ETNS group on Day 6, compared with up-regulation in survivors on both days; (3) up-regulation in the non-survivor groups on both days, compared with down-regulation in survivors on both days; and, (4) up-regulation in the non-survivor groups on Day 6, compared with down-regulation in survivors on both days (Figure 2A). In general, the overall pattern of these separate clusters is one of significant and opposing regulation of expression when comparing NHPs that survived EBOV infection to NHPs that did not survive infection, regardless of treatment.

Analysis of the overall gene set using Ingenuity Pathway Analysis (IPA; Ingenuity Systems) identified several statistically significant cellular and molecular functions, including cell death and survival, cellular growth and proliferation, infectious disease, cell cycle, and immune cell trafficking. In addition, we probed the 4 gene clusters to determine if there were any cluster-specific pathways or networks that were functionally enriched. We found that the four main clusters were associated with: (1) molecular transport, cell-mediated immune response, and cell development; (2) cellular growth and proliferation, drug metabolism, cell death and survival, cell-to-cell signaling and interaction, and cell cycle; (3) cell-to-cell signaling and interaction, hematological system development and function, and cellular growth and proliferation; and, (4) cellular compromise, and cell morphology (Figure 2A). In particular, we observed that many of the genes in Cluster 2 were associated with cell metabolism, suggesting that this process is down-regulated in non-survivors when compared to survivors.

Cluster 3 is also heavily associated with immune and inflammatory responses, and several genes are expressed predominantly by immune cells. These functional annotations are consistent with previous studies assessing host immune response to EBOV infection using microarrays [22,24]. For example, we observed up-regulation of genes associated with innate immune response, regulation of cytokine and chemokine production, regulation of apoptosis, and interferon response, as reported in previous studies [22,24]. The observed regulation of these genes is consistent with the hypothesis that non-survivors experience severe dysregulation of the immune response, as opposed to survivors, which maintain a normal level of expression [22,24].

To evaluate the classification of our gene set, we hierarchically clustered individual arrays ([37]; *complete linkage* method; Figure 2B). The resulting dendrogram has two major branches, in which NHPs that survive EBOV infection (green) cluster separately from those that did not (black; Figure 2B), indicating that the expression profiles of survivors is distinguishable from the expression profile of non-survivors. Correspondingly, this gene set allowed us to accurately distinguish between the survivors and the combined non-survivor groups (permutation test, one-tailed $P<0.0001$); however, we found no significant difference in the expression profiles when comparing the two non-survivor groups. Additionally, this gene set is capable of perfectly classifying survivor and non-survivor groups (Figure 2C). As in Figure 1, we used leave-one-out cross-validation to evaluate whether survivors could be distinguished from non-survivors. We found that cross-validation was able to correctly classify samples with 100% accuracy, i.e. no survivors were mistakenly classified as non-survivors, or vice versa (AUC = 1.00; Figure 2C). These two tests confirm that the expression profiles of survivors are distinguishable from the expression profile of non-survivors; however, the expression profile of untreated non-survivors (EO) is indistinguishable from that of treated non-survivors (ETNS). Finally, we compared this list of 238 genes against the previously identified list of 20 probes which were highly predictive for distinguishing between survivors and non-survivors. Of the 20 probes, we found that 17 (85%) also appeared in the set of 238 genes. These results confirm that the 238 genes are predictive for distinguishing between treatment groups and survival outcomes, and that the predictive power is comparable to the list of 20 highly predictive genes.

Survival-associated genes are transcriptionally interconnected

To further investigate how these survival-associated genes are functionally associated, we determined if the genes were transcriptionally related, *e.g.* by being participants in a major pathway or signaling network, or by having a common upstream transcriptional regulator. We used IPA to probe the set of 238 genes in order to identify transcriptional regulators that were common to a large number of genes in our dataset.

This analysis identified four transcription factors whose downstream targets were statistically over-represented in our network: CCAAT/enhancer-binding protein alpha (CEBPA; Fisher's exact test, $P<0.0001$), tumor protein 53 (p53; $P<0.001$), megakaryoblastic leukemia 1 (MKL1; $P<0.0001$) and myocardin-like protein 2 (MKL2; $P<0.0001$). The probability of finding at least as many up-stream transcriptional regulators in a gene set of 245 probes is extremely unlikely, suggesting that this result is not due to random chance (re-sampling, $P\approx0.01$). The transcription factors and their targets are shown in Figure 3, which shows genes colored according to the difference in mean expression between NHPs that survived EBOV infection and those that did not. This network of transcription factors and their downstream targets is highly interconnected, with several transcription factors sharing targets; in particular, MKL1 and MKL2 appear to co-regulate all their targets (Figure 3). As would be expected for transcription factors that are post-translationally activated, the four transcription factors were not themselves significantly differentially regulated in our dataset. For each transcription factor, we compared the expression of the downstream targets in NHPs which survived infection to NHPs that did not survive infection, to determine if downstream targets correlated with survival.

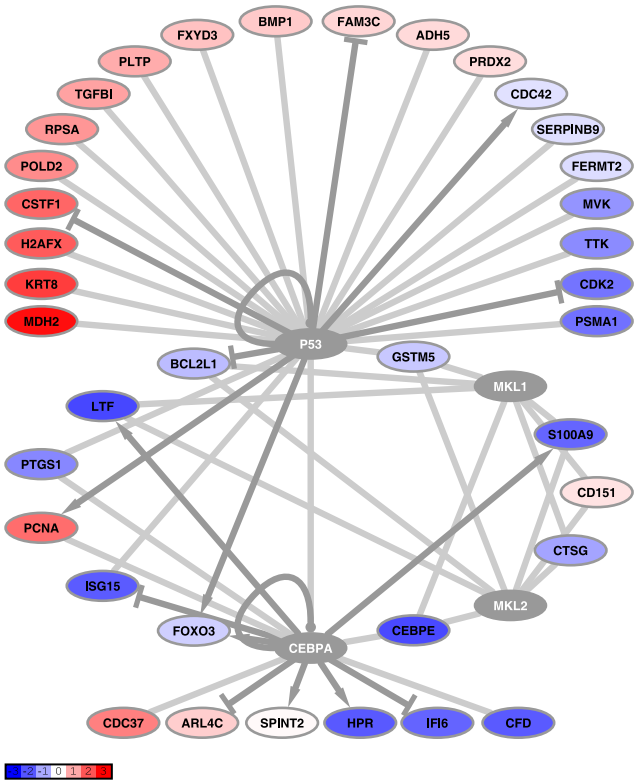

Figure 3. Survival-associated genes are transcriptionally interconnected. A network illustrating transcriptional connections between 37 genes which distinguish between "EBOV-infected, Treated Survivors" (ETS), "EBOV-infect, Treated Non-Survivors" (ETNS), and NHPs that did not receive anticoagulant treatment ("EBOV Only"; see *Materials and Methods*). Nodes are labeled with the gene name, edges indicate regulation of expression from transcription factor to target gene; arrows indicate up-regulation, bars indicate down-regulation, circles indicate up- and down-regulation, and unmarked, light grey edges indicate unknown regulation. Nodes are colored according to the difference in mean expression between the ETS and "Non-Survivor" (NS; ETNS and EO) groups; blue indicates lower expression values in the ETS group compared to the NS group, red indicates higher expression values in the ETS group compared to the NS group (scale indicated). Nodes that are colored grey (*e.g.* p53) are master transcriptional regulators which were not identified as being differentially expressed by our protocol.
doi:10.1371/journal.pntd.0003061.g003

CCAAT/enhancer binding protein alpha (CEBPA)-regulated genes are down-regulated in NHPs that survive EBOV infection when compared to non-survivors

Of the 238 genes associated with a transcriptional pattern correlated with survival, 13 genes are transcriptionally regulated by CEBPA (Figure 4). Of these, 2 probes (CEBPE, and IFI6) distinguished NHPs that survived infection ("EBOV-infected,

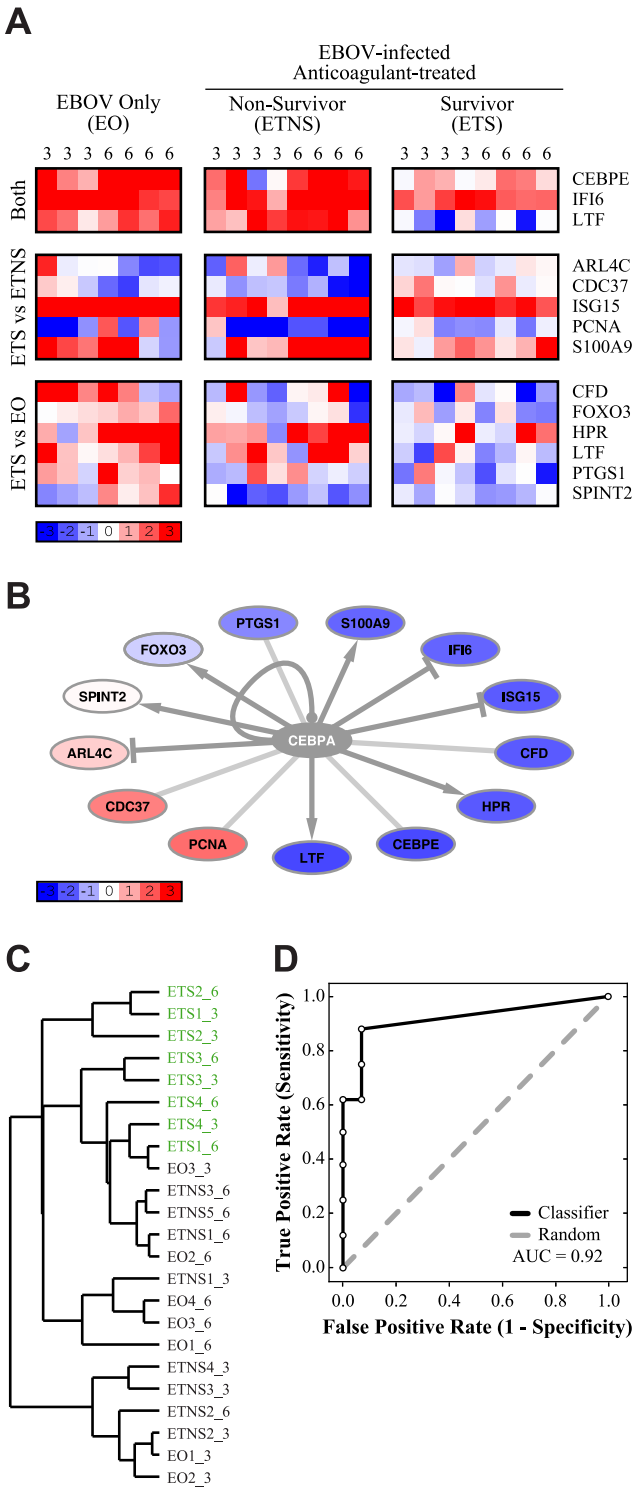

Figure 4. CCAAT/enhancer binding protein alpha (CEBPA)-regulated genes are down-regulated in NHPs that survive EBOV infection when compared to non-survivors. (**A**) A heatmap depicting the gene expression values of 14 probes (13 genes) which are transcriptionally regulated by CEBPA. The days post-infection are labeled at the top of each column. Probes are grouped according to which comparison identified them, listed to the left. Gene names are listed to the right. Blue indicates down-regulation, and red indicates up-regulation, of genes compared to pre-infection baseline, white indicates no change in expression (scale indicated). (**B**) A network illustrating transcriptional connections between the 14 genes and CEBPA. Blue indicates lower expression values in "EBOV-infected, Treated Survivors" (ETS), compared to "Non-Survivors" (NS), red indicates higher expression values in the ETS group compared to the NS group (scale indicated). Grey nodes (*e.g.* CEBPA) were not identified as being differentially expressed by our protocol. (**C**) A dendrogram depicting hierarchical clustering of individual samples, using the gene set of 14 probes. ETS samples are indicated in green; samples are labeled "groupX_Y", where X is the sample number and Y is the days post-infection. (**D**) A ROC curve depicting how accurately this gene set can classify survivors and non-survivors, based on leave-one-out cross-validation (AUC = 0.92). This gene set is indicated in black, and a random ROC curve is indicated in grey.
doi:10.1371/journal.pntd.0003061.g004

Treated Survivors"; ETS) from those that did not ("Non-Survivors"; NS). Both probes had lower expression values in the ETS group. Five probes distinguished between survivors and "EBOV-infected, Treated Non-Survivors" (ETNS); 3 of these probes had higher expression values in survivors (ARL4C, CDC37, and PCNA), and 2 had lower expression values in survivors (ISG15 and S100A9). Six probes distinguished between survivors and the "EBOV Only" (EO) group, of which 4 had lower expression values in survivors (CFD, FOXO3, HPR, and PTGS1). One probe, SPINT2, distinguished between survivors and the EO group only on Day 6. One gene, LTF, was identified by two non-identical probes; one probe distinguished survivors from both non-survivor groups, whereas the other probe only distinguished between survivors and the "EBOV Only" group. In both cases, the probe had lower expression values in survivors (Figure 4A).

To confirm this expression pattern, we used RT-PCR to examine a subset of the CEBPA-regulated genes (data not shown). We examined 11 of the 13 genes by comparing the changes in expression on Days 3 and 6 in the microarray dataset to the RT-PCR dataset (see *Materials and Methods*). The RT-PCR dataset reflected the trends observed in the microarray dataset with above 60% accuracy in all three treatment groups (EO: 68.2%; ETNS: 77.3%; ETS: 88.6%). For CEBPA-regulated genes, the RT-PCR confirms the expression trends observed in the microarray dataset.

The observed expression patterns of the downstream targets are consistent with decreased transcriptional activity of CEBPA (Figure 4B), when compared to a previous study which identified transcriptional targets of CEBPA (GEO accession GSE2188, [60]). Although there is a clear pattern of expression that suggests an underlying biological mechanism (*i.e.* down-regulation of CEBPA), this set of probes alone is insufficient to statistically distinguish between NHPs that survived EBOV infection and those that did not. Hierarchical clustering of CEBPA target expression patterns in individual arrays reveals a dendrogram in which the survivor and non-survivor groups are not discrete (Figure 4C; ETS indicated in green). There is some similarity between the expression profiles of survivors (green) and non-survivors, especially treated non-survivors (ETNS), which appears to be driven by within-animal responses, and not by treatment or survival (Figure 4C). However, we found that leave-one-out cross-validation was able to correctly classify samples with high accuracy, i.e. very few survivors were mistakenly classified as non-survivors (AUC = 0.92; Figure 4D). This result demonstrates that, when considering the CEBPA signature as a classifier, survivors are distinguishable from non-survivors, but with less accuracy than the full gene sets identified in Figures 1 and 2.

Tumor protein 53 (p53) regulates transcription of some survival-associated genes

We identified 26 probes, corresponding to 26 genes, which are transcriptionally regulated by p53 (Figure 5A). Of these, 2 probes

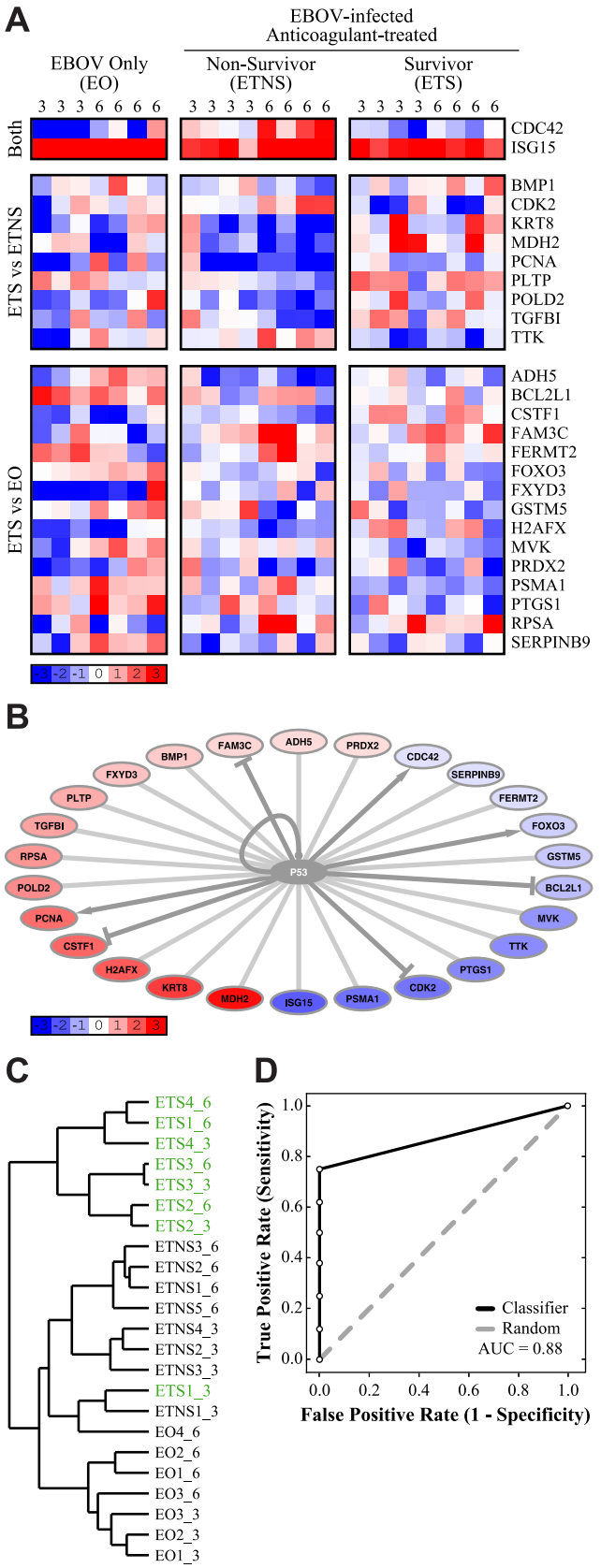

Figure 5. Tumor protein 53 (p53) regulates transcription of some survival-associated genes. (**A**) A heatmap depicting the gene expression values of 26 probes (26 genes) which are transcriptionally regulated by p53. The days post-infection are labeled at the top of each column. Probes are grouped according to which comparison identified them, listed to the left. Gene names are listed to the right. Blue indicates down-regulation, and red indicates up-regulation, of genes compared to pre-infection baseline, white indicates no change in expression (scale indicated). (**B**) A network illustrating transcriptional connections between the 26 genes and p53. Blue indicates lower expression values in "EBOV-infected, Treated Survivors" (ETS), compared to "Non-Survivors" (NS), red indicates higher expression values in the ETS group compared to the NS group (scale indicated). Grey nodes (*e.g.* p53) were not identified as being differentially expressed by our protocol. (**C**) A dendrogram depicting hierarchical clustering of individual samples, using the gene set of 26 probes (permutation test, one-tailed $P<0.0001$). ETS samples are indicated in green; samples are labeled "groupX_Y", where X is the sample number and Y is the days post-infection. (**D**) A ROC curve depicting how accurately this gene set can classify survivors and non-survivors, based on leave-one-out cross-validation (AUC = 0.88). This gene set is indicated in black, and a random ROC curve is indicated in grey.
doi:10.1371/journal.pntd.0003061.g005

(CDC42 and ISG15) distinguished NHPs that survived EBOV infection ("EBOV-infected, Treated Survivors"; ETS) from those that did not ("Non-Survivors"; NS); in addition, both probes had lower values in survivors. Nine probes distinguished between survivors and "EBOV-infected, Treated Non-Survivors" (ETNS), of which 7 had higher expression in survivors (BMP1, KRT8, MDH2, PCNA, PLTP, POLD2, and TGFBI), and 2 had lower expression values in survivors (CDK2 and TTK). Fifteen probes distinguished between survivors and the "EBOV Only" (EO) group, of which 7 had higher expression values in survivors (ADH5, CSTF1, FAM3C, FXYD3, H2AFX, PRDX2, and RPSA), and 8 had lower expression values in survivors (BCL2L1, FERMT2, FOXO3, GSTM5, MVK, PSMA1, PTGS1, and SERPINB9; Figure 5B). However, the observed expression patterns of these targets were not consistent with any clearly defined activation or repression of p53, suggesting that the transcription factor may be differently regulated in different cell types within the PBMC population.

To confirm this expression pattern, we used RT-PCR to examine a subset of the P53-regulated genes (data not shown). We examined 20 of the 26 genes (76.9%) by comparing the changes in expression on Days 3 and 6 in the microarray dataset to the RT-PCR dataset (see *Materials and Methods*). The three treatment groups had comparable levels of agreement between the RT-PCR dataset and the microarray dataset (EO: 70%; ETNS: 72.5%; ETS: 80%). For P53-regulated genes, the RT-PCR confirms the expression trends observed in the microarray dataset.

In a comparison to an independently derived dataset examining the effects of p53 dosage, we found that approximately 50% of our p53 targets had an expression pattern concordant with p53 activation, whereas the other 50% had discordant expression patterns which suggested inhibition or down-regulation of p53 (GEO accession GSE11547, Hosako *et al.*). Because a consensus pattern of expression could not be established, we cannot draw conclusions about the activity of p53 in NHPs that survived EBOV infection when compared to non-survivors. Despite this lack of consensus regarding the activity of p53 in different treatment groups, hierarchical clustering revealed that this set of probes is able to distinguish survivors and non-survivors with some accuracy (Figure 5C; ETS indicated in green; permutation test, one-tailed $P<0.0001$). The observed dendrogram has three major branches: (i) an ETS cluster, which is clearly separated from the other branches of the tree; (ii) an ETNS cluster with two misclassified arrays, which are EO and ETS samples; and, (iii) an EO cluster. This suggests that the gene set is sufficient to distinguish survivors

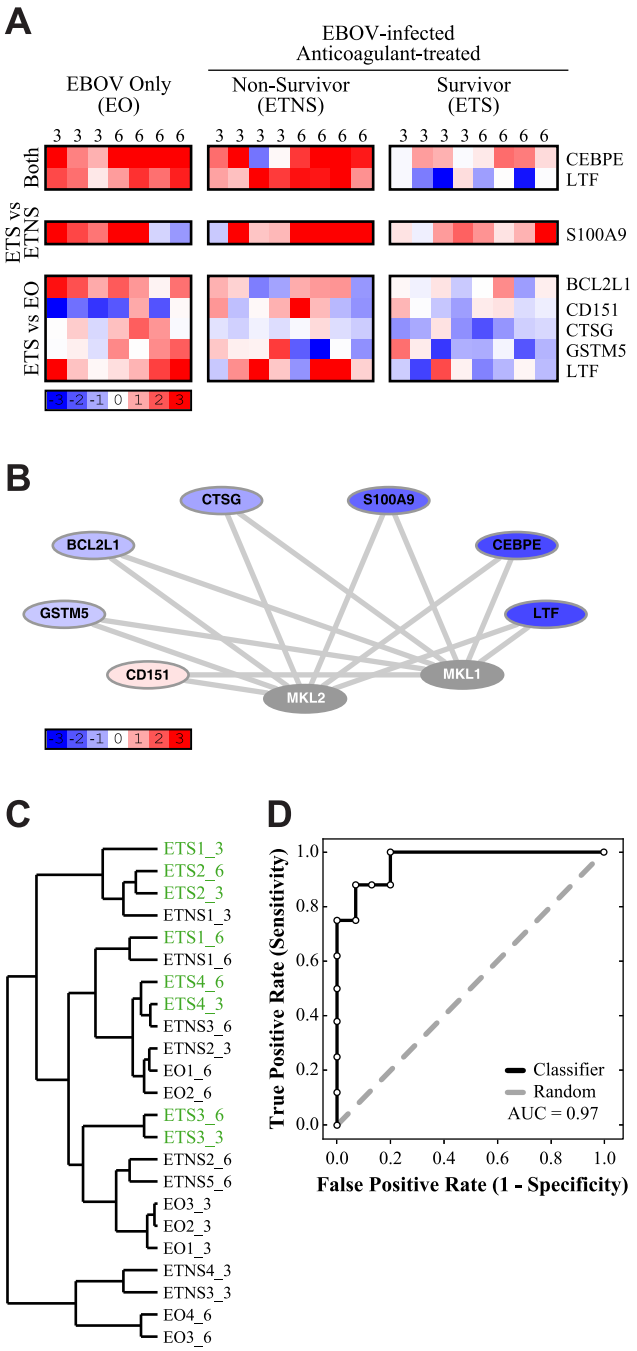

Figure 6. Megakaryoblastic leukemia 1 (MKL1) and myocardin-like protein 2 (MKL2) jointly regulate some survival-associated genes. (**A**) A heatmap depicting the gene expression values of 8 probes (7 genes) which are jointly transcriptionally regulated by MKL1 and MKL2. The days post-infection are labeled at the top of each column. Probes are grouped according to which comparison identified them, listed to the left. Gene names are listed to the right. Blue indicates down-regulation, and red indicates up-regulation, of genes compared to pre-infection baseline, white indicates no change in expression (scale indicated). (**B**) A network illustrating transcriptional connections between the 7 genes and MKL1/2. Blue indicates lower expression values in "EBOV-infected, Treated Survivors" (ETS), compared to "Non-Survivors" (NS), red indicates higher expression values in the ETS group compared to the NS group (scale indicated). Grey nodes (*e.g.* MKL1) were not identified as being differentially expressed by our protocol. (**C**) A dendrogram depicting hierarchical clustering of individual samples, using the gene set of 7 probes. ETS samples are indicated in green; samples are labeled "groupX_Y", where X is the sample number and Y is the days post-infection. (**D**) A ROC curve depicting how accurately this gene set can classify survivors and non-survivors, based on leave-one-out cross-validation (AUC = 0.97). This gene set is indicated in black, and a random ROC curve is indicated in grey.
doi:10.1371/journal.pntd.0003061.g006

from non-survivors, and secondarily can also distinguish the two non-survivor groups from one another. The case of a survivor array being misclassified as a non-survivor array is understandable, given that the array sample is from Day 3, when gene expression differences are relatively small between all different treatment groups. We found that leave-one-out cross-validation was able to correctly classify samples with some accuracy (AUC = 0.88; Figure 5D), although the classification was not as strong as the full dataset.

Megakaryoblastic leukemia 1 (MKL1) and myocardin-like protein 2 (MKL2) jointly regulate some survival-associated genes

Finally, we identified 8 probes, corresponding to 7 genes, whose transcription is jointly regulated by MKL1 and MKL2 (Figure 6). Of these, 2 probes (CEBPE and LTF) distinguished NHPs that survived EBOV infection ("EBOV-infected, Treated Survivors"; ETS) from those that did not ("Non-Survivors"; NS). Notably, CEBPE is also co-regulated by another member in the CCAAT-enhancer binding protein family, CEBPA; therefore, it is included in both gene sets. In addition, 1 probe (S100A9) distinguished between survivors and "EBOV-infected, Treated, Non-Survivors" (ETNS), and 5 probes distinguished between survivors and the "EBOV Only" group (EO; BCL2L1, CD151, CTSG, GSTM5, and LTF). All genes had lower expression values in the survivors than non-survivors (Figure 6A), with the exception of CD151, which exhibited significant down-regulation in the EO group. Importantly, MKL1 and MKL2 appear to co-regulate the full set of genes (Figure 6B). Notably, the majority of the genes regulated by MKL1 and MKL2 are also regulated by CEBPA (CEBPE, LTF and S100A9) or p53 (BCL2L1 and GSTM5), suggesting that the regulatory pattern exhibited by these genes could be confounded by regulatory activity from additional transcription factors.

We confirmed the expression pattern of the 7 genes regulated by MKL1 and MKL2 by RT-PCR (data not shown). We examined all genes co-regulated by these transcription factors by comparing the changes in expression on Days 3 and 6 in the microarray dataset to the RT-PCR dataset (see *Materials and Methods*). For both non-survivor groups, 71.4% of genes exhibited the same expression trends in both the microarray dataset and the RT-PCR dataset. For survivors, 85.7% of genes exhibited the same expression trends in both the microarray dataset and the RT-PCR dataset. For MKL1 and MKL2-regulated genes, the RT-PCR confirms the expression trends observed in the microarray dataset.

A previous study found that a double-knockout of MKL1 and MKL2 increases expression of CEBPE, CTSG, GSTM5, LTF and S100A9, suggesting that MKL1 and MKL2 may jointly act as repressors for these genes under natural conditions [61]. In contrast, a double-knockout of MKL1 and MKL2 decreases expression of BCL2L1 and CD151, suggesting that MKL1 and MKL2 activate transcription of these genes [61]. The expression pattern observed with this gene set suggests that MKL1 and MKL2 may be up-regulated or activated in NHPs that survive EBOV infection, compared to those that do not. Although this

observed expression pattern is consistent with what would be observed if MKL1 and MKL2 were activated, this set of probes is too small to distinguish between survivors and non-survivors groups (Figure 6C; ETS indicated in green). A hierarchically-clustered dendrogram has two major branches, which are each interspersed with EO, ETNS and ETS samples; this suggests that the overall expression profile of survivors and non-survivors is not statistically distinguishable. Although the dendrogram suggests that the overall expression profile in each array is incapable of distinguishing survivors and non-survivors, evaluation of individual gene contributions using leave-one-out cross-validation shows that this gene set is capable of classifying samples with high accuracy (AUC = 0.97; Figure 6D).

Discussion

These results show the potential of high-throughput transcriptional studies for identifying putative markers of survival following EBOV infection. In particular, we identified a minimal survival-associated gene set that accurately distinguished survival outcome following post-infection anticoagulant treatment of non-human primates (NHPs) infected with EBOV. We identified 20 genes that were characterized by significant, coherent and opposing expression patterns when comparing survivors and non-survivors. Several of these genes exhibit differential regulation as early as 3 days post-infection, prior to the appearance of clinical symptoms of EBOV infection; this early differential regulation is especially important for the identification of early-stage biomarkers to distinguish disease outcomes.

Importantly, several of these genes are associated with different viral infections [46–50]. Proteins such as ILF3 and RUVBL2 are known to suppress viral replication in other viruses [44,45,53], and we observe that their expression is higher in survivors than non-survivors. Notably, an isoform of ILF3 is known to bind EBOV protein VP35, suppressing the function of the viral polymerase [54]. This suggests a mechanism of action in which survivors may up-regulate the transcription of certain genes, *e.g.* ILF3, in order to suppress viral replication.

We also observed that microRNA 122 (miR-122) is down-regulated in survivors compared to non-survivors, suggesting that inhibition of miR-122 activity increases survival following EBOV infection. To date, there have been no studies investigating whether miR-122 interacts with the EBOV genome, but it is well-documented that miR-122 binds the Hepatitis C virus genome to support the replication of this virus [55,56]. Comparison of the putative binding motifs of miR-122 [55] to the consensus sequence of EBOV (Mayinga, Zaire, 1976) [62] reveals multiple potential binding sites in the viral genome (data not shown). This suggests that miR-122 is worth further investigation as a regulator of EBOV infection.

Though a minimal set of 20 genes could separate survivors from non-survivors, we were interested in also studying the general host response to EBOV infection, and to determine if there was a survival-associated transcriptional profile. We identified 238 genes that accurately distinguished treatment groups and survival outcomes following EBOV infection. Functional annotation of these 238 genes confirmed that this gene set was comparable with previously published studies of EBOV infection [21,22,24,27]. In particular, the expression pattern that we observe for IL6, in which non-survivors are significantly more up-regulated in early stages of infection than survivors, is supported by similar changes in protein concentration reported in previous studies [7,9,25]. There is also a pattern of significant up-regulation of genes associated with immune response in non-survivors, but not in survivors, consistent with the hypothesis that non-survivors exhibit severe dysregulation of the inflammatory response [21,22,24,27]. Importantly, our work highlights the utility of using a minimal survival-associated gene set to identify individual genes correlated with survival following EBOV infection, which is not possible when assessing global transcriptional responses in the host, as in previous work [22,24].

When we compare our data to the results of a study of survival-associated biomarkers in human samples from the Sudan virus (SUDV) outbreak [25], we find notable similarities. Similar to this study, we observed significant up-regulation chemokines and cytokines, such as CCL3, CXCL10, IL1RN, IL6 and TNF, throughout infection. This study reported that ferritin was a good correlate of hemorrhage and death in response to SUDV infection [25]. We observed down-regulation of ferritin throughout infection, although this pattern was not correlated with survival on a transcriptional level. However, we find that another iron-binding protein, lactotransferrin (LTF) is highly correlated with survival outcome in EBOV-infected NHPs. This similarity suggests that iron modulation may play an important role in regulating filovirus infection, especially in relation to coagulopathies and hemorrhage, and that our biomarkers merit further study in a human system.

We identified 3 transcriptional modules which were significantly enriched in the gene set: (i) CCAAT/enhancer-binding protein alpha (CEBPA); (ii) tumor protein 53 (p53); and (iii) megakaryoblastic leukemia 1 (MKL1) and myocardin-like protein 2 (MKL2). Previous studies have shown that p53 plays a crucial role during viral infection, which invariably disrupts normal cell cycle processes, in a variety of DNA and RNA viruses [63–66]. In particular, p53 is known to be associated with the Type I interferon response and has been previously reported to enhance viral-induced apoptosis in other infections [65,67]. In our examination of p53-regulated genes, we found that several are associated with regulation of apoptosis (e.g. BCL2L1 [68], CDC42 [69], CDK2 [70], FOXO3 [71], and PCNA [72]). This may suggest a role for p53 as a mediator of apoptosis following EBOV infection. However, due to a lack of a consensus pattern of expression, we are unable to determine the underlying regulation of p53 in this dataset, and therefore cannot draw conclusions about the activity of p53 in NHPs that survived EBOV infection when compared to non-survivors.

Interestingly, there is no consensus as to how CCAAT/enhancer binding proteins (such as CEBPA) function in general during viral infection, but they have been previously implicated in promoting the replication of some viruses. For example, CEBP binding sites exist in the Human immunodeficiency virus (HIV) genome, and CEBPs are required for the replication and regulation of HIV [73–75] and Simian immunodeficiency virus [76]. Similarly, physical binding and interactions have been observed between CEBPs and the proteins of Hepatitis B virus [77,78], Epstein-Barr virus [79], and HIV [80]. In contrast, CEBPs have been known to down-regulate or inhibit replication of T-cell leukemia virus [81,82] and some human papillomaviruses [83]. Our results suggest that strong CEBP responses are correlated with poorer prognosis following EBOV infection. We hypothesize that CEBP-regulated genes may contribute to the inflammatory response to infection, or to the dysregulation of coagulation.

Our studies are the first to suggest a role for MKL1 and MKL2 in viral infection, although roles for both proteins were recently identified in megakaryocyte differentiation and platelet formation [61]. Because dysregulation of coagulation is a common characteristic of EBOV infection, it is possible that MKL1 and MKL2

regulate coagulation in response to EBOV challenge. Indeed, we observe that the downstream targets of MKL1 and MKL2 exhibit an expression profile consistent with up-regulation of MKL1 and MKL2 in survivors, compared to non-survivors. This implies that survivors increase the regulation of coagulation processes, potentially avoiding the typical coagulopathies associated with late-stage EBOV infection. However, the majority of genes that are regulated by MKL1 and MKL2 are also regulated by CEBPA or p53, suggesting that the regulation observed is not due to the transcriptional activity of MKL1 and MKL2 alone. Despite this, these genes display a strong expression profile that is consistent with up-regulation of MKL1 and MKL2 when compared to a previous study [61], suggesting that survivors are able to recover in part due to normal MKL1 and MKL2 function.

It is important to note that we did not find a single unique gene that distinguished between survival outcomes of EBOV-infected NHPs, suggesting that survival following anticoagulant treatment is driven by a complex set of transcriptional responses. In addition, gene sets and pathways we have identified are associated with survival following anticoagulant treatment, and are therefore specific to this condition. We also stress that the observed results are in EBOV-infected NHPs, and our findings and conclusions may not be applicable to additional viral infections, although infection-specific signatures may exist. Under these conditions, we identify several complex transcriptional responses that clearly differentiate between survivors and non-survivors following EBOV infection. In particular, we observe several survival-associated profiles that are driven by specific upstream transcriptional regulators (*e.g.* CEBPA, p53, and MKL1/MKL2). Notably, these transcription factors have not been previously associated with EBOV infection, and would not have been identified without pathway analysis, due to lack of differential regulation. In particular, the ability of a small set of 20 genes to distinguish between survival outcomes suggests that they could potentially serve as biomarkers of disease outcome. Our results demonstrate that classification of treatment groups or disease outcome can be accomplished with a small gene set, which can be useful for identifying individual transcriptional markers associated with survival following anticoagulant treatment of EBOV infection.

Supporting Information

Figure S1 This figure illustrates the overall process of microarray analysis described in this paper, including sample collection, microarray processing and normalization, and analysis of the data (as described in *Materials and Methods*).
(PNG)

Table S1 This table lists the 20 genes from the minimal survival-associated gene set (as described in *Materials and Methods*). The table lists the Gene Symbol, Gene Entrez ID, a brief description of the gene name, and a short list of gene functions as described by the Gene Ontology "function" category.
(XLS)

Table S2 This table lists 241 genes from the general survival-associated transcriptional profile (as described in *Materials and Methods*). The table lists the Gene Symbol, Gene Entrez ID, a brief description of the gene name, and a short list of gene functions as described by the Gene Ontology "function" category.
(XLS)

Acknowledgments

We would like to thank John Coller and the staff of the Stanford Functional Genomics Facility for supplying us with the human cDNA microarrays used for this study.

Footnotes: Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

Author Contributions

Conceived and designed the experiments: JYY SG JHC YX ANH LEH. Performed the experiments: JYY ANH JBG. Analyzed the data: SG JYY. Contributed reagents/materials/analysis tools: JBG KHR TWG LEH. Contributed to the writing of the manuscript: SG JHC.

References

1. (2011) Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses; King AMQ, Adams, M.J., Carstens, E.B., and Lefkowitz, E.J., editor. San Diego: Elsevier Academic Press.
2. Feldmann H, Geisbert TW (2011) Ebola haemorrhagic fever. Lancet 377: 849–862.
3. Feldmann H, Jones S, Klenk HD, Schnittler HJ (2003) Ebola virus: from discovery to vaccine. Nat Rev Immunol 3: 677–685.
4. Feldmann H, Jones SM, Daddario-DiCaprio KM, Geisbert JB, Stroher U, et al. (2007) Effective post-exposure treatment of Ebola infection. PLoS Pathog 3: e2.
5. Geisbert TW, Daddario-Dicaprio KM, Geisbert JB, Reed DS, Feldmann F, et al. (2008) Vesicular stomatitis virus-based vaccines protect nonhuman primates against aerosol challenge with Ebola and Marburg viruses. Vaccine 26: 6894–6900.
6. Geisbert TW, Geisbert JB, Leung A, Daddario-DiCaprio KM, Hensley LE, et al. (2009) Single-injection vaccine protects nonhuman primates against infection with marburg virus and three species of ebola virus. J Virol 83: 7296–7304.
7. Geisbert TW, Hensley LE, Jahrling PB, Larsen T, Geisbert JB, et al. (2003) Treatment of Ebola virus infection with a recombinant inhibitor of factor VIIa/tissue factor: a study in rhesus monkeys. Lancet 362: 1953–1958.
8. Geisbert TW, Lee AC, Robbins M, Geisbert JB, Honko AN, et al. (2010) Postexposure protection of non-human primates against a lethal Ebola virus challenge with RNA interference: a proof-of-concept study. Lancet 375: 1896–1905.
9. Hensley LE, Stevens EL, Yan SB, Geisbert JB, Macias WL, et al. (2007) Recombinant human activated protein C for the postexposure treatment of Ebola hemorrhagic fever. J Infect Dis 196 Suppl 2: S390–399.
10. Jones SM, Feldmann H, Stroher U, Geisbert JB, Fernando L, et al. (2005) Live attenuated recombinant vaccine protects nonhuman primates against Ebola and Marburg viruses. Nat Med 11: 786–790.
11. Richardson JS, Yao MK, Tran KN, Croyle MA, Strong JE, et al. (2009) Enhanced protection against Ebola virus mediated by an improved adenovirus-based vaccine. PLoS One 4: e5308.
12. Sullivan NJ, Sanchez A, Rollin PE, Yang ZY, Nabel GJ (2000) Development of a preventive vaccine for Ebola virus infection in primates. Nature 408: 605–609.
13. Geisbert TW, Hensley LE, Larsen T, Young HA, Reed DS, et al. (2003) Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. Am J Pathol 163: 2347–2370.
14. Geisbert TW, Young HA, Jahrling PB, Davis KJ, Larsen T, et al. (2003) Pathogenesis of Ebola hemorrhagic fever in primate models: evidence that hemorrhage is not a direct effect of virus-induced cytolysis of endothelial cells. Am J Pathol 163: 2371–2382.
15. Hensley LE, Jones SM, Feldmann H, Jahrling PB, Geisbert TW (2005) Ebola and Marburg viruses: pathogenesis and development of countermeasures. Curr Mol Med 5: 761–772.
16. Flint SJ, Enquist LW, Racaniello VR, Skalka AM (2009) Principles of Virology. SciELO Espana.
17. Alvarez CP, Lasala F, Carrillo J, Muniz O, Corbi AL, et al. (2002) C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. J Virol 76: 6841–6844.
18. Harty RN, Brown ME, Wang G, Huibregtse J, Hayes FP (2000) A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding. Proc Natl Acad Sci U S A 97: 13871–13876.
19. Timmins J, Schoehn G, Ricard-Blum S, Scianimanico S, Vernet T, et al. (2003) Ebola virus matrix protein VP40 interaction with human cellular factors Tsg101 and Nedd4. J Mol Biol 326: 493–502.
20. Gupta M, Mahanty S, Ahmed R, Rollin PE (2001) Monocyte-derived human macrophages and peripheral blood mononuclear cells infected with ebola virus secrete MIP-1alpha and TNF-alpha and inhibit poly-IC-induced IFN-alpha in vitro. Virology 284: 20–25.
21. Hensley LE, Young HA, Jahrling PB, Geisbert TW (2002) Proinflammatory response during Ebola virus infection of primate models: possible involvement of the tumor necrosis factor receptor superfamily. Immunol Lett 80: 169–179.

22. Rubins KH, Hensley LE, Wahl-Jensen V, Daddario DiCaprio KM, Young HA, et al. (2007) The temporal program of peripheral blood gene expression in the response of nonhuman primates to Ebola hemorrhagic fever. Genome Biol 8: R174.
23. Villinger F, Rollin PE, Brar SS, Chikkala NF, Winter J, et al. (1999) Markedly elevated levels of interferon (IFN)-gamma, IFN-alpha, interleukin (IL)-2, IL-10, and tumor necrosis factor-alpha associated with fatal Ebola virus infection. J Infect Dis 179 Suppl 1: S188–191.
24. Yen JY, Garamszegi S, Geisbert JB, Rubins KH, Geisbert TW, et al. (2011) Therapeutics of Ebola hemorrhagic fever: whole-genome transcriptional analysis of successful disease mitigation. J Infect Dis 204 Suppl 3: S1043–1052.
25. McElroy AK, Erickson BR, Flietstra TD, Rollin PE, Nichol ST, et al. (2014) Ebola hemorrhagic fever: novel biomarker correlates of clinical outcome. J Infect Dis: doi: 10.1093/infdis/jiu088.
26. Sanchez A (2001) Ebola Viruses: Wiley Online Library.
27. Reed DS, Hensley LE, Geisbert JB, Jahrling PB, Geisbert TW (2004) Depletion of peripheral blood T lymphocytes and NK cells during the course of ebola hemorrhagic Fever in cynomolgus macaques. Viral Immunol 17: 390–400.
28. Bosio CM, Aman MJ, Grogan C, Hogan R, Ruthel G, et al. (2003) Ebola and Marburg viruses replicate in monocyte-derived dendritic cells without inducing the production of cytokines and full maturation. J Infect Dis 188: 1630–1638.
29. Bray M, Geisbert TW (2005) Ebola virus: the role of macrophages and dendritic cells in the pathogenesis of Ebola hemorrhagic fever. Int J Biochem Cell Biol 37: 1560–1566.
30. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80.
31. Smyth GK (2005) Limma: linear models for microarray data; Gentleman R CV, Dudoit S, Irizarry R, and Huber W., editor. New York: Springer.
32. Team RDC (2010) R: A language and environment for statistical computing. In: Computing RFfS, editor. Vienna, Austria.
33. Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, et al. (2007) A comparison of background correction methods for two-colour microarrays. Bioinformatics 23: 2700–2707.
34. Smyth GK, Speed T (2003) Normalization of cDNA microarray data. Methods 31: 265–273.
35. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30: 207–210.
36. de Hoon MJ, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. Bioinformatics 20: 1453–1454.
37. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95: 14863–14868.
38. Saldanha AJ (2004) Java Treeview–extensible visualization of microarray data. Bioinformatics 20: 3246–3248.
39. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498–2504.
40. Medigeshi GR, Hirsch AJ, Brien JD, Uhrlaub JL, Mason PW, et al. (2009) West nile virus capsid degradation of claudin proteins disrupts epithelial barrier function. J Virol 83: 6125–6134.
41. Kato N, Ji G, Wang Y, Baba M, Hoshida Y, et al. (2005) Large-scale search of single nucleotide polymorphisms for hepatocellular carcinoma susceptibility genes in patients with hepatitis C. Hepatology 42: 846–853.
42. Lee JW, Liao PC, Young KC, Chang CL, Chen SS, et al. (2011) Identification of hnRNPH1, NF45, and C14orf166 as novel host interacting partners of the mature hepatitis C virus core protein. J Proteome Res 10: 4522–4534.
43. Shin HJ, Kim SS, Cho YH, Lee SG, Rho HM (2002) Host cell proteins binding to the encapsidation signal epsilon in hepatitis B virus RNA. Arch Virol 147: 471–491.
44. Gomila RC, Martin GW, Gehrke L (2011) NF90 binds the dengue virus RNA 3′ terminus and is a positive regulator of dengue virus replication. PLoS One 6: e16687.
45. Wang P, Song W, Mok BW, Zhao P, Qin K, et al. (2009) Nuclear factor 90 negatively regulates influenza virus replication by interacting with viral nucleoprotein. J Virol 83: 7850–7861.
46. Adams WC, Bond E, Havenga MJ, Holterman L, Goudsmit J, et al. (2009) Adenovirus serotype 5 infects human dendritic cells via a coxsackievirus-adenovirus receptor-independent receptor pathway mediated by lactoferrin and DC-SIGN. J Gen Virol 90: 1600–1610.
47. Berlutti F, Pantanella F, Natalizi T, Frioni A, Paesano R, et al. (2011) Antiviral properties of lactoferrin–a natural immunity molecule. Molecules 16: 6992–7018.
48. Valenti P, Antonini G (2005) Lactoferrin: an important host defence against microbial and viral attack. Cell Mol Life Sci 62: 2576–2587.
49. Valimaa H, Tenovuo J, Waris M, Hukkanen V (2009) Human lactoferrin but not lysozyme neutralizes HSV-1 and inhibits HSV-1 replication and cell-to-cell spread. Virol J 6: 53.
50. Zheng Y, Zhang W, Ye Q, Zhou Y, Xiong W, et al. (2012) Inhibition of Epstein-Barr virus infection by lactoferrin. J Innate Immun 4: 387–398.
51. Li Y, Masaki T, Yamane D, McGivern DR, Lemon SM (2013) Competing and noncompeting activities of miR-122 and the 5′ exonuclease Xrn1 in regulation of hepatitis C virus replication. Proc Natl Acad Sci U S A 110: 1881–1886.
52. Sion-Vardy N, Lasarov I, Delgado B, Gopas J, Benharroch D, et al. (2009) Measles virus: evidence for association with lung cancer. Exp Lung Res 35: 701–712.
53. Kakugawa S, Shimojima M, Neumann G, Goto H, Kawaoka Y (2009) RuvB-like protein 2 is a suppressor of influenza A virus polymerases. J Virol 83: 6429–6434.
54. Shabman RS, Leung DW, Johnson J, Glennon N, Gulcicek EE, et al. (2011) DRBP76 associates with Ebola virus VP35 and suppresses viral polymerase function. J Infect Dis 204 Suppl 3: S911–918.
55. Jopling CL, Schutz S, Sarnow P (2008) Position-dependent function for a tandem microRNA miR-122-binding site located in the hepatitis C virus RNA genome. Cell Host Microbe 4: 77–85.
56. Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P (2005) Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science 309: 1577–1581.
57. Ambroise C, McLachlan GJ (2002) Selection bias in gene extraction on the basis of microarray gene-expression data. Proc Natl Acad Sci U S A 99: 6562–6566.
58. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286: 531–537.
59. Simon R, Radmacher MD, Dobbin K, McShane LM (2003) Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. J Natl Cancer Inst 95: 14–18.
60. Gery S, Gombart AF, Yi WS, Koeffler C, Hofmann WK, et al. (2005) Transcription profiling of C/EBP targets identifies Per2 as a gene implicated in myeloid leukemia. Blood 106: 2827–2836.
61. Smith EC, Thon JN, Devine MT, Lin S, Schulz VP, et al. (2012) MKL1 and MKL2 play redundant and crucial roles in megakaryocyte maturation and platelet formation. Blood 120: 2317–2329.
62. Volchkov VE, Volchkova VA, Chepurnov AA, Blinov VM, Dolnik O, et al. (1999) Characterization of the L gene and 5′ trailer region of Ebola virus. J Gen Virol 80 (Pt 2): 355–362.
63. Lazo PA, Santos CR (2011) Interference with p53 functions in human viral infections, a target for novel antiviral strategies? Rev Med Virol 50: doi: 10.1002/rmv.696.
64. Collot-Teixeira S, Bass J, Denis F, Ranger-Rogez S (2004) Human tumor suppressor p53 and DNA viruses. Rev Med Virol 14: 301–319.
65. Turpin E, Luke K, Jones J, Tumpey T, Konan K, et al. (2005) Influenza virus infection increases p53 activity: role of p53 in cell death and viral replication. J Virol 79: 8802–8811.
66. Marques JT, Rebouillat D, Ramana CV, Murakami J, Hill JE, et al. (2005) Down-regulation of p53 by double-stranded RNA modulates the antiviral response. J Virol 79: 11105–11114.
67. Munoz-Fontela C, Macip S, Martinez-Sobrido L, Brown L, Ashour J, et al. (2008) Transcriptional role of p53 in interferon-mediated antiviral immunity. J Exp Med 205: 1929–1938.
68. Hagenbuchner J, Ausserlechner MJ, Porto V, David R, Meister B, et al. (2010) The anti-apoptotic protein BCL2L1/Bcl-xL is neutralized by pro-apoptotic PMAIP1/Noxa in neuroblastoma, thereby determining bortezomib sensitivity independent of prosurvival MCL1 expression. J Biol Chem 285: 6904–6912.
69. Thomas A, Giesler T, White E (2000) p53 mediates bcl-2 phosphorylation and apoptosis via activation of the Cdc42/JNK1 pathway. Oncogene 19: 5259–5269.
70. Huang H, Regan KM, Lou Z, Chen J, Tindall DJ (2006) CDK2-dependent phosphorylation of FOXO1 as an apoptotic response to DNA damage. Science 314: 294–297.
71. Zhang X, Tang N, Hadden TJ, Rishi AK (2011) Akt, FoxO and regulation of apoptosis. Biochim Biophys 1813: 1978–1986.
72. Witko-Sarsat V, Mocek J, Bouayad D, Tamassia N, Ribeil JA, et al. (2010) Proliferating cell nuclear antigen acts as a cytoplasmic platform controlling human neutrophil survival. J Exp Med 207: 2631–2645.
73. Henderson AJ, Calame KL (1997) CCAAT/enhancer binding protein (C/EBP) sites are required for HIV-1 replication in primary macrophages but not CD4(+) T cells. Proc Natl Acad Sci U S A 94: 8714–8719.
74. Henderson AJ, Connor RI, Calame KL (1996) C/EBP activators are required for HIV-1 replication and proviral induction in monocytic cell lines. Immunity 5: 91–101.
75. Schwartz C, Catez P, Rohr O, Lecestre D, Aunis D, et al. (2000) Functional interactions between C/EBP, Sp1, and COUP-TF regulate human immunodeficiency virus type 1 gene transcription in human brain cells. J Virol 74: 65–73.
76. Mohan M, Aye PP, Borda JT, Alvarez X, Lackner AA (2008) CCAAT/enhancer binding protein beta is a major mediator of inflammation and viral replication in the gastrointestinal tract of simian immunodeficiency virus-infected rhesus macaques. Am J Pathol 173: 106–118.
77. Choi BH, Park GT, Rho HM (1999) Interaction of hepatitis B viral X protein and CCAAT/enhancer-binding protein alpha synergistically activates the hepatitis B viral enhancer II/pregenomic promoter. J Biol Chem 274: 2858–2865.
78. Mahe Y, Mukaida N, Kuno K, Akiyama M, Ikeda N, et al. (1991) Hepatitis B virus X protein transactivates human interleukin-8 gene through acting on nuclear factor kB and CCAAT/enhancer-binding protein-like cis-elements. J Biol Chem 266: 13759–13763.

79. Wu FY, Wang SE, Chen H, Wang L, Hayward SD, et al. (2004) CCAAT/enhancer binding protein alpha binds to the Epstein-Barr virus (EBV) ZTA protein through oligomeric interactions and contributes to cooperative transcriptional activation of the ZTA promoter through direct binding to the ZII and ZIIIB motifs during induction of the EBV lytic cycle. J Virol 78: 4847–4865.
80. Ambrosino C, Ruocco MR, Chen X, Mallardo M, Baudi F, et al. (1997) HIV-1 Tat induces the expression of the interleukin-6 (IL6) gene by binding to the IL6 leader RNA and by interacting with CAAT enhancer-binding protein beta (NF-IL6) transcription factors. J Biol Chem 272: 14883–14892.
81. Hivin P, Gaudray G, Devaux C, Mesnard JM (2004) Interaction between C/EBPbeta and Tax down-regulates human T-cell leukemia virus type I transcription. Virology 318: 556–565.
82. Grant C, Nonnemacher M, Jain P, Pandya D, Irish B, et al. (2006) CCAAT/enhancer-binding proteins modulate human T cell leukemia virus type 1 long terminal repeat activation. Virology 348: 354–369.
83. Bauknecht T, Shi Y (1998) Overexpression of C/EBPbeta represses human papillomavirus type 18 upstream regulatory region activity in HeLa cells by interfering with the binding of TATA-binding protein. J Virol 72: 2113–2124.