

## Rapid Molecular Strategy for Filovirus Detection and Characterization<sup>▽</sup>

Junhui Zhai,<sup>1†</sup> Gustavo Palacios,<sup>1†</sup> Jonathan S. Towner,<sup>2</sup> Omar Jabado,<sup>1</sup> Vishal Kapoor,<sup>1</sup>  
Marietjie Venter,<sup>3</sup> Allen Grolla,<sup>4</sup> Thomas Briese,<sup>1</sup> Janusz Paweska,<sup>5</sup> Robert Swanepoel,<sup>5</sup>  
Heinz Feldmann,<sup>4,6</sup> Stuart T. Nichol,<sup>2</sup> and W. Ian Lipkin<sup>1\*</sup>

*Jerome L. and Dawn Greene Infectious Disease Laboratory, Mailman School of Public Health, Greene Infectious Disease Laboratory, New York, New York<sup>1</sup>; Special Pathogens Branch, Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia<sup>2</sup>; Department of Medical Virology, University of Pretoria/NHLS Tshwane Academic Division, Pretoria, South Africa<sup>3</sup>; Special Pathogens Unit, National Institute for Communicable Diseases, Sandringham, South Africa<sup>4</sup>; Special Pathogens Program, National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada<sup>5</sup>; and Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba, Canada<sup>6</sup>*

Received 11 September 2006/Returned for modification 7 October 2006/Accepted 11 October 2006

**Filoviruses have the capacity to cause lethal outbreaks of hemorrhagic fever in primates. Here we present a simple consensus reverse transcription-PCR method for filovirus recognition and characterization and demonstrate its utility with all known filovirus strains. Phylogenetic assignment is achieved by automated web-based sequence analysis of amplification products.**

Since filoviruses were first recognized in 1967, more than 1,800 cases of human infection have been reported, with mortality rates as high as 90%. Interest in filoviruses has increased with the globalization of travel and trade, and the attendant potential for their appearance in new contexts, and the concern that highly pathogenic viruses may be exploited as biological weapons. Current treatment options for infected individuals are only supportive; nonetheless, early accurate diagnosis is critical for the control of contagion and is likely to become an important factor in clinical outcome as new management strategies come into practice. PCR-based assays have become fundamental tools in clinical diagnosis and outbreak control (6, 10, 13). A major challenge to the utility of these assays is the diversity of filoviral genomes. To address this challenge we developed a consensus PCR (cPCR) method that utilizes a cocktail of specific primers in a one-step reverse transcription-PCR (RT-PCR). The resulting assay is less sensitive to sequence drift than assays currently in use. This method has been joined to a program for sequence analysis of cPCR product that allows automated speciation.

The primers were designed by using Greene SCPrimer software (<http://scprimer.cpmc.columbia.edu/SCPrimerApp.cgi>). The program uses a greedy algorithm to identify the most conserved sequences and create the minimum set of primers needed for amplification of all sequences in the alignment. Primers are selected within standard design constraints ( $T_m = 55$  to 65°C, GC content = 40 to 60%, no hairpins); degenerate positions are introduced in cases where necessary (Table 1).

The assay was standardized using synthetic DNA standards containing Zaire ebolavirus (ZEBOV), Sudan ebolavirus

(SEBOV), and Lake Victoria marburgvirus (MARV) L-gene sequences cloned into pGEM-T-Easy vector system (Promega, Fermantas, Lithuania). The PCR parameters (MgCl<sub>2</sub> concentration, annealing temperature, primer concentration, and enzymatic system) were optimized individually. We found 2.5 mM MgCl<sub>2</sub>, 58°C, 0.5 μM concentrations of each primer, and the QIAGEN One-Step RT-PCR kit (QIAGEN, Valencia, CA) to be optimal.

Synthetic RNA standards were generated from linearized target sequence plasmids using T7 polymerase (mMessage mMachine; Invitrogen, Carlsbad, CA). After quantitation by UV spectrometry, RNA was serially diluted in 2.5 μg of yeast tRNA (Sigma)/ml and analyzed by filovirus consensus one-step RT-PCR (RT: 50°C for 30 min; PCR: 94°C for 12 min to inactivate the polymerase antibodies, followed by 94°C for 30 s, 58°C for 50 s, and 72°C for 90 s for 35 cycles, with a final extension for 7 min). The sensitivity of the assay was 50 to 500 RNA copies per assay, where each assay represents 100 μl of blood.

Nineteen filovirus strains were tested, including MARV strains Musoke, Popp, Voegel, Angola, Ozolin, 05DRC99, 07DRC99, 09DRC99, and RAVN; Reston ebolavirus (REBOV) strain Pennsylvania; Cote d'Ivoire ebolavirus (CIEBOV); SEBOV strains Gulu, Boniface, Maleo, and Yambio; and ZEBOV strains Mayinga, Zaire95, Gabon 1994, and Gabon 1996. All extracts yielded products of appropriate length. No amplification products were detected in extracts of uninfected cells or cells infected with Lassa fever, Crimean-Congo hemorrhagic fever, or Rift Valley fever viruses.

Performance with clinical materials was evaluated using blood from 13 human victims of filoviral hemorrhagic fevers, including five cases of ZEBOV from the 1995 Kikwit outbreak in the Democratic Republic of the Congo (DRC), four cases of MARV VHF from the 2000 Durba outbreak in the DRC, and four cases of MARV VHF from the 2005 Uige outbreak in Angola. Infections had been previously diagnosed through virus isolation or RT-PCR. All samples yielded positive results. RNA was extracted from 200 μl of blood. After dissolution in

\* Corresponding author. Mailing address: Jerome L. and Dawn Greene Infectious Disease Laboratory, Mailman School of Public Health, Columbia University, 722 West 168th St., Rm. 1801, New York, NY 10032. Phone: (212) 342-9031. Fax: (212) 342-9044. E-mail: wil2001@columbia.edu.

† J.Z. and G.P. contributed equally to this study.

▽ Published ahead of print on 1 November 2006.

TABLE 1. Primers for filovirus consensus RT-PCT<sup>a</sup>

Forward primer	Sequence	Reverse primer	Sequence	Gene
Greene-Filo-U12683-A	TATTCTCYCTACAAAAGCATTGGG	Greene-Filo-L13294-A	GCTTCTGCGAGTGTGGACATT	L
Greene-Filo-U12683-B	TATTTTCCATTCAAAAACACTGGG	Greene-Filo-L13294-B	GCTTCACAAAGTGTGGACATT	L
Greene-Filo-U12683-C	TATTTTCAATCCAAAAGCACTGGG	Greene-Filo-L13294-C	GCTTCGCAGAGGGTTGGACATT	L
Greene-Filo-U12683-D	TATTCTCTGTTCAAAAACATTGGG	Greene-Filo-L13294-D	GCCTCACATAAAGTTGGACATT	L

<sup>a</sup> Primer positions refer to ZEBOV strain Mayinga sequences (AF086833).

10 µl of water, 5 µl of total RNA were reverse transcribed using random hexamer primers. The resulting cDNA was dried to facilitate shipment to Columbia University. Prior to PCR amplification, the cDNA was reconstituted in 5 µl of water.

Templates for dideoxy sequencing were obtained by excising the band of interest after size fractionation of PCR products by 1% agarose gel electrophoresis. A set of 40 filovirus L-gene sequences (593 nucleotides [nt]) comprising 18 newly sequenced strains and 22 previously available in GenBank (July

2006) were aligned by using the program CLUSTAL X (version 1.83) (11). Phylogenetic analysis was performed by using the best model of nucleotide substitution (according to Modeltest [8], HKY 85 [3], with gamma distribution correction = 0.304). Programs from the MEGA package (version 3) (4) were used to produce phylogenetic trees, reconstructed through the neighbor-joining method. The statistical significance of a particular tree topology was evaluated by bootstrap resampling of the sequences 1,000 times. The phylogenetic tree obtained is shown

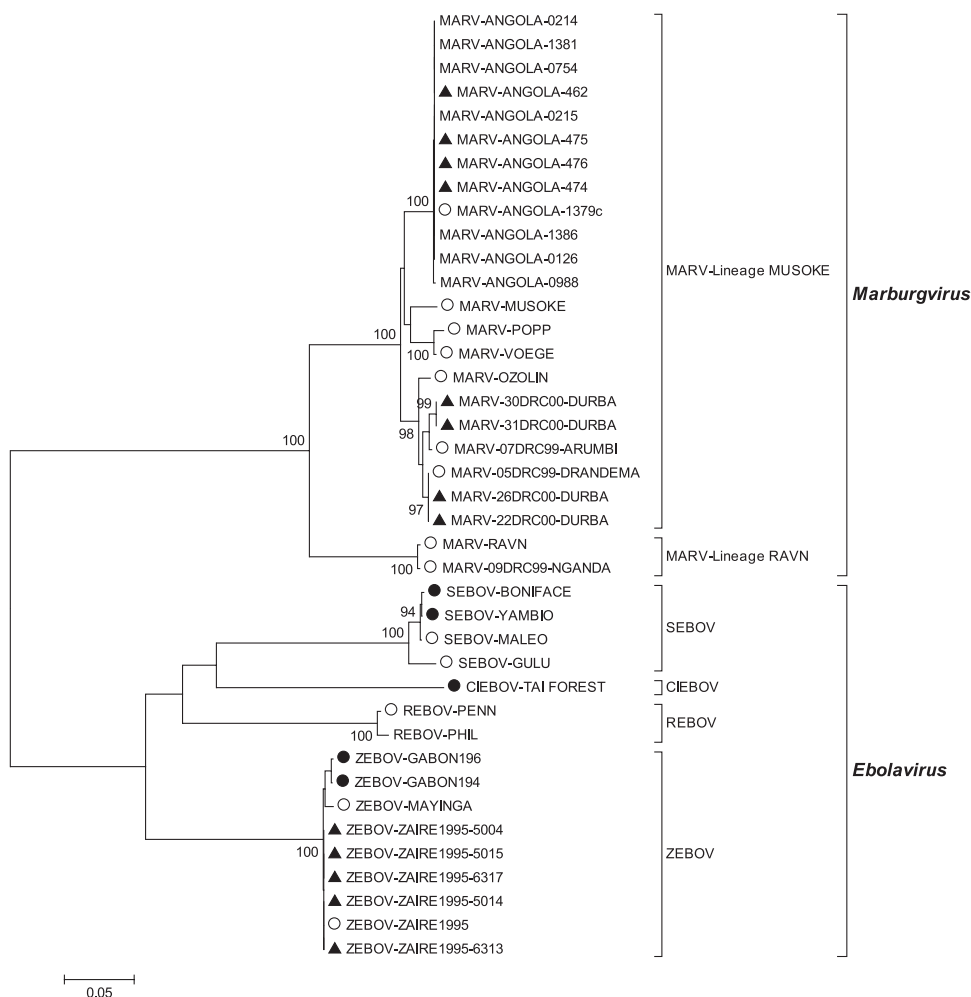


FIG. 1. Phylogenetic analysis of filoviral L-gene sequences. Forty filoviral RNA polymerase nucleotide sequences (regions comprising 593 nt) were aligned with CLUSTAL W. Phylogenetic analysis was performed by using the best model of nucleotide substitution according to Modeltest, HKY 85, with correction for gamma distribution = 0.3040, using the neighbor-joining method to reconstruct the phylogenetic tree (MEGA version 3.0 software package). Statistical significance was estimated by bootstrap analysis with 1,000 pseudoreplicate data sets. Strains are denoted by genus, species, and lineage. Black triangles indicate clinical samples detected and characterized in the present study. Open circles indicate virus strains used for assay validation. Black circles indicate virus isolates not previously sequenced in this 593-nt region.

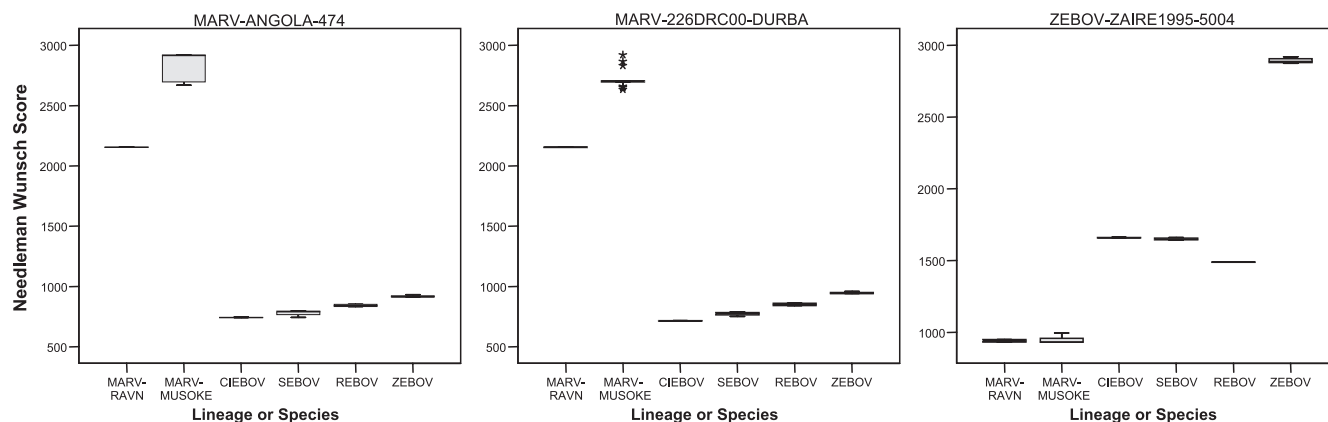


FIG. 2. Pairwise analysis of three virus sequences randomly chosen from the clinical samples obtained in the present study. A pairwise Needleman-Wunsch (NW) score was calculated for MARV-226DRC00, MARV species MUSOKE lineage; MARV-ANGOLA-474, MARV species, MUSOKE lineage; and ZEBOV-ZAIRE 1995-5004, ZEBOV species, against all other members of the database. Graphs plot the 75th-to-25th percentile range of the NW score (box), outliers (\*), maximum ( $\ominus$ ), and minimum ( $\oplus$ ) NW score values for each of the three viruses with respect to individual filoviral lineages.

in Fig. 1. The analysis not only allowed the classification within known filovirus species, MARV, CIEBOV, SEBOV, REBOV, and ZEBOV, but also confirmed the existence of lineages (MUSOKE and RAVN), a finding consistent with established genetic relationships (12).

A pairwise sequence comparison was done within filovirus sequences to assess the potential for establishing a simple program for classification of filoviral species and lineages similar to programs previously built for mumps (7), dengue serotype 1 (2), and adenoviruses (1). Pairwise comparisons were done by global alignment using the Needleman Wunsch algorithm (5), implemented by a program from EMBOSS (the European molecular biology open software suite) (9). This analysis allowed the determination of species and lineage in all clinical samples (Fig. 2). The validity of the method was confirmed by analysis of variance, comparing the scores of sequence comparisons within species or lineages to comparisons between genotypes. The automated program for filoviral analysis is available for public use at <http://www.greeneidlab.columbia.edu>. Sequence submissions are retained in the database; each use results in incorporation of filoviral sequences introduced since it was last queried.

The cPCR assay described here is not the first reported for filovirus surveillance. Sanchez et al. established an assay in 1999 that continues to serve successfully in laboratory and field applications (10). However, the second-generation cPCR assay we describe here has the advantage of yielding a 640-nt product (including primer sequence) that can be sequenced for automated phylogenetic analysis, allowing a more accurate placement of newly identified filovirus-positive samples relative to existing species, lineages, and strains. The performance of the assay has been validated with all known filovirus strains, including CIEBOV, a species for which there is as yet no published L-gene sequence. The database that is the foundation for automated phylogenetic analysis is continually updated ensuring currency of sequence information. Use of this technology is anticipated to provide insight into the appearance and distribution of known and novel filoviral species and lineages and to enhance rapid response to these and other emerging pathogens.

This study was supported by National Institutes of Health awards AI51292, AI056118, AI55466, and US4AI57158 (Northeast Biodefense Center-Lipkin) and by the Ellison Medical Foundation.

#### REFERENCES

- Casas, I., A. Avellon, M. Mosquera, O. Jabado, J. E. Echevarria, R. H. Campos, M. Rewers, P. Perez-Brena, W. I. Lipkin, and G. Palacios. 2005. Molecular identification of adenoviruses in clinical samples by analyzing a partial hexon genomic region. *J. Clin. Microbiol.* **43**:6176-6182.
- Domingo, C., G. Palacios, O. Jabado, N. Reyes, M. Niedrig, J. Gascon, M. Cabrerizo, W. I. Lipkin, and A. Tenorio. 2006. Use of a short fragment of the C-terminal E gene for detection and characterization of two new lineages of dengue virus 1 in India. *J. Clin. Microbiol.* **44**:1519-1529.
- Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**:160-174.
- Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**:1244-1245.
- Needleman, S. B., and C. D. Wunsch. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**:443-453.
- Palacios, G., T. Briese, V. Kapoor, O. Jabado, Z. Liu, M. Venter, J. Zhai, N. Renwick, A. Grolla, T. W. Geisbert, C. Drosten, J. Towner, J. Ju, J. Paweska, S. T. Nichol, R. Swanepoel, H. Feldmann, P. B. Jahrling, and W. I. Lipkin. 2006. MassTag polymerase chain reaction for differential diagnosis of viral hemorrhagic fever. *Emerg. Infect. Dis.* **12**:692-695.
- Palacios, G., O. Jabado, D. Cisterna, F. de Ory, N. Renwick, J. E. Echevarria, A. Castellanos, M. Mosquera, M. C. Freire, R. H. Campos, and W. I. Lipkin. 2005. Molecular identification of mumps virus genotypes from clinical samples: standardized method of analysis. *J. Clin. Microbiol.* **43**:1869-1878.
- Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**:817-818.
- Rice, P., I. Longden, and A. Bleasby. 2000. EMBOSS: the European molecular biology open software suite. *Trends Genet.* **16**:276-277.
- Sanchez, A., T. G. Ksiazek, P. E. Rollin, M. E. Miranda, S. G. Trappier, A. S. Khan, C. J. Peters, and S. T. Nichol. 1999. Detection and molecular characterization of Ebola viruses causing disease in human and nonhuman primates. *J. Infect. Dis.* **179**(Suppl. 1):S164-S169.
- Thompson, J., T. Gibson, F. Plewniak, F. Jeanmougin, and D. Higgins. 1997. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**:4876-4882.
- Towner, J. S., M. L. Khristova, T. K. Sealy, M. J. Vincent, B. R. Erickson, D. A. Bawiec, A. L. Hartman, J. A. Comer, S. R. Zaki, U. Stroher, F. Gomes da Silva, F. del Castillo, P. E. Rollin, T. G. Ksiazek, and S. T. Nichol. 2006. Marburgvirus genomics and association with a large hemorrhagic fever outbreak in Angola. *J. Virol.* **80**:6497-6516.
- Towner, J. S., P. E. Rollin, D. G. Bausch, A. Sanchez, S. M. Crary, M. Vincent, W. F. Lee, C. F. Spiropoulou, T. G. Ksiazek, M. Lukwiya, F. Kaducu, R. Downing, and S. T. Nichol. 2004. Rapid diagnosis of Ebola hemorrhagic fever by reverse transcription-PCR in an outbreak setting and assessment of patient viral load as a predictor of outcome. *J. Virol.* **78**:4330-4341.