Molecular diagnostic and genetic characterization of highly pathogenic viruses: application during Crimean–Congo haemorrhagic fever virus outbreaks in Eastern Europe and the Middle East


To cite this version:


HAL Id: pasteur-03516032
https://pasteur.hal.science/pasteur-03516032
Submitted on 7 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License
Molecular diagnostic and genetic characterization of highly pathogenic viruses: application during Crimean–Congo haemorrhagic fever virus outbreaks in Eastern Europe and the Middle East

C. Filippone¹,², P. Marianneau³, S. Murri⁴, N. Mollard⁴, T. Avsic-Zupan⁵, S. Chinikar⁶, P. Després⁷, V. Caro⁸, A. Gessain¹,², N. Berthet¹,² and N. Tordo⁴,⁹

¹) Institut Pasteur, Unit of Epidemiology and Pathophysiology of Oncogenic Viruses, Paris, 2) Department of Virology, Institut Pasteur, URA CNRS URA 3015, Paris, 3) Virology Unit, Laboratoire de Lyon, ANSES, Lyon, 4) Institut Pasteur, Unit of Biology of Viral Emerging Infection, National-WHO-OIE Reference Centre for Viral Haemorrhagic Fevers, Lyon, France, 5) Medical Faculty, Institute of Microbiology and Immunology, Ljubljana, Slovenia, 6) Institut Pasteur of Iran, Arboviruses and Viral Haemorrhagic Fevers Laboratory, National Reference Laboratory, Teheran, Iran, 7) Institut Pasteur, Unit of Molecular Interactions Flavivirus–Host, WHO-OIE Reference Centre for Viral Haemorrhagic Fevers, 8) Institut Pasteur, Genotyping of Pathogens and Public Health and 9) Institut Pasteur, Antiviral Strategies Unit, WHO-OIE Reference Centre for Viral Haemorrhagic Fevers, Paris, France

Abstract

Several haemorrhagic fevers are caused by highly pathogenic viruses that must be handled in Biosafety level 4 (BSL–4) containment. These zoonotic infections have an important impact on public health and the development of a rapid and differential diagnosis in case of outbreak in risk areas represents a critical priority. We have demonstrated the potential of a DNA resequencing microarray (PathogenID v2.0) for this purpose. The microarray was first validated in vitro using supernatants of cells infected with prototype strains from five different families of BSL-4 viruses (e.g. families Arenaviridae, Bunyaviridae, Filoviridae, Flaviviridae and Paramyxoviridae). RNA was amplified based on isothermal amplification by Phi29 polymerase before hybridization. We were able to detect and characterize Nipah virus and Crimean–Congo haemorrhagic fever virus (CCHFV) in the brains of experimentally infected animals. CCHFV was finally used as a paradigm for epidemics because of recent outbreaks in Turkey, Kosovo and Iran. Viral variants present in human sera were characterized by BLASTN analysis. Sensitivity was estimated to be 10⁵–10⁶ PFU/mL of hybridized cDNA. Detection specificity was limited to viral sequences having ∼13–14% of global divergence with the tiled sequence, or stretches of ∼20 identical nucleotides. These results highlight the benefits of using the PathogenID v2.0 resequencing microarray to characterize geographical variants in the follow-up of haemorrhagic fever epidemics; to manage patients and protect communities; and in cases of bioterrorism.

Keywords: Crimean–Congo haemorrhagic fever virus, differential diagnosis, microarray, viral haemorrhagic fevers, viral zoonoses

Original Submission: 13 June 2012; Revised Submission: 24 August 2012; Accepted: 3 September 2012

Editor: E. Gould

Article published online: 14 December 2012
Clin Microbiol Infect 2013; 19: E118–E128
10.1111/1469-0691.12075

 Corresponding author: N. Tordo, Antiviral Strategies Unit, WHO Collaborative Centre for Arboviruses and Viral Haemorrhagic Fevers, OIE Reference Laboratory for RVFV and CCHFV, Institut Pasteur, 25 rue du Dr Roux, 75724 – Paris – Cedex 15, France
E-mail: ntordo@pasteur.fr

Introduction

Viruses recognized as highly pathogenic for humans must be manipulated in a Biosafety level 4 (BSL–4) laboratory. They include viruses associated with encephalitis and respiratory infections, such as recently emerged members of the genus Henipavirus, family Paramyxoviridae and haemorrhagic fever viruses in the families Arenaviridae, Filoviridae, Bunyaviridae and Flaviviridae [1]. Infections with these viruses lead to a wide spectrum of clinical outcomes, from flu-like and malaria-like symptoms to vascular complications that may cause death [1,2]. Most members of the genus Flavivirus (family Flaviviridae) are arthropod-borne, as are those of the family Bunyaviridae, except for the genus Hantavirus which is rodent-borne or insectivore-borne [2,3]. Viruses of the family Arenaviridae are also rodent-borne [2]. Those of the genus Henipavirus have bat
reservoirs but may also infect humans through contact with infected horses or pigs [4]. Recent data indicate that bats are also probable reservoirs and vectors for viruses of the family Filoviridae [5,6]. Interhuman transmission and nosocomial infections also contribute to spreading the diseases [2,7].

Development of vaccines to prevent infection by these emerging zoonotic viruses is limited and only ribavirin has been used as an efficacious treatment for several of them [1], so early, rapid and specific diagnosis is critically important for disease control. At-risk areas should possess the necessary facilities and equipment, as well as rapid tests, to be prepared for public health emergencies [2–8]. Accurate diagnoses have traditionally relied on specific serological and virological analyses, which include western blotting, ELISA, immunofluorescence staining, genome detection by PCR and quantitative PCR, and ultimately, virus isolation [9–13]. Molecular methods are rapid and specific but are limited by the high genetic variability among different viral strains. To overcome this limitation, macroarray and microarray technology platforms have been developed to detect and identify a large number of pathogens in a single assay [14–20]. Long oligonucleotide probes have been used previously for the detection of viruses associated with haemorrhagic fevers [16]. Low-density macroarrays allowed different variants of Crimean–Congo haemorrhagic fever virus (CCHFV) to be rapidly detected [17], but were complicated by a requisite reverse transcription (RT–) PCR step. High-density resequencing microarrays not only detect pathogens but also determine nucleic acid sequences to single base-pair resolution. A large panel of viral genome sequences from different geographical origins can be characterized in a single test. The high-density resequencing DNA microarray, PATHOGENID v2.0, has been shown to be useful for rapid diagnosis during emerging viral infections, such as the 2009 influenza pandemic [18], and was useful for genotyping members of the family Rhabdoviridae [19].

Here, we used the PATHOGENID v2.0 microarray to detect highly pathogenic viruses. We first validated the microarray with in vitro samples by analysing supernatants from cells infected by prototype virus strains and variants belonging to five families of BSL-4 agents (Arenaviridae, Bunyaviridae, Filoviridae, Flaviviridae, Paramyxoviridae). We then evaluated its performance during a health emergency situation by testing human sera from CCHFV outbreaks in Turkey (2009), Kosovo (2001) and Iran (2009). CCHFV belongs to the genus Nairovirus, family Bunyaviridae and has the largest geographic distribution among haemorrhagic fever viruses [21,22]. Zoonotic infection occurs either directly through its vectors, which are various tick species from the genus Hyalomma, or indirectly through contact with infected livestock. Hospital environments are also vulnerable to inter-human transmissions [23]. CCHFV infection is associated with several clinical outcomes, some of which can become life threatening [22]. CCHFV outbreaks or sporadic cases have occurred in Mauritania [24], Iran [10], Turkey [25], Kosovo [26] and Sudan [23].

Materials and Methods

Ethics statement
This work includes a retrospective study on 12 human sera from clinical specimens submitted to France National-WHO-OIE Reference Centres for diagnosis during CCHF epidemics in Kosovo, Turkey and Iran.

The collection of the remaining samples to be used for scientific purpose was declared to and approved by the Comité de Protection des Personnes, Ile-de-France I and the French Research Ministry (no. DC 2011-1471) according to French regulations.

Animal experimental methods were approved by the Région Rhône Alpes Ethics Committee (France).

Viruses
Viral strains and geographical variants (Table 1 and Table 2) were cultured and isolated in permissive Vero-E6 cells as previously described [11,27]. To simulate the complexity of clinical samples, we pooled RNA samples from different Vero-E6 cell cultures that had each been infected by a single virus. Twelve pooled RNA samples of one to three viruses each were prepared. For Junin virus (family Arenaviridae) and Sin Nombre virus (genus Hantavirus, family Bunyaviridae) synthetic cDNA sequences (Eurofins MWG Operon, Ebersberg, Germany) were used as templates for the amplification step.

Human sera from CCHFV outbreaks
Sera from 12 infected humans were collected during CCHFV outbreaks (2003–09) in the Balkans (five from Kosovo, 2001 and two from Turkey, 2009) and the Middle East (five from Iran, 2009).

Animal biopsies
One non-human primate, a New World squirrel monkey (Saimiri sciureus) was experimentally infected intravenously with 10^7 PFU UM-MC1 Malaysian isolate of Nipah virus [28] as previously described [29]. It was imported from a breeding colony in French Guiana and housed in the BSL-4 animalcare facility in Lyon. The animal was observed daily for signs of disease onset; disease symptoms appeared at day 10 and lasted for 3 days before the moribund monkey was humanely euthanized. A brain biopsy was taken at necropsy and frozen at −80°C.
In another experiment, ten newborn Swiss mice were intracranially inoculated with 20 μL CCHFV (i.e. 10^3 PFU) each in the BSL-4 animal-care facility in Lyon. Seven days after infection, mice were euthanized. Brain was collected, crushed in phosphate-buffered saline 1/10 weight/volume, and clarified by centrifugation for 15 min at 600 g before storage at −80°C.

RNA extraction
RNA extraction was performed using the QIAamp Viral RNA Mini Kit (Qiagen Inc., Valencia, CA, USA) as previously described [11]. For BSL-4 viruses, the cell lysis step was carried out at the Jean Mérieux BSL-4 Laboratory (Lyon, France) according to the validated BSL-4 procedure.

Amplification of viral RNA
Extracted viral RNAs were reverse transcribed into cDNA using SuperScript III reverse transcriptase (Invitrogen Inc., Carlsbad, CA, USA) then amplified by the whole transcriptome amplification (WTA) approach in the presence of random hexamer primers. An optimized protocol based on isothermal amplification by the Phi29 polymerase was applied to the QuantiTect Whole Transcriptome Kit (Qiagen) as previously described [30].

Quantitative RT-PCR and PCR
Quantitative RT-PCR and PCR amplifications of CCHFV sequences present in infected cell supernatants or human sera were performed in a Light-Cycler Instrument (Roche Applied

<table>
<thead>
<tr>
<th>TABLE 1. Sequences of the RNA-dependent RNA polymerase genes of highly pathogenic viruses tiled on the PATHOGENID v2.0 microarray</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Family</strong></td>
</tr>
<tr>
<td>ARENAVIRIDAE</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>BUNYAVIRIDAE</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Phlebovirus</td>
</tr>
<tr>
<td>FLAVIVIRIDAE</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>FILOVIRIDAE</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>PARAMYXOVIRIDAE</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Location referred to the NS5 gene.
*Location referred to the entire genome.
Sciences, Basel, Switzerland) [31]. Treated samples were: (i) extracted RNA, (ii) cDNA obtained following reverse transcription of extracted RNA using random primers, and (iii) WTA products obtained following amplification by Phi29 polymerase.

Hybridization to PATHOGENID v2.0 microarray and data analysis
The PATHOGENID v2.0 microarray is the second generation of a microarray developed through a collaboration between Affymetrix and Institut Pasteur [19,30]. It was designed to detect 949 genes, including 126 different viral sequences [18,19], 18 of which correspond to highly pathogenic viral agents (Table 1).

The entire microarray experimental procedure is summarized in Figure 1. Total cDNA (20–25 μL in 25 μL) that had been amplified from 100 μL of cell culture supernatant or from 25 μL of a serum sample was fragmented, labelled and hybridized overnight at 45°C to the PATHOGENID v2.0...
The array was then washed and scanned according to instructions provided by Affymetrix. Results were analysed using GENECHIP OPERATING SOFTWARE version 4.0 (GCOS), GENECHIP SEQUENCE ANALYSIS SOFTWARE version 4.0 (GSEQ), and the ABACUS algorithm [32].

The call rate value (the percentage of nucleotides identified by the microarray) obtained from each sample hybridized on the microarray was used to determine the degree of hybridization of that sample and to compare it with that of other samples. All the obtained sequences were exported into a FASTA-formatted file and then subjected to BLASTN analysis to identify viral variants.

After scanning and analysis, all the chips were destroyed according to BSL-4 waste guidelines.

**Direct sequencing**

All specimens used either for the validation steps of the PATHOGENID v2.0 microarray or for clinical investigation of the outbreaks, were sequenced directly. To analyse the CCHFV strains, classical, nested or semi-nested PCR were performed to amplify the region tiled on the microarray, e.g. the 531 bases of the L segment encoding the RNA-dependent RNA polymerase. Degenerate primer design and sequence analysis were performed using MacVECTOR software (MacVector Inc., Cary, NC, USA). Primer position refers to the L genome segment of the prototype CCHFV strain (IbAr10200):

- fw2645 (5′-TGCTCWTYYATTTCCCTGTC-3′); rev3269 (5′-TNACACRTTGGGTGACA-3′); fw2576 (5′-GGGGAAA TAAAGCAGACCA-3′); rev3371 (5′-TCYGTAAAGCATT CATTRCCT-3′). The PCR fragments were purified by ultrafiltration before sequencing (Millipore, Billerica, MA, USA). Sequencing was performed using a BigDye Terminator v1.1 cycle sequencing kit (Applied Biosystems, Carlsbad, CA, USA) and purified by ethanol precipitation. Sequence chromatograms from both strands were obtained on an automated sequence analyser ABI3730XL (Applied Biosystems) with the PCR primers. The percentage of sequence divergence was calculated for each sample by determining the number of mutations relative to the prototype sequence tiled on the microarray.
Phylogenetic analysis
A phylogenetic analysis of CCHFV sequences was performed by the neighbour-joining method using BioNUMERICS software for Windows (version 5.1, Applied-Maths, Sint-Martens-Latem, Belgium). The sequences used for this purpose were: (i) all the RNA-dependent RNA polymerase L sequences available in GenBank, (ii) the sequences obtained by direct sequencing and (iii) the sequences obtained from the microarray results.

Results
We used the high-density PATHOGENID v2.0 resequencing microarray to detect and identify a number of different highly pathogenic viruses. This work was divided into two parts: (i) a validation step, in which we used supernatants from cultured cells infected with viral strains that matched the prototype probes tiled on the microarray and their variants, and (ii) an exploratory step, in which we used human sera from CCHFV outbreaks to evaluate the potential of the microarray to be used in public health emergencies.

Detection and differential diagnosis of viral prototype strains and their variants
We assessed whether the PATHOGENID v2.0 microarray could be used to correctly detect and identify different viruses present in a single sample designed to resemble a complex biological specimen that might occur in nature or the laboratory (i.e. screening a pool of samples). Hence, total RNA was extracted from the supernatants of cells that had been infected with a single viral strain. Then, pools of RNA from up to three different supernatants were made to resemble likely combinations that might coexist in the same geographical area or animal host (Table 2). For two viruses that were absent from our laboratory collection (Junin virus and Sin Nombre virus), two plasmids encompassing the synthetic sequences tiled on the microarray were introduced and Sin Nombre virus sequence (Josiah, Sierra Leone) detected the heterologous Alkhurma virus (call rate: 91.3%). Among the family Arenaviridae, the tiled Popp strain of Marburg marburgvirus was as efficient for detection of the Musoke strain (call rate: 98.2%) as for the homologous strain (call rate: 98.4%). The tiled Reston ebolavirus and Zaire ebolavirus sequences allowed detection of the homologous species (call rate: 98.6% and 94.8%, respectively) and an additional variant from Gabon 2001 (call rate: 91.3%). Among the family Flaviviridae, the tiled sequences from the Kyasanur Forest disease virus KFD P 9605 and tick-borne encephalitis virus Neudoerfl detected the heterologous Alkhurma virus (call rate: 73.9%) and Omsk haemorrhagic fever virus Balangul (call rate: 41.7%), respectively. Finally for the genus Henipavirus (family Paramyxoviridae) Nipah virus Malaysia and Hendra virus Australia were perfectly detected by the homologous sequence (call rates: 99.6% and 99.4%, respectively).

In summary, the PATHOGENID v2.0 resequencing microarray very efficiently detected: (i) prototype virus strains and the two synthetic probes with excellent call rates (>97%); (ii) variants with high call rates similar to those of prototype strains (e.g. Lassa virus Guinea, CCHFV Mauritania, Tchoupitoulas virus, Marburg marburgvirus Musoke, Zaire ebolavirus Gabon); (iii) variants with moderate call rates (e.g. 41.6%, for Omsk haemorrhagic fever virus). Variants with low call rates (30.9% for CCHFV from China, 28.7% for Lassa virus from Ivory Coast) were also detected, although less significantly. Finally, highly divergent variants were not detected by the microarray (data not shown).

Application of the microarray to CCHFV outbreaks
We next evaluated the ability of the microarray to detect viruses in human serum samples that were collected during virus outbreaks. CCHFV was chosen as an example because this virus has emerged several times in recent years, particularly in Eastern Europe (Balkan region) and the Middle East. We used (i) sera from 12 CCHFV-positive patients from recent outbreaks in Turkey (2009), Kosovo (2001) and Iran...
(2009); and (ii) four CCHFV strains (Nigeria, Mauritania, Turkey and China) grown in cell culture (Table 2). We sequenced 531 bp of the polymerase gene of each strain/isolate and constructed a phylogenetic tree that also included all the CCHFV sequences available in GenBank (Fig. 2). Phylogenetic analysis distinguished five genetic clusters, as has been previously described [26,33,34]. Two clusters are in Africa: one is spread from western (Mauritania, Senegal and Nigeria) to southern Africa (South African Republic) and includes the Nigerian and Mauritanian sequences; the other is restricted to Equatorial Africa (Congo, Uganda). A Eurasian cluster spreads from Kosovo/Turkey northward to Russia. A Middle East cluster comprises samples from Oman, Iraq, Pakistan, Tajikistan to China. Concerning the viruses present in the human sera we tested, the Iranian viruses formed a distinct branch in the Middle East cluster whereas those from Kosovo and Turkey segregated in two sub-branches of the Eurasian cluster: one together with the Kosovo Hoti strain (Kosovo 423, 426, 429 and Turkey 090139); the other with the Turkey Kelkit06 and 200310849 strains (Turkey 090139 and Kosovo 427).

The microarray clearly detected three out of the four CCHFV reference strains, the China strain being only poorly detected (call rate 30.9% but no BLASTN confirmation). It also allowed the geographical characterization of five out of the 12 CCHFV serum samples: two samples from Turkey and three from Kosovo, all belonging to the Eurasian cluster (Table 3). The two remaining Kosovo samples and all five samples from Iran were not detected.

To determine why these samples were not detected, we characterized the viral genetic material at each step of the detection process (Table 3). We used quantitative PCR to precisely measure the amounts of specific viral genetic material present before and after RNA amplification. The amount of viral RNA in each original sample was comparable to or slightly higher than (±10^1 maximum) the amount of specific cDNA after random priming. This indicated that reverse transcription did not substantially affect the amount of specific viral genetic material.

![FIG. 2. Phylogenetic Tree. Phylogenetic analysis of Crimean–Congo haemorrhagic fever virus (CCHFV) variants was performed using a 531-bp sequence in the CCHFV L segment encoding RNA-dependent RNA polymerase (position 2717–3248) and the neighbour-joining method with BioNÚMERICS software for Windows (version 5.1, Applied Maths). Sequences were: (i) retrieved from GenBank (L); (ii) experimentally obtained from supernatants of CCHFV-infected cell cultures or from infected human serum (Seq); or (iii) the results output from the microarray (Chip) (*).](image-url)

©2012 The Authors
Clinical Microbiology and Infection ©2012 European Society of Clinical Microbiology and Infectious Diseases, CMI, 19, E118-E128
Open access under CC BY-NC-ND license.
material. In contrast, the WTA isothermal amplification of cDNA by Phi29 polymerase significantly increased this amount. The increase from the original amount of RNA in the sample to the cDNA after WTA was between $1.08 \times 10^5$ PFU/mL to $3.46 \times 10^7$ PFU/mL, with a mean increase of $3.72 \times 10^6$ PFU/mL. Interestingly, the lower amount of amplified cDNA detected by the microarray was $2.8 \times 10^5$ PFU/mL for the Kosovo sample 429 (Kosovo samples 423 and 426 with 3.5 \times 10^5 PFU/mL and 8.1 \times 10^4 PFU/mL, respectively, were not detected). On the other hand, the Turkey sample 090139 with $9.9 \times 10^5$ PFU/mL was not detected. As Kosovo and Turkey samples are equally divergent from the tiled sequence (~10%, Table 3), the sensitivity detection limit of the microarray must therefore be estimated between $10^5$ and $10^6$ PFU/mL of amplified cDNA.

Limited genetic material did not explain why the five samples from Iran were not detected, because the amounts of amplified cDNA hybridized on the microarray ($8.9 \times 10^4$ to $5.0 \times 10^5$ PFU/mL) were all well above the $10^7/10^8$ PFU/mL detection limit (Table 3). The China strain was poorly detected despite the presence of sufficient material hybridized (5.0 \times 10^8 PFU/mL). Therefore, a degree of divergence of about 13.7–14.7% from the tiled sequence (Nigerian IbAr10200 strain) is the specificity detection limit of the microarray.

For the sequences detected by the microarray, call rate values were between 29.2% (Kosovo 429) and 70.8% (Kosovo 427), which were globally lower than those obtained from the infected cell supernatants (62.9–99.6%). Nevertheless, the BLASTN analysis allowed the geographical origin of the different isolates to be assessed with a precision dependent on the quality of the call rate. Sequences from samples having call rates >70% were precisely segregated into their specific sub-cluster in the phylogenetic tree along with sequences obtained by their direct sequencing (Fig. 2). This is the case for the Nigeria and Mauritania strains and for the Kosovo 427 serum (Eurasian sub-cluster). The only difference consisted of a longer branch on the tree that was proportional to the number of nucleotides undetermined by the microarray. For the Kosovo 429 and Turkey 090137 sera, which yielded lower call rates (29.2% and 45.7%, respectively), the analysis nevertheless specified that they belonged in the Eurasian cluster.

---

**TABLE 3.** Quantitative evaluation of the different steps of the microarray procedure for the detection of clinical serum samples from Crimean–Congo haemorrhagic fever virus (CCHFV) outbreaks

<table>
<thead>
<tr>
<th>Sample</th>
<th>RNA (PFU/mL)a</th>
<th>cDNA (PFU/mL)a</th>
<th>WTA (PFU/mL)a</th>
<th>Call rateb (%)</th>
<th>BLASTN (Homologous Strain)</th>
<th>Divergencec versus homologous strain (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nigeria (supernatant)</td>
<td>5.9 \times 10^4</td>
<td>8.5 \times 10^4</td>
<td>1.0 \times 10^10</td>
<td>99.6</td>
<td>IbAr10200</td>
<td>0.2</td>
</tr>
<tr>
<td>Mauritania (supernatant)</td>
<td>8.1 \times 10^4</td>
<td>1.8 \times 10^5</td>
<td>2.8 \times 10^11</td>
<td>98.0</td>
<td>ArD399554</td>
<td>1.9</td>
</tr>
<tr>
<td>Turkey (supernatant)</td>
<td>6.2 \times 10^4</td>
<td>1.5 \times 10^5</td>
<td>1.4 \times 10^11</td>
<td>62.9</td>
<td>Turkey2003/0849/Keikits06</td>
<td>10.7</td>
</tr>
<tr>
<td>China (supernatant)</td>
<td>9.1 \times 10^4</td>
<td>2.2 \times 10^5</td>
<td>5.0 \times 10^8</td>
<td>30.9</td>
<td>—</td>
<td>13.7</td>
</tr>
<tr>
<td>Turkey 090137 (serum)</td>
<td>1.2 \times 10^4</td>
<td>2.6 \times 10^4</td>
<td>5.6 \times 10^4</td>
<td>45.7</td>
<td>—</td>
<td>10.15</td>
</tr>
<tr>
<td>Turkey 090139 (serum)</td>
<td>7.6 \times 10^4</td>
<td>2.4 \times 10^4</td>
<td>9.9 \times 10^8</td>
<td>33.5</td>
<td>—</td>
<td>10</td>
</tr>
<tr>
<td>Kosovo 422 (serum)</td>
<td>1.2 \times 10^4</td>
<td>2.4 \times 10^4</td>
<td>6.1 \times 10^4</td>
<td>39.0</td>
<td>Europa</td>
<td>ND</td>
</tr>
<tr>
<td>Kosovo 423 (serum)</td>
<td>2.3 \times 10^4</td>
<td>1.9 \times 10^5</td>
<td>3.5 \times 10^4</td>
<td>22.9</td>
<td>—</td>
<td>9.3</td>
</tr>
<tr>
<td>Kosovo 426 (serum)</td>
<td>1.1 \times 10^4</td>
<td>1.4 \times 10^5</td>
<td>8.1 \times 10^4</td>
<td>ND</td>
<td>ND</td>
<td>9.8</td>
</tr>
<tr>
<td>Kosovo 427 (serum)</td>
<td>1.8 \times 10^4</td>
<td>1.3 \times 10^5</td>
<td>1.9 \times 10^4</td>
<td>70.8</td>
<td>Turkey2003/0849/Keikits06</td>
<td>9.8</td>
</tr>
<tr>
<td>Kosovo 429 (serum)</td>
<td>2.6 \times 10^4</td>
<td>7.0 \times 10^4</td>
<td>2.8 \times 10^3</td>
<td>29.2</td>
<td>Europa</td>
<td>9.7</td>
</tr>
<tr>
<td>Iran 397 (serum)</td>
<td>7.0 \times 10^4</td>
<td>8.5 \times 10^4</td>
<td>2.5 \times 10^7</td>
<td>23.5</td>
<td>—</td>
<td>14.1</td>
</tr>
<tr>
<td>Iran 402 (serum)</td>
<td>3.2 \times 10^4</td>
<td>1.8 \times 10^5</td>
<td>3.9 \times 10^6</td>
<td>26.2</td>
<td>—</td>
<td>14.7</td>
</tr>
<tr>
<td>Iran 406 (serum)</td>
<td>1.8 \times 10^4</td>
<td>2.9 \times 10^4</td>
<td>3.6 \times 10^6</td>
<td>24.0</td>
<td>—</td>
<td>14.5</td>
</tr>
<tr>
<td>Iran 407 (serum)</td>
<td>4.0 \times 10^4</td>
<td>3.2 \times 10^4</td>
<td>5.0 \times 10^6</td>
<td>28.4</td>
<td>—</td>
<td>14.7</td>
</tr>
<tr>
<td>Iran 409 (serum)</td>
<td>4.7 \times 10^4</td>
<td>1.9 \times 10^5</td>
<td>8.9 \times 10^6</td>
<td>21.5</td>
<td>—</td>
<td>14.7</td>
</tr>
</tbody>
</table>

---

**TABLE 4.** Quantitative evaluation of the different steps of the microarray procedure for viral detection in brain samples from experimentally infected animals

<table>
<thead>
<tr>
<th>Sample</th>
<th>RNA (PFU/mL)a</th>
<th>cDNA (PFU/mL)a</th>
<th>WTA (PFU/mL)a</th>
<th>Call rateb (%)</th>
<th>BLASTN (Homologous Strain)</th>
<th>Divergencec versus homologous strain (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCHFV IbAr10200 (mouse brain)</td>
<td>1.9 \times 10^5</td>
<td>2.0 \times 10^7</td>
<td>1.9 \times 10^5</td>
<td>82.6</td>
<td>IbAr10200</td>
<td>1.9</td>
</tr>
<tr>
<td>Nipah virus UMMC1 (monkey brain)</td>
<td>3.08 \times 10^4</td>
<td>4.4 \times 10^7</td>
<td>2.3 \times 10^7</td>
<td>60.9</td>
<td>UMM-MC1</td>
<td>ND</td>
</tr>
</tbody>
</table>

CCHFV, Crimean–Congo haemorrhagic fever virus; ND, not done.

Divergencec versus homologous strain (%) was calculated as the percentage of divergence of each sample versus the homologous strain (IbAr10200 strain).

---

©2012 The Authors
Clinical Microbiology and Infection ©2012 European Society of Clinical Microbiology and Infectious Diseases, CMI, 19, E118-E128
Open access under CC BY-NC-ND license.
Application of the microarray to infected animal brain
The capacity of the microarray to detect viruses in animal samples was tested (Table 4). CCHFV was detected in the brain of newborn mice experimentally infected intracranially for the purpose of virus isolation. The amplification of the cDNA by Phi29 was even more efficient than for human serum samples (increase ratio of $10^2$ from the original RNA, $10^3$ from the cDNA), which indicates that the complexity of genetic material of the sample did not impair WTA amplification. However, the call rate was lower than for supernatants of cells infected with the same CCHFV IbAr10200 strain (82.6% versus 99.6%), suggesting a higher background for hybridization (Table 3).

In addition, the neurotropic Nipah virus was detected in the brain of a monkey moribund upon an experimental intravenous infection. As observed above, the complexity of genetic material in the sample did not significantly affect the amplification of viral material (increase ratio of $5.2 \times 10^3$ from the cDNA) but generated a higher background for hybridization (call rate 60.9% versus 99.6%) compared with that obtained with cell supernatant infected with the same viral strain (Table 3).

Discussion

Highly pathogenic viruses are endemic in developing countries where their impact on public health is especially important in light of the absence of efficacious treatments and vaccines [1]. Occasionally, they can be brought into the developed world by travellers and could be misused for bioterrorism. These viruses produce haemorrhagic fevers, encephalitis or respiratory symptoms, but their aetiology is hard to establish in the absence of specific clinical symptoms. Hence, rapid differential diagnosis during outbreaks represents a critical public health priority.

Among the molecular techniques used in clinical and field diagnosis, (RT)-PCR is considered a reference standard because of its versatility and rapid turnover. However, it may also be limited by pitfalls such as the genetic variability of the viral isolates or doubtful aetiology requiring the design of a battery of specific or degenerated primers, etc. Under these conditions, DNA microarray technology offers the advantage of performing a differential diagnosis in a single test. It has already proven effective for pathogen detection and epidemiological studies [14,20,35]. The GreenChip 60-mer oligonucleotide array provided a good level of sensitivity for the diagnosis of different infections including viral haemorrhagic fevers, but was problematic because it required correction of probe intensities and subtraction of the negative control [16]. The resequencing microarray approach rapidly identifies virus variants while simultaneously characterizing their genome sequences [20]. The confidence levels of these data depend on the virus’ similarity to a tiled reference sequence [36]. It is a promising diagnostic alternative for RNA viruses which have high levels of genetic variability [15,19,37]. The PATHOGENID v2.0 resequencing microarray has precisely identified the geographic origin of virus isolates, which is crucial for monitoring an epidemic or a pandemic [18]. It has also been used to help in genotyping of viruses for taxonomic purposes [19]. In our study, we evaluated the ability of this microarray to detect variants of highly pathogenic BSL-4 viruses from the families Arenaviridae, Bunyaviridae, Flaviviridae, Filoviridae and Paramyxoviridae. We first validated its spectrum in differential diagnosis, then explored its potential in sensitivity and specificity for use with human serum samples from CCHFV outbreaks in Eastern Europe (Balkan region) and the Middle East.

Validation was performed using different types of samples (i.e. cell supernatants, human sera and animal brain) at different degrees of complexity and divergence from the tiled sequence. In single analyses containing multiple virus types, the microarray was able to identify specific viruses among pathogens that produce similar symptoms, and to discriminate between variants of different origins. This is crucial for clinical management of outbreaks that may involve viruses, bacteria or parasites [1,2,9]. The ~48-h procedure required to complete the assay may appear less rapid than classical PCR-based methods. However, when a differential diagnosis is needed for an unknown aetiology, the PATHOGENID v2.0 microarray might be competitive because it does not require (i) designing specific primers for all potential etiologic agents, (ii) setting up the corresponding PCR assays, and (iii) performing the sequence and bioinformatic analyses.

Crimean–Congo haemorrhagic fever virus was chosen as a model infectious agent with which to test the microarray because it has a widespread geographic distribution [8,10,21,22] and substantial genetic diversity [24,25,33,34,38–40]. The Nigerian strain (from the African cluster [41] tiled on the microarray: (i) perfectly detected variants of the same African cluster (e.g. Mauritania, call rate: >97%); (ii) correctly identified viruses of the Eurasian cluster, which are about 9% divergent (e.g. Turkey, call rate: 70%); and (iii) weakly detected viruses from the Middle East cluster (China, call rate: 30.9%). Analysis of human sera from recent epidemics in Kosovo, Turkey and Iran clearly demonstrated the utility of this microarray for detection and characterization by phylogenetic analysis of viruses circulating during outbreaks (Table 3 and Fig. 2). For example, it showed that variants from two sub-groups of the Eurasian cluster were co-circulating in the Balkan region (Kosovo/Turkey), which confirmed previous observations [42].
Using quantitative RT-PCR [28], the microarray sensitivity limit was estimated to be between $10^5$ and $10^6$ PFU/mL of hybridized cDNA per sample. As the mean amplification ratio from the original RNA to the cDNA after WTA was $3.72 \times 10^4$ (considering all samples) and $1.75 \times 10^4$ (only considering serum samples), the detection limit of the microarray is between $10^1$ and $10^2$ equivalent PFU of original viral RNA per mL of serum. This compares favourably with the sensitivity limit of the quantitative RT-PCR method described by Wölfel et al. [31], which detected 1 164 copies/mL of plasma. The specificity limit in terms of divergence from the tiled L segment sequence (531 nucleotides) was estimated at about 13–14%, a value exceeded by the undetected Iranian samples, and approached by the poorly detected China strain (13.7%), which lacked significant match by BLASTN analysis. The CCHFV isolate from Kosovo 429 and the Lassa virus strain AV, despite their low call rates (29.2% and 28.7%, respectively) were detected by BLASTN because they share, respectively, stretches of 21 and 25 consecutive nucleotides with the tiled sequence. This was not the case for the CCHFV isolates Iranian 407 (28.4%) and Turkey 09139 (33.5%) with similar call rates (28.4% and 33.5%, respectively) but sharing no stretches longer than 11 and 16 nucleotides. This indicates that the microarray may preferentially identify sequences that have stretches of ~20 consecutive nucleotides identical to the tiled sequence, regardless of the overall similarity.

Apart from human samples, the potential of the microarray was also preliminarily tested in animal organ material. It was able to detect viruses in brain samples from experimentally infected animals. This has been demonstrated not only in mouse brain that was intentionally infected intracranially for virus isolation, but also in moribund Saimiri sciureus infected intravenously with the neurotropic Nipah virus. In both cases, the amplification of the viral sequences was not affected by the complexity of brain genetic material but a higher background was observed during the hybridization step (lower call rates).

Taking all results together, there is still room to enlarge the spectrum of pathogen detection by increasing the capacity of the microarray. This would allow not only the detection of all currently known isolates but also the discovery of new ones with reliable sequence information. To this purpose, the next-generation panvirological microarray, PATHOGENID v3.0, will include additional CCHFV sequences from the Middle East, Greece and Asian clusters, as well as geographical variants of families Filoviridae (Bundibugyo ebolavirus, Sudan ebolavirus and Ivory Coast strain), Arenaviridae (Ippy virus, Mopeia virus, Mobala virus and Tacaribe virus), Paramyxoviridae (Tioman virus) and Bunyaviridae (Prospect Hill virus). This improved covering of the sequence space will allow detection of new emerging viruses substantially divergent from the tiled sequences.

Acknowledgements

We thank all members of the ‘Programme Transversal de Recherche’ PTR no. 426-DEVA at the Institut Pasteur, Paris: Felix Rey, Jean-Claude Manuguerra, Hervé Bourhy and Michèle Bouloy. We acknowledge the Foundation Mérieux, Lyon, for the availability of the Affymetrix station. We thank the ICAREB (Marie-Noelle Ungeheuer, Sylvie Gerbaud) and the PIRC at Institut Pasteur for assistance in complying with ethics guidelines. Animal experiments were conducted in the BioSafety Level 4 facilities (‘Jean Mérieux’ Laboratory, Lyon, France).

Funding

This research was financially supported by PTR no 426-DEVA, by Grant no. UC1 A1062613 (Kennedy) from the US National Institute of Allergy and Infectious Diseases, National Institutes of Health, by Total-Institut Pasteur Foundation and by three European programmes including CCHFV studies: (i) the FP7 Coordination and Support Action ‘ARBO-ZOONET’ no. 211757, an International Network for Capacity Building for the Control of Emerging Vector-Borne Zoonotic Diseases, particularly CCHF, Rift Valley Haemorrhagic Fever and West-Nile Fever (Coordination: Michèle Bouloy); (ii) the FP7 Large-Scale Integrating Research Project ‘EMPERIE’ no. 223498, an European Management Platform for Emerging and Re-emerging Infectious Disease Entities (Coordination: Albertus Osterhaus); (iii) the FP7 Large-Scale Integrating Research Project ‘EDENEXT’ no. 261504 on Biology and control of vector-borne infections in Europe (Coordination: Renaud Lancelot). Claudia Filippone is financially supported by a postdoctoral grant from Ile de France (DIM) and by the FP7 ‘EMPERIE’. The platform ‘Genotyping of Pathogens and Public Health’ is supported in part by the Institut de Veille Sanitaire (Saint-Maurice, France).

Transparency Declaration

The authors have no conflicts of interest to declare.

Supporting Information

Additional Supporting Information may be found in the online version of this article:
Table S1. Raw sequences in FASTA format obtained following hybridization on the Pathogen IDv2.0 microarray of amplified viral RNA obtained from (A) cellular supernatants, (B) human sera and (C) animal brain. The sequences are listed following the same order of Tables 3, 4 and 5.

References