Systems Immunology Reveals Markers of Susceptibility to West Nile Virus Infection

Feng Qian,* Gautam Goel,b Hailong Meng,c Xiaomei Wang,a Fuping You,a Lesley Devine,a Khadir Raddassi,f Melissa N. Garcia,f Kristy O. Murray,c Christopher R. Bolen,d* Renaud Gaujoux,k Shai S. Shen-Orr,f David Hafler,d Erol Fikrig,a,d Ramnik Xavier,b Steven H. Kleinstein,c,g Ruth R. Montgomeryb

Section of Rheumatology,c Interdepartmental Program in Computational Biology and Bioinformatics, Section of Infectious Disease,a Department of Laboratory Medicine,c Department of Neurology,f Department of Pathology,f and Howard Hughes Medical Institute,b Yale School of Medicine, New Haven, Connecticut, USA; Baylor College of Medicine, Houston, Texas, USA; Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, Massachusetts, USA; Department of Immunology, Rappaport Institute of Medical Research, Faculty of Medicine and Faculty of Biology, Technion, Haifa, Israel

West Nile virus (WNV) infection is usually asymptomatic but can cause severe neurological disease and death, particularly in older patients, and how individual variations in immunity contribute to disease severity is not yet defined. Animal studies identified a role for several immunity-related genes that determine the severity of infection. We have integrated systems-level transcriptional and functional data sets from stratified cohorts of subjects with a history of WNV infection to define whether these markers can distinguish susceptibility in a human population. Transcriptional profiles combined with immunophenotyping of primary cells identified a predictive signature of susceptibility that was detectable years after acute infection (67% accuracy), with the most prominent alteration being decreased IL1B induction following ex vivo infection of macrophages with WNV. Deconvolution analysis also determined a significant role for CXCL10 expression in myeloid dendritic cells. This systems analysis identified markers of pathogenic mechanisms and offers insights into potential therapeutic strategies.

Individual variations in immune status and function determine responses to infection and contribute to disease severity and outcome. Recent advances in high-throughput and bioinformatics technology now allow detailed analysis of complex interactions to generate a systems-level understanding of disease susceptibility (1–4). Definition of the biological signatures underlying immune responsiveness requires quantification of elements of innate and adaptive immune responses, a cohort exhibiting a range of clinical outcomes, and a sufficient sample size to accommodate natural variations in responsiveness. We have undertaken a comprehensive profile of stable characteristics of individual immune cell frequencies and gene expression (5, 6) to define markers that identify key mechanisms of resistance and susceptibility to virus infection.

West Nile virus (WNV) is a mosquito-borne, enveloped, positive-strand RNA virus belonging to the family Flaviviridae, which includes yellow fever virus, hepatitis C virus, and dengue virus (7). The emergence of WNV in North America was first documented in 1999 in New York City, and over the past decade, WNV has become established throughout the United States and has spread into Canada, Mexico, and the Caribbean. CDC reports indicate infections of >31,000 people, including >1,200 fatalities, and the cumulative incidence of WNV infection may reach 3 million people (7, 8). WNV patients exhibit considerable variation in clinical responses to infection, and there are currently no FDA-approved treatments available. Infants, the immunocompromised, and elderly individuals are more susceptible to neurological involvement that may result in death (7, 8). Although encephalitis generally occurs in only a small subset of patients, 50% of reported cases in 2012 were of the more severe forms of neuroinvasive disease, including encephalitis (7).

Control of WNV infection by the immune system is multifactorial, including viral recognition receptors (Toll-like receptors [TLRs] and RIG-I-like receptors [RLRs]), control of the permeability of the blood-brain barrier, and both innate and adaptive immune determinants. Certain HLA types, chemokines, and interferon pathway elements are associated with a risk of more severe outcomes in humans, and multiple pathways have been investigated in murine models (7, 9). In particular, the severity of WNV infection is associated with genetic polymorphisms in the interferon response pathway elements 2′-5′-oligoadenylate synthetase 1 (OAS), interferon regulatory factor 3 (IRF3), and MX-1 (7), and upregulation of type I interferons is critical for immediate antiviral defense pathways and to generate an effective adaptive T cell- and B cell-mediated sustained immune response (10, 11). Despite these advances, and certain age-related susceptibilities

Received 25 July 2014 Returned for modification 15 September 2014 Accepted 17 October 2014

Address correspondence to Ruth R. Montgomery, ruth.montgomery@yale.edu.
* Present address: Feng Qian, State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, China; Christopher R. Bolen, Stanford University, Beckman Center, Stanford, California, USA.
† F.G., G.G., and H.M. contributed equally to this work.
‡ H.K. and R.R.M. contributed equally to this work.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/CVI.00508-14.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.
doi:10.1128/CVI.00508-14
(12–14), the mechanisms that lead to severe infection remain poorly defined (7).

We have recently shown that antibody levels were not significantly different between subjects with a history of asymptomatic infection and those with severe infection and that subjects with a history of severe infection had significantly lower levels of serum interleukin-4 (IL-4) (15). Identifying this difference suggests that sensitization is required for severe disease.

In the present study, we hypothesized that innate immune responses are associated with severity of WNV infection in humans. To test this hypothesis, we studied a cohort of blood donors who had been diagnosed with either asymptomatic or severe WNV infection. The goal of this study was to identify molecular and cellular signatures associated with susceptibility and resistance to severe WNV infection. Through bioinformatic analysis of gene expression profiling and automated multidimensional flow cytometry analyses, we identify molecular signatures defining individual immune responses. We then validated these signatures in a second, independent cohort of blood donors.

**Materials and Methods**

**Human subjects.** Blood was obtained with written informed consent under guidelines approved by the Human Investigations Committees of the University of Texas Health Science Center, Baylor College of Medicine, and Yale University School of Medicine. Donors had no acute illness and took no antibiotics or nonsteroidal anti-inflammatory drugs at the time of sampling (14). Previous history of WNV infection was determined clinically and validated by positive immunoblot results (12, 17). Donors included asymptomatic individuals and patients with severe disease (none with mild disease), defined as invasive neurological illness such as encephalitis or meningitis (Table 1). Asymptomatic donors were identified by a rapid nucleic acid test at the blood bank or by immunoblotting (12). Subjects in the disease severity groups were not statistically different with regard to age, gender, or race in this study. Blood samples from both recruiting sources were collected in cell preparation tubes (Becton Dickinson and Co., Franklin Lakes, NJ), centrifuged within 2 h of collection, and processed the next day. Each donor was not assessed by each assay, and samples were randomly chosen over >2 years for assays under study at the time of recruitment. Complete data from this study will be made available through the NIAID ImmPort data repository (study identification number SDY58).

**Preparation of blood cells.** Peripheral blood mononuclear cells (PBMCs) were suspended in RPMI 1640 supplemented with 20% (vol/vol) human serum (Lonza, MD), 100 U/ml penicillin, and 100 µg/ml streptomycin (Invitrogen, CA) (12). PBMCs were used in suspension at 2 × 10^7 cells/ml or plated at 5 × 10^6 cells/35-mm well. Nonadherent cells were washed away after 2 h, and adherent monocytes were incubated for 6 to 8 days to mature into macrophages, as described previously (12, 18).

**WNV strains and infections.** Virulent WNV (CT-2741), provided by John Anderson, Connecticut Agricultural Experiment Station, New Haven, CT (19), was passaged in Vero cells, and viral PFU were quantified by plaque assays (12). CT-2741 has been shown to be lethal in murine models of WNV infection (20, 21). WNV studies were conducted in a biosafety level 3 facility licensed by the State of Connecticut and Yale University. Primary PBMCs and macrophages from WNV donors were incubated in medium alone (mock), infected with WNV (multiplicity of infection [MOI] of 1), or stimulated with poly(I·C) (50 µg/ml; InvivoGen, San Diego, CA) for 24 h. Quantitative PCR (qPCR) of the WNV envelope gene from the samples for microarray analysis validated efficient viral infection (data not shown).

**Flow cytometry.** For multiparameter robotic fluorescence-activated cell sorter (FACS) analysis, fresh PBMCs were frozen in 90% fetal bovine serum (FBS) containing 10% dimethyl sulfoxide (DMSO) and stored in liquid N2 for batched analysis. For assays, cells were thawed, washed, and incubated in wells with 8 antibody panels (5 panels defined by the HIPC (22)) by using a custom-programmed BioMek robotic platform, and labeling was detected using an LSR Fortessa instrument (BD Biosciences). For intracellular cytokine studies, fresh PBMCs (1 × 10^6 cells/96-well plate) in a solution containing RPMI–10% FBS and antibiotics were incubated, as described previously (14), for 6 h in medium alone or with ligands for TLR4 (lipopolysaccharide [LPS] [10 ng/ml]),TLR7/8 (R848 [10 µM]; Invivogen), and TLR3 [poly(I·C) (50 µg/ml)], with brefeldin A (10 µg/ml; Sigma) for the last 3 h. PBMCs were labeled at 4°C as described previously (14), fixed, and frozen at −80°C for batched analysis. For assays, cells were thawed; permeabilized; labeled for alpha interferon (IFN-α) (Antigenix America, NY), tumor necrosis factor alpha (TNF-α), and IL-6 (BD Biosciences, CA); and detected using an LSR II instrument (BD Biosciences); and analyzed by using FlowJo software (Tree Star, OR), as described previously (14). Statistical tests were two tailed, and a P value of <0.05 was considered significant. Multivariable analyses were performed by using SAS version 9.2 (SAS Institute, Cary, NC), and bivariate analyses were performed by using the Prism 4.03 biostatistics package (Graphpad, CA). Differences noted remained significant after adjusting for age, gender, and recruitment site.

**RNA preparation, qPCR, microarrays, and Nanostring analysis.** Total RNA was harvested from PBMCs and macrophages by using the RNeasy minikit (Qiagen, CA), and cDNA was synthesized by using the AffinityScript Multi Temperature cDNA synthesis kit (Stratagene, TX). Primers and probes for qPCRs were obtained from Applied Biosystems or as described previously (14). Batched samples were amplified for 60 cycles (iCycler; Bio-Rad, CA), and duplicate values were normalized to β-actin values. Levels of WNV in samples from disease groups were not significantly different by PCR. Total RNA from paired macrophage and PBMC samples was assessed by using the Illumina HumanHT-12 v4 Beadchip.
whole-human genome expression array at the Keck Biotechnology Resource Laboratory of Yale University. RNA quality for Nanostring analysis was verified by an RNA integrity number (RIN) of >7 (average value of 9.5 for 180 samples). Nanostring nCounter analysis was performed at the David H. Murdock Research Institute (DHMRI) (Kannapolis, NC).

Microarray analysis. Raw expression data were normalized by using the quantile method provided by the lumi package in R/Bioconductor (23). The batch effect was included as an independent variable in the linear model of LIMMA (24). Differentially expressed genes were identified by an absolute fold change and by a statistically significant change in expression, as determined by LIMMA, using a Benjamini-Hochberg false discovery rate (FDR). Responding genes with an absolute fold change of ≥1.5 and an FDR cutoff of <0.05 were applied for both PBMCs and macrophages. For selecting genes for Nanostring analysis from PBMCs, an absolute fold change of ≥1.25 and an FDR cutoff of <0.25 between the disease groups for baseline or stimulated gene expression were used; for selecting genes from macrophages, an absolute fold change of ≥1.5 and an FDR cutoff of <0.05 were used.

Normalization of Nanostring panels. Data sets from the Nanostring platform were normalized with 12 invariant genes selected from 50 housekeeping genes; they represent both lower and higher expression levels and were the most invariant genes across cell type and treatment conditions. Quality control measures and normalization procedures were implemented according to the manufacturer’s instructions. Differential expression was determined by LIMMA using a Benjamini-Hochberg FDR cutoff of <0.05.

Development of a susceptibility classifier. A support vector machine (SVM) model was constructed to predict the clinical outcome of WNV infection (asymptomatic versus severe) by using Nanostring gene expression fold changes (mock versus ex vivo WNV infected) as features. The SVM was implemented by using the R package kernlab with a radial basis function (RBF) kernel and automatic sigma estimation (25). The accuracy of the model was computed by using 3-fold cross-validation, which was repeated 1,000 times. For each repeat, subjects with asymptomatic infection and those with severe infection were randomly divided into 3 groups of approximately equal size. Using data from the two training folds, signature genes were identified through differential expression analysis (P < 0.05), and these genes were used to construct an SVM model, which was then evaluated on the testing fold to estimate accuracy. The significance of this approach was evaluated by comparing the distribution of accuracies (over 1,000 cross-validation repeats) with that obtained by applying the same procedure to permuted data (i.e., with subject phenotypes randomly shuffled for each repeat) using a t test. The final susceptibility classifier was built by constructing an SVM model using differentially expressed genes (P < 0.05) from the full data set. To evaluate the poly(1-C) response measured by Nanostring analysis, we measured accuracy using the same SVM model developed for the WNV response. Significance was determined by comparing this accuracy to a null distribution created by using this same model on permuted data (i.e., with subject phenotypes randomly shuffled for each repeat). To evaluate the poly(1-C) response measured by microarray analysis, the SVM model was rebuilt by using microarray expression data restricted to the genes included in the WNV response SVM model. In this case, accuracy and significance were estimated by using cross-validation and permutation tests, as described above.

Deconvolution analysis. Deconvolution-derived estimates of cell subset-specific differential expression were performed through the CellMix package (26) on the log2-transformed Nanostring data by using cell-specific significance analysis of microarrays (csSAM) methodology (27). csSAM integrates sample composition heterogeneity (e.g., as measured by flow cytometry) and global gene expression into a linear regression framework to detect differential gene expression within each constituting cell type. Briefly, differential expression between two groups of samples is detected for each cell type by (i) jointly estimating average cell type-specific expression profiles within each group separately and (ii) performing a permutation test based on significance analysis of microarrays (SAM) statistics of the group differences in each gene and cell type to provide an FDR of true differential expression (27). Standard differential analysis on global gene expression was performed by using SAM, as implemented in the csSAM package, using default values, which computes SAM statistics and FDRs for each gene on median-centered data.

Microarray data accession number. The data files for this study have been deposited in the GEO database under GEO series accession number GSE46681.

RESULTS

To identify factors associated with susceptibility to severe WNV infection, we enrolled two independent cohorts of study participants (n = 90 total) with a defined history of WNV infection. Both groups of subjects were recruited from the same sites according to identical recruitment criteria and stratified into asymptomatic (n = 41) and severe (neurological) (n = 49) disease groups, as defined by CDC clinical criteria for neurological involvement and/or laboratory testing of viral load at the time of infection (17). Importantly, we did not include unexposed healthy controls whose anti-WNV efficiency is undefined or any mild (nonneurological) cases of WNV. Participants were healthy at the time of blood donation, were not different with regard to gender or age, and were predominantly white (88.9%) (Table 1). To identify characteristic individual differences in the immune state that were likely to exist at the time of WNV infection, subjects were enrolled considerably after resolution of the acute response (for severe infection, 5.9 ± 3.1 years after recovery from infection, with a range of 0.3 to 9.8 years; for asymptomatic infection, 0.7 ± 1.2 years, with a range of 0.2 to 8.1 years), long after viral RNA is undetectable (28). Potential susceptibility markers were derived from a multisystem analysis of gene expression profiling, immunophenotyping, and in vitro stimulation experiments (Fig. 1).

Differences in gene expression and pathway activity between stratified patient cohorts. To investigate altered cell states that may be associated with a differential response to infection, we collected primary PBMCs and macrophages from an initial cohort WNV subjects with asymptomatic (n = 21) and severe (n = 18) infection and assessed global gene expression using microarray analysis (Fig. 1B). When we previously compared baseline gene expression levels in PBMCs from this cohort by using loose criteria, we identified 105 genes that displayed altered expression levels at a false discovery rate (FDR) of 25% (15), but we did not identify any statistically significant differences between disease groups using a more stringent cutoff of a 5% FDR. As previous studies have shown that stimulating cells ex vivo can identify differences that are not apparent at baseline (29), we infected PBMCs and macrophages with WNV ex vivo for 24 h (MOI = 1). The viral burden was quantified by qPCR of the WNV envelope gene and was not different between subjects with a history of asymptomatic disease and those with severe disease (P > 0.1), making the two groups comparable. We identified many changes to global gene expression following WNV infection (1,564 genes at an FDR of <0.05) (Fig. 1B), including both cell type-specific and shared anti-WNV gene expression, and the majority of genes differentially expressed were upregulated (see Fig. S1 in the supplemental material). Notably, only one gene (TME1M38) was differentially regulated after WNV infection in PBMCs (upregulated with an average log2-fold change of 0.60 for subjects with a history of asymptomatic infection and downregulated with an average log2-fold change of 0.24 for subjects with a history of severe infection); no genes in mac-
Phages were differentially associated with a history of asymptomatic or severe WNV infection after correction for multiple testing (FDR $\leq 0.05$) (see Table S1 in the supplemental material). While our single-gene analysis did not detect many significant genes, manual inspection of the data identified several biologically interesting candidates among the most differentially expressed genes (see Table S1 in the supplemental material).

To increase statistical power to detect biologically relevant patterns, we conducted a transcriptional pathway analysis designed to detect coordinated, but possibly small, changes in sets of related genes. Using quantitative set analysis for gene expression (QuSAGE [30]), we identified KEGG pathways that were significantly differentially activated between subjects with a history of asymptomatic infection and those with severe infection, including 8 KEGG pathways in PBMCs (FDR $< 0.05$) (Fig. 2) but none in macrophages (FDR $< 0.05$). The two pathways with the greatest significance and difference between groups were chemokine signaling and NOD-like receptor signaling pathways, which were both significantly higher in subjects with a history of asymptomatic infection (pathway activity 0.12- and 0.21-fold higher in asymptomatic individuals, respectively). We also saw a statistically significant although smaller change in the focal adhesion pathway, which was lower in the group with severe infection. Chemokines and adhesion molecules have both been recognized as being relevant in severe infection with WNV for mediating the entry of WNV into the brain (7). Thus, subjects with a history of...
asymptomatic infection and those with severe WNV infection were associated with distinct transcriptional responses to WNV infection in vitro.

Immunophenotypic and functional differences between cells from stratified patient cohorts. In our transcriptional analysis, we found that a larger number of targets were identified from PBMCs than from macrophages, most likely reflecting complex interactions that occur in the mixed cell population and perhaps suggesting that these cell types may account for more of the variation in clinical responses. To examine differences in cell subset frequency that may contribute to susceptibility to WNV infection between groups, we employed multiparameter flow cytometry of PBMCs to profile samples from both cohorts of subjects that had sufficient cells (Fig. 1D), which included 30 subjects with a history of severe infection and 21 asymptomatic subjects (Fig. 1E). Samples were quantified by using 8 panels of antibody markers, including 5 designed by the HIPC (22), to cover a broad array of immune cell types (T cells, B cells, NK cells, Th17 cells, monocytes/dendritic cells [DCs], and Tregs). Importantly, repeated samples were highly consistent across time for 7 donors whose PBMCs were drawn twice at a 6-month interval (average R of 0.94) (Fig. 3A), suggesting that these measurements are a stable indication of the state of an individual’s immune system (at least over several months), and we thus used them as a proxy for the immune state at the time of WNV exposure. We also found broad overall consistency in subset frequencies across subjects (Fig. 3B), consistent with data from previous studies of cell subsets (5), which might be expected, as all subjects were healthy at the time of recruitment. Nevertheless, 14 of the 88 cell subsets compared between the severely infected and asymptomatic disease groups were identified as having significant differences reflecting both innate and adaptive cell types (P < 0.05) (Fig. 3C). Of note, subjects with a history of severe disease had lower levels of plasmacytoid DCs (pDCs) (CD123+) (P < 0.05), which are critical for antiviral responses. In the profiling of T cells, severely infected subjects showed elevated levels of regulatory T cells (P < 0.01), although these cells were previously shown to have a protective role in acute infection in murine models and in acute asymptomatic WNV patients and patients with mild WNV infection (possibly mediated by Tim3+ Tregs [31, 32]). Although these individual subset differences did not remain significant after correcting for multiple-hypothesis testing (FDR > 0.05), these results raise the possibility that susceptibility to severe disease correlates with an altered PBMC composition.

Having identified alterations in subset frequencies of monocytes and pDCs of the innate immune system with potential roles in resistance to infection with WNV (Fig. 3C), we sought to quantify functional differences in initial viral recognition and innate immune responses in these cells. We stimulated primary immune cells of subjects from both disease groups with ligands of Toll-like receptors that mimic viral ligands and bacterial LPS, a ligand for TLR4. Compared to asymptomatic donors, we detected reduced levels of TNF from monocytes (Fig. 3D) and of IFN-α from pDCs (Fig. 3E) as well as elevated levels of TNF and IL-6 from myeloid DCs (mDCs) (Fig. 3F and G) of donors with a history of severe infection. Thus, a history of severe disease was also associated with
Identification of WNV susceptibility genes using a focused Nanostring panel. To further explore candidate susceptibility genes identified in microarrays in the first cohort, we designed a focused gene panel for Nanostring and tested samples from the independently enrolled second subject cohort (also used for immunophenotyping) (Fig. 1F). As Nanostring has an improved dynamic range compared with microarrays, we included probes for the most promising candidates from our microarray single-gene differential expression analysis of WNV disease groups. Comparison of microarrays of untreated samples at baseline under loose FDR conditions showed a total of 140 genes that were altered between asymptomatic and severely infected subjects, and these were chosen for Nanostring analysis (131 genes in PBMCs [FDR < 0.25] and 9 genes in macrophages [FDR < 0.45]) (see Table S2 in the supplemental material). Of these, 85 genes were part of our previously reported 105-gene signature that was based on the same FDR cutoff but looser fold change criteria on the same samples (15). Many genes were equivalent at baseline between cohorts, with differential expression being evident only in response to perturbations such as infection with WNV. One example is ELF4, a positive regulator for IFN production that we recently identified which plays a role in signaling through the MAVS-TBK1 complex (34). The level of induction of ELF4 by WNV was significantly higher in macrophages of subjects with severe infection than in macrophages of asymptomatic subjects, which may contribute to more severe infection through increases...
in the expression levels of multiple cytokines (fold induction by qPCR of 2.5 versus 1.7; \( P < 0.04; n = 15/\text{group} \)) (see Fig. S2 in the supplemental material). Our Nanostring panel also includes genes chosen from other sources, such as markers that have been shown in the murine model to play a role in the dissemination of WNV in vivo or in anti-WNV immune responses (MMP9 [35], SOCS2 [12], IFITM3 [36], and CXCL1 [37]). We included genes that we identified through transcriptome sequencing (RNA-Seq) analysis (18) as being critical for resistance to infection by WNV (e.g., CLEC5a and TNFAIP), IL-8 responses (15), and IL-4-induced genes (e.g., CIORF122, ALOX5, CXCR7, DNAJB5, LILRA5, MARCKSL1, TPST1, UBE2Q1, and ZSWIM4), since we have shown that the level of IL-4 is lower in serum of severely infected WNV patients than in asymptomatic subjects (15). In addition, we included 36 candidate genes in our focused gene panel based on prior biological knowledge, such as those identified by genomic studies in murine models and human genome-wide association studies (GWAS) (e.g., SEMA7A, CASP12, RFC1, ANPEP, and SCN1A) (7) as well as those with known antiviral involvement (e.g., IRF3, OAS1, and MX-1). In order to normalize the gene expression levels across chips, 50 “housekeeping” genes were also included based on their near-constant expression levels across cell types and conditions in our microarray studies. Overall, we selected 467 genes to generate a Nanostring panel for quantitative monitoring of gene expression (see Table S2 in the supplemental material).

The genes selected for the focused Nanostring panel represent both known and novel gene targets (Fig. 1C; see also Table S2 in the supplemental material). The relevance of these genes to the differential outcomes of WNV infection was investigated in the independent second cohort of stratified WNV subjects (\( n = 19 \) severe; \( n = 16 \) asymptomatic), consisting of subjects assessed for immunophenotyping but not included in the microarray experiments used to design the gene panel (Fig. 1D). PBMC and macrophage samples were infected with WNV ex vivo as for the microarray analysis, and results from the microarray and Nanostring platforms showed excellent concordance (\( R^2 > 0.8 \)) (Fig. 4A and data not shown). As expected, the fold changes observed by Nanostring analysis were often higher than those observed by microarray analysis (both upregulated and downregulated) due to the increased dynamic range of the platform. Nanostring analysis identified 7 genes that were significantly differentially expressed in PBMCs between the disease groups after infection with WNV (FDR < 0.05) (IL2RA, OLR1, ZC3H12A, KCTD12, TLR8, IFNG, and SLAMF7) (see Fig. S3A in the supplemental material). In addition, one gene, IL1B (\( P < 1.2e^{-5} \)), was significantly differentially expressed in macrophages between disease groups after infection with WNV (FDR < 0.05) (see Fig. S3B in the supplemental material).

We next asked whether the gene expression fold changes following WNV infection in macrophages could predict the clinical phenotype of subjects (asymptomatic and with severe infection). Using a cross-validation approach, we demonstrated that a support vector machine (SVM) model constructed by using genes that meet a differential expression threshold \( P \) value of <0.05 could predict phenotypes of individual subjects with 67% accuracy. Although only a moderate value, the predictive ability was highly significant (\( P < 2.2e^{-16} \)), thus demonstrating that our perturbation profiling approach can distinguish clinically relevant responders in the absence of ongoing infection. A similar analysis of PBMC responses had an average cross-validation accuracy of 67%. Overall, this analysis identified a susceptibility signature including 59 genes in PBMCs (see Fig. S3A in the supplemental material) and 21 genes in macrophages (see Fig. S3B in the supplemental material) based on the in vitro response of blood samples to WNV stimulation. As expected, the genes in these susceptibility signatures were able to separate the disease groups by principal component analysis (PCA) (Fig. 4B and data not shown).

While discovery of a susceptibility signature offers an important screening tool, in vitro infection with WNV is not feasible in most clinical settings. Thus, we examined whether differences in susceptibility could be distinguished in our cohorts by stimulation of macrophages with poly(I:C), the ligand for TLR3 and a model ligand for viral infections (38). To accomplish this, samples from our second cohort were stimulated in parallel with poly(I:C) for testing by Nanostring analysis (\( n = 16 \) asymptomatic; \( n = 19 \) severe) (see Table S2 in the supplemental material). We found that...
the SVM model constructed by using the WNV infection response could also predict individual-subject phenotypes from the poly(I·C) response (61% accuracy; \(P = 2.2 \times 10^{-16}\)). Furthermore, when we examined the poly(I·C) response from samples in our first cohort measured by microarray analysis, these same genes could also be used to distinguish the disease groups (63% average 3-fold cross-validation accuracy; \(P = 2.2 \times 10^{-16}\)), even though none of these genes were found to be individually differentially expressed.

Integration of cell subset frequencies and transcriptional profiles. Since PBMC gene expression profiles represent a mixture of cell types, standard differential expression analysis may miss cell subset-specific differences. In our study, the same samples from the second cohort were assessed by multiparameter flow and Nanostring transcriptional profiling (Fig. 1D), allowing rare and valuable integration of high-dimensional data for correlation with the clinical outcomes of our stratified cohort. To search in depth for such differences, we employed a deconvolution approach to identify genes with different baseline expression levels between the asymptomatic \((n = 16)\) and severe \((n = 19)\) infection groups in the second cohort within specific cell types (27). Indeed, while standard analysis did not detect any differences, deconvolution analysis identified significant differences in baseline expression levels for CXCL10 in mDCs (FDR < 0.05) (Fig. 5). CXCL10 was also significantly upregulated in PBMCs and macrophages following WNV infection (FDR < 0.001). Three other genes were also identified as being moderately differentially expressed (FDR < 0.2) in a cell subset-specific manner: both TNFRSF9 and GBP1 were upregulated among asymptomatic subjects in mDCs, and the RGS18 level was higher in CD8 T cells of severe subjects. Thus, along with altered innate cell type frequencies, we identified genes with significantly altered expression in mDCs when we compared subjects with a history of asymptomatic WNV infection and those with severe WNV infection.

DISCUSSION

High-dimensional multisystem profiling approaches have been successful in determining immunological outcomes such as vaccine responsiveness (1–4). Here, our multisystem study ex vivo of primary cells from a stratified cohort has revealed determinants that contribute to a susceptible profile. Our studies permit an informed prediction of critical factors in antiviral responses, and such translational studies are essential, as animal models may not reflect pathophysiological responses (39). The methods employed here, including validated multiparameter immunophenotyping and genome-wide transcriptional profiling, provide data for discriminating between susceptible and resistant WNV subjects. The relevance of the genes identified here is heightened, as the critical differences were identified in a healthy convalescent cohort based on severity at the time of illness and were evident from the stable characteristics of individual immune cell frequencies and gene expression (5,6). We identified a predictive gene signature of susceptibility (67% accuracy) with a high level of significance to distinguish anti-WNV response patterns. The distinguishing susceptibility elements were not apparent at baseline but were revealed by responses to stimulation both in transcriptional profiles and in functional cellular responses.

Of the genes identified in the susceptibility signature, while some are well characterized, several were not previously known to be involved in anti-WNV responses. Critical anti-WNV factors identified here in primary human cells (7, 10, 11, 31) include factors such as the IFITM protein family, IFI27 from our signature, that contribute broadly to cellular resistance to WNV and dengue virus (36,40). Chemokines and adhesion molecules play a critical role in access to the brain in severe neurological infection by WNV (7). Notably, our study also highlights the importance of production of CXCL10, which leads to the control of WNV in the murine model (41). CXCL10 is decreased in dendritic cells from older
donors in response to WNV infection ex vivo (14) and is lower in macrophages of a cohort of subjects who fail to clear infection with the related flavivirus hepatitis C virus (29).

The most prominent alteration revealed in our studies of primary cells from a stratified cohort is a key role for IL1B signaling in the control of WNV in the central nervous system (CNS) (42) and in the production of CXCL10, mediated by polyubiquitination of the antiviral transcription factor IRF3 (43), which may play a role in the differential expression of CXCL10 in our cohorts. Furthermore, in the murine model, IL-1 receptor signaling is critical for the activation of both dendritic cell and adaptive CD4+ T cell responses and the control of viral load in the CNS (44). Central to this activation is the bioactive sphingolipid mediator S1P, which is induced upon immune cell activation (45) and which is in turn inactivated by SGPP2 (sphingosine-1-phosphate phosphatase 2), another of the genes in our susceptibility signature. S1P receptor ligation has recently been shown to diminish the cytokine storm produced after influenza virus infection and holds promise for therapeutic intervention in acute infection (46).

In combination with results of GWAS studies, previously identified age-related susceptibility (12, 14), and experimental models (7), the genes newly identified here illuminate elements of an altered immune state that leads to susceptibility to WNV. Another signature gene, SEMA4A, a class IV semaphorin expressed widely in the body and mainly on DCs in immune cells which is known to stimulate T cells and promote Th cell differentiation (47), was not previously identified in anti-WNV responses. WNV-infected subjects with a history of severe infection have reduced numbers of Th1/Th17 cells and Tregs (31), consistent with reduced SEMA4A levels (31). The related class VII semaphorin family member SEMA7A (CD108) was previously shown to diminish WNV infection through reduced viral load and levels of TNF-α mediated by transforming growth factor β1 (TGF-β1)/Smad6 signaling leading to diminished blood-brain barrier permeability and lower levels of viral entry to the brain (21).

The ability to characterize a person’s immune state to predict the response to infection is an attainable goal, and here, we demonstrate such an approach by identifying a signature associated with susceptibility to WNV. Our cohort of human subjects provides many essential elements for this in-depth investigation, including significantly different clinical responses and sufficient availability of experimental material from healthy subjects. While many characteristic individual immune parameters are stable and reflect a baseline immune status (5, 6), our subjects were not examined prior to infection, our studies do not address acute infection, and the ex vivo gene expression differences may not be causally related to clinical outcome. While we cannot exclude that responses in severe subjects may result from WNV infection or effects of persistent antigen (17, 48), our findings of differential innate cell function suggest that naive responses are being examined. Finally, it will be valuable to examine IL-1 responses across different populations, as IL-1 is known to vary in genetic association studies and, as a key regulator of the inflammatory process, may be critical to resistance to viral infections (49).

Through systems analysis of primary cells and bioinformatics analysis, we have identified global changes in virally induced gene expression and identified novel determinants of susceptibility to WNV. Our development of a focused gene panel creates a resource for investigations of other pathogens, in particular for related viruses such as hepatitis C virus and dengue virus. These “systems” investigations promote the identification of predictors of disease course and complications that are particularly notable for being detected in the absence of ongoing infection. The availability of such multifaceted interrogation of reliably curated patient cohorts with data-sharing and data-mining techniques should accelerate the identification of critical elements of immune resistance and mechanisms of pathogenesis for targeted therapeutic decision-making.

ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health (HHS grants N27220110019C, U19AI089992, AI091816, and AI057229) and the Gillson Longenbaugh Foundation.

S.S.S.-O. is a Taub fellow, R.G. is a Lady Davis fellow, and E.F. is an investigator of the Howard Hughes Medical Institute.

We are grateful to Stephanie Argraves, Shi Chen, Barbara Siconolfi, and Sui Tsang for expert research support and to the Yale HIPC team for insightful discussions.

REFERENCES


25. Durrr AM, Robinette ML, Klein RS. 2013. IL-1R1 is required for...


