Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Chapter from the book *Viral Genomes - Molecular Structure, Diversity, Gene Expression Mechanisms and Host-Virus Interactions*

Downloaded from: http://www.intechopen.com/books/viral-genomes-molecular-structure-diversity-gene-expression-mechanisms-and-host-virus-interactions

Interested in publishing with InTechOpen?
Contact us at book.department@intechopen.com
Application of a Microarray-Based Assay for the Study of Genetic Diversity of West Nile Virus

Andriyan Grinev, Zhong Lu, Vladimir Chizhikov and Maria Rios
Center for Biologics Evaluation and Research, US Food and Drug Administration USA

1. Introduction

1.1 Molecular virology and epidemiology of West Nile virus

West Nile virus (family Flaviviridae, genus Flavivirus, WNV) is a small, enveloped, single stranded, positive RNA genome virus. WNV is a member of the Japanese encephalitis serogroup, which includes St Louis encephalitis virus (SLEV), Japanese encephalitis virus (JEV), Murray Valley encephalitis virus (MVEV), Kunjin virus (KUNV), and Usutu virus (USUV), which have all been shown to cause disease in humans. The virion consists of an envelope and prM-M dimers surrounding an icosahedral capsid of approximately 50 nm in size (Beasley, 2005). The WNV genomic RNA is approximately 11 kb in length, and contains 10 genes within a single open reading frame (ORF) that encodes for a single polyprotein flanked by 5’ and 3’ untranslated regions (UTR). The approximately 3430 amino acid WNV polyprotein is processed by cellular proteases and by the viral NS2B-NS3 protease into 3 structural and 7 non-structural proteins (NS) (Fig. 1).

![Fig. 1. Scheme of WNV genome and virion composition. The 11 kb positive RNA genome contains a single ORF encoding the 3 structural proteins that form the virus particle and the 7 non-structural proteins required for virus replication and immune evasion.](image-url)
The structural proteins i.e., capsid (C), premembrane-membrane (prM-M), and envelope (E), interact with the viral genomic RNA and with the host cell membrane to assemble viral particles. The structural proteins are not only essential for virion assembly and release, but they are also the major targets for virus neutralizing antibodies. The seven viral nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) are all necessary for genome replication (Khromykh et al., 2000). NS1 is a secreted glycoprotein implicated in immune evasion (Schlesinger, 2006). NS2A plays a role in virus assembly as well as inhibiting IFN-β promoter activation (Leung et al., 2008; Mackenzie et al., 1998). NS3 contains an ATP-dependent helicase, and in conjunction with the NS2B protein, functions as a serine protease, which is required for virus polyprotein processing (Chappell et al., 2005; Clum et al., 1997; Falgout et al., 1991). NS4A is responsible for a rapid expansion and modification of the endoplasmic reticulum that helps establish replication domains (Khromykh et al., 1998; Mackenzie et al., 1998). NS4B blocks the IFN response (Evans et al., 2007; Munoz-Jordan et al., 2005). NS5 is a methyltransferase and RNA-dependent RNA polymerase (Beasley, 2005; Egloff et al., 2002).

The untranslated regions (UTR) are involved in translation and viral RNA replication and likely play an important role in genome packaging. Both the 5' UTR and the 3' UTR in the WNV genome form highly conserved secondary and tertiary structures, some elements of which are similar among mosquito-borne flaviviruses. The cyclization of the flavivirus genome is necessary for viral RNA replication. In addition to base pairing between 5'-3' UAR and 5'-3' CS specific sequences involved in cyclization, a third stretch of nucleotides was identified to form a double-stranded region between the 5' and 3' UTRs (Friebe & Harris, 2010). Different functional regions have been described inside the 5' UTR and 3' UTR of flaviviruses based on such factors as nucleotide content, degree of sequence conservation, occurrence of repeated sequence motifs, and predicted secondary structure (Gritsun & Gould, 2007; Markoff, 2003; Proutski et al., 1997; Tajima et al., 2006). The 5' end of the WNV genomic RNA has a type I cap structure (m7GpppAmp) mediating cap-dependent translation. The 5' UTR contains two functional elements, the stem-loop A (SLA) and capsid-coding region hairpin (cHP) essential for RNA replication. The 3' UTR is generally divided into three regions based on the differences in the level of conservation: (1) the variable region is located immediately after the ORF; (2) the intermediate region has a moderate level of conservation and contains several hairpin motifs; (3) the conserved 3'-terminal region contains a cyclization sequences and stable stem-loop structure (Bryant et al., 2005; Markoff, 2003). These regions are believed to contain sequences that confer identity of the flaviviruses as demonstrated by attempts to exchange portions of the 3' UTR between WNV and dengue virus (DENV) that resulted in chimeric viruses which were unable to replicate (Yu et al., 2008).

WNV is maintained in nature by transmission between mosquitoes and birds, but it can also infect humans, other mammals (Beasley, 2005; Petersen & Marfin, 2002) and reptiles (Klenk et al., 2004) by mosquito bite (Fig.2). *Culex* spp. mosquitoes are the main vectors of WNV, although the virus has also been found in at least 43 other mosquito species (Granwehr et al., 2004; Higgs et al., 2004; Petersen et al., 2001). WNV can be transmitted vertically and overwinter in hibernating female mosquitoes, providing the mechanisms for viral persistence and reemergence each spring (Nasci et. al, 2001). WNV has spread within many bird species, including crows, magpies, and jays, house sparrows, house finches, grackles, and others representing 63 species, 30 families and 14 orders (Kramer & Bernard, 2001). They are all primarily competent reservoirs for WNV infection. By contrast, mammals...
including humans and horses are “dead-end” hosts in this enzootic cycle. They do not develop prolonged high-level viremia, so the concentration of the virus in blood is insufficient to infect a feeding mosquito. Most human infections are asymptomatic (~80%). The severity of symptomatic cases ranges from flu-like illness (~20% of infections) to severe neurological disease (~1%) (Hayes & Gubler, 2006). Additional modes of transmission were identified in 2002, including human-to-human by blood transfusion, breast-feeding, transplacental transmission, and by organ transplants extending the impact of WNV to blood safety and other areas of public health worldwide (Austgen et al., 2004; Pealer et al., 2003; Sbrana et al., 2005).

Fig. 2. Scheme of WNV transmission cycle. The maintenance of WNV in nature depends on an enzootic cycle involving many avian and mosquito species. Humans and other animals are incidental hosts that can become infected by WNV-infected mosquito bites. The virus can be also transmitted from human to human by blood transfusion and by solid organ transplantation.

Historically, since its isolation in Uganda in 1937, WNV outbreaks occurred in Africa, Europe, the Middle East, and caused a rare and mild febrile illness in humans and horses. A significant geographical expansion occurred starting from 1999 when the virus was introduced into North America. First detected in the U.S. in 1999, WNV has become endemic, causing yearly summer outbreaks. In 2002, the virus spread westward and the number of reported human cases increased dramatically. The North American epidemics of 2002 and 2003 represent the largest WNV outbreaks ever reported. WNV now is the most widespread arbovirus in the world (Kramer et al., 2008). The spread of the virus and intensity of the outbreak was correlated with the appearance of a new genotype with higher virulence and ability to disseminate in mosquitoes (Beasley, 2005). From the first outbreak in 1999 through 2010, WNV is estimated to have infected ~4 million humans in the US, causing over 30,000 serious illnesses, including 12,729 neuroinvasive disease cases with 1,206 deaths reported to the CDC (http://www.cdc.gov/ncidod/dvbid/westnile/). The virus has also been detected in the continental U.S. and in several areas of Canada, Central and South America, and the Caribbean. The persistence of WNV indicates that it has become endemic.
in the Western Hemisphere (Kramer et al., 2008). The pattern of yearly reoccurring outbreaks in North America differs from that of sporadic outbreaks observed in Europe and Africa. However, in recent years, WNV epidemics in humans and horses have become more frequent in several Southern European countries, and these epidemics could potentially be associated with an emergence of new viral genotypes (Chevalier et al. 2011). The speed with which the virus spread over the world triggered great interest and prompted a detailed investigation of the genetic evolution of the virus in search of the cause of its rapid adaptability.

Based on phylogenetic analyses, WNV has been initially divided into two major genetic lineages (Lanciotti et al., 2002). Lineage 1 included viruses circulating in Europe, Israel, United States, India, Russia, and Australia, while the Lineage 2 contained strains that circulated in sub-Saharan Africa and Madagascar. Lineage 1 was further divided into 3 sub-clades: 1a (including strains from Africa, Europe, US, Middle East, and Russia), 1b (Kunjin strain from Australia), and 1c (India) (Lanciotti et al., 2002). WNV has now been reported to have at least five distinct lineages based on phylogenetic analysis of all known full WNV genome sequences of viral isolates that correlates well with the geographical points of their isolation from various regions around the world including the U.S. (Davis et al., 2005; Herring et al., 2007; Grinev et al., 2008a; McMullen, 2011), Europe and Mediterranean (Parreira et al., 2007), and Africa (Botha et al., 2008).

In 2001 a new WNV genotype, named WN02, emerged in the US. The new genotype became prevalent in 2002, eventually displacing the ancestor genotype NY99, which is believed to have been introduced to the New World from the Middle East (Davis et al., 2005; Herring et al., 2007; Lanciotti et al., 1999). When compared to the WN99 genotype, the WN02 genotype possesses a few fixed silent nucleotide mutations and one amino acid substitution in the E protein (E-V159A). The highest rate of nucleotide sequence divergence among viruses isolated from 2002-2010 varies in the range of 0.4% - 0.6% (Davis et al., 2005; Ebel et al., 2004; Grinev et al., 2008a; McMullen et al., 2011). It is noteworthy that 80% of the nucleotide changes are observed in the structural regions represented by U<->C transitions; 75% among them are silent mutations (Grinev et al., 2008a). The possible explanation for the rapid displacement of the WN99 genotype by the new dominant genotype WN02 is due to the ability of the new viruses to more efficiently proliferate in domestic mosquitoes (Jerzak et al., 2005; Moudy et al., 2007). Phylogenetic analysis of modern WNV isolates demonstrates the existence of at least two subtypes of WN02 genotype co-circulating in North America. In addition to the common E-V159A amino acid substitution two other substitutions have become fixed in the significant part of North American WNV population: NS4a-A85T and NS5-K318R. Positive selection of these two amino acid substitutions potentially could impact viral fitness, phenotype and virulence (McMullen et al., 2011). As indicated by sequence analysis of new WNV genetic variants isolated in different areas of the US, the virus continues to diverge from the precursor isolate. Thus, changes in the WNV genome and viral proteins have the potential to negatively affect the sensitivity of screening and diagnostic assays currently used for virus detection, and to impact the development of vaccines and potential antiviral therapeutic agents. Therefore, development of new methods able to rapidly detect the emergence of WNV genetic variants is critical for epidemiological surveillance.
1.2 Basics of nucleotide microarrays

The monitoring and surveillance of pathogens is highly dependent on the capability of detection technology to simultaneously monitor multiple genomic signatures specific for different genetic variants of the pathogen. One of the approaches that enable this type of analysis is microarray technology. Generally, nucleotide microarrays are microscopic slides loaded with hundreds or thousands of pathogen-specific probes (DNA fragments or synthetic oligonucleotides) which can specifically hybridize with the target molecules to produce either quantitative (gene expression) or qualitative (diagnostic) data. DNA microarray technology provides an opportunity to perform parallel nucleic acid hybridization with a large number of immobilized oligonucleotides on a small surface area. Microarrays have the unique potential to simultaneously detect and identify defined pathogens, as well as to detect mutations within the complete viral genomes and target areas of a pathogen’s genome. It provides a significant advantage for the field of clinical microbiology and molecular epidemiological studies. Printed microarrays historically were the first arrays utilized for detection of mutations in many research laboratories and they are so-called because of the “printing” of the probes onto the surface of a glass microscope slide (Fig. 3).

![Fig. 3. Schematic outline of a printed microarray experiment. Microarray probes are spotted onto a surface of glass slide. In this example, the target sample and quality control sample (QC) with different fluorescent labeling were mixed and hybridized to the microarray probes. The efficiency of hybridization is monitored through measurement of the fluorescent signal from each spot by using a laser microarray scanner equipped with two lasers: 632 nm and 543 nm for excitation of Cy5 and Cy3 dyes respectively. The obtained fluorescent images are analyzed using specific computer software.](www.intechopen.com)
Chemically activated glass slides are commonly used for microarrays because they permit irreversible attachment of microarray probes, allowing for efficient hybridization kinetics between probes and analyzed targets, and have a low background fluorescence level (Cheung et al., 1999). Printed arrays can be produced using double-stranded DNA (dsDNA) fragments or oligonucleotides. For dsDNA microarrays, the probes usually consist of PCR products (amplicons) obtained using primers designed from a known genomic sequence or shotgun library clones. The double-stranded probes must be denatured prior to hybridization, either using a special printing buffer or after immobilization on the glass surface (Tomiuk et al., 2001). These microarrays containing relatively long 200-800 bp dsDNA probes usually demonstrate a high sensitivity but are not very useful for detection of minor genetic differences (e.g., single nucleotide mutations) between the probe and target nucleic acid (Hager et al., 2006). In contrast to dsDNA microarray probes, the length of oligonucleotide probes usually ranges from 20 to 80 nucleotides. Use of shorter probe lengths increases the microarray specificity and enables efficient detection of minor genetic changes between the probe and target (Chou et al., 2004). For more efficient attachment of oligonucleotide probes to the functional groups on the surface of chemically activated glass slides (usually aldehyde, epoxy, and succinimide groups), the 5' or 3' end of the probes contain primary amino groups introduced during chemical synthesis.

Microarray technology has been used to study gene expression in clinical and biological samples, detect and genotype pathogens (Honma et al., 2007; Wade et al., 2004), detect single base pair mismatches (Anthony et al., 2003; Hacia et al., 1999), design genomic maps (Roerig et al., 2005), and study viral evolution (Cherkasova et al., 2003).

The methods based on combination of initial PCR amplification of target genetic material followed by hybridization of amplicons with specific microarray oligonucleotide probes allowed for reconstitution of instant mutation profiles and determination of evolutionary divergence of individual viral isolates (Neverov et al., 2006). The microarrays consisting of multiple individual short oligoprobes were shown to be an efficient and sensitive genetic method for detection of single point mutations in viral and bacterial genomes (Chizhikov et al., 2002; Grinev et al., 2008b; Laassri et al., 2003, 2005, 2007; Volokhov et al., 2002). Microarray assays can also help simultaneously detect and identify the genotype and strain of common food-borne viruses without using PCR (Chen et al., 2011). In general, microarray technology can be easily implemented for detection and genotyping of any pathogen. Recently a pan-Microbial Detection Array was designed to detect all known viruses and bacteria (Gardner et al., 2010). Although further improvements, optimizations, and automation are still needed to fully implement the microarray technique in routine research and clinical practices, the potential role of these robust technologies in rapid diagnostics of multiple viral and bacterial pathogens is indisputable (Miller, 2009).

2. Material and methods

2.1 Plasma samples

The microarray development and evaluation study included the analysis of a total of 34 plasma specimens from blood donor units identified as positive for WNV by nucleic acid tests used to screen blood donations (Table 1).
Table 1. WNV isolates used for microarray assay validation. FFP indicates fresh frozen plasma sample. Passage P1 indicates first isolation in Vero cells; P2 and P3 indicates subsequent virus passages. Isolates in boldface were used for the full genome array validation.

These samples were collected in different geographic locations of the continental U.S. from the 2002-2009 epidemic seasons under IRB approved informed consent. In addition to 23 previously published isolates, which were used for structural region investigation (Grinev et al., 2008b), 11 WNV isolates from 2007 and 2009 were used to conduct microarray analyses of their full genomes in order to detect emerged genetic differences in comparison with that of the reference WNV strain NY99.
2.2 Viral isolates

West Nile virus isolation from tested plasma samples was performed using Vero cells. Vero cells were plated in T75 flasks and grown to 85% confluence in EMEM (GIBCO BRL, Gaithersburg, MD, USA) supplemented with 5% fetal bovine serum (FBS) (HyClone, Logan, UT) and 10µg/mL of penicillin/streptomycin (GIBCO). For viral isolation, growth medium was removed, 500 µl of each plasma sample were added to individual flasks and the total volume was adjusted to 5 ml with fresh medium. Vero cells were incubated with the viral inoculum for 2 hours, either at room temperature under gentle rocking or at 37°C with mixing every 10-15 min. After incubation, 10 ml of fresh medium were added, and the cultures were additionally incubated at 37 °C in 5% CO₂ and observed daily under phase microscopy for gross morphological degeneration i.e., cytopathic effect (CPE). Supernatants were harvested when extensive CPE was observed. Harvested supernatants were centrifuged to remove cell debris and aliquots were frozen at -80 ºC until further analysis.

2.3 RNA extraction

Total RNA was extracted from 1-3 ml of plasma samples with Trizol reagent (Invitrogen, Carlsbad, CA), according to the manufacturer protocol with additional step of ethanol precipitation. Viral RNA from 140 µl of infected Vero cell culture supernatants was extracted by using the QiaAMP viral RNA extraction kit (Qiagen, Valencia, CA), according to the manufacturer protocol. Each RNA sample was dissolved in 60 µl of RNase-free water and stored at -80°C.

2.4 Reverse transcription

Reverse transcription was performed in 20-40 µl reaction volume at 47°C for 2 h using a mixture of specific reverse primers (Table 2) and SuperScript III (Invitrogen, Carlsbad, CA) reverse transcription system according to the manufacturer's instructions. 2-3 µl of the reaction mixture were used for the subsequent DNA amplification.

2.5 PCR amplification

PCR fragments covering the entire structural region of the FDA-Hu2002 plasma sample were amplified by semi-nested PCR (Figure 4) from cDNA using the Hi-Fidelity PCR system (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Primers used for PCR amplification are shown in Table 2.

cDNA was amplified in the first round of PCR in a GeneAmp 9700 thermocycler (Applied Biosystems Inc., Foster City, CA) using the following protocol: denaturation at 94 °C for 30 s., 35 cycles of 30 s at 94 °C, 30 s at 50 °C, and 2 min at 68 °C. The final extension was carried out at 68 °C for 7 min. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen Valencia, CA), according to the manufacturer’s protocol. 5 µl of the 1st round PCR product was used for the 2nd round of amplification with reverse primers containing the T7 RNA polymerase promoter sequence at the 5’ ends tagged to the WNV sequence using the cycling program described above and the protocol for Hi-Fidelity PCR kit.
Table 2. Forward and reverse primers used for PCR amplification of different regions of WNV genome. Numbers in brackets indicates the primer position in the NY99 genome. Reverse primers marked as ‘m’ contained the T7 RNA polymerase promoter sequence.

<table>
<thead>
<tr>
<th>Set</th>
<th>Primers name and sequence</th>
</tr>
</thead>
</table>
| F1-1300 | F1  5’-AGTGTCCGCTGAGATGTCGAC (1-24)  
R1300  5’-TGTCGCTGCAATTCGCCAAGCACG (1300-1319)  
R1300m  5’-GCCTAAAGCGACTTACCTAGGGGATGCTGATATGACCCA |
| F980-R2000 | F980  5’-CTTGGAAATGCAAGACCCAAGAC (976-995)  
R2000m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F1690-R2490 | F1690  5’-GAGACGTTAATGGAAGTGTGA (1675-1694)  
R2490  5’-CCTACGGTCTCCACACATGGTGTGACCCGAAGCCC (2685-2685)  
R2490m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F2340-R3470 | F2340  5’-TCCTGGAGCAATGCTCTCAGAT (2326-2345)  
R3470  5’-CTCGATCGTTCATACACATACCA (3455-3467)  
R3470m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F3330-R4120 | F3330  5’-GAGACGTTAATGGAAGTGTGA (3334-3356)  
R4120  5’-CATACGCGACTTGGTCTCCTTT (4115-4136)  
R4120m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F4070-R4950 | F4070  5’-CTCCTGGATGGTGATGAAAATGG (4054-4076)  
R4950  5’-CGCTGTTGATGCTCTGACCCGAATGG (4933-4952)  
R4950m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F4810-R5650 | F4810  5’-CGCTGTCGACCCATACCTTG (4801-4818)  
R5650  5’-GCCCTCGTATCAGAATGTCCA (5644-5664)  
R5650m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F5510-R6430 | F5510  5’-AGCAITGCGACGAAAGGATGA (5491-5510)  
R6430  5’-TAATGTCGCTGCTGACCCGAGACTAC (6412-6435)  
R6430m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F6290-R8770 | F6290  5’-CGACCGGAGGTGTCTTGATGG (6288-6311)  
R8770  5’-CGCTCGGCTCGCTACCTTCCA (6751-6770)  
R8770m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F6690-R7550 | F6690  5’-CCCTCTCATCCAGCCGAGAAC (6675-6692)  
R7550  5’-GCCCTCGTATCAGAATGTCCA (7543-7562)  
R7550m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F7420-R8260 | F7420  5’-CCACACCCATCATCGCAAG (7409-7427)  
R8260  5’-CGTCAGGGACGACCTTCTTCCA (8260-8279)  
R8260m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F8170-R9050 | F8170  5’-CATAGGACGATGCTGGGCCTC (8155-8174)  
R9050  5’-CTCTCTTCCCATCACTGTGTAATAGC (9035-9059)  
R9050m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F8920-R9810 | F8920  5’-CAGGCTTGGTGCTCATGTT (8906-8924)  
R9810  5’-GAACCTGCTGCGCAATCATACC (9794-9814)  
R9810m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F9750-R10630 | F9750  5’-CTCCCTATGCTATGTCAAAAGTT (9734-9755)  
R10630  5’-GGTCCTGCTCGACCCGAGGG (10619-10638)  
R10630m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
| F10550-R11029 | F10550  5’-TGAGTGAGATGCTGGTCCTGGCC (10537-10556)  
R11029  5’-AGATGGCTGCTGCTGCAACACAG (11005-11029)  
R11029m  5’-GCCTAAATGACCTACTATAGGAGGAGGGTGCATGATCAAGCC |
Viral RNA samples isolated from Vero cell culture supernatants were amplified using forward and T7 tagged reverse primers (Table 2) using the OneStep RT-PCR Kit (Qiagen Valencia, CA), according to the manufacturer’s Q-Solution protocol with the 40 cycle program as recommended. PCR products were separated by electrophoresis in 0.8% agarose gel prepared in 1×TAE buffer containing 0.2 mg/ml of ethidium bromide. The stained DNA fragments were excised under UV light. PCR products were purified using the MinElute Gel Extraction Kit (Qiagen Valencia, CA), according to the manufacturer’s protocol, and stored at -20 °C. DNA fragments were used for the subsequent RNA synthesis using T7 RNA polymerase.

2.6 Preparation of fluorescently labeled WNV RNA for hybridization

The scheme for preparation of fluorescently labeled RNA samples is shown in Figure 4. Single-stranded RNA (ssRNA) samples used for hybridization were synthesized by T7 polymerase-driven transcription of the PCR products using the MEGA script T7 High Yield Transcription Kit (Ambion, Austin, TX) according to the manufacturer’s instructions. The MICROMAX ASAP RNA Labeling Kit (Perkin Elmer, Boston, MA) was used to incorporate Cy3 fluorophore into the ssRNA molecules. Fluorescently labeled ssRNA samples were purified from unincorporated dye using the Centrisep Spin Columns (Princeton Separations, Adelphia, NJ), dried under vacuum, and solubilized in the MICROMAX Hybridization Buffer III at a final concentration of 0.5-1.0 µM. The Cy5 antisense-QC oligonucleotide was prepared by 5’-end labeling with indocarbocyanine (Cy5)-dCTP during synthesis. The Cy5-antisense QC oligonucleotide was purified by high performance liquid chromatography (HPLC).

2.7 Microarray oligoprobe design and microchip fabrication

Oligonucleotide probes (oligoprobes) were designed on the basis of the nucleotide sequence of the reference strain NY99 (GenBank accession no.: AF196835) using the OligoScan software. A total of 1274 oligoprobes overlapping by half-lengths, with melting temperatures around 50°C, were designed for microarray-based detection of single point mutations in the entire WNV genome. Each oligoprobe spotting mixture contained 20 µM specific oligoprobe and 1 µM quality control (QC) oligonucleotide in 1× printing buffer (150 mM sodium phosphate, pH 8.5). The probes were used to print five identical arrays (each array contained triplicate set of oligoprobes) per each amine-binding glass slide (CodeLink, Amersham Biosciences, Piscataway, NJ) using a contact microspotting robot PIXSYS 5500 (Cartesian Technologies, Inc.) and a ChipMaker microspotting device equipped with CMP-7 pins delivering approximately 2 to 3 nl of a spotting mixture per spot (Tele-Chem International Inc.). Normally, the size of spots did not exceed 250 µm in diameter. After printing slides were processed and stored according to the manufacturer’s protocol. Thus, each microarray slide could be used for simultaneous analysis of five RNA samples. The array layout is shown in Fig. 4. The quality of each lot of printed slides was tested by hybridizing the last printed slide with Cy5-labeled antisense QC oligonucleotide.

2.8 Hybridization conditions

Hybridization between microarray oligoprobes and fluorescently labeled ssRNA samples was performed for 1 hour at 50°C. Before hybridization, Cy3-labeled ssRNA samples was
mixed with the Cy5-QC oligonucleotide at a molar ratio of 10 to 1, followed by denaturing at 95°C for 2 min and rapid chilling on ice. The area of each array on the slide was covered with a separate 20×9-mm glass cover-slip containing two rails of 35µm (±10µm) in height (Erie Scientific, Portsmouth, NH) that created a tiny “hybridization chamber” between the rail-elevated cover-slip and microarray slide. Each hybridization mixture (approximately 5-10 µl) was gently loaded by pipette into this space; the slides were placed in a hybridization cassette (Tele-Chem International, Inc., Sunnyvale, CA) to minimize sample evaporation and hybridized for 1 hour at 50°C. After hybridization, the slides were washed twice for 5 min with 2× SSC with 0.1% SDS pre-warmed to 50°C, then once for 2 min with 0.2× SSC buffer and once for 1 min with 0.1× SSC buffer all at room temperature followed by centrifugation at 1000 rpm for 5 min to remove any traces of the buffer.

Fig. 4. Scheme of preparation of Cy3 fluorescently labeled RNA samples. The Cy3 image shows the design of array and the results of hybridization experiment. Each microarray slide contained five identical arrays for analysis of five different RNA samples including the reference WNV isolate NY99 and four clinical samples (S1-S4).

2.9 Microarray scanning and image analysis

The fluorescent images of processed microarray slides were generated using GenePix 4100 (Axon Instruments) and ScanArray 5000 (Perkin Elmer) scanners equipped with two lasers.
operating at 632 nm (for excitation of Cy5 dye) and 543 nm (for excitation of Cy3 dye). Images were analyzed using ScanArray Express 2.1 (Perkin Elmer) and GenePix 3.0 (Axon Instruments) software. Each spot was defined by manual positioning of a grid over the array image. Atypical and empty spots were manually flagged and excluded from further analysis. Background fluorescence readings obtained from the region surrounding each spot were subtracted, and the net value of the Cy3 fluorescence signal from each oligoprobe was divided by the Cy5 signal value from the QC probe of the same spot to minimize the effect of spot size variation on resulting data. Data files generated by ScanArray Express and GenePix software were exported into MS Excel. To identify positions in the WNV genome where mutations occurred, the intensities of fluorescent signals from each array spot obtained for tested WNV isolates were compared with that of the reference NY99 isolate. The fluorescent signal ratio values between the reference NY99 and the tested sample were normalized using a linear regression model. A signal intensity ratio threshold from reference spots, specific to each microarray printing lot, was defined as an average ratio plus two standard deviation values. Any spots showing a ratio greater than the threshold value for a particular printing lot (i.e. considerably lower signal than reference isolate NY99) potentially indicated the presence of a mutation in the genomic area covered by the oligoprobe of that spot.

2.10 Validation of the microarray results

Validation of the microarrays was performed by comparing microarray results with the sequencing data obtained for the tested WNV isolate. Direct sequencing of PCR products was described previously by Grinev et al. (2008a).

3. Results and discussion

3.1 Optimization of the WNV microarray assay

The objective of this study was to develop and assess the feasibility of a DNA microarray approach for rapid throughput detection of spontaneous nucleotide mutations in the genome of WNV. A microarray containing a set of 1274 oligoprobes overlapping by half of their lengths (overlapping factor = 2) was developed and evaluated by assessing the ability of microarray to detect all identified mutations in 34 clinical isolates of WNV that were previously sequenced in our laboratory (Fig. 5).

In fact, optimization of an oligonucleotide microarray assay is a multi-parametric task. Thus, the fluorescent signal from the microarray oligoprobes hybridized to the fluorescently labeled target is known to depend on several factors including the sequence of oligoprobes, the character of mutations, and the position of the mismatched nucleotide in the oligoprobe, as well as the propensity of the RNA hybridization target to form a secondary structure (Relogio et al., 2002; Liu et al., 2005; Naiser et al., 2008). The goal of the optimization process was to determine the hybridization conditions which would enable the sensitive and specific hybridization of the target RNA to the vast majority of oligoprobes composing the WNV microarray. In our study, we optimized the temperature and time of hybridization, the stringency of post-hybridization washing conditions, the image scanning settings (PMT and laser power) to achieve the most efficient discrimination for each probe. It should be noted that the different oligoprobes have slightly different melting temperatures with the
hybridization target that also contributes to the difference in sensitivity-specificity balance for each oligoprobe and its targets.

Fig. 5. Scheme of the WNV microarray assay. Five arrays containing 1274 oligoprobes overlapped by half-lengths that covered the entire genome of WNV are shown. The Cy3 hybridization images show the layout of the printed arrays, each oligoprobe was spotted in triplicate. The analysis of the hybridization profile of WNV isolate ID140 from the 2007 U.S. epidemic is shown as a chart of hybridization signal ratios (y-axis) between the reference isolate NY99 and the tested isolate. Oligoprobe numbers are shown on the x-axis.

The oligoprobes were designed to have similar thermodynamic characteristics to ensure uniform hybridization signals from all microarray probes. The design of the microarray probes was performed using the OligoScan software, which is capable of selecting multiple oligonucleotide probes with similar thermodynamic features. The microarray contained overlapping oligoprobes of 15-26 bp in length with melting temperatures around 50°C, which was previously shown to be optimal for the detection of single nucleotide mutations and deletions-insertions (Laassri et al., 2005, 2007). The predicted melting temperatures of the oligoprobes varied from 49.7°C to 52.6°C. Therefore, an optimization of hybridization temperature that would provide efficient hybridization and high specificity for each spot on the array is required. We evaluated three hybridization temperatures: 47, 50 and 53°C. During the optimization process we determined that hybridization and washing of the microarrays at 50°C resulted in better discrimination between perfect matches and mismatches for most of the spots of the array.

The synthetic 5'-aminated oligoprobes were printed on CodeLink Activated slides, previously shown to be suitable slides for detection of single nucleotide mismatches (Laassri et al., 2007). Five arrays containing triplicate sets of oligoprobes were printed on each slide.
to allow for simultaneous analysis of four target samples in each hybridization experiment. One array on each separate slide was always hybridized with the fluorescently labeled RNA prepared using the reference WN-NY99 strain while other four arrays were used for hybridization with fluorescently labeled RNAs prepared from the WNV isolates to be tested. To assess the reproducibility of microarray fabrication, all hybridizations were repeated twice using different slides. The hybridization temperature and time, post-hybridization washing and the detector setting were optimized. The MS Excel worksheet was developed as a result of analysis of each scan image and data obtained from triplicates of each oligoprobe averaged by the median. The values from each microarray spot were normalized by the signal from the quality controls. The normalized signals from the reference array were then divided by the relevant signals obtained from the sample array. The occurrence of mismatches between oligoprobe and fluorescently labeled RNA target resulted in significant reduction of the hybridization signal in comparison with a perfect matching pair. Thus, the monitoring of the ratios of hybridization signals from unknown samples and the reference samples could be used as a tool for detection of spontaneous mutations in the target region of the WNV genome. The scheme of microarray experiments is shown on Figure 5.

DNA microarray technology provides an opportunity to perform parallel nucleic acid hybridization with a large number of immobilized oligonucleotides on a small surface area.

DNA microarrays containing short oligonucleotide probes (15-25 nt) provide a greater discrimination power compared to microarrays composed of larger oligonucleotides or PCR-amplified DNA fragments. The strongest signal ratios between perfect matched and mismatched sequences were observed when mutations were located near the center of an oligoprobe, and the shortest probes always had better discriminatory power (Urakawa et al., 2003). Although the microarray hybridization method provides limited information about the position of mutations in the analyzed genomic region when compared to sequencing, it has the significant advantage of allowing the identification of “hot spots” where random mutations occur within short (a few nucleotides) areas. It also may allow for the detection of those mutations even when they occur at relatively low levels (up to 1%) as in the case of mixtures of quasispecies that cannot be detected by traditional direct sequencing methods (Cherkasova et al., 2003; Leberre et al., 2007). The efficiency of microarrays for identification and discrimination of closely related bacteria and viruses has been previously demonstrated (Chizhikov et al., 2002; Hsia et al., 2007; Laassri et al., 2003; Nordström et al., 2005; Volokhov et al., 2002; Wade et al., 2004). The use of oligonucleotide microchips for screening of random mutations is based on the ability of microarrays to identify the presence of single-nucleotide mutations in the hybridization template (Hacia et al., 1999; Urakawa et al., 2003).

3.2 Optimization of hybridization probes

In general, single-base-pair discrimination can be achieved by optimization of hybridization or array washing conditions. However, optimization of washing conditions does not guarantee the equal efficiency of mismatch detection for all designed oligopropbes, which is likely to be caused by nature and position of the particular mutation or even formation of the secondary and tertiary structures in the hybridization mixture. The increase of size of the RNA template tends to increase the complexity of the secondary structure of the RNA molecule, and may affect the efficiency of hybridization of the RNA template with the microarray oligoprobes. Therefore, the RNA template may be subdivided into shorter
overlapping templates covering the full length of the larger template. The use of a hybridization RNA target of approximately 2500 nucleotides covering the entire structural region of WNV resulted in a twofold reduction of some hybridization signals when compared to the use of a mixture of three overlapping RNA templates in the equivalent molar concentrations. The typical results of this experiment are shown in Figure 6.

Fig. 6. The analysis of microarray hybridization is shown as a chart that contains the oligoprobe numbers on the x-axis. The y-axis shows the signal ratio values between the isolate NY99 and two isolates from 2005: BSL13-05; and GCTX2-05. The results for hybridization using a mixture of three overlapping RNAs covering the whole structural region of WNV and 5’UTR are shown in blue. The results for hybridization using the one long RNA covering the whole structural region of WNV and 5’UTR are shown in red.

Consequently, mixtures of three RNA templates (instead of one long RNA target) were normally used for hybridization with microarray to ensure the high efficiency of mutation discrimination. The initial use of in-house printed arrays represented a convenient model for the development and optimization of mutation-detecting microarrays. However, the low densities of in-house printed arrays usually force the use of a substantial number of slides for a single hybridization experiment. It is quite laborious and relatively expensive. After optimization of the microarray system including the set of oligoprobes and hybridization conditions, a high density array can be prepared using a well developed system like the widely used GeneChip produced by Affymetrix, which relies on in situ synthesis of all oligonucleotides covering the entire WNV genome in a single array. In contrast to the printed oligonucleotide arrays described above, the oligonucleotide probes for high-density arrays are synthesized directly on the surface of the microarray, which is usually a small quartz wafer. Because in situ-synthesized probes normally form a very tiny spots, multiple overlapping probes for each target may be included to improve sensitivity, specificity, and statistical accuracy. On the other hand, the use of a single array and mixture of RNA probes
in a single experiment requires additional optimization of hybridization conditions. We found a loss of some signals while using a template composed of a mixture of 15 overlapping RNAs covering the entire WNV genome, probably due to formation of double stranded molecules between complementary parts of the WNV genome such as the cyclization sequences in the 5’ and 3’ UTRs and tight secondary structure formation. The usage of a number of shorter RNAs produced by multiplex PCR of a target region as a hybridization template might reduce this negative effect and improve the signal, as would chemical defragmentation of the long RNA probe.

### 3.3 Evaluation of genetic stability of West Nile Virus after isolation in Vero cell culture

Our study required WNV isolation from human plasma samples in Vero cell cultures because of the low concentration of WNV and limited volume of starting material available from many of the specimens. It was shown previously that genetic changes in flaviviruses can be induced by consecutive passages in Vero cell cultures. For example, multiple passage of Dengue virus (DENV), which is closely related to WNV, in Vero cells has resulted in emergence of mutants with amino acid changes in E and occasionally in prM, but not in C. Some nucleotide mutations were detected after the first five passages (Lee et al., 1997). Further studies showed that three nucleotide mutations (two in E and one in NS1 proteins) that resulted in two amino acid alterations (one each in E and NS1) emerged in the DENV genome after 20 continuous passages in Vero cells. Additional passage of the virus for 30 passages caused four nucleotide changes (two each in E and NS1) that resulted in three amino acid substitutions (one in E and two in NS1) (Chen et al., 2003).

In order to investigate whether viral isolation using three consecutive passages in Vero cells resulted in genetic changes in the WNV genome, we cultivated the isolate FDA/HU-02 and compared the sequence and microarray results for the three passages (P1, P2 and P3). Total RNA samples were isolated from aliquots of the original plasma and from each of the passages 1-3. RNA samples were reverse transcribed and amplified by PCR using WNV-specific primers designed to cover the complete structural region of the virus, followed by PCR product purification, direct sequencing, and microarray hybridization. We found no significant changes in fluorescent signal ratio values obtained in one slide microarray hybridization experiment for each original plasma sample and the RNA isolated from each of the Vero cell culture passages. We also performed comparative sequence analysis of the aforementioned viral samples. There was no difference between the genomic sequence obtained from the isolate FDA-Hu2002 (P1) and two passages (P2 and P3) when compared to the genomic sequence obtained from RNA extracted from the original plasma sample. These results are in good concordance with the genetic stability data previously published for the chimeric Dengue and Yellow Fever-Dengue vaccine candidates passaged 10-20 times in Vero cells (Guirakhoo et al., 2004; Butrapet et al., 2006). Therefore we assume that WNV isolates generated from human specimens by a few passages in Vero cells represented the original virus, and that the structural region was not changed during a low number of serial passages in Vero cell cultures.

### 3.4 Evaluation of oligonucleotide microarray using clinical WNV isolates

The ability of microarray to detect mutations in the target regions was evaluated by testing 34 previously sequenced WNV isolates obtained in the course of the 2002-2009 US epidemics.
Table 3. Fixed nucleotide mutations and ratios of hybridization signals normalized against the reference isolate NY99. Unique mutations are not shown.

23 WNV isolates from 2002-2005 epidemic seasons we used to evaluate the first array containing oligoprobes covering the structural region. Fluorescent signal ratios for all mutations ranged from 4.4 to 85.5 depending on the position of the mismatch within the oligoprobe and the character of the mismatched nucleotide (Grinev et al., 2008b). Table 3 shows the positions of identified nucleotide mutations and ratios of hybridization signals from the respective hybridization templates normalized against the reference isolate NY99 for 24 fixed mutations determined in the complete genomes of isolates from 2007 and 2009 epidemics. All 11 completely sequenced isolates from this study shared 12 nucleotide mutations including a non-silent mutation in Env T1442C. Ten more common mutations were detected in 4 isolates ARC140-07, BSL5-09, BSL6-09 and BSL11-09. The signal ratios for these mutations varied in the range 3.1-73.9. The fluorescence signal ratios produced by single nucleotide mutations in all hybridization experiments exceeded the experimentally determined cut-off threshold value ranging from 0.2 to 2.1 when compared to that of completely matched pairs. The ratio value over 2.1 can be considered as an indication on the potential mutation in a specific genomic region pointing to the need for additional analysis of this region by sequence analysis. This approach has the advantage of substantially reducing the numbers and length of sequences required for proper surveillance studies. The results of our study showed that the WNV microarray was able to unambiguously detect all mutations in the viral genome previously identified by routine sequencing analysis.
4. Conclusion

Viral adaptation through fixation of spontaneous mutations is an important factor potentially associated with reoccurrence of WNV outbreaks in the New World. The emergence of new genetic variants of WNV raise issues of public health importance because they may affect the sensitivity of both screening and diagnostic assays, as well as the development of vaccines and drugs. We have developed and optimized a WNV microarray assay, which enabled simple monitoring of WNV genetic variability and rapid detection of any nucleotide mutation within the entire viral genome. Our microarray system potentially can serve as a high throughput, rapid and effective approach for the identification of WNV mutations, and characterization of circulating WNV genetic variants.

5. Acknowledgments

We would like to thank Drs Caren Chancey, German Anez Gutierrez and Robert Duncan for helpful discussion and review of the manuscript, Dr Majid Laassri for technical assistance, and Dr Konstantin Chumakov for OligoScan software.

6. References


Application of a Microarray-Based Assay for the Study of Genetic Diversity of West Nile Virus

Kingdom. Transactions of the Royal Society of Tropical Medicine and Hygiene, Vol.98, No.2, (February 2004), pp. 82–87, ISSN 0035-9203


Viruses are small infectious agents that can replicate only inside the living cells of susceptible organisms. The understanding of the molecular events underlying the infectious process has been of central interest to improve strategies aimed at combating viral diseases of medical, veterinary and agricultural importance. Some of the viruses cause dreadful diseases, while others are also of interest as tools for gene transduction and expression and in non-polluting insect pest management strategies. The contributions in this book provide the reader with a perspective on the wide spectrum of virus-host systems. They are organized in sections based on the major topics covered: viral genomes organization, regulation of replication and gene expression, genome diversity and evolution, virus-host interactions, including clinically relevant features. The chapters also cover a wide range of technical approaches, including high throughput methods to assess genome variation or stability. This book should appeal to all those interested in fundamental and applied aspects of virology.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: