


# A Simple Screening Approach To Prioritize Genes for Functional Analysis Identifies a Role for Interferon Regulatory Factor 7 in the Control of Respiratory Syncytial Virus Disease

Jacqueline U. McDonald,<sup>a</sup> Myrsini Kaforou,<sup>b</sup> Simon Clare,<sup>c</sup> Christine Hale,<sup>c</sup> Maria Ivanova,<sup>a\*</sup> Derek Huntley,<sup>d</sup> Marcus Dorner,<sup>e</sup> Victoria J. Wright,<sup>b</sup> Michael Levin,<sup>b</sup> Federico Martinon-Torres,<sup>f</sup> Jethro A. Herberg,<sup>b</sup>  John S. Tregoning<sup>a</sup>

Mucosal Infection and Immunity Group, Section of Virology, Imperial College London, St. Mary's Campus, London, United Kingdom<sup>a</sup>; Section of Paediatrics, Imperial College London, St. Mary's Campus, London, United Kingdom<sup>b</sup>; Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, United Kingdom<sup>c</sup>; Imperial College Centre for Integrative Systems Biology and Bioinformatics, Imperial College London, United Kingdom<sup>d</sup>; Molecular Virology, Section of Virology, Imperial College London, St. Mary's Campus, London, United Kingdom<sup>e</sup>; Department of Paediatrics, Hospital Clínico Universitario de Santiago, Santiago de Compostela, Spain<sup>f</sup>

**ABSTRACT** Greater understanding of the functions of host gene products in response to infection is required. While many of these genes enable pathogen clearance, some enhance pathogen growth or contribute to disease symptoms. Many studies have profiled transcriptomic and proteomic responses to infection, generating large data sets, but selecting targets for further study is challenging. Here we propose a novel data-mining approach combining multiple heterogeneous data sets to prioritize genes for further study by using respiratory syncytial virus (RSV) infection as a model pathogen with a significant health care impact. The assumption was that the more frequently a gene is detected across multiple studies, the more important its role is. A literature search was performed to find data sets of genes and proteins that change after RSV infection. The data sets were standardized, collated into a single database, and then panned to determine which genes occurred in multiple data sets, generating a candidate gene list. This candidate gene list was validated by using both a clinical cohort and *in vitro* screening. We identified several genes that were frequently expressed following RSV infection with no assigned function in RSV control, including *IFI27*, *IFIT3*, *IFI44L*, *GBP1*, *OAS3*, *IFI44*, and *IRF7*. Drilling down into the function of these genes, we demonstrate a role in disease for the gene for interferon regulatory factor 7, which was highly ranked on the list, but not for *IRF1*, which was not. Thus, we have developed and validated an approach for collating published data sets into a manageable list of candidates, identifying novel targets for future analysis.

**IMPORTANCE** Making the most of “big data” is one of the core challenges of current biology. There is a large array of heterogeneous data sets of host gene responses to infection, but these data sets do not inform us about gene function and require specialized skill sets and training for their utilization. Here we describe an approach that combines and simplifies these data sets, distilling this information into a single list of genes commonly upregulated in response to infection with RSV as a model pathogen. Many of the genes on the list have unknown functions in RSV disease. We validated the gene list with new clinical, *in vitro*, and *in vivo* data. This approach allows the rapid selection of genes of interest for further, more-detailed

Received 23 April 2016 Accepted 26 May 2016 Published 28 June 2016


**Citation** McDonald JU, Kaforou M, Clare S, Hale C, Ivanova M, Huntley D, Dorner M, Wright VJ, Levin M, Martinon-Torres F, Herberg JA, Tregoning JS. 2016. A simple screening approach to prioritize genes for functional analysis identifies a role for interferon regulatory factor 7 in the control of respiratory syncytial virus disease. *mSystems* 1(3):e00051-16. doi:10.1128/mSystems.00051-16.

**Editor** Sergio Baranzini, University of California, San Francisco

**Copyright** © 2016 McDonald et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to John S. Tregoning, john.tregoning@imperial.ac.uk.

\*Present address: Maria Ivanova, Janssen Pharmaceuticals, The Netherlands.

 From great complexity comes forth simplicity: combining data sets to find novel gene targets for respiratory syncytial virus (RSV) research

studies, thus reducing time and costs. Furthermore, the approach is simple to use and widely applicable to a range of diseases.

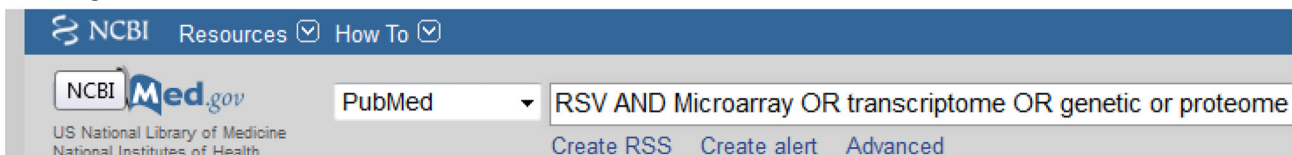
**KEYWORDS:** host response, immunity, respiratory syncytial virus, viral immunity

The interpretation of large data sets—big data—is one of the challenges of modern biology (1). Several powerful approaches have been developed to derive functional correlates from these large data sets, but they all have some limitations. Analysis of correlated or functionally related groups of genes *en bloc* simplifies analysis (2, 3); however, this approach loses gene level detail, particularly for genes with unknown roles. Systems biology approaches to identify key genes within pathways have been applied in vaccination (4) and infection (5) studies, but there is still a requirement to select individual genes for further analysis. High-throughput screenings enable the rapid identification of gene functions in an *in vitro* context (6–8), but these screenings only investigate the role of genes in the context of individual cells, not in relation to the system as a whole. Recently, programs of work have been developed to systematically target every gene in the mouse genome to define its function (9, 10); nevertheless, cost and ethical considerations require a focused selection of the targets of interest. The plethora of data available makes the prioritization of genes for further analysis challenging and often requires specialized skill sets and costly software. We propose a novel and simple approach to integrate published data sets to rapidly identify genes for their functions in the control of infection by using freely available software.

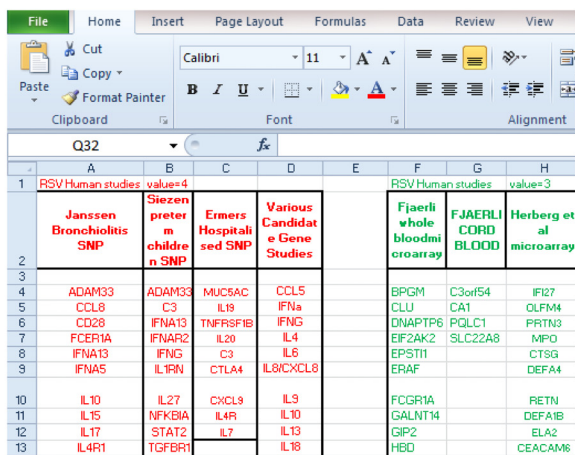
We used this novel approach to identify genes involved in the host response to respiratory syncytial virus (RSV) infection as a proof of principle. RSV is a ubiquitous infection in early life and a significant cause of disease (11). While the majority of children are infected with RSV during infancy, only a small proportion (2%) require hospitalization, of whom many have known risk factors, including prematurity, congenital heart disease, or immunodeficiency. However, the majority of hospitalized children (73 to 85%) have no known risk factor (12, 13). This phenotypic variability in host response may reflect the role of host genetic polymorphisms in protection against or potentiation of severe disease (14). Severe RSV disease is associated with perturbation of normal airway function in the lower respiratory tract, but the events leading to the perturbation of airway function after RSV infection are not clear. Indeed, the cause of disease may be heterogeneous, with virus-induced cell death causing disease in some infants and excess local inflammation having a role in others. The broader cellular immune response to RSV has been well dissected, with protective and pathogenic roles assigned to many cell types, including macrophages (15, 16) and NK (17, 18) and T cells (19, 20), but the molecular immune profile has not been fully explored. A number of studies in the last decade have published host omics profiles of RSV disease, identifying signatures of RSV infection, which enable discrimination between RSV and other respiratory viral infections (21, 22). Published associations of genes with the response to RSV are derived from a diverse range of systems, both *in vitro* and *in vivo*, in humans and mice, and based on RNA, DNA, and protein data (see Table S1 in the supplemental material). However, while these studies have identified genes that change in response to infection, they have not defined the functional roles of individual genes in RSV disease.

The aims of the present study were to integrate multiple published data sets to prioritize the genes associated with RSV infection and to dissect their functions in RSV disease. To achieve these goals, we employed a combination of *in silico*, *in vitro*, and *in vivo* approaches. Our conclusions were consistently supported by a new clinical study. Integrating multiple studies in this fashion increases confidence in the role of the genes identified. Utilizing this approach we identified *IRF7* as a key gene in the control of RSV. Here, we demonstrate that integration of published omic data sets with high-throughput studies generates insights into the genetic control of infection.

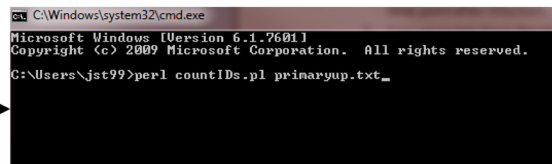
## Step 1. Literature search



## Step 2. Curate published gene lists into single excel database



## Step 3. Run countIDs in Perl



## Step 4. Generate candidate list

Weighted Score	Gene
24	IFI27
23	IFIT3
18	GBP1
16	IFI44L
15	OAS3, IFI44, CXCL10
14	HP
12	ADAM33, OAS2, IL1RN, IFIT2, HERC5, ARG1, ISG15

## Step 5. Validate:

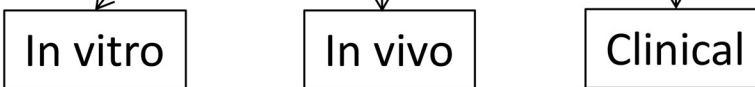


FIG 1 Flowchart of the gene selection method used in this study.

## RESULTS

### Meta-analysis of RSV data sets reveals functional pathways in the control of RSV.

We developed a novel approach to mine the mass of published gene and protein profile data in order to prioritize genes for functional profiling (Fig. 1). A literature search was performed to identify studies using omics tools to analyze the response to RSV infection. Data were collected from multiple published studies of RSV disease, selecting studies that included accessible lists of genes and/or proteins that were detectable following RSV infection of either humans or mice (see Table S1 in the supplemental material) or identified as significant in genetic studies or genome-wide association studies. Because of the heterogeneity of the approaches, when multiple data sets were available in a single study, we focused on the 24 h after infection time points in primary infection (rather than reinfection) and set an arbitrary cutoff of a 2-fold increase or decrease in gene expression for transcriptomic studies. Studies were collated in a single database, and we then used a custom Perl script ([countIDs](https://sourceforge.net/projects/countids/); <https://sourceforge.net/projects/countids/>) to parse the file to find genes that were present in multiple studies. Genes were ranked by frequency of occurrence and weighted on the basis of the types of studies they appeared in: genetic association studies, *ex vivo* human studies, or *in vitro* human or mouse studies, with more weighting on the human than the murine studies. This subjective weighting score was based on perceived relevance to human infection. Weighting reflected the nature of the input study, not specific data layers, which were treated equally; genes, mRNA, and

Downloaded from <http://msystems.asm.org/> on November 10, 2016 by guest

**TABLE 1** Genes most frequently upregulated following RSV infection<sup>a</sup>

Weighted score	Gene(s)
24	<i>IFI27</i>
23	<i>IFIT3</i>
18	<i>GBP1</i>
16	<i>IFI44L</i>
15	<i>OAS3, IFI44, CXCL10</i>
14	<i>HP</i>
12	<i>ADAM33, OAS2, IL-1RN, IFIT2, HERC5, RSAD2/Viperin, ARG1, ISG15</i>
11	<i>CXCL11, STAT2, TNF, FCER1A, IRF7, FCGR1A, CA1, STAT1</i>
10	<i>IFN-<math>\gamma</math>, IFIT1, CCL5, EPST11, MX2</i>
9	<i>CEACAM6, C3, CCL8, CXCL9, TRAC, IFI35, MX1, MPO, LCN2, OTOF</i>
8	<i>VDR, BPGM, IFI6, IL-10, ANXA3, OLFM4, SAMHD1, SERPING1, DEFA1, IFN-A13, RTP4, NOS2A, AIM2, Jun, OASL, GBP4</i>
7	<i>SCGB1A1, ISG20, CHI3L1, OAS1A, MUC5AC, PRIC285, IL-6</i>
6	<i>IFITM3, IL-20, MMP8, DEFA4, OASL2, ATF3, CXCL2, TMCS, TF, HBBP1, FCGR1B, IFIH1, CXCL8, CCL4, IL-15, PRF1, ALAS2, NFKBIA, MSP, ELA2, KLRD1, IL-7, MMP9, CD14</i>
5	<i>DEFA3, THOC4, CEACAM8, IFIT5, LAMP3, ERAF, IFN-<math>\alpha</math>, ALDH1A1, LGALS9, GZMB, LTF, CCL7, HBM, OAS1, CTNNAL1, WARS, LY6E, HBD, IGTP, S100A12, PSMB8, DHX58, IFI1, IFI47, CCL2, PSMB9, GNLY</i>
4	<i>EIF2AK2, MYD88, BATE2, CMPK2, GMPR, LILRB4, OSMR, IIGP1 TNFSF13B, DAXX, HLA-G, HSPA8, IL-18BP, NMI, HLA-B, SAMD9L, CD177, IIGP2, LAP3, GBP2, USP18, PLAC8, MS4A6D, FCGR1, IFITM1, PARP9, AIF1, SERPINA3G, IFI202B</i>

<sup>a</sup>Genes were collated from multiple studies of RSV. A cutoff of a 2-fold increase in expression, compared to the reference group in the study from which the data were collated, was used when available. Genes were weighted as follows on the basis of the studies from which they were collated: human genetic studies, 4; human *in vivo* microarray studies, 3; human *in vitro* microarray studies, 2; mouse studies, 1. After weighting, genes were analyzed for multiple hits by a custom Perl script.

proteins were given equal weight. A list of candidate genes was then generated (Table 1). Using the weighted analysis, the genes that were most commonly reported as being upregulated were *IFI27*, *IFIT3*, *GBP1*, *IFI44L*, *OAS3*, *IFI44*, and *CXCL10* (Table 1). No functional roles in the control of RSV infection have been previously described for these genes. Fewer genes were downregulated after RSV infection, and they were less uniformly represented between studies; the downregulated genes included *CLC*, *NDUFS1*, and *PFDN5* (Table 2). For comparison, an unweighted analysis was also performed (see Tables S2 and S3 in the supplemental material); this analysis identified similar patterns of genes, with interferon the gene for (IFN) regulatory factor 7 (*IRF7*) relatively higher in the unweighted analyses than in the weighted analyses. Thus, we can take published data sets and condense them into a single list. Interestingly despite the heterogeneous nature of the input studies, we identified a large overlap in the genes identified.

To visualize the genes and interactions, we plotted the list by using the Ingenuity software platform (Fig. 2). Genes with scores of  $\geq 6$  were included. The analysis identified a mixture of secreted factors (cytokines and chemokines) and intracellular factors (transcription factors and IFN-stimulated genes [ISGs]). On the basis of our previous experience and the published literature, it was interesting that extracellular proteins upregulated in response to RSV are often associated with enhanced disease (15, 23–25), whereas intracellular proteins are associated with disease control (26). To simplify the presentation of the interactions, we focused on the interactions of the top 16 upregulated genes and the top 5 downregulated genes, looking at direct interactions that have been observed experimentally. The main observation from this was that *IRF7*, which was found to be upregulated in nine independent studies (7, 27–32),

**TABLE 2** Genes most frequently downregulated following RSV infection<sup>a</sup>

Weighted score	Gene(s)
9	<i>CLC</i>
6	<i>NDUFS1, PFDN5</i>
5	<i>RTN1, CAT, FCER1A, TSPAN8, ALOX15, GPR56, KLRB1</i>
4	<i>XRCC5, LMNA, UBD, CCT3, HSPA8, GARS</i>

<sup>a</sup>Genes were collated from multiple studies of RSV. A cutoff of a 2-fold decrease in expression was used when available. Genes were weighted as follows on the basis of the studies from which they were collated: human genetic studies, 4; human *in vivo* microarray studies, 3; human *in vitro* microarray studies, 2; mouse studies, 1. After weighting, genes were analyzed for multiple hits by a custom Perl script.



**TABLE 3** Canonical pathways<sup>a</sup>

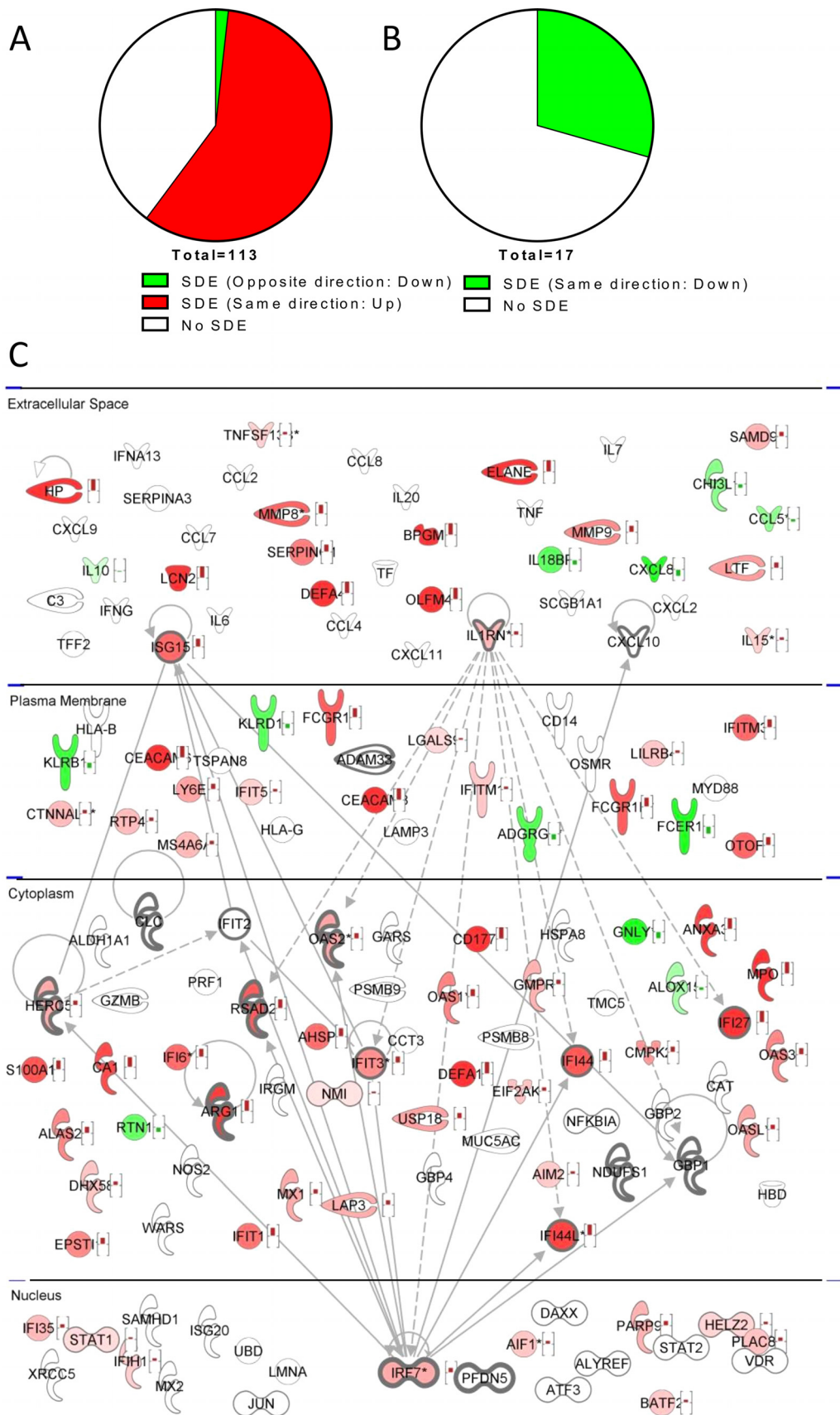
Ingenuity canonical pathway	Molecules
IFN signaling	OAS1, IFIT1, IFN- $\gamma$ , IFITM1, STAT1, IFN-A1/IFN-A13, IFIT3, STAT2, MX1, IFI35, IFITM3, PSMB8
Activation of IRF by cytosolic pattern recognition receptors	Jun, DHX58, STAT2, IFIT2, IL-6, NFKBIA, IRF7, STAT1, TNF, IFN-A1/IFN-A13, ISG15, IFIH1, IL-10
Communication between innate and adaptive immune cells	IL-15, TNFSF13B, IFN- $\gamma$ , CCL5, CXCL10, HLA-G, IL-6, CXCL8, IL-1RN, TNF, IFN-A1/IFN-A13, IL-10, HLA-B, CCL4
Role of hypercytokinemia/hyperchemokine in pathogenesis of influenza	IL-15, IL-1RN, IFN- $\gamma$ , CCL5, CXCL10, TNF, IFN-A1/IFN-A13, CCL2, CCL4, IL-6, CXCL8
Role of pattern recognition receptors in recognition of bacteria and viruses	EIF2AK2, OAS1, IFN- $\gamma$ , C3, MYD88, CCL5, OAS2, IL-6, CXCL8, IRF7, TNF, IFN-A1/IFN-A13, OAS3, IFIH1, IL-10
Granulocyte adhesion and diapedesis	CXCL9, CCL8, CCL7, CCL5, MMP9, CXCL10, CXCL2, CXCL8, CXCL11, IL-1RN, TNF, CCL2, MMP8, CCL4
Agranulocyte adhesion and diapedesis	CXCL9, CCL8, CCL7, CCL5, MMP9, CXCL10, CXCL2, CXCL8, CXCL11, IL-1RN, TNF, CCL2, MMP8, CCL4
Role of cytokines in mediating communication between immune cells	IL-15, IL-1RN, IFN- $\gamma$ , TNF, IFN-A1/IFN-A13, IL-20, IL-10, IL-6, CXCL8
Differential regulation of cytokine production in intestinal epithelial cells by IL-17A and IL-17F	IFN- $\gamma$ , CCL5, TNF, CCL2, IL-10, CCL4, LCN2
Dendritic cell maturation	FCGR1B, IL-15, MYD88, FCGR1A, STAT2, IL-6, NFKBIA, IL-1RN, STAT1, TNF, IFN-A1/IFN-A13, IL-10, HLA-B

<sup>a</sup>Ingenuity pathway analysis was applied to the top-scoring genes.

blood gene expression data from patients infected with RSV. We compared the transcriptomic profile of children hospitalized because of infection with RSV to that of age-matched healthy controls by using microarrays. The list of genes derived from the literature (Table 1) was compared to the genes that were significantly differentially expressed (SDE) in children infected with RSV versus healthy controls, and an overlap of 73 out of 130 genes was observed. Of the genes identified as upregulated in the literature-derived list, 66 were observed to be significantly upregulated in the clinical study, 45 were not SDE, and only 2 were SDE but in the opposite direction (Fig. 3A). The list of downregulated gene from the literature was smaller but also had a smaller proportion of agreement with the clinical study—only 5 of 17 genes were present in both studies (Fig. 3B). These data were then mapped onto the network built by using the literature-derived list (Fig. 3C), demonstrating that many of the top-ranking genes from the literature-derived list, including *IRF7*, *OAS2*, *RSAD2*, *HERC5*, *ISG15*, *IFI44*, *IL-1RN*, *ARG1*, and *IFIT3*, were in agreement between the two methods.

**Validation of the gene list by *in vitro* assay.** Of note, a number of the genes that we identified as commonly upregulated have no known role in the control of RSV infection. We wished to screen the genes identified for their effects on RSV infection by the flow cytometry-based screening method described by Schoggins et al. (7). We screened 39 ISGs identified by our *in silico* screening, and we also included receptors and transcription factors identified as upstream regulators by pathway analysis in previous studies (*IL-28RA*, *IRF1*, *IRF2*, *SOCS1*, *SOCS2*, *STAT3*, *TLR3*, and *TLR7*). The HEp-2 epithelial cell line was used, as it represents the cell lineages that RSV first encounters during an infection. HEp-2 cells were transduced with lentiviral vectors expressing each ISG and red fluorescent protein (RFP) in the same vector prior to infection with RSV expressing green fluorescent protein (GFP) (33). PKR, IFI6, and OASL overexpression reduced the GFP expression level (and therefore infection) by more than 75% of the control (Fig. 4). Of the genes identified with the most hits in the *in silico* studies (score of >12), the reductions in infectivity were as follows: *IFI27*, 65.6%  $\pm$  26.6%; *IFIT3*, 62.5%  $\pm$  10.6%; *IFI44L*, 67.3%  $\pm$  39.7%; *GBP1*, not in panel; *OAS3*, 66.3%  $\pm$  36.0%; *IFI44*, 74.5%  $\pm$  25.5%; *ISG15*, 58.3%  $\pm$  11.5%. *IRF7* overexpression led to a 76.7%  $\pm$  11.7% reduction in RSV replication, which demonstrates that, in addition to being centrally located in the predicted gene networks from *in silico* analysis, *IRF7* has a role in the control of RSV infection. The *in vitro* data support a role in viral control for the genes identified by our novel screening method.

**Validation of the gene list by an *in vivo* infection model.** The overarching aim of this study was to identify new genes of interest for further study. The informatic and



**FIG 3** Validation of bioinformatic screening of a patient cohort. SDE genes from the clinical cohort were compared with the literature-derived gene list. Shown are overlaps between the gene lists expressed as pie charts, with  
(Continued)

Downloaded from http://msystems.asm.org/ on November 10, 2016 by guest

*in vitro* analyses identified *IRF7* as being involved in the response to RSV, and it has not previously been studied in the context of RSV infection. To validate our screening approach, we compared RSV infection in mice deficient in *IRF7*, a gene on our list, with one that is associated with antiviral responses but is not on our list (*IRF1*). *IRF7*<sup>-/-</sup> and *IRF1*<sup>-/-</sup> mice were intranasally infected with  $5 \times 10^5$  PFU of RSV-A (A2 strain) and monitored daily for weight loss for 7 days postinfection. Cohorts of mice were sacrificed on days 4 and 7 postinfection to quantify viral burdens and immunological changes over the course of the challenge. Mice were compared to wild-type controls on the same background. *IRF7*<sup>-/-</sup> mice showed significant weight loss on days 6 and 7 postinfection compared to their wild-type littermates ( $P < 0.01$ ) (Fig. 5A). There was no difference in weight loss between *IRF1*<sup>-/-</sup> and wild-type controls (Fig. 5B). RSV loads were significantly greater in both *IRF7*<sup>-/-</sup> and *IRF1*<sup>-/-</sup> mice on day 4 postinfection ( $P < 0.05$ , Fig. 5C and D) but not on day 7 postinfection. Cellular infiltrate was quantified over the course of infection, which showed a significant increase in total cells resident in the lungs on day 7 postinfection in *IRF7*<sup>-/-</sup> mice but not in *IRF1*<sup>-/-</sup> mice ( $P < 0.05$ , Fig. 5E and F). Flow cytometry revealed an increase in all cellular subpopulations in *IRF7*<sup>-/-</sup> mice relative to those of wild-type mice on day 7 postinfection. In particular, the total numbers of NK cells in the lungs were significantly higher ( $P < 0.05$ ) (Fig. 5G and H); there was no significant difference in the *IRF1*<sup>-/-</sup> mice. Analysis of inflammatory cytokines present in the lungs revealed differences on day 7 postinfection (Fig. 5I and J), with significantly higher levels of IL-1 $\beta$  in the lungs of *IRF1*<sup>-/-</sup> mice than in those of wild-type controls ( $P < 0.05$ ).

## DISCUSSION

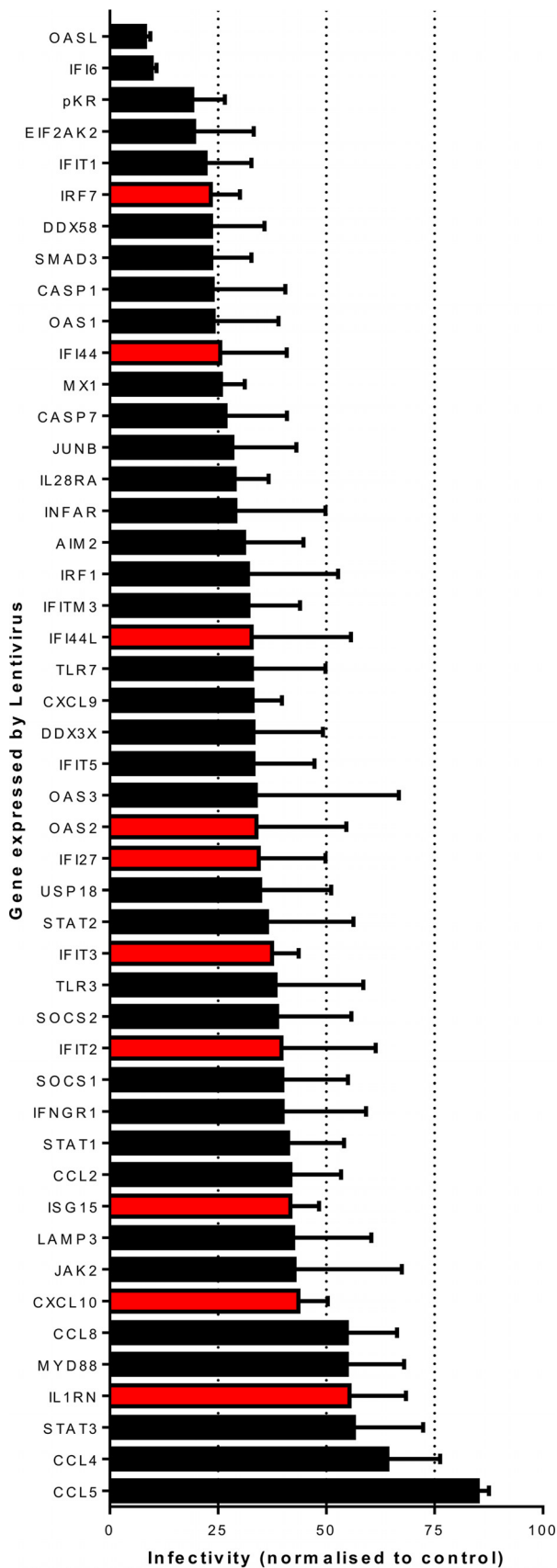
Here we used a novel integrative approach to identify and characterize genes that are upregulated in response to RSV infection for further analysis. Previous studies have explored genetic signatures to discriminate RSV infection from other viral infections (21, 22). Our approach enabled the identification of relevant genes in a hypothesis-free fashion, identifying genes with both known and unknown functions. We used a novel algorithm that permitted the integration of multiple large genetic, gene expression, and protein data sets to identify genes consistently upregulated after RSV infection across multiple model systems. Using this approach, we were able to distill multiple heterogeneous studies into a single list of candidate molecules, generate testable hypotheses, and then demonstrate functional relevance. While we focused on RSV, this approach is broadly applicable to other pathogens for which large sets of gene expression data are available, and the data mining program is available as an open-source program.

We identified a number of genes with no known function in RSV disease as potential targets for future investigation, including *IFI27*, *IFIT3*, *IFI44L*, *GBP1*, *OAS3*, and *IFI44*. Several of these have reported roles from other infections, but a role in RSV infection has not been reported. *IFI27* (also called ISG12) plays a proinflammatory role by inducing the nuclear export of an anti-inflammatory nuclear receptor, NR4A1 (34), and recently has been related to proliferation and human epidermal cell cycling (35); *GBP1* is a GTPase with a possible role in actin remodeling (36); *IFI44* is antiproliferative; *OAS3* interacts with RNase L (37); and ISG15 is a ubiquitin-like modifier (38) that has been shown to reduce RSV viral growth *in vitro* (39). However, some of the most frequently upregulated genes, for example, *IFI44L*, have no assigned molecular function. We took some of these genes forward into an *in vitro* assay and observed that there was a partial reduction of viral replication with all of the top hits tested (*GBP1* was not tested, as it

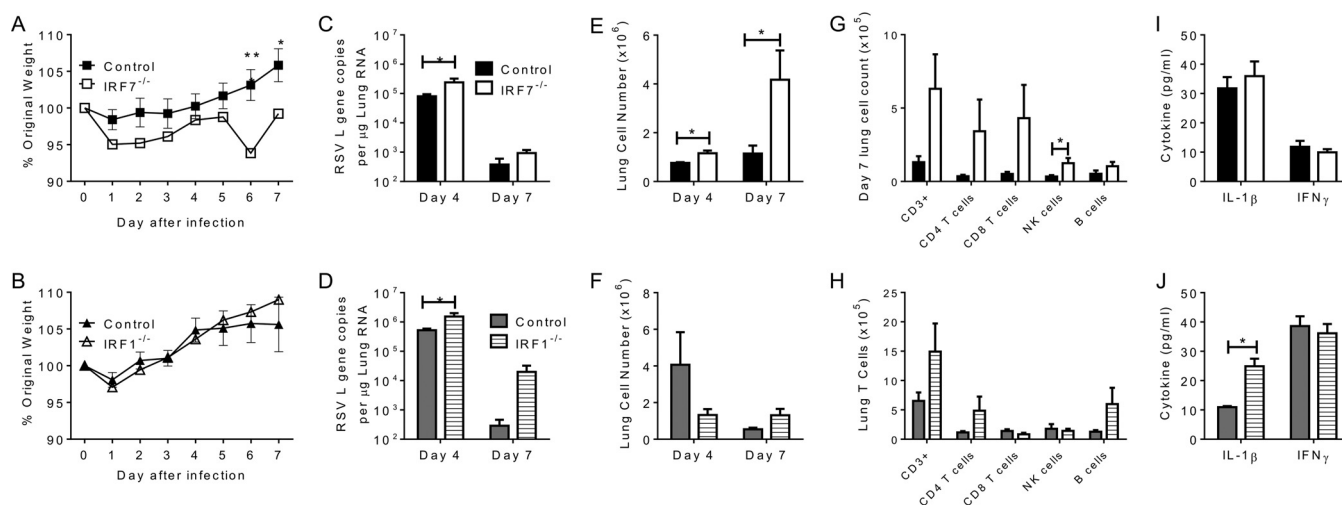
### Figure Legend Continued

directionality of agreement indicated for genes on the literature-derived list that are upregulated (A) or downregulated (B). Also shown are relative expression data from RSV-infected patients overlaid on the gene network derived from the literature-derived list (C). Upregulated genes are red, downregulated genes are green, genes identified in bioinformatic study but not clinical study are white, shading represents differential expression, and genes on the literature-derived list are indicated by bold outlining.





**FIG 4** Flow cytometry confirmation of inhibitory functions of genes identified *in silico*. HEp-2 cells were transduced with lentiviral vectors expressing genes of interest identified by *in silico* screening, (Continued)



**FIG 5** IRF7, but not IRF1, is important in the control of RSV infection. *IRF7*<sup>-/-</sup> (A, C, E, G, I) or *IRF1*<sup>-/-</sup> (B, D, F, H, I) mice were infected with  $5 \times 10^5$  PFU of RSV strain A2 and compared to wild-type controls on the same background. Mice were weighed daily, and weight changes were recorded as percentages of the original weight (A, B). Lungs were excised, and viral loads were calculated by quantitative PCR on days 4 and 7 postinfection (C, D). Total lung cell counts (E, F) were calculated, along with the total numbers of CD3, CD4, and CD8 (T); CD19 (B); and DX5<sup>+</sup> (NK) (D, I) cells measured in the lungs by flow cytometry on day 7 postinfection. Levels of the inflammatory cytokines IL-1 $\beta$  and IFN- $\gamma$  in the lungs (E, J) were measured by enzyme-linked immunosorbent assay on day 7 postinfection. Results are mean values  $\pm$  the standard error of the mean ( $n > 5$ ). Statistical significance was assessed by Student *t* test (\*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ).

was not in the lentiviral panel). We focused on IRF7 for the *in vivo* studies because it gave a strong knockdown *in vitro* and was central to the predicted *in silico* network. IRF7 is an amplificatory molecule responding to pattern recognition receptor detection of viral infection, inducing a further cascade of IFNs (40), and is identified as the master regulator of type I IFN-dependent immune responses (41). Previous studies have shown a role for IRF7 in human metapneumovirus (42) and influenza virus (43) infections. A recent study has demonstrated a role for IRF7 in the upregulation of RIG-I in response to RSV infection *in vitro* (44), but the present study is the first to demonstrate a central role for IRF7 in the control of RSV infection both *in vivo* and *in vitro*.

There are some limitations to this study. First, expression profiling of cells in the peripheral blood has limitations in terms of representativeness of responses in the respiratory epithelium, which RSV infects, and most of the studies are based on peripheral blood signatures. Because of the heterogeneous nature of the available databases and publications about genes associated with RSV infection, we had to make inclusion decisions that led to a slight skewing of the gene list. Where multiple gene sets were available, we chose to include genes upregulated at 24 h after infection, which may have skewed the gene set to the IFN- $\alpha/\beta$  response. Where reported, a cutoff change of 2-fold was used, but these data were not reported in several studies; likewise, not all published data had gene lists that could be incorporated into the present study. We chose a system to weight the data for the analysis giving priority to genetic association studies and *in vivo* human data over *in vitro* and mouse data. This weighting score was based on a subjective decision about the perception of the relevance of different data types to human infection. While it oversimplifies the differences both between and within different study types, once the data have been collated, other scoring systems could easily be applied to the same metadata because of the simplicity of the analysis tools used. It is of note that weighted and unweighted analyses gave similar lists. This demonstrates the power of this tool, because it can be adapted to

#### Figure Legend Continued

and 24 h later, the cells were infected with RSV expressing GFP. Cells were harvested at 48 h postinfection, and expression relative to that of control lentivirus-transfected wells was assessed. Each bar represents the mean value of three experiments  $\pm$  the standard error of the mean. Red bars represent the top upregulated genes on the literature-derived list.

different questions, integrating heterogeneous data sets; furthermore, the simplicity of the approach means that this can be performed quickly and easily. The *in vitro* screening contained only ISGs, which restricted the analysis of genes that were identified but not in this family; for example, GBP1. Finally, there was only a limited disease phenotype in the control mice because they were C57BL/6 mice, which are relatively resistant to RSV infection; these mice were used to match the gene knockout animals. One limiting factor in the data mining, in our experience, is a lack of standardization of published data sets; different papers have gene and protein lists in different file formats with different nomenclatures, and many only had lists in tables. Therefore, transcription of the data had to be done by hand. A more uniform approach to these data sets would enable more studies to be included in meta-analyses.

Understanding more about the functions of the genes that are most commonly upregulated following RSV infection may give us insight into pathways to control disease after infection. A number of papers have proposed genetic signatures of RSV infection; here, we show that the genes that are commonly upregulated as a result of RSV infection are characteristic of inflammation and viral control. There is an ongoing debate as to whether inflammation or virus-mediated pathology is the primary cause of disease after RSV infection. In the this context, one of the striking features of the pathway analysis was a divergence in the effects on disease outcome of the gene products localized to the extracellular space and those found in the nucleus and cytoplasm. Animal studies suggest that the extracellular proteins enhance disease following RSV infection (45) by increasing inflammation, while the intracellular proteins reduce disease by decreasing viral replication. In support of this, we have recently shown that antiviral gene *IFITM3* is important in the prevention of RSV infection (26, 46). In the *in vitro* studies, overexpression of *IFITM3* led to a  $68\% \pm 20.6\%$  reduction in RSV replication. Furthermore, overexpression of the chemokines CCL4, CCL5, and CCL8 had little effect on viral replication. This suggests that boosting the antiviral response without increasing inflammation would be a good strategy to control RSV disease. One potential target to achieve this would be the type I IFN response, which, if boosted, may increase the transcription of antiviral genes. However, recent studies have shown that in RSV (47) and influenza virus (48) infections, type I IFN contributes to inflammation and disease after viral infection, suggesting that there is a sweet spot of IFN production and either too little or too much can lead to a disease state. Targeting the host response may be particularly beneficial, as it is not necessarily specific to the pathogen and is less likely to induce antimicrobial resistance.

There are several other approaches for integrating heterogeneous data for immunology research (49). Selection of the tool to use depends on the desired outcome, ranging from analysis of T cell receptor and antibody repertoire analysis to network analysis and visualization. Some tools enable the integration of gene expression data with DNA variation (eQTL) (50) or against the epigenetic status of the same gene (51), both of which enable greater understanding of the processes underpinning gene regulation and expression in the immune response. For higher-level analysis, data can be merged by using network analysis tools to see novel interactions between genes (52). These are sophisticated models underpinned by statistical techniques requiring specialized skill sets and data analysis software to be performed rigorously. Our approach is not. It requires no specialist informatic skills or software (for a how-to guide, see Text S1 in the supplemental material); it uses Excel (but would work with any spreadsheet program) and a free Perl script. Additionally, it performs a different function prioritizing genes for further study. This tool generates a hypothesis-free candidate list that can be investigated further. That the genes were biologically plausible and affected RSV infection *in vitro* and *in vivo* validates it as a downselection tool. In conclusion, here we describe a novel literature data mining approach for candidate gene prioritization. It has the benefit of simplicity and is broadly applicable to a range of infectious diseases.

## MATERIALS AND METHODS

**Meta-analysis of papers and *in silico* analysis.** We selected papers published prior to July 2015 by using the search terms “RSV AND Microarray OR transcriptome OR genetic or proteome” in PubMed; other studies from previous literature searches were also included (see Table S1 in the supplemental material). Papers were not included if the data set described in the study with fold change and time point data were not easily accessible. Individual genes were included for analysis if they were reported to have a >2-fold difference from the reference group in the specific study. Gene lists were harvested from the published literature and collated in a single excel file. For analysis, genes were given subjective weighting scores based on the types of studies from which they were collated, as follows: human genetic studies, 4; human *in vivo* transcriptomic studies, 3; human *in vitro* studies, 2; murine studies, 1. A custom Perl script, countIDs (available at <https://sourceforge.net/projects/countids/>), was written to parse the gene list file generated to assign a weighted score to each gene, generating an output gene list. Figure 1 is a flowchart of the process; for a step-by-step guide to the approach, see Text S1 in the supplemental material. For visualization, genes with assigned scores of  $\geq 6$  were included for Ingenuity Pathway Analysis (IPA; Qiagen).

**Validation in a clinical cohort.** We established a case-control group comprising 27 RSV patients and 80 healthy controls. Whole blood (2.5 ml) was collected into PAXgene blood RNA tubes (PreAnalytiX; Germany), incubated for 2 h, frozen at  $-20^{\circ}\text{C}$  within 6 h of collection, and then stored at  $-80^{\circ}\text{C}$ . RNA was extracted with PAXgene blood RNA kits (PreAnalytiX; Germany) according to the manufacturer’s instructions. The integrity and yield of the total RNA were assessed with an Agilent 2100 Bioanalyzer and a NanoDrop 1000 spectrophotometer. After quantification and quality control, biotin-labeled cRNA was prepared with Illumina TotalPrep RNA amplification kits (Applied Biosystems) from 500 ng of RNA. Labeled cRNA was hybridized overnight to HumanHT-12 v4 Expression BeadChip arrays (Illumina). After washing, blocking, and staining, the arrays were scanned with an Illumina BeadArray Reader according to the manufacturer’s instructions. With Genome Studio software, the microarray images were inspected for artifacts and quality control parameters were assessed. Data were analyzed in the R language and environment for statistical computing (version 3.1.2) (53). Mean raw intensity values for each transcript were transformed to a logarithmic scale (base 2), corrected for local background intensities, and normalized by robust spline normalization. We identified the transcripts that were SDE between the RSV-infected children and the healthy control group with an adjusted *P* value of  $<0.05$  by using a linear model for transcript expression. The functions *lmFit* and *eBayes* in the R package *limma* were used to calculate statistics.

**Validation *in vitro*.** ISG library creation and screening were performed with a modified version of the assay described previously (7). A total of  $2 \times 10^5$  HEp-2 cells was seeded into 96-well plates overnight prior to transfection with  $10^5$  of the individual RFP-ISG lentiviruses in Dulbecco’s modified Eagle’s medium supplemented with 20 mM HEPES and 4 mg/ml Polybrene by spinoculation at 1,000 rpm for 1 h. Twenty-four hours later, the cells were infected with GFP-expressing RSV (RSV-GFP) (33), and 24 h after that, the cells were harvested for analysis by flow cytometry. Live/dead discrimination was performed by addition of the LIVE/DEAD Fixable Aqua Dead Cell Stain kit (Molecular Probes) prior to the acquisition of data on an LSRFortessa (BD). Viral infection was determined on the basis of the percentage of GFP-RFP double-positive cells, and the relative infectivity in each well was normalized to the samples transfected with control lentivirus only. Data were analyzed with CyAn ADP Summit 4.3.

**Validation *in vivo* with gene knockout mice.** Sex-matched 8- to 10-week-old wild-type C57BL/6N *IRF1*<sup>tm1a(EUCOMM)Wtsi</sup> and *IRF7*<sup>tm1(KOMP)Wtsi</sup> mice (54) (Wellcome Trust Sanger Institute) were maintained in accordance with United Kingdom Home Office regulations and United Kingdom Animals (Scientific Procedures) Act 1986 under project license PPL 80/2596. Animals were supplied with food and water *ad libitum* and monitored daily for signs of illness. Founder mice were phenotyped through pipelines at the Wellcome Trust Sanger Institute as described previously (9, 10).

**RSV *in vivo* studies.** RSV strain A2 (from P. Openshaw, Imperial College London) was grown in HEp-2 cells, and viral titers were determined by plaque assay. Mice were infected intranasally with  $5 \times 10^5$  PFU in a volume of 100  $\mu\text{l}$  under isoflurane anesthesia. They were weighed daily to monitor disease severity. Their lungs were removed, the smaller lobe was snap-frozen in liquid nitrogen for RNA extraction, and the remainder was homogenized by passage through 100- $\mu\text{m}$  cell strainers (Falcon). Red blood cells in the lung sample were lysed in ammonium chloride buffer, and the remaining cells were resuspended in RPMI medium with 10% fetal calf serum. Viable cell numbers were determined by trypan blue exclusion, and lung cell types were differentiated by flow cytometry as described previously (55). In brief, cells were suspended in Fc block (anti-CD16/32; BD) in phosphate-buffered saline–1% bovine serum albumin and stained with surface antibodies CD3-fluorescein isothiocyanate (BD, Oxford United Kingdom), CD4-allophycocyanin (APC; BD), CD8-APC Alexa 750 (Invitrogen, Paisley, United Kingdom), NK1.1-peridinin chlorophyll protein-Cy5.5 (BD), and CD19-eFluor 450 (eBioscience, Hatfield, United Kingdom). Cells were run on a BD FACS Aria II. Singlet lymphocyte cells were defined on the basis of their size, side scatter, and doublet discrimination, and then their immune phenotypes were analyzed on the basis of cell surface markers. RSV loads were measured by quantitative reverse transcription-PCR for the RSV L gene with primers and probes previously described (19), and L gene copy numbers were determined with an RSV L gene standard and expressed relative to the amount (micrograms) of lung RNA. The cytokines interleukin-1 $\beta$  (IL-1 $\beta$ ) and IFN- $\gamma$  in the lungs were quantified with DuoSet kits from R&D Systems.

**Statistical analysis.** Statistical analysis was performed with weighted Student *t* tests in GraphPad Prism 6.0.

**Microarray data accession number.** The microarray data from the clinical study have been uploaded to the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>) and assigned accession number [GSE80179](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80179).

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/mSystems.00051-16>.

Text S1, DOCX file, 0.01 MB.

Table S1, DOCX file, 0.1 MB.

Table S2, DOCX file, 0.01 MB.

Table S3, DOCX file, 0.01 MB.

## ACKNOWLEDGMENTS

We thank Alex Bukreyev (UTMB) and Peter Collins (NIH) for the RSV-GFP, John Schoggins and Charles Rice for the ISG library, and Gordon Dougan for advice and discussion. The data from the clinical cohort were collected by the IRIS Consortium: Imperial College London (United Kingdom; Stuart Gormley, Jethro A. Herberg, Myrsini Kaforou, Michael Levin; Hannah Shailes, and Victoria J. Wright), Micropathology Ltd. (United Kingdom; Colin Fink and Ed Sumner), Hospital Clínico Universitario de Santiago (Spain; Miriam Cebej Lopez, Antonio Salas, Antonio Justicia Grande, Irene Rivero, Alberto Gómez Carballa, Jacobo Pardo Seco, José María Martín Sánchez, Lorenzo Redondo Collazo, Carmen Rodríguez-Tenreiro, Lucia Vilanova Trillo, and Federico Martínón Torres), and Southampton (United Kingdom; Saul Faust, Jenni McCorkill, and Sanjay Patel).

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement 115308 (BioVacSafe), resources of which are composed of a financial contribution from the European Union Seventh Framework Programme (FP7/2007-2013) and an EFPIA members' in-kind contribution. Animal resources are funded by the Wellcome Trust via grant 098051.

J.T. and S.C. designed the research. J.M., C.H., J.T., and S.C. performed the *in vivo* studies. J.T., S.C., and J.H. wrote the paper. M.I. and M.D. performed the *in vitro* studies. M.K., J.H., V.W., F.M.T., and M.L. performed and analyzed the clinical studies. D.H., J.M., and M.K. performed the *in silico* studies.

## REFERENCES

- Marx V. 2013. Biology: the big challenges of big data. *Nature* **498**: 255–260. <http://dx.doi.org/10.1038/498255a>.
- Xiong Q, Ancona N, Hauser ER, Mukherjee S, Furey TS. 2012. Integrating genetic and gene expression evidence into genome-wide association analysis of gene sets. *Genome Res* **22**:386–397. <http://dx.doi.org/10.1101/gr.124370.111>.
- Chaussabel D, Baldwin N. 2014. Democratizing systems immunology with modular transcriptional repertoire analyses. *Nat Rev Immunol* **14**: 271–280. <http://dx.doi.org/10.1038/nri3642>.
- Querec TD, Akondy RS, Lee EK, Cao W, Nakaya HI, Teuwen D, Pirani A, Gernert K, Deng J, Marzolf B, Kennedy K, Wu H, Bennouna S, Oluoch H, Miller J, Vencio RZ, Mulligan M, Aderem A, Ahmed R, Pulendran B. 2009. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nat Immunol* **10**:116–125. <http://dx.doi.org/10.1038/ni.1688>.
- Litvak V, Ratushny AV, Lampano AE, Schmitz F, Huang AC, Raman A, Rust AG, Bergthaler A, Aitchison JD, Aderem A. 2012. A FOXO3-IRF7 gene regulatory circuit limits inflammatory sequelae of antiviral responses. *Nature* **490**:421–425. <http://dx.doi.org/10.1038/nature11428>.
- Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, Habjan M, Binder M, Stefanovic A, Eberle CA, Goncalves A, Bürckstümmer T, Müller AC, Fauster A, Holze C, Lindsten K, Goodbourn S, Kochs G, Weber F, Bartenschlager R, Bowie AG, Bennett KL, Colinge J, Superti-Furga G. 2012. Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature* **487**: 486–490. <http://dx.doi.org/10.1038/nature11289>.
- Schoggins JW, MacDuff DA, Imanaka N, Gainey MD, Shrestha B, Eitson JL, Mar KB, Richardson RB, Ratushny AV, Litvak V, Dabelic R, Manicassamy B, Aitchison JD, Aderem A, Elliott RM, García-Sastre A, Racaniello V, Snijder EJ, Yokoyama WM, Diamond MS, Virgin HW, Rice CM. 2014. Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity. *Nature* **505**:691–695. <http://dx.doi.org/10.1038/nature12862>.
- Karlas A, Machuy N, Shin Y, Pleissner KP, Artarini A, Heuer D, Becker D, Khalil H, Ogilvie LA, Hess S, Mäurer AP, Müller E, Wolff T, Rudel T, Meyer TF. 2010. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* **463**:818–822. <http://dx.doi.org/10.1038/nature08760>.
- Ayadi A, Birling MC, Bottomley J, Bussell J, Fuchs H, Fray M, Gailus-Durner V, Greenaway S, Houghton R, Karp N, Leblanc S, Lengger C, Maier H, Mallon AM, Marschall S, Melvin D, Morgan H, Pavlovic G, Ryder E, Skarnes WC, Selloum M, Ramirez-Solis R, Sorg T, Teboul L, Vasseur L, Walling A, Weaver T, Wells S, White JK, Bradley A, Adams DJ, Steel KP, Hrabec de Angelis M, Brown SD, Herault Y. 2012. Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project. *Mamm Genome* **23**:600–610. <http://dx.doi.org/10.1007/s00335-012-9418-y>.
- White JK, Gerdin AK, Karp NA, Ryder E, Buljan M, Bussell J, Salisbury J, Clare S, Ingham NJ, Podrini C, Houghton R, Estabel J, Bottomley JR, Melvin DG, Sunter D, Adams NC, Sanger Institute Mouse Genetics Project, Tannahill D, Logan DW, Macarthur DG, Flint J, Mahajan VB, Tsang SH, Smyth I, Watt FM, Skarnes WC, Dougan G, Adams DJ, Ramirez-Solis R, Bradley A, Steel KP. 2013. Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell* **154**:452–464. <http://dx.doi.org/10.1534/genetics.112.146647>.
- Tregoning JS, Schwarze J. 2010. Respiratory viral infections in infants:

- causes, clinical symptoms, virology, and immunology. *Clin Microbiol Rev* **23**:74–98. <http://dx.doi.org/10.1128/CMR.00032-09>.
12. Murray J, Bottle A, Sharland M, Modi N, Aylin P, Majeed A, Saxena S, Medicines for Neonates Investigator Group. 2014. Risk factors for hospital admission with RSV bronchiolitis in England: a population-based birth cohort study. *PLoS One* **9**:e89186. <http://dx.doi.org/10.1371/journal.pone.0089186>.
  13. García CG, Bhore R, Soriano-Fallas A, Trost M, Chason R, Ramilo O, Mejias A. 2010. Risk factors in children hospitalized with RSV bronchiolitis versus non-RSV bronchiolitis. *Pediatrics* **126**:e1453-1460. <http://dx.doi.org/10.1542/peds.2010-0507>.
  14. Drysdale SB, Milner AD, Greenough A. 2012. Respiratory syncytial virus infection and chronic respiratory morbidity—is there a functional or genetic predisposition? *Acta Paediatr* **101**:1114–1120. <http://dx.doi.org/10.1111/j.1651-2227.2012.02825.x>.
  15. Pribul PK, Harker J, Wang B, Wang H, Tregoning JS, Schwarze J, Openshaw PJ. 2008. Alveolar macrophages are a major determinant of early responses to viral lung infection but do not influence subsequent disease development. *J Virol* **82**:4441–4448. <http://dx.doi.org/10.1128/JVI.02541-07>.
  16. Kolli D, Gupta MR, Sbrana E, Velayutham TS, Chao H, Casola A, Garofalo RP. 2014. Alveolar macrophages contribute to the pathogenesis of hMPV infection while protecting against RSV infection. *Am J Respir Cell Mol Biol* **51**:502–515. <http://dx.doi.org/10.1165/rcmb.2013-0414OC>.
  17. Harker JA, Yamaguchi Y, Culley FJ, Tregoning JS, Openshaw PJ. 2014. Delayed sequelae of neonatal respiratory syncytial virus infection are dependent on cells of the innate immune system. *J Virol* **88**:604–611. <http://dx.doi.org/10.1128/JVI.02620-13>.
  18. Li F, Zhu H, Sun R, Wei H, Tian Z. 2012. Natural killer cells are involved in acute lung immune injury caused by respiratory syncytial virus infection. *J Virol* **86**:2251–2258. <http://dx.doi.org/10.1128/JVI.06209-11>.
  19. Tregoning JS, Yamaguchi Y, Harker J, Wang B, Openshaw PJ. 2008. The role of T cells in the enhancement of respiratory syncytial virus infection severity during adult reinfection of neonatally sensitized mice. *J Virol* **82**:4115–4124. <http://dx.doi.org/10.1128/JVI.02313-07>.
  20. Welliver TP, Garofalo RP, Hosakote Y, Hintz KH, Avendano L, Sanchez K, Velozo L, Jafri H, Chavez-Bueno S, Ogra PL, McKinney L, Reed JL, Welliver RC, Sr. 2007. Severe human lower respiratory tract illness caused by respiratory syncytial virus and influenza virus is characterized by the absence of pulmonary cytotoxic lymphocyte responses. *J Infect Dis* **195**:1126–1136. <http://dx.doi.org/10.1086/512615>.
  21. Herberg JA, Kaforou M, Gormley S, Sumner ER, Patel S, Jones KD, Paulus S, Fink C, Martinon-Torres F, Montana G, Wright VJ, Levin M. 2013. Transcriptomic profiling in childhood H1N1/09 influenza reveals reduced expression of protein synthesis genes. *J Infect Dis* **208**:1664–1668. <http://dx.doi.org/10.1093/infdis/jit348>.
  22. Mejias A, Dimo B, Suarez NM, Garcia C, Suarez-Arrabal MC, Jartti T, Blankenship D, Jordan-Villegas A, Ardura MI, Xu Z, Banchereau J, Chaussabel D, Ramilo O. 2013. Whole blood gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection. *PLoS Med* **10**:e1001549. <http://dx.doi.org/10.1371/journal.pmed.1001549>.
  23. Harker JA, Godlee A, Wahlsten JL, Lee DC, Thorne LG, Sawant D, Tregoning JS, Caspi RR, Bukreyev A, Collins PL, Openshaw PJ. 2010. Interleukin 18 coexpression during respiratory syncytial virus infection results in enhanced disease mediated by natural killer cells. *J Virol* **84**:4073–4082. <http://dx.doi.org/10.1128/JVI.02014-09>.
  24. Matthews SP, Tregoning JS, Coyle AJ, Hussell T, Openshaw PJ. 2005. Role of CCL11 in eosinophilic lung disease during respiratory syncytial virus infection. *J Virol* **79**:2050–2057. <http://dx.doi.org/10.1128/JVI.79.4.2050-2057.2005>.
  25. Tregoning JS, Pribul PK, Pennycook AM, Hussell T, Wang B, Lukacs N, Schwarze J, Culley FJ, Openshaw PJ. 2010. The chemokine MIP1alpha/CCL3 determines pathology in primary RSV infection by regulating the balance of T cell populations in the murine lung. *PLoS One* **5**:e9381. <http://dx.doi.org/10.1371/journal.pone.0009381>.
  26. Everitt AR, Clare S, McDonald JU, Kane L, Harcourt K, Ahras M, Lall A, Hale C, Rodgers A, Young DB, Haque A, Billker O, Tregoning JS, Dougan G, Kellam P. 2013. Defining the range of pathogens susceptible to Ifitm3 restriction using a knockout mouse model. *PLoS One* **8**:e80723. <http://dx.doi.org/10.1371/journal.pone.0080723>.
  27. Wong TM, Boyapalle S, Sampayo V, Nguyen HD, Bedi R, Kamath SG, Moore ML, Mohapatra S, Mohapatra SS. 2014. Respiratory syncytial virus (RSV) infection in elderly mice results in altered antiviral gene expression and enhanced pathology. *PLoS One* **9**:e88764. <http://dx.doi.org/10.1371/journal.pone.0088764>.
  28. Ravi LI, Li L, Sutejo R, Chen H, Wong PS, Tan BH, Sugrue RJ. 2013. A systems-based approach to analyse the host response in murine lung macrophages challenged with respiratory syncytial virus. *BMC Genomics* **14**:190. <http://dx.doi.org/10.1186/1471-2164-14-190>.
  29. Hennus MP, Janssen R, Pennings JL, Hodemaekers HM, Kruijsen D, Jansen NJ, Meyaard L, van Vught AJ, Bont LJ. 2012. Host response to mechanical ventilation for viral respiratory tract infection. *Eur Respir J* **40**:1508–1515. <http://dx.doi.org/10.1183/09031936.00177111>.
  30. Ioannidis I, McNally B, Willette M, Peoples ME, Chaussabel D, Durbin JE, Ramilo O, Mejias A, Flaño E. 2012. Plasticity and virus specificity of the airway epithelial cell immune response during respiratory virus infection. *J Virol* **86**:5422–5436. <http://dx.doi.org/10.1128/JVI.06757-11>.
  31. Pennings JL, Schuurhof A, Hodemaekers HM, Buisman A, de Rond LC, Widjoatmodjo MN, Luytjes W, Kimpen JL, Bont L, Janssen R. 2011. Systemic signature of the lung response to respiratory syncytial virus infection. *PLoS One* **6**:e21461. <http://dx.doi.org/10.1371/journal.pone.0021461>.
  32. Janssen R, Pennings J, Hodemaekers H, Buisman A, van Oosten M, de Rond L, Oztürk K, Dormans J, Kimman T, Hoebee B. 2007. Host transcription profiles upon primary respiratory syncytial virus infection. *J Virol* **81**:5958–5967. <http://dx.doi.org/10.1128/JVI.02220-06>.
  33. Zhang L, Peoples ME, Boucher RC, Collins PL, Pickles RJ. 2002. Respiratory syncytial virus infection of human airway epithelial cells is polarized, specific to ciliated cells, and without obvious cytopathology. *J Virol* **76**:5654–5666. <http://dx.doi.org/10.1128/JVI.76.11.5654-5666.2002>.
  34. Papac-Milicevic N, Breuss JM, Zaujec J, Ryban L, Plyushch T, Wagner GA, Fenzl S, Dremsek P, Cabaravdic M, Steiner M, Glass CK, Binder CJ, Uhrin P, Binder BR. 2012. The interferon stimulated gene 12 inactivates vasculoprotective functions of NR4A nuclear receptors. *Circ Res* **110**:e50–e63. <http://dx.doi.org/10.1161/CIRCRESAHA.111.258814>.
  35. Hsieh WL, Huang YH, Wang TM, Ming YC, Tsai CN, Pang JH. 2015. IFI27, a novel epidermal growth factor-stabilized protein, is functionally involved in proliferation and cell cycling of human epidermal keratinocytes. *Cell Prolif* **48**:187–197. <http://dx.doi.org/10.1111/cpr.12168>.
  36. Ostler N, Britzen-Laurent N, Liebl A, Naschberger E, Lochnit G, Ostler M, Forster F, Kunzelmann P, Ince S, Supper V, Praefcke GJ, Schubert DW, Stockinger H, Herrmann C, Stürzl M. 2014. Gamma interferon-induced guanylate binding protein 1 is a novel actin cytoskeleton remodeling factor. *Mol Cell Biol* **34**:196–209. <http://dx.doi.org/10.1128/MCB.00664-13>.
  37. Ibsen MS, Gad HH, Thavachelvam K, Boesen T, Després P, Hartmann R. 2014. The 2'-5'-oligoadenylate synthetase 3 enzyme potentially synthesizes the 2'-5'-oligoadenylates required for RNase L activation. *J Virol* **88**:14222–14231. <http://dx.doi.org/10.1128/JVI.01763-14>.
  38. Zhao C, Denison C, Huibregtse JM, Gygi S, Krug RM. 2005. Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. *Proc Natl Acad Sci U S A* **102**:10200–10205. <http://dx.doi.org/10.1073/pnas.0504754102>.
  39. González-Sanz R, Mata M, Bermejo-Martín J, Álvarez A, Cortijo J, Melero JA, Martínez I. 2016. ISG15 is upregulated in respiratory syncytial virus infection and reduces virus growth through protein ISGylation. *J Virol* **90**:3428–3438. <http://dx.doi.org/10.1128/jvi.02695-15>.
  40. Ning S, Pagano JS, Barber GN. 2011. IRF7: activation, regulation, modification and function. *Genes Immun* **12**:399–414. <http://dx.doi.org/10.1038/gene.2011.21>.
  41. Honda K, Yanai H, Negishi H, Asagiri M, Sato M, Mizutani T, Shimada N, Ohba Y, Takaoka A, Yoshida N, Taniguchi T. 2005. IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature* **434**:772–777. <http://dx.doi.org/10.1038/nature03464>.
  42. Spahn KM, Loh Z, Lynch JP, Ullah A, Zhang V, Baturcam E, Werder RB, Khajornjiraphan N, Rudd P, Loo YM, Suhrbier A, Gale M, Jr., Upham JW, Phipps S. 2014. IRF-3, IRF-7, and IPS-1 promote host defense against acute human metapneumovirus infection in neonatal mice. *Am J Pathol* **184**:1795–1806. <http://dx.doi.org/10.1016/j.ajpath.2014.02.026>.
  43. Crotta S, Davidson S, Mahlakoiv T, Desmet CJ, Buckwalter MR, Albert ML, Staeheli P, Wack A. 2013. Type I and type III interferons drive redundant amplification loops to induce a transcriptional signature

- in influenza-infected airway epithelia. *PLoS Pathog* **9**:e1003773. <http://dx.doi.org/10.1371/journal.ppat.1003773>.
44. **Fang L, Choudhary S, Tian B, Boldogh I, Yang C, Ivanciuc T, Ma Y, Garofalo RP, Brasier AR.** 2015. Ataxia telangiectasia mutated kinase mediates NF- $\kappa$ B serine 276 phosphorylation and interferon expression via the IRF7-RIG-I amplification loop in paramyxovirus infection. *J Virol* **89**:2628–2642. <http://dx.doi.org/10.1128/JVI.02458-14>.
  45. **Openshaw PJ, Tregoning JS.** 2005. Immune responses and disease enhancement during respiratory syncytial virus infection. *Clin Microbiol Rev* **18**:541–555. <http://dx.doi.org/10.1128/CMR.18.3.541-555.2005>.
  46. **Zhang W, Zhang L, Zan Y, Du N, Yang Y, Tien P.** 2015. Human respiratory syncytial virus infection is inhibited by IFN-induced transmembrane proteins. *J Gen Virol* **96**:170–182. <http://dx.doi.org/10.1099/vir.0.066555-0>.
  47. **Goritzka M, Durant LR, Pereira C, Salek-Ardakani S, Openshaw PJ, Johansson C.** 2014. Alpha/beta interferon receptor signaling amplifies early proinflammatory cytokine production in the lung during respiratory syncytial virus infection. *J Virol* **88**:6128–6136. <http://dx.doi.org/10.1128/JVI.00333-14>.
  48. **Davidson S, Crotta S, McCabe TM, Wack A.** 2014. Pathogenic potential of interferon  $\alpha\beta$  in acute influenza infection. *Nat Commun* **5**:3864. <http://dx.doi.org/10.1038/ncomms4864>.
  49. **Kidd BA, Peters LA, Schadt EE, Dudley JT.** 2014. Unifying immunology with informatics and multiscale biology. *Nat Immunol* **15**:118–127. <http://dx.doi.org/10.1038/ni.2787>.
  50. **Shabalin AA.** 2012. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**:1353–1358. <http://dx.doi.org/10.1093/bioinformatics/bts163>.
  51. **Pepke S, Wold B, Mortazavi A.** 2009. Computation for ChIP-seq and RNA-seq studies. *Nat Methods* **6**:S22–S32. <http://dx.doi.org/10.1038/nmeth.1371>.
  52. **Arazi A, Pendergraft WF III, Ribeiro RM, Perelson AS, Hacohen N.** 2013. Human systems immunology: hypothesis-based modeling and unbiased data-driven approaches. *Semin Immunol* **25**:193–200. <http://dx.doi.org/10.1016/j.smim.2012.11.003>.
  53. **R Development Core Team.** 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
  54. **Skarnes WC, Rosen B, West AP, Koutourakis M, Bushell W, Iyer V, Mujica AO, Thomas M, Harrow J, Cox T, Jackson D, Severin J, Biggs P, Fu J, Nefedov M, de Jong PJ, Stewart AF, Bradley A.** 2011. A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**:337–342. <http://dx.doi.org/10.1038/nature10163>.
  55. **Russell RF, McDonald JU, Ivanova M, Zhong Z, Bukreyev A, Tregoning JS.** 2015. Partial attenuation of respiratory syncytial virus with a deletion of a small hydrophobic gene is associated with elevated interleukin-1 $\beta$  responses. *J Virol* **89**:8974–8981. <http://dx.doi.org/10.1128/JVI.01070-15>.