MCP Papers in Press. Published on November 15, 2017 as Manuscript RA117.000310

Multiplexed biomarker panels discriminate Zika and Dengue virus infection in humans

Guang Song[1,§], Hee-Sool Rho[1,§], Jianbo Pan[2,§], Pedro Ramos[3], Ki-Jun Yoon[4,5], Freddy A. Medina[6], Emily M. Lee[7], Daniel Eichinger[3], Guo-li Ming[4,5,8,9,10], Jorge L. Muñoz-Jordan[6]*, Hengli Tang[7]*, Ignacio Pino[3]*, Hongjun Song[4,5,8,9,10]*, Jiang Qian[2]*, Heng Zhu[1]*

[1]Department of Pharmacology & Molecular Sciences; Johns Hopkins School of Medicine, Baltimore, MD 21205;

[2]Department of Ophthalmology; Johns Hopkins School of Medicine, Baltimore, MD 21205;

[3]CDI Laboratories, Inc. Mayaguez, PR 00682;

[4]Institute for Cell Engineering, Johns Hopkins School of Medicine, Baltimore, MD 21205;

[5]Department of Neurology, Johns Hopkins School of Medicine, Baltimore, MD 21205;

[6]Centers for Disease Control and Prevention, Dengue Branch, San Juan, PR 00920;

[7]Department of Biological Science, Florida State University, Tallahassee, FL 32306;

[8]The Solomon H. Snyder Department of Neuroscience, Johns Hopkins School of Medicine, Baltimore, MD 21205;

[9]Department of Neuroscience and Mahoney Institute for Neurosciences, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

[10]The Institute for Regenerative Medicine, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

*Correspondence should be addressed to: ckq2@cdc.gov (JLM); ignacio.pino@cdi-lab.com (IP); shongju1@jhmi.edu (HS); jiang.qian@jhmi.edu (JQ); and hzhu4@jhmi.edu (HZ).

§These authors contributed to this work equally.

Disclaimer: The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Running title: Biomarker for Zika and Dengue virus infection

Character count: 6,021

Keywords: Protein array, Zika virus, Dengue virus, Serological biomarker.

Abbreviations

ZIKV, Zika virus;

DENV, Dengue virus;

GBS, Guillain-Barré syndromes;

RT-PCR, Reverse transcription polymerase chain reaction;

ELISA, Enzyme-linked immunosorbent assay;

PRNT, Plaque reduction neutralization test;

ORFs, Open reading frames;

GST, Glutathione S-transferases;

ROC, Receiver operating characteristic;

MAC-ELISA, IgM antibody capture enzyme-linked immunosorbent assay.

Summary

Zika virus (ZIKV) and dengue virus (DENV) are closely related flaviviruses that cause widespread, acute febrile illnesses, notably microcephaly for fetuses of infected pregnant women. Detecting the viral cause of these illnesses is paramount to determine risks to patients, counsel pregnant women, and help fight outbreaks. A combined diagnostic algorithm for ZIKV and DENV requires Reverse transcription polymerase chain reaction (RT-PCR) and IgM antibody detection. RT-PCR differentiates between DENV and ZIKV infections during the acute phases of infection, but differentiation based on IgM antibodies is currently nearly impossible in endemic areas. We have developed a ZIKV/DENV protein array and tested it with serum samples collected from ZIKV- and DENV-infected patients and healthy subjects in Puerto Rico. Our analyses reveal a biomarker panel that are capable of discriminating ZIKV and DENV infections with high accuracy, including Capsid protein from African ZIKV strain MR766, and other 5 pair of proteins (NS1, NS2A, NS3, NS4B and NS5) from ZIKV and DENV respectively. Both sensitivity and specificity of the test for ZIKV from DENV are around 90%. We propose that the ZIKV/DENV protein array will be used in future studies to discriminate patients infected with ZIKV from DENV.

Introduction

The rapid, global spread of Zika virus (ZIKV), its association with congenital ZIKV and Guillain-Barré syndromes (GBS), and the demonstration that it is sexually transmitted, have raised unprecedented public health concerns(1). Indeed, ZIKV infection is now known to be a cause of microcephaly and severe brain abnormalities and has been linked to other birth defects (https://www.cdc.gov/zika/pdfs/zika-key-messages.pdf). It has become one of the greatest concerns for pregnant women, who can pass ZIKV to their developing fetus if infected during pregnancy(2). Therefore, it is critical to determine the potentially causal virus in an individual who presents with suspicion of infection. However, the arrival of ZIKV in the Americas was preceded by decades of endemic transmission of dengue virus (DENV), which has a similar clinical presentation that makes it difficult for clinicians to accurately distinguish(3).

Diagnostic algorithms for ZIKV and DENV often include RNA nucleic acid testing (e.g., RT-PCR) and IgM detection tests, according to guidelines that can be found in references for physicians(4-6). RT-PCR can detect and differentiate ZIKV and DENV early in infection(7). However, later in infection when RT-PCR assays are less sensitive, or in asymptomatic pregnant women, differentiating these two diseases requires specific serology-based diagnostic methods, which is challenging due to immunologic cross-reactivity among the flaviviruses(3). Enzyme-linked immunosorbent assay (ELISA) tests of IgM have recently become available for determining recent or remote ZIKV exposure, but do not completely differentiate ZIKV from DENV(8). ZIKV IgM positive results by these tests require confirmation by a plaque reduction neutralization test (PRNT), a cumbersome and lengthy laboratory method with some utility for differentiating flaviviruses in populations with low exposure(9), but not in populations with high incidence of DENV and ZIKV, such as Puerto Rico(5, 10). To counsel potentially infected pregnant women, establish the cause of congenital syndrome, or make informed decisions to treat patients with GBS or hemorrhagic disease, it is important to determine the pathogen responsible for the infection. Recently, a DENV protein microarray were fabricated through gene

synthesis and successfully applied to discriminate patients infected by DENV from other related pathogens, or uninfected patients(11), suggesting a protein array containing proteomes of ZIKV and DENV would be an effective method to differentiate patients infected by ZIKV or DENV.

In this study, we developed a ZIKV/DENV protein array and applied it to discover a biomarker panel that can specifically distinguish ZIKV and DENV infections in humans. By comparing the IgM serum profiles obtained with our protein array with “gold standard” IgM-confirmed RT-PCR ZIKV-positive and DENV-positive samples, we discovered and validated a biomarker panel that could distinguish ZIKV from DENV infections with high accuracy.

Experimental Procedures

Serum Samples

All Serum samples were obtained from the Dengue Surveillance System sample collection at Puerto Rico under guidelines approved by the CDC institutional review board (IRB). Patients did not provide verbal or written consent for this study. Samples were de-linked from patient identifiers according to CDC's IRB. Serum samples used for elaboration in the discovery stage and validation panels were listed in Supplemental Figure 1. ZIKV specimens were collected during the 2016 epidemic. DENV specimens were collected in years 2010-2012, prior to the arrival of ZIKV in Puerto Rico. Of the 49 dengue IgM positive samples by CDC IgM antibody capture enzyme-linked immunosorbent assay (MAC-ELISA), 33 were from patients with confirmed DENV RT-PCR positive results by the CDC DENV-RT-PCR(7). The 33 Zika IgM positive samples were dengue IgM negative and collected from patients with Zika RT-PCR positive results by the CDC TRIOPLEX test(8). Serum samples were obtained by standard procedures(12) and stored at -20°C until use.

Viral cDNA preparation

To fabricate a ZIKV/DENV protein array, two ZIKV strains (i.e., African MR766 and Cambodia strain FSS13025), and two DENV sero-types (i.e., DENV-1 TH-S-man Hawaii strain and DENV-2 NGC Thailand strain) were selected, respectively. Lysates of virus-infected mosquito cells were used to extract total RNA and 1 μg of total RNA was converted to cDNA using Superscript III (Thermo Fisher Scientific) for PCR templates.

Primer design and PCR conditions

Primer sets were designed for amplifying the full-length ORFs and *attB1* and *attB2* sequences at the 5´and 3'-ends of each primer were added to clone PCR amplicons into Gateway Entry vector pDONR221 (Thermo Fisher Scientific) by gateway recombination. To use the expression vector pEGHA for expression and purification of N-terminal GST fusion proteins, a stop codon (TAA) was added between the gene and the *attB2* gene specific sequence in the reverse primers. Four ORFs, 2K, NS2B, NS3, and NS4A, were highly conserved between ZIKV strains, so those were only cloned from the MR-766 strain so as to minimize the redundancy of viral proteins on the arrays. With regard to capsid proteins encoded by ZIKV, we also included two capsid protein C genes (C) with an additional propeptide of 14 amino acids in addition to the capsid protein C (prC). The list of all primers used in this experiment can be found in Supplemental Table S3. PCR amplification was carried out with 1 μL of cDNA in 50 μL PCR reaction using Phusion High-Fidelity DNA polymerase (New England BioLabs), and a final concentration of 10 pmol for each primer using the following PCR conditions: 98°C for 30 s; 5 cycles of 98°C for 10 s, 50-55°C for 30 s, 72°C for 30-100 s; 20 cycles of 98°C for 10 s, 65°C for 30 s, 72°C for 30-100 s; 72°C for 5 min. PCR products were run on 1% agarose gels and visualized by ethidium bromide staining. PCR amplicons with correct sizes were directly used to clone into Gateway Entry vector pDONR221 using BP recombinase (Thermo Fisher Scientific).

Gateway cloning and restriction digestion

Cloning of all viral ORFs was performed using the Gateway recombination reaction. Both BP and LR recombination reactions were done in 5 μL reactions including 300 ng of pDONR221 or pEGHA vectors, 3 μL of PCR amplicon, and 1 μL of BP or LR clonase mix (Thermo Fisher Scientific). BP and LR recombination reactions were incubated for 2-6 hr at room temperature and all 5 μL of BP and LR reaction mixtures chemically transformed to DH5a competent cells. Four single colonies were picked for each ORF and each clone was verified by both restriction enzyme digestion and Sanger sequencing analyses. Initially, the cloning was validated by *BsrG*I restriction enzyme digestion to measure the insert sizes from two single colonies. Those BP clones with expected sizes were confirmed by sequencing. All viral ORF entry clones were verified as lacking amino acid mutations by comparing to reference genome sequences of each virus strain. To express and purify individual viral proteins, LR recombination reactions were carried out to shuttle the inserts of sequence-verified Entry clones into the yeast expression vector pEGHA. The LR recombination reaction was performed as described above in the BP recombination reaction, and four LR clones were examined by restriction enzyme *BsrG*I digestion. Two verified LR clones were transformed to yeast strain Y258 to express GST fusion proteins under the control of galactose-inducible *GAL1* promoter.

ZIKV and DENV protein expression and purification

All recombinant pEGHA plasmids carrying confirmed ZIKV and DENV ORFs were transformed to yeast strain Y258 using a standard method. Each recombinant ZIKV and DENV protein was expressed as an N-terminal GST fusion under the induction of galactose and purified using glutathione sepharose affinity chromatography, as described previously(13). In brief, yeast clones were grown in small volumes of SC-Ura/glucose liquid medium overnight as primary cultures. Saturated yeast cultures were transferred in a 1:50 ratio to a large culture of SC-

Ura/raffinose until the O.D. 600 reached 1.0. Next, galactose was added to the culture at a final concentration of 2% to induce expression of the recombinant proteins for 4-6 hours at 30°C. Subsequently, the harvested yeast cell pellets were mechanically lysed in lysis buffer with zirconia beads. After pelleting the cell debris with centrifugation at 3500 g for 10 min at 4°C, the supernatant was transferred to a new tube with prewashed glutathione beads and incubated overnight at 4°C to capture the GST fusion proteins. The glutathione beads were washed six times with both high salt (50 mM HEPES at pH7.5 with 500 mM NaCl, 1mM EGTA, 10% glycerol, and 0.1% BME) and low salt (50 mM HEPES at pH7.5 with 100 mM NaCl, 1mM EGTA, 10% glycerol, and 0.1% BME) washing buffers successively. Finally, GST-tagged individual ZIKV and DENV proteins were eluted into printing buffer (50 mM HEPES at pH7.5 with 100 mM NaCl, 30% glycerol, 40 mM glutathione, and 0.03% Triton-100).

Fabrication of ZIKV/DENV protein arrays

All of the eluted proteins, together with positive controls (50 ng/μL hIgG, 50 ng/μL hIgM) and negative controls (900 ng/μL BSA, 150 ng/μL GST, and elution buffer) were arrayed into 384-well plates, and then printed onto SuperEpoxy (ArrayIt, U.S.A.) glass slides with an Ultra Marathon II Printer (ArrayJet, UK). Each ZIKV/DENV protein and the controls were printed in duplicate, in a 2X7 subarray format per slide, which allowed for testing 14 serum samples at once. The quality of the ZIKV/DENV protein array was confirmed by probing with anti-GST antibodies.

Removal of cross-reactive IgGs from serum samples

Human serum samples were first depleted of IgG using anti-human IgG (Fc-specific)-conjugated agarose beads (Sigma - A3316). The anti-human IgG agarose beads were equilibrated with blocking buffer (5% BSA in TBST), and 0.1 ml (bed volume) of pre-equilibrated agarose beads

were transferred to each well of a 96 DeepWell™ filter plate. Ten microliters of each human serum sample were diluted 20-fold with blocking buffer. Before diluted serum samples were added to the agarose beads, the 96 DeepWell™ plates were centrifuged at 1,000 x g at room temperature for 1 min. Diluted serum samples were added directly to the wells containing anti-human IgG agarose beads, incubated at RT on an orbital shaker for 30 min with shaking (180 rpm), and centrifuged at 1,000 x g at RT for 2 min to collect the IgM enriched flow-through. The flow-through samples were diluted 1:400 in blocking buffer prior to being added to the microarrays.

Serum profiling assays on the ZIKV/DENV protein arrays

Protein microarrays containing 14 identical subarrays were mounted with an adhesive press-to-seal silicone isolator (Grace Bio-Labs - 665217) to create 14 individual chambers. Each chamber was blocked with blocking buffer and incubated at 37°C on an orbital shaker at 60 rpm for 1 hr. The blocking buffer was replaced with individual pre-absorbed and diluted serum samples and incubated for 1 hr at 37°C with shaking. The array slide was mounted face up in a 50 ml conical tube and centrifuged using a fixed angle rotor (Sorvall SS-34) to completely remove sample liquid, and then washed three times for 10 min in 1X TBST. Silicone isolators were carefully removed and the slides were incubated with Cy3-labeled goat anti-human IgG Fc-specific (Sigma - C2571) and Alexa 647-labeled goat anti-human IgM Fc5μ fragment-specific (Jackson ImmunoResearch - 109-606-129), each diluted 1:1000 in blocking buffer, for 1 hr at RT, washed three times for 10 min in 1X TBST, rinsed once in double-distilled water, spun dry at 800 rpm for 2 min, and scanned using a GenePix® 4000B scanner (Molecular Devices, Sunnyvale, CA) using 700 as a constant PMT gain value on both channels.

Data analysis for assays performed on virus focused array

Data analysis was performed using a method similar to the one described in our previous studies(14). First, the median values of the foreground (F_{ij}) and background (B_{ij}) intensity at a given protein spot (i,j) on the protein arrays were extracted. The signal intensity (R_{ij}) of each protein spot was defined as F_{ij}/B_{ij}. Since each protein was printed in duplicate on an array, R_{ij} was averaged for each protein as R_p. The signal value for each protein was normalized by dividing the median value of negative controls for each sample. For each homologue protein, a ratio of R_{ZIKV+} / R_{DENV+} as a new feature was also added in this study.

To distinguish ZIKV positive samples from DENV+ samples using ZIKV proteins (R_{ZIKV+} and/or R_{ZIKV+} / R_{DENV+}), *P* values obtained from the two-sided Student's t-test, where the differential ZIKV proteins must be up-regulated for ZIKV positive samples compared to DENV+ samples, were calculated and adjusted as false discovery rates using the GenePattern platform(15). The optimal cutoff value for each candidate was evaluated with the highest discriminant ability, which is defined as

$$Discriminant\ ability = \frac{Sensitivity + Specificity}{2}$$

And the sensitivity and specificity were calculated for each protein using the following formula,

$$Sensitivity = \frac{TP}{TP + FN}$$

and

$$Specificity = \frac{TN}{FP + TN},$$

Where the true positives (TP) refer to ZIKV-infected patients correctly classified as positives. The true negatives (TN) are the DENV-infected patients correctly classified as negatives. The false positives (FP) are DENV-infected incorrectly classified as positives, and the false negatives (FN) are ZIKV-infected patients incorrectly classified as negatives.

To further improve the classification power, we attempted to identify and validate combinatorial biomarker panels using the discovery and validation cohorts, respectively. To do so, we exhaustively evaluated the performance for all possible combinations between two and six features. First, we employed a binary scoring system to convert the actual signal intensity of each protein to either 1 or 0, such that 1 represented signal intensity > the optimal cut off value and 0 otherwise. Next, we evaluated the performance of every possible combination in the discovery cohort. For a given combination of n proteins, the sum of the binary scores of the n proteins was assigned to each serum sample as a summary score. If the summary score of a sample was greater than k ($1 \leq k \leq n$), the sample was called positive. The sensitivity and specificity at the best discriminant value were recorded for each combination. Finally, we identified the combination and its k value with the best discriminant ability. The method was also applied to the validation cohort to test the identified combinatorial biomarker.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

Results

Fabrication of a ZIKV/DENV protein array

To fabricate a ZIKV/DENV protein array for serological assays, we first set out to clone the open reading frames (ORFs) encoded by two ZIKV strains (Asian and African lineage), and by DENV serotypes 1 and 2 (Figure 1A). Using cDNAs of ZIKV- and DENV-infected mosquito cells as templates, we PCR-amplified and cloned 26 ZIKV and 24 DENV full-length ORFs into the Gateway Entry vector (Figure 1B, C). The success of the cloning was first examined by BsrGI restriction digestion, followed by Sanger sequencing analysis. Inserts of sequence-confirmed Entry clones were then shuttled into a yeast expression vector pEGHA via LR recombination (16). The confirmed constructs were transformed into the budding yeast S. cerevisiae for protein purification.

Using our previously established protocol(16), the viral proteins were individually purified from yeast as N-terminally tagged Glutathione S-transferase (GST) fusion proteins and their quality and quantity were examined using Coomassie staining. A total of 23 ZIKV- and 25 DENV-encoded proteins were purified, as N-terminally tagged GST fusions with sufficient yield and >80% purity as judged by Coomassie staining (Figure 1D; Table 1). All of the purified proteins were then spotted onto epoxy-coated glass slides to form a ZIKV/DENV array. To achieve assays of high throughput, 14 identical subarrays in a 2X7 arrangement that was 96-well compatible were spotted on a single glass slide. As internal controls, we also spotted human IgG, IgM, GST, and a Cy3/Cy5 mixture on each subarray, as well as BSA at many locations. The total number of viral antigens on the array was 48 (Table 1). The quality of this protein array was examined by using anti-GST probing (Figure 1E).

Optimization of IgM-based serum profiling

To determine which viral proteins were better at discriminating immunoglobulin molecules from ZIKV or DENV-infected patients, we first optimized the serological assay conditions by

systematically testing a variety of parameters, including different titers of serum samples, binding and washing conditions in combination with variations in temperature, and array-scanning settings. Here, we chose one ZIKV-positive serum sample and one DENV-positive serum sample, which were confirmed with both RT-PCR and IgM detection tests. Because it is beneficial to simultaneously detect both IgM and IgG binding profiles, we employed Cy5-labeled anti-human IgM and Cy3-labeled anti-human IgG secondary antibodies as detection reagents. Although we quickly determined the best serum titer, washing and scanning conditions, we initially observed that several of the ZIKV-positive serum samples tested showed poor anti-IgM reactivity to the ZIKV proteins on the arrays, raising the possibility that the ZIKV epitopes might be masked in these patients by cross-reactive anti-DENV IgGs from a previous infection. Indeed, more than 90% of the adult population in Puerto Rico has been infected at least once by DENV(17). We also observed that in initial conditions, several DENV-positive sera could readily recognize some of the ZIKV proteins, and vice versa. To improve the IgM signal intensity reduced by cross-reactive antibodies, we tested three methods in parallel: 1) depletion of serum IgG with anti-human IgG Fc-specific beads, 2) serum IgM enrichment with mannan-binding protein beads, and 3) IgM purification with anti-human IgM agarose (see Experimental Procedures for details). On the basis of the serum profiles (Supplemental Figure 1), the IgG depletion method produced the best results as it dramatically improved the IgM-dependent signals, while reducing IgG-dependent signals to a minimum.

Identification of discriminative biomarkers

The goal of this study is to unambiguously discriminate anti-ZIKV and anti-DENV human antibodies in patients. We collected 82 specimens positive for either anti-ZIKV or anti-DENV antibodies, confirmed by RT-PCR testing. We randomly selected 2/3 of the samples for the discovery phase and used the remaining samples for the validation phase (Supplemental Table S1).

Implementing the IgG depletion step in the optimized assay protocol, we performed the serum profiling assays on the ZIKV/DENV protein arrays with 22 ZIKV positive and 32 DENV positive serum samples in a biomarker discovery phase screening. As expected, the anti-human IgM channels showed overall stronger signal intensity than that observed in the anti-human IgG channel (Supplemental Figure 2). This was particularly true for some ZIKV proteins, such as nonstructural proteins 1 (NS1), RNA-directed RNA polymerase NS5 (NS5), and Small envelope protein M proteins (M) (see NS5 examples in Figure 2). These observations indicated that there are IgM antibodies that have the potential to specifically recognize ZIKV and DENV infections in the majority of the samples.

To quantitatively analyze the binding signals of the anti-IgM serum profiling, we first extracted the median values of the foreground and background signals of each protein spot. Because each protein was printed in duplicate, we took the average of the median signals for each pair of the proteins to represent their signals. Next, we used the median foreground over background (F/B) ratios as the normalized signal values for each protein on a subarray (i.e., each serum sample). To enable fair signal comparison across different subarrays, we used the signals obtained from empty spots located at the same relative positions on each subarray as negative controls and performed the cross-array normalization.

Combinatorial biomarkers

Once assay conditions were optimized, we were able to identify the best ZIKV and DENV protein combinations to discriminate ZIKV- and DENV-infected individuals by comparing the IgM serum profiles against each virus protein. For each protein, we compared the signal values between the positive and negative groups and calculated the specificity values for each protein. Here we defined the sensitivity as the ratio between the number of samples we correctly called ZIKV positive and the total number of confirmed ZIKV positive samples. Similarly, the specificity was calculated by taking the ratio between the numbers of samples we correctly called DENV

positive and the total number of confirmed DENV positive samples. The overall performance was evaluated using a receiver operating characteristic (ROC) curve. Using this analysis, we identified the top proteins with the best prediction power. For example, NS4A encoded by ZIKV MR776 was identified with the best AUC value of 0.741, which corresponds to a specificity of 75% (Supplemental Table S2). Other top proteins are ZIKV capsid protein C (prC), NS2B and NS3 proteins (Supplemental Table S2).

To further improve the discriminative power, we implemented the use of signal ratios between some pairs of ZIKV and DENV homologous proteins. For example, when the NS5 protein encoded by the ZIKV African strain was used alone, the AUC reached to 0.71. Similarly, the NS5 encoded by DENV-2 showed an AUC value of 0.60 (Figure 2a). However, when the signal ratio between these two homologous proteins was used, the ability to differentiate ZIKV from DENV infection improved significantly (AUC=0.80) (Figure 2B).

We then sought to determine whether use of a combination of the homologous protein ratios and individual proteins would improve the assay performance. We employed an exhaustive search for all the combinations (up to 6 features) among all the proteins on the array. The best performance was obtained from a panel of biomarkers that includes 5 ratios (NS1, NS2A, NS3, NS4B and NS5) and one individual protein (African strain Capsid). The specificity of the test for ZIKV is 90.6% (Figure 3).

Validation of the identified biomarker panel

To validate the biomarker panel identified in the discovery phase, we employed an independent patient sample cohort, in which 11 were ZIKV positive and 17 were DENV positive. Using the same experimental protocol and computational methods, we determined that the specificity of the biomarker panel in this validation cohort was 88.2%, confirming that this biomarker panel could distinguish ZIKV and DENV infections with high specificity (Figure 3).

Discussion

We have identified a biomarker panel for highly specific anti-ZIKV and -DENV IgM detection for potential diagnostic use. We observed that the performance of this biomarker panel was significantly better than that of its individual components (Figure 2). Indeed, the same observations have been reported in biomarker discovery efforts for other human autoimmune diseases(18-20). The success of this IgM-based biomarker panel was also dependent on the removal of cross-reacting IgGs, which could compete off IgMs via binding to the epitopes of ZIKV proteins. Notably, in a recent report Zhang et al. failed to identify IgM-based biomarkers for diagnosis of ZIKV infections owing to the cross-reactivity(21). We foresee that this protein array-based approach could also be used to identify IgG-based biomarkers via a similar antibody depletion step. For example, to remove anti-DENV IgGs that are cross-reactive to ZIKV antigens, serum samples could be incubated with purified DENV proteins on a column before probing with the ZIKV/DENV arrays.

The biomarker panel identified in this study effectively discriminates ZIKV from DENV infections. Therefore, this array platform may be readily applied to identify IgM-based biomarkers that can distinguish ZIKV- or DENV-infected patients from healthy subjects. Because of the high-content nature of protein arrays, we could further expand this array by adding additional proteins encoded by other relevant viruses, such as yellow fever virus and chikungunya virus.

According to the Guidance for US Laboratories Testing for Dengue Virus and Zika Virus(22, 23), both Dengue and Zika can be diagnosed acutely using RT-PCR. The challenge in discriminating these viruses lies in testing specimens in the convalescent stage where RT-PCR is no longer sufficiently sensitive and indirect methods such as serology testing for antibodies is employed, but cross reactivity in the tests is observed. Diagnostic serology for these viruses is used in patients that show up at clinics or hospitals more than one week after the onset of symptoms. More than 95% of dengue- and Zika-infected patients tested for up to 60 days have

detectable IgM antibody(12), and even beyond 12 weeks according to CDC' new guidance(22). Also, it has been reported that IgM can be detected in infected patients for up to more than six months after infection(24). Anyway, all of the specimens used in our test were re-confirmed to be IgM positive in tests performed by the CDC, especially for the a few samples from patients with long time post onset (i.e. 80 days). On the other hand, both DENV and ZIKV are RNA viruses and highly variable, but their genomes display enough conservation to allow the development of serological tests for their diagnosis. For example, the first ZIKV strain identified in Puerto Rico during 2015 displays a 90% nucleotide similarity to the first ZIKV strain discovered in Uganda in 1947. More importantly, there is even less of a difference at the amino acid level. This means that when using an approach such as ours that uses whole virus protein, the likelihood of the test working for many years at the same sensitivity is high despite the variable nature of these viruses.

Our current procedure takes roughly 4.5 hrs to assay 14 samples/slide, which can be further improved by shortening the entire procedure to <60 min with further optimization and by increasing the throughput to 40 samples/slide, because many fewer proteins will be needed for the diagnosis. Furthermore, once the standard operation protocol is established, the entire procedure can be fully automated, which will greatly facilitate large scale screens in hospitals and blood banks. Finally, these biomarker panels can be potentially incorporated into a lateral flow device to facilitate multiplex, point-of-care diagnosis of ZIKV and DENV infections.

Acknowledgement

We thank Kimberly Christian for comments. This work was partially supported by NIH grants R01GM111514 (to H.Z & J.Q.), U19AI131130 (to G-l.M, & H.T.), R35NS097370 (to G-l.M.), R37NS047344 (to H-j.S.), and FSU Zika seed funding (to H.T.).

Conflict of Interest

Dr. Zhu is a co-founder and shareholder of CDI Laboratories, Inc. and Dr. Qian is a consultant of CDI Laboratories, Inc..

Author Contributions

G.S., H.R., I.P., J.Q. F.M., J.M. and H.Z. conceived and designed the experiments. F.M. and J.M. collected serum samples and supported construction for ZIKV/DENV serum associated experiment. E.M.L. and H.T. prepared virial cDNA and designed all the primers. G.S., H.R. constructed all the ZIKV/DENV ORF clones and purified recombinant proteins. G.S. and P.R. fabricated ZIKV/DENV protein arrays and performed the serum hybridization. J.P. performed biostatistics analysis. D.E. and G.M. provided experimental materials. J.M., I.P., H.S., J.Q., and H.Z. interpreted the data and wrote the manuscript. J.M., H.T., I.P. and H.Z. coordinated and oversaw the study.

References

1. Lazear, H. M., and Diamond, M. S. (2016) Zika Virus: New Clinical Syndromes and Its Emergence in the Western Hemisphere. Journal of virology 90, 4864-4875

2. Petersen, L. R., Jamieson, D. J., Powers, A. M., and Honein, M. A. (2016) Zika Virus. N Engl J Med 374, 1552-1563

3. Priyamvada, L., Quicke, K. M., Hudson, W. H., Onlamoon, N., Sewatanon, J., Edupuganti, S., Pattanapanyasat, K., Chokephaibulkit, K., Mulligan, M. J., Wilson, P. C., Ahmed, R., Suthar, M. S., and Wrammert, J. (2016) Human antibody responses after dengue virus infection are highly cross-reactive to Zika virus. Proceedings of the National Academy of Sciences of the United States of America 113, 7852-7857

4. Organization, P. A. H. (accessed may 5th 2017) Dengue: guidelines for patient care in the Region of the Americas. 2. ed. http://iris.paho.org/xmlui/handle/123456789/31207

5. (2016) Announcement: Guidance for U.S. Laboratory Testing for Zika Virus Infection: Implications for Health Care Providers. MMWR Morb Mortal Wkly Rep 65, 1304

6. Shukla, S., Hong, S. Y., Chung, S. H., and Kim, M. (2016) Rapid Detection Strategies for the Global Threat of Zika Virus: Current State, New Hypotheses, and Limitations. Front Microbiol 7, 1685

7. Santiago, G. A., Vergne, E., Quiles, Y., Cosme, J., Vazquez, J., Medina, J. F., Medina, F., Colon, C., Margolis, H., and Munoz-Jordan, J. L. (2013) Analytical and clinical performance of the CDC real time RT-PCR assay for detection and typing of dengue virus. PLoS Negl Trop Dis 7, e2311

8. Administration, U. S. F. a. D. (accessed may 5th 2017) Emergency Use Authorizations https://www.fda.gov/MedicalDevices/Safety/EmergencySituations/ucm161496.htm - zika

9. Shan, C., Xie, X., Ren, P., Loeffelholz, M. J., Yang, Y., Furuya, A., Dupuis, A. P., 2nd, Kramer, L. D., Wong, S. J., and Shi, P. Y. (2017) A Rapid Zika Diagnostic Assay to Measure Neutralizing Antibodies in Patients. EBioMedicine 17, 157-162

10. Rabe, I. B., Staples, J. E., Villanueva, J., Hummel, K. B., Johnson, J. A., Rose, L., Mts, Hills, S., Wasley, A., Fischer, M., and Powers, A. M. (2016) Interim Guidance for Interpretation of Zika Virus Antibody Test Results. MMWR Morb Mortal Wkly Rep 65, 543-546

11. Qi, H., Zhou, H., Czajkowsky, D. M., Guo, S., Li, Y., Wang, N., Shi, Y., Lin, L., Wang, J., Wu, and Tao, S. C. (2017) Rapid Production of Virus Protein Microarray Using Protein Microarray Fabrication through Gene Synthesis (PAGES). Mol Cell Proteomics 16, 288-299

12. Paz-Bailey, G., Rosenberg, E. S., Doyle, K., Munoz-Jordan, J., Santiago, G. A., Klein, L., Perez-Padilla, J., Medina, F. A., Waterman, S. H., Gubern, C. G., Alvarado, L. I., and Sharp, T. M. (2017) Persistence of Zika Virus in Body Fluids - Preliminary Report. N Engl J Med

13. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) Global analysis of protein activities using proteome chips. Science 293, 2101-2105

14. Pan, J., Song, G., Chen, D., Li, Y., Liu, S., Hu, S., Rosa, C., Eichinger, D. J., Pino, I., Zhu, H., Qian, J., and Huang, Y. (2017) Identification of serological biomarkers for early diagnosis of lung cancer using a protein array-based approach. Mol Cell Proteomics

15. Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J. P. (2006) GenePattern 2.0. Nature genetics 38, 500-501

16. Zhu, J., Liao, G., Shan, L., Zhang, J., Chen, M. R., Hayward, G. S., Hayward, S. D., Desai, P., and Zhu, H. (2009) Protein array identification of substrates of the Epstein-Barr virus protein kinase BGLF4. Journal of virology 83, 5219-5231

17. Mohammed, H., Tomashek, K. M., Stramer, S. L., and Hunsperger, E. (2012) Prevalence of anti-dengue immunoglobulin G antibodies among American Red Cross blood donors in Puerto Rico, 2006. Transfusion 52, 1652-1656

18. Hu, C. J., Pan, J. B., Song, G., Wen, X. T., Wu, Z. Y., Chen, S., Mo, W. X., Zhang, F. C., Qian, J., Zhu, H., and Li, Y. Z. (2017) Identification of Novel Biomarkers for Behcet Disease Diagnosis Using Human Proteome Microarray Approach. Mol Cell Proteomics 16, 147-156

19. Huang, W., Hu, C., Zeng, H., Li, P., Guo, L., Zeng, X., Liu, G., Zhang, F., Li, Y., and Wu, L. (2012) Novel systemic lupus erythematosus autoantigens identified by human protein microarray technology. Biochemical and biophysical research communications 418, 241-246

20. Vermeulen, N., de Beeck, K. O., Vermeire, S., Van Steen, K., Michiels, G., Ballet, V., Rutgeerts, P., and Bossuyt, X. (2011) Identification of a novel autoantigen in inflammatory bowel disease by protein microarray. Inflamm Bowel Dis 17, 1291-1300

21. Zhang, B., Pinsky, B. A., Ananta, J. S., Zhao, S., Arulkumar, S., Wan, H., Sahoo, M. K., Abeynayake, J., Waggoner, J. J., Hopes, C., Tang, M., and Dai, H. (2017) Diagnosis of Zika virus infection on a nanotechnology platform. Nat Med 23, 548-550

22. Infection, G. f. U. L. T. f. Z. V., and July 24 https://www.cdc.gov/zika/laboratories/lab-guidance.html.

23. Testing, L. G. a. D. https://www.cdc.gov/dengue/clinicallab/laboratory.html.

24. Prolonged IgM Antibody Response in People Infected with Zika Virus: Implications for Interpreting Serologic Testing Results for Pregnant Women. http://www.idsociety.org/Breaking_News_and_Alerts/CDC_HAN/Prolonged_IgM_Antibody

_Response_in_People_Infected_with_Zika_Virus__Implications_for_Interpreting_Serologic_Testing_Results_for_Pregnant_Women/

Figure Legends

Figure 1. Fabrication of a ZIKV/DENV protein array.

(A) Flow chart of ZIKV/DENV protein array fabrication. (B) Examples of successful PCR amplifications of ZIKV ORFs using cDNAs of ZIKV-infected mammalian cells. (C) Examples of ZIKV ORFs cloned into the Gateway system. Upper panel: plasmid DNA extracted from single colonies after the Gateway BP reactions was digested with *BsrG*I to release the inserts and subjected to gel electrophoresis. Lower panel: plasmid DNA extracted from single colonies after the Gateway LR reactions was digested with *BsrG*I to release the inserts and subjected to gel electrophoresis. (D) Coomassie stains of purified ZIKV and DENV proteins. ZIKV and DENV proteins were expressed and purified from yeast cells as N-terminal GST-fusion proteins. Purified proteins were subjected to SDS PAGE electrophoresis and the quantity and quality of the purified proteins were examined with Coomassie stain. (E) Images of the ZIKV/DENV protein arrays. On each epoxy-coated slide, 14 identical subarrays were spotted in a 2X7 arrangement on a glass slide and visualized with anti-GST probing.

Figure 2. The ratio of signals produced with homologous proteins produced the best results to distinguish ZIKV from DENV infection.

(A) Examples of box plot analysis of NS5. The intensity distribution of ZIKV positive and DENV positive samples with NS5 from ZIKV MR766 (left two boxes) and signal ratio between these two proteins (right two boxes). (B) The receiver operating characteristic (ROC) curves for NS5 proteins and NS5 ratio between ZIKV and DENV.

Figure 3. Biomarker panel with 6 features.

Orange and light blue lines represent samples scored as positives and negatives, respectively. A sample was predicted as ZIKV positive when at least 2 features in this panel scored positive. Values of sensitivity and specificity in both discovery and validation stages are listed below.

Table

Table 1. Purified virus proteins.

*Protein	Cambodia	MR766	DEN 1	DEN 2
prC	Y	Y	Y	Y
C	Y		Y	
prM	Y	Y	Y	Y
M	Y	Y	Y	Y
E	Y	Y	Y	Y
NS1	Y	Y	Y	Y
NS2A	Y	Y	Y	Y
NS2B		Y	Y	Y
NS3		Y	Y	Y
NS4A		Y	Y	Y
NS4A+2K		Y	Y	
2K+NS4B	Y	Y	Y	
NS4B	Y	Y	Y	Y
NS5	Y	Y	Y	Y

* “Y” means the protein is available on the ZIKV/DENV protein array.

Figures

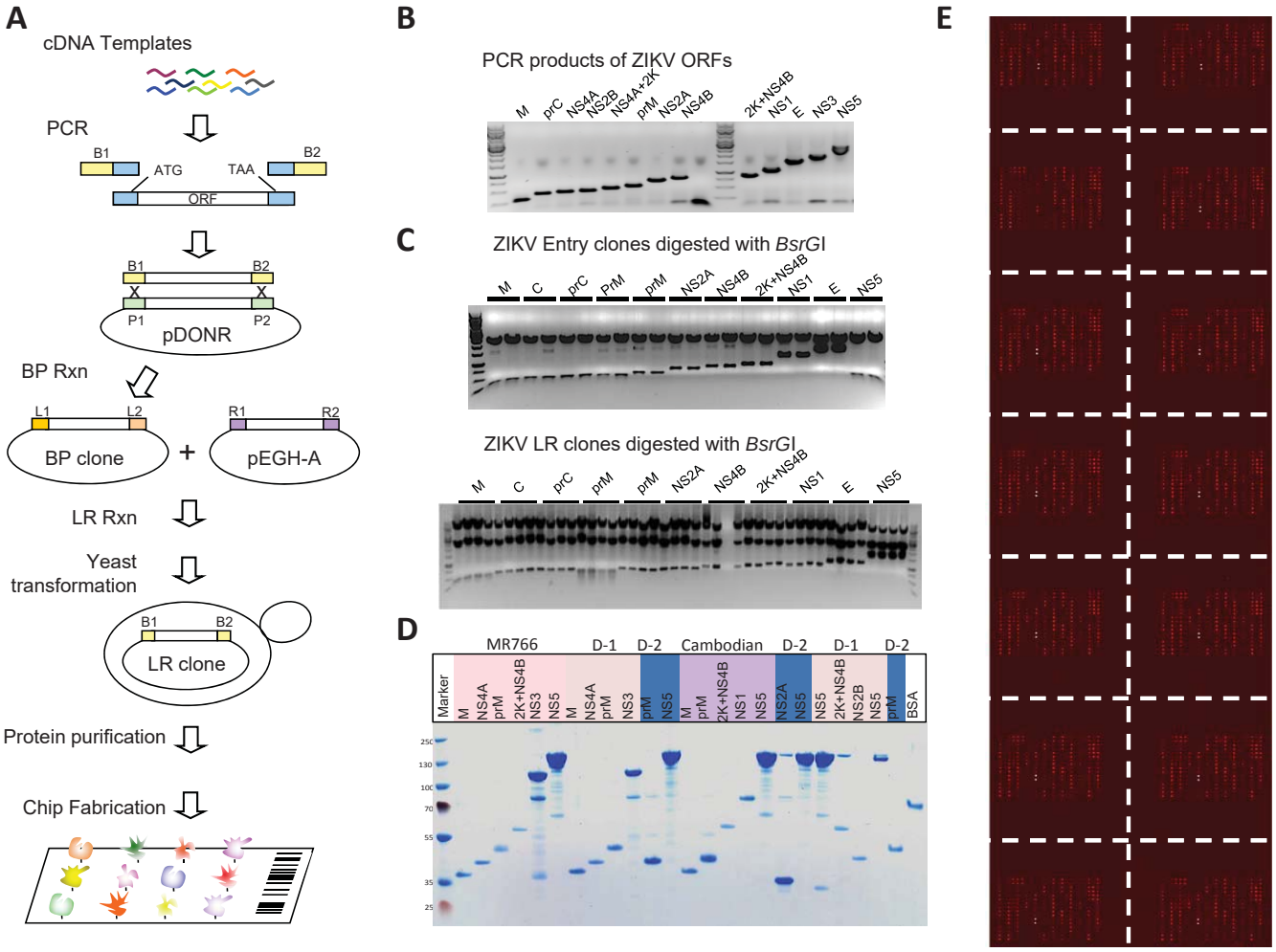

Figure 1

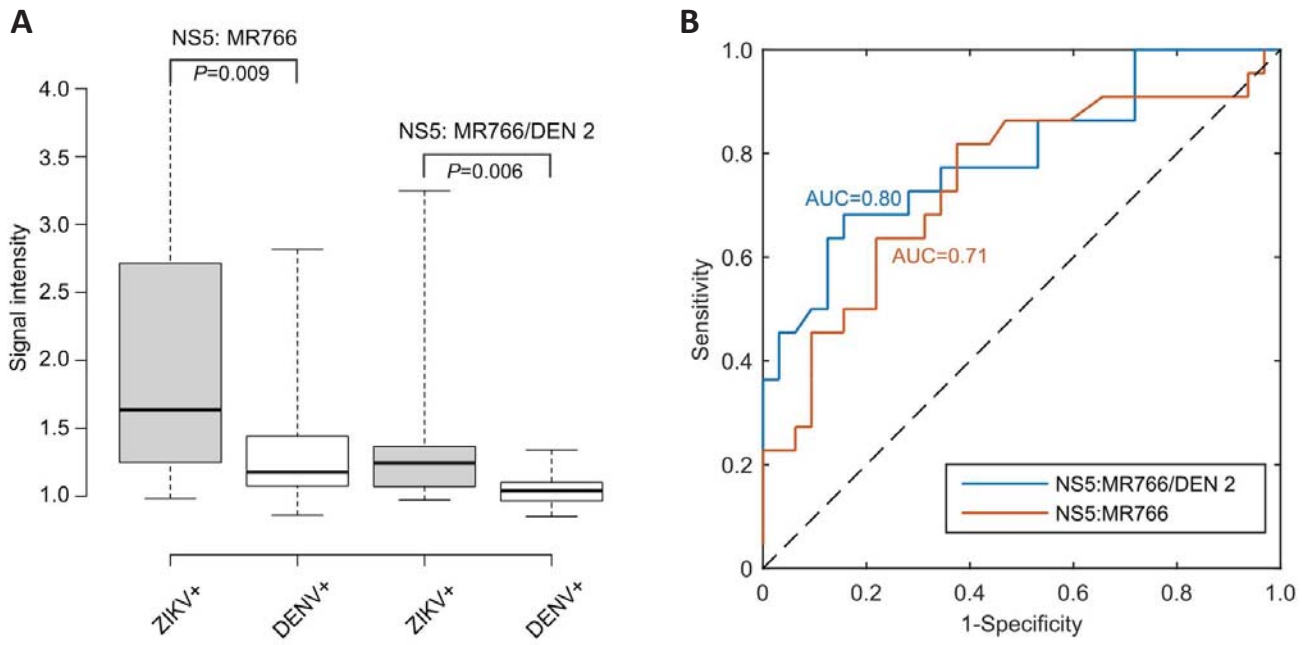

Figure 2

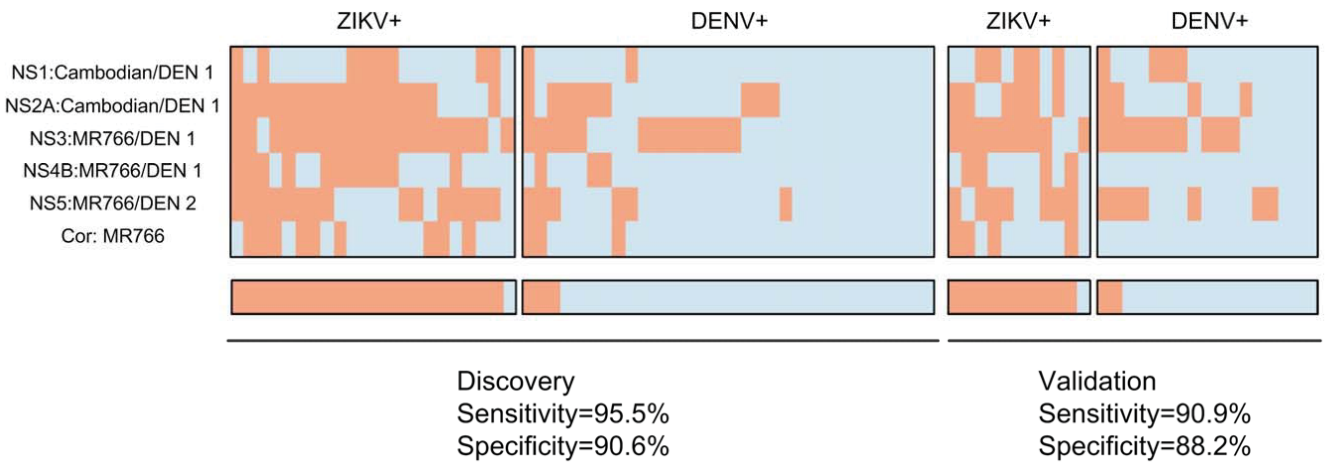

Figure 3