

# Viral Metagenomic Analysis of Cerebrospinal Fluid from Patients with Acute Central Nervous System Infections of Unknown Origin, Vietnam

Nguyen To Anh, Le Nguyen Truc Nhu, Nguyen Thi Thu Hong, Tran My Phuc, Pham Thi Thanh Tam, Dang Thao Huong, Tran Tuan Anh, Xutao Deng, Ho Dang Trung Nghia, Tran Thua Nguyen, Nguyen Van Hung, Nguyen Dac Thuan, Pham Thi Hong Phuong, Nguyen Van Vinh Chau, Stephen Baker, Eric Delwart, Guy Thwaites, Le Van Tan, for the VIZIONS Consortium<sup>1</sup>

Central nervous system (CNS) infection is a serious neurologic condition, although the etiology remains unknown in >50% of patients. We used metagenomic next-generation sequencing to detect viruses in 204 cerebrospinal fluid (CSF) samples from patients with acute CNS infection who were enrolled from Vietnam hospitals during 2012–2016. We detected 8 viral species in 107/204 (52.4%) of CSF samples. After virus-specific PCR confirmation, the detection rate was lowered to 30/204 (14.7%). Enteroviruses were the most common viruses detected ( $n = 23$ ), followed by hepatitis B virus (3), HIV (2), molluscum contagiosum virus (1), and gemycircularvirus (1). Analysis of enterovirus sequences revealed the predominance of echovirus 30 (9). Phylogenetically, the echovirus 30 strains belonged to genogroup V and VIIb. Our results expanded knowledge about the clinical burden of enterovirus in Vietnam and underscore the challenges of identifying a plausible viral pathogen in CSF of patients with CNS infections.

**W**orldwide, the annual incidence of acute encephalitis in nonoutbreak settings during 1983–2000 ranged from 0.07 to 12.6 cases/100,000 population (1). According to the World Health Organization, meningitis caused 379,000 deaths and encephalitis caused 150,000 deaths globally in 2015 (2). As a consequence,

central nervous system (CNS) infection is a leading cause of years lived with disability in low-income countries (3).

More than 100 known pathogens can cause CNS infections (1). However, the distribution of CNS infection pathogens is geographically dependent and has been shaped by the emergence of novel viruses. In Asia, Nipah virus and enterovirus A71 have been recognized as emerging neurotropic pathogens over the past few decades. In 1999, West Nile virus arrived in the United States and since then has established endemic circulation (4).

Despite recent advances in molecular diagnostics, especially sensitive virus-specific PCR, encephalitis cases of unknown origin remain a substantial problem. Worldwide,  $\approx 50\%$  of patients with CNS infections have no etiology identified (1,5,6).

Over the past decade, metagenomic next-generation sequencing (mNGS) has emerged as a sensitive hypothesis-free approach for detection of pathogens (especially viruses) in clinical samples (7). However, in resource-limited settings like Southeast Asia and Vietnam, a limited number of mNGS studies

Author affiliations: Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam (N.T. Anh, L.N.T. Nhu, N.T.T. Hong, T.M. Phuc, P.T.T. Tam, D.T. Huong, T.T. Anh, H.D.T. Nghia, G. Thwaites, L.V. Tan); Hue Central Hospital, Hue City, Vietnam (T.T. Nguyen); Dak Lak General Hospital, Ban Me Thuot City, Vietnam (N.V. Hung); Khanh Hoa General Hospital, Nha Trang City, Vietnam (N.D. Thuan); Dong Thap General Hospital, Dong Thap, Vietnam (P.T.H. Phuong); Vitalant Research Institute, San Francisco, California, USA (X. Deng, E. Delwart); University of California Department of Laboratory Medicine,

San Francisco (X. Deng, E. Delwart); Pham Ngoc Thach University, Ho Chi Minh City (H.D.T. Nghia); Hospital for Tropical Diseases, Ho Chi Minh City (N.V.V. Chau); University of Cambridge Institute of Therapeutic Immunology and Infectious Disease, Cambridge, UK (S. Baker); Centre for Tropical Medicine, Nuffield Department of Medicine, University of Oxford, Oxford, UK (S. Baker, G. Thwaites)

DOI: <https://doi.org/10.3201/eid2701.202723>

<sup>1</sup>Members are listed at the end of this article.

examining known and unknown viruses in cerebrospinal fluid (CSF) samples from patients with CNS infections have been conducted, even though in this tropical region of the world, novel viruses are likely to emerge (P. Zhou et al., unpub. data, <https://doi.org/10.1101/2020.01.22.914952>), and diverse CNS infection pathogens have been documented. Collectively, improving our knowledge about viral causes of CNS infections is essential for clinical management and development of intervention strategies. In this study, by using a mNGS approach, we set out to search for known and unknown viruses in CSF samples collected from patients in Vietnam with CNS infections of unknown causes who were enrolled in a hospital-based surveillance study conducted during 2012–2016.

## Materials and Methods

### Clinical Study and Selection of CSF Samples for mNGS Analysis

The study used CSF samples collected from patients with suspected CNS infection enrolled in a hospital-based surveillance program conducted in Vietnam during December 2012–October 2016 (5). The study was conducted as part of the Vietnam Initiative on Zoonotic Infections (VIZIONS) project (5), and patient recruitment was carried out at 7 provincial hospitals across Vietnam. After collection, as per the study protocol, all CSF samples were tested for a range of pathogens by using the diagnostic work-up of the clinical study (Appendix Table 1, <https://wwwnc.cdc.gov/EID/article/27/1/20-2723-App1.pdf>). The remaining volume of the CSF samples were stored at  $-80^{\circ}\text{C}$  for further testing.

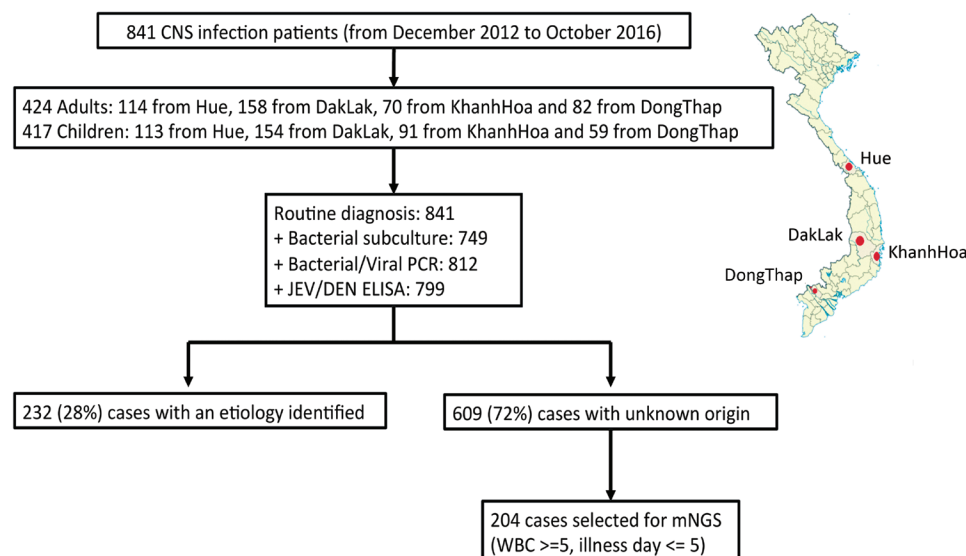
We focused our metagenomic analysis on patients of unknown origin from 4 provincial hospitals in central (Hue and Khanh Hoa), highland (Dak Lak), and southern (Dong Thap) Vietnam (Figure 1), representing 3 distinct geographic areas in Vietnam. To increase the chance of detecting a virus in the CSF samples, we only selected patients with CSF leukocyte counts  $\geq 5$  cells/ $\text{mm}^3$  and an illness duration  $\leq 5$  days.

### mNGS Assay

mNGS assay was carried out as previously described (8). Before viral nucleic acid (NA) isolation, 100  $\mu\text{L}$  of each CSF sample was treated with Turbo DNase (Ambion, Life Technology, ThermoFisher, <https://www.thermoFisher.com>) and RNase I enzyme (Ambion). Then viral NA was isolated using a QIAamp viral RNA kit (QIAGEN GmbH, <https://www.qiagen.com>), and recovered in 50  $\mu\text{L}$  of elution buffer provided with the extraction kit. Double-stranded DNA was synthesized from the isolated viral NA by using a set of 96 nonribosomal primers (FR26RV–Endoh primers) and then was randomly amplified by using the FR20RV primer (5'-GCCGGAGCTCTGCAGATATC-3'). Finally, the amplified product was subjected to a library preparation step by using Nextera XT sample preparation kit (Illumina, <https://www.illumina.com>), following the manufacturer's instructions, and sequenced by using a MiSeq reagent kit, version 3 (600 cycles) (Illumina) in a MiSeq platform (Illumina).

### mNGS Data Analysis

Potential viral reads were identified by using an in-house viral metagenomic pipeline running on a 36-node Linux cluster as described previously (9). In



**Figure 1.** Flowchart overview of diagnostic results for study of patients with suspected central nervous system infections admitted to 4 of 7 provincial hospitals, Vietnam, December 2012–October 2016. Inset map indicates places where samples were collected (red dots).

brief, after duplicate reads and reads belonging to human or bacterial genomes were filtered out, the remaining reads were assembled *de novo*. The resulting contigs and singlet reads were then aligned against a customized viral proteome database by using an approach based on BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Next, the candidate viral reads were aligned against a nonredundant nonvirus protein database to remove any false-positive reads (i.e., reads with expected values higher than those against viral protein databases). Any virus-like sequence with an expected value  $\leq 0.00001$  was considered a significant hit. Finally, a reference-based mapping approach (Genious 8.1.5; Biomatters, <https://www.geneious.com>) was used to assess the levels of identity and genome coverage of the corresponding viruses.

### PCR Confirmatory Testing of mNGS Results

PCR assays were conducted to confirm mNGS hits for each specific virus identified from the viral metagenomic pipeline. Depending on availability of CSF, the PCR confirmations were performed either on leftover NA or newly extracted NA. A viral mNGS result was considered positive only if it was subsequently confirmed by PCR analysis of the original NA samples. The nucleotide sequences of primers and probes used for PCR confirmatory testing are shown in Appendix Table 2 (8).

### Serotype Identification and Phylogenetic Analysis

For enterovirus serotype determination based on the obtained sequences generated by viral mNGS, we used a publicly available genotyping tool (10). To determine the relationship between enterovirus strains we sequenced and global strains, we first performed pairwise alignment by using the ClustalW tool in Geneious 8.1.5, and then reconstructed a maximum-likelihood phylogenetic tree by using IQ Tree 1.4.3 (11). A similar phylogenetic approach was used for other viruses. The generated sequences of this study were submitted to GenBank (accession no. PRJNA561465).

### Ethics

The study was approved by the corresponding institutional review board of local hospitals in Vietnam, where the patients were enrolled, and the Oxford Tropical Ethics Committee. Informed consent was obtained from each study participant or a legal guardian.

### Results

#### CSF Samples Available for mNGS Analysis

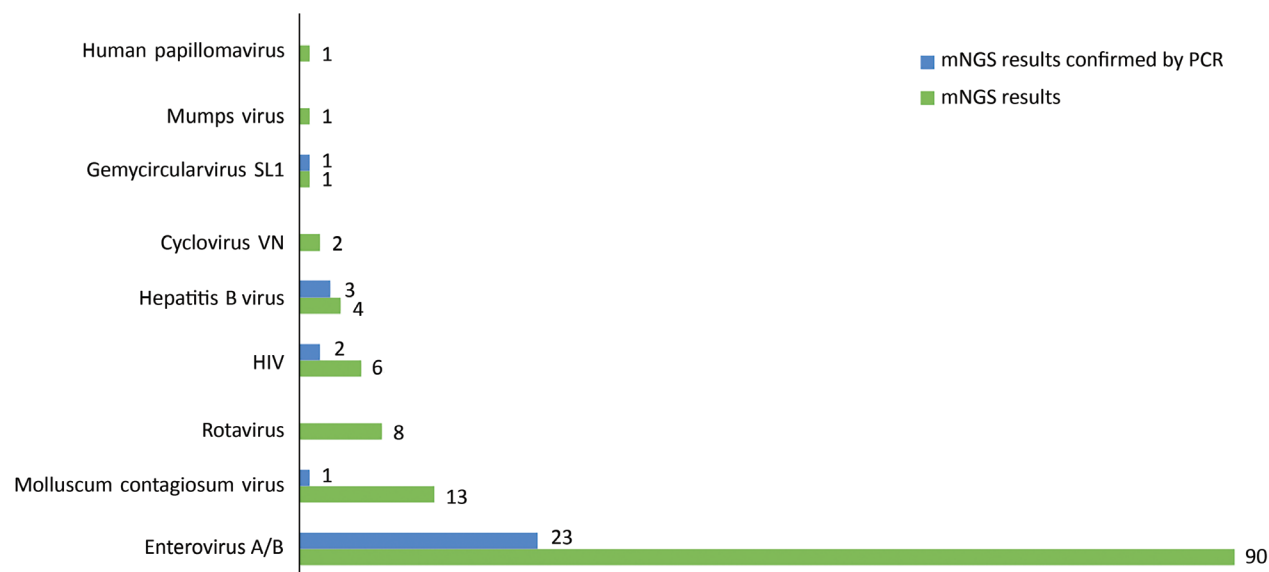
From the clinical study described previously, a total of 841 patients were enrolled from Hue, Khanh Hoa, Dak Lak, or Dong Thap provincial hospitals. Of these, 609/841 (72%) patients had no etiology identified. The etiologic profiles of the patients in

**Table.** Baseline characteristics and clinical data of patients with acute central nervous system infections enrolled for mNGS analysis of CSF samples, Vietnam, December 2012–October 2016\*

Characteristic	Patients with unknown cause enrolled for mNGS, n = 204	Patients with mNGS negative, n = 174	Patients with enterovirus detected, n = 23	p value†
Sex				
M	135 (66)	114 (65.5)	15 (65)	
F	69 (34)	60 (34.5)	8 (35)	
Age, y, median (range)	20.5 (0–92)	24 (0–92)	13 (2–27)	0.005
Location				
Hue	37 (18)	28 (16)	9 (39)	
Dak lak	98 (48)	87 (50)	10 (43.5)	
Khanh Hoa	28 (14)	22 (13)	4 (17.5)	
Dong Thap	41 (20)	37 (21)	0	
3-d fever (at enrollment or preceding 3 d)				
Fever	148 (72.5)	126 (72.4)	17 (74)	0.054
Temperature, C°, median (range)	39 (37.5–42.0)	39 (37.5–42.0)	38.5 (38.0–40.5)	
Fever with unknown temperature	29 (14.2)	22 (12.6)	6 (26)	
No fever	20 (9.8)	19 (11)	0	
Unknown	7 (3.5)	7 (4)	0	
Outcome				
Death or discharge to die	22 (11)	22 (12.6)	0	
Discharge with complete recovery	108 (53)	86 (49.4)	18 (78.3)	
Discharge with incomplete recovery	35 (17)	31 (17.8)	2 (8.7)	
Transfer to another hospital	34 (16.5)	30 (17.2)	3 (13)	
Other (patient request)	3 (1.5)	3 (1.7)	0	
Unknown	2 (1)	2 (1.3)	0	
CSF white cells, cells/mm <sup>3</sup> median (min–max)	88.5 (5–40,000)	71.5 (5–40,000)	110 (8–1200)	0.343

\*Values are no. (%) except as indicated, CSF, cerebrospinal fluid; mNGS, metagenomic next-generation sequencing.

†Statistical comparisons were performed for groups of patients with mNGS-negative results and enterovirus detected, by Mann-Whitney test.



**Figure 2.** Number of cerebrospinal fluid samples with detected viruses by metagenomic next-generation sequencing and then confirmed by virus-specific PCR or reverse-transcription PCR, Vietnam, December 2012–October 2016. Samples were collected from patients with suspected central nervous system infection. For human papillomavirus, confirmatory testing was not performed because of the unavailability of a PCR assay.

whom a pathogen was detected will be reported separately. Of the patients in whom a pathogen was not identified, 204 met our selection criteria, and their CSF samples were included for viral mNGS analysis (Figure 1).

#### Baseline Characteristics of the Included Patients

The baseline characteristics and outcome of the 204 study patients are described in Table 1. Male patients were predominant. A substantial proportion of the patients were seriously ill; fatal outcome was recorded in 22 (11%), whereas incomplete recovery was recorded in 17% ( $n = 35$ ) and deterioration (reflected by being transferred to other hospitals) in 16.5% ( $n = 34$ ).

#### General Description of mNGS Results

A total of 204 CSF samples were subjected to 3 NGS runs, and 108 million reads were obtained (median number of reads per sample 445,412 [range 430–908,890]). Of these, viral reads accounted for 0.64% ( $n = 692,731$ ; median number of reads per sample 2,001 [range 4–268,933]). Excluding common contaminants and commensal viruses such as torque teno virus, which are not reported in this article, sequences related to a total of 8 distinct viral species were identified in 107/204 (52.4%) patients. These

viruses are either known to be infectious to humans (e.g., enteroviruses, rotavirus, mollusum contagiosum virus [MCV], human papillomavirus, HIV, and hepatitis B virus [HBV]) or are without evidence of human infections besides previous detection in sterile human samples (e.g., cyclovirus-VN and gemyrcircularvirus) (Figure 2).

#### mNGS Result Assessment by Specific PCR Analysis

After virus-specific PCR confirmatory testing, the proportion of patients in whom a virus was found by mNGS was reduced from 53% (108/204) to 14.7% (30/204). Accordingly, the number of virus species was reduced from 8 to 5 (Figure 2); enteroviruses were the most common virus detected, accounting for 11.3% (23/204) of the included patients, followed by HBV ( $n = 3$ ), HIV ( $n = 2$ ), gemyrcircularvirus, and MCV (1 each) (Figure 2). Because of the focus of our study and the unavailability of the PCR assays, confirmatory testing for human papillomavirus was not performed.

#### Characteristics of the 23 Enterovirus-Infected Patients

All 23 enterovirus-infected patients were admitted to hospitals from the central or highland areas (Table), and none were from Dong Thap Province. Male patients were slightly predominant, accounting for

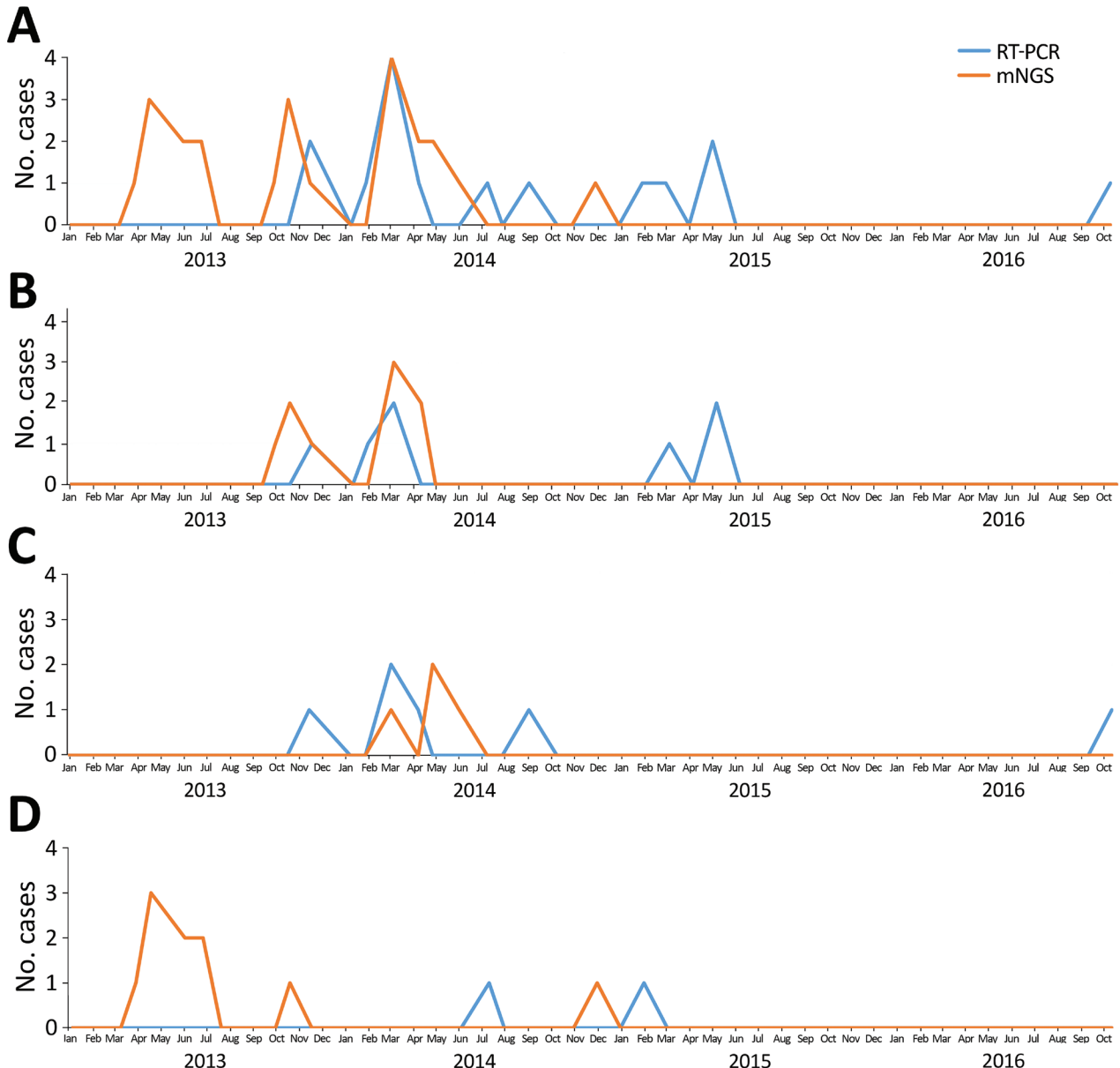
56%. Notably, the enterovirus-infected patients were younger than those who were mNGS-negative (Table). At discharge, incomplete recovery or transfer to other hospitals because of disease deterioration were recorded in 21.7% (Table).

Enterovirus cases were not detected during January 2015–December 2016. During 2013 and 2014, two main peaks were observed during March–July and September–December (Figure 3, panel A); cases from Dak Lak and Khanh Hoa contributed to the first peak (Figure 3, panels B and C), and cases from Khanh Hoa

and Hue contributed to the second (Figure 3, panels C and D). The general baseline characteristics of patients with HBV, HIV, gemycircularvirus, and MCV are shown in Appendix Table 3.

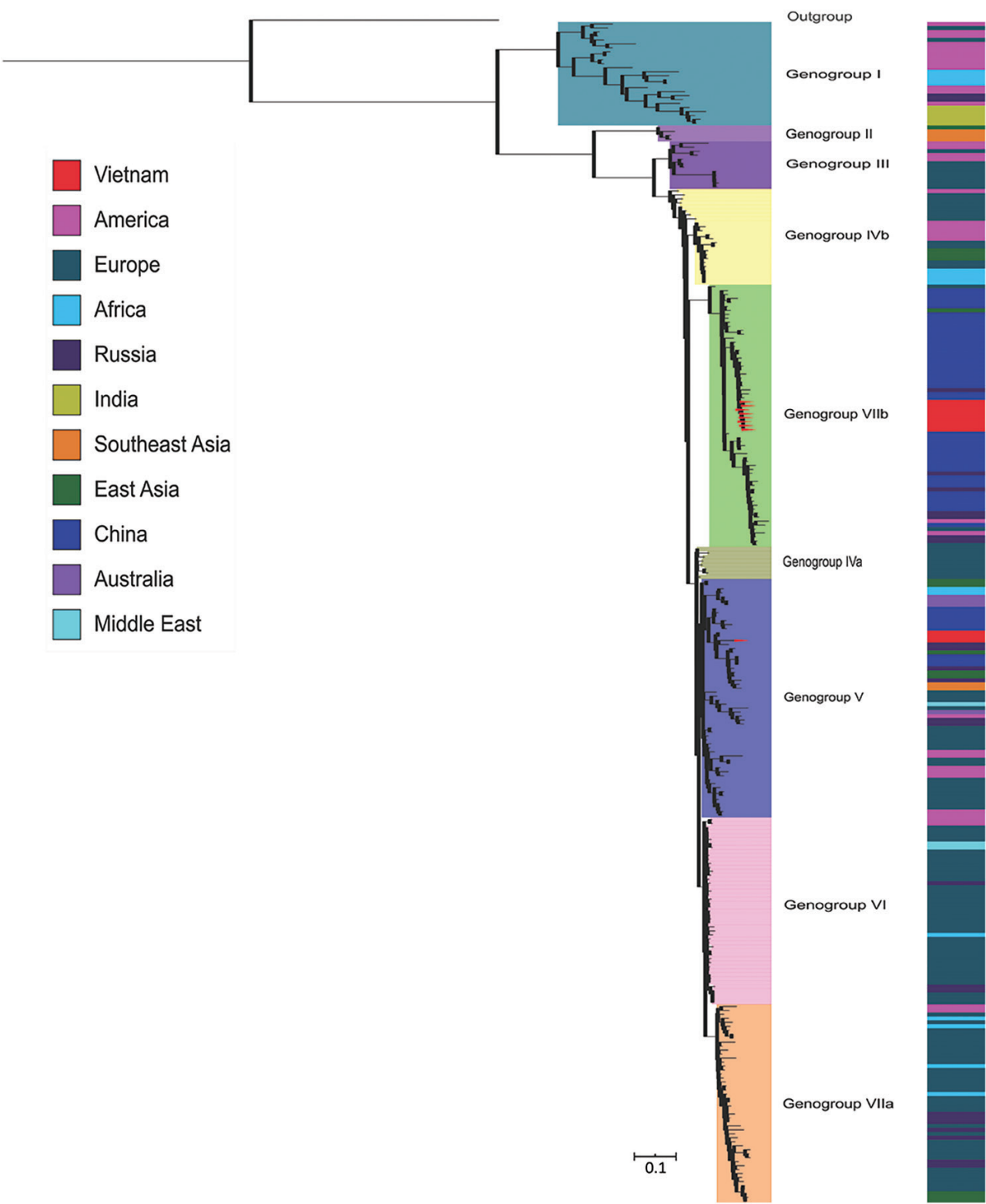
### Genetic Characterization of Enteroviruses and Gemycircularvirus

mNGS generated sufficient sequence information for an enterovirus serotyping assessment in 11/23 cases. Subsequently, results of serotyping analysis based on the NGS sequences showed that echovirus 30 (E30)



**Figure 3.** Temporal distribution of enterovirus cases detected from cerebrospinal fluid samples of patients with suspected central nervous system infection by metagenomic next-generation sequencing and RT-PCR, Vietnam, December 2012–October 2016. Enterovirus RT-PCR results were obtained from the original study. RT-PCR, reverse transcription PCR. A) Combined data from 3 provinces; B) data from Hue province; C) data from Khanh Hoa province; D) data from Dak Lak province.





**Figure 4.** Phylogenetic tree of 298 complete viral protein 1 sequences of echovirus 30 (876 nt) isolated from cerebrospinal fluid samples of patients with suspected central nervous system infection, Vietnam, December 2012–October 2016. The inner color strip indicates 7 genogroups. The outer color strip indicates different countries of echovirus 30 isolates included in the tree. The outgroup is echovirus 21 Farina. The E30 sequences generated by metagenomic next-generation sequencing are highlighted in red.

was the most common serotype detected ( $n = 9$ , 39% of enteroviruses), followed by enterovirus A71 and enterovirus B80 (1 each, 4.3%). Phylogenetically, the 9 E30 strains sequenced in our study belonged to 2 distinct genogroups, V and VIIb, and showed close relationship with E30 strains circulating in Russia and elsewhere in Asia, including China (Figure 4).

In addition to enterovirus sequences, a gemycircularvirus genome was obtained from a 12-year-old boy. Phylogenetic analysis revealed that this gemycircularvirus strain was closely related to a gemycircularvirus species previously found in CSF sample from a patient with a CNS infection of unknown origin in Sri Lanka (12); the level of amino acid identities between the 2 strains were 98.79% for replication-coding sequences and 99.3% for capsid protein-coding sequences.

## Discussion

We describe a viral mNGS investigation characterizing the human virome in CSF of 204 patients in Vietnam with suspected CNS infection of unknown origin. We successfully detected 4 human viral pathogens (enteroviruses, HIV, HBV, and MCV) and 1 virus species (gemycircularvirus) of unknown tropism and pathogenicity in a total of 30 (14.7%) patients. Most patients therefore remained without a known etiology, underscoring the ongoing challenge in identifying a plausible viral pathogen in CSF of patients with CNS infections.

Enteroviruses were the most common viruses, found in 11.3% (23/204) of all analyzed patients (Figure 2), most of whom were children and young adults. This age distribution of enterovirus-infected patients is consistent with observational data from a previous report from Vietnam (6), although the median age was slightly higher compared with data from other countries (13,14). Geographically, all the enterovirus-infected patients were admitted to hospitals from central and highland Vietnam, and none was from southern Vietnam. The underlying mechanism determining this observed spatial pattern of enterovirus-positive cases in this study remains unknown. Our sampling timescale perhaps was not long enough to capture the circulation of enteroviruses in Dong Thap Province. Enteroviruses were previously reported as a leading cause of CNS infection across central and southern Vietnam (6,15,16). Collectively, our findings suggest that reverse transcription PCR (RT-PCR) testing for enteroviruses should be considered in children and young adults with CNS infections.

Of the detected enteroviruses, E30 was the most common serotype. E30 is a well-known pathogen of pediatric aseptic meningitis worldwide (17).

Phylogenetically, at global scale, E30 belongs to 2 different lineages with distinct patterns of circulation and spread, 1 with a global distribution and the other with geographic restriction within Asia (17). The cocirculation of 2 E30 lineages in Vietnam suggests that E30 was imported into Vietnam on at least 2 occasions. Our analyses thus also contribute to the body of knowledge about the genetic diversity of E30 strains circulating in Vietnam.

The detection of bloodborne viruses such as HBV and HIV is unlikely to have a direct link with patients' neurologic symptoms, although HBV has previously been reported in CSF of patients with CNS infections of unknown origin (18). The detection of HIV in CSF might have been a consequence of traumatic tap occurring during the lumbar puncture, as reflected by the high number of red blood cells in 1 of 2 HIV-positive CSF samples (data not shown). However, neuroinvasion of HIV has also been reported (19). Likewise, the pathogenic potential of a gemycircularvirus genome requires further investigation, although the detection of the gemycircularvirus genome in CSF has been reported in several papers (12,18,20). The detection of MCV and papillomavirus in CSF might result from contamination of viral skin flora during lumbar puncture.

Similar to previous reports about discrepancy between mNGS and conventional diagnostic testing (8,18,21), our observations found that most mNGS-positive results were not confirmed by subsequent viral RT-PCR assays, especially the sensitive enterovirus-specific RT-PCR with a limit of detection of  $\approx 9$  copies/reaction (22). Such results could be attributable to bleedover (also called index hopping) of indices from reads of 1 sample into reads of another sample co-sequenced on the same Illumina run (R. Sinha et al., unpub. data, <https://doi.org/10.1101/125724>). Applying double indexes, which was not used in our study, has been shown to substantially reduce, but not eliminate, the cross-contamination phenomenon between samples in the same run.

Our study has some limitations. First, as outlined previously, we did not employ a double unique index combination strategy per sample as part of the sequencing procedure. The well-known index hopping phenomenon possibly explains the high discrepancy between confirmatory PCR and mNGS results (21,23,24) and emphasizes the usefulness of dual indexing and including no template controls. As such, we pragmatically chose to verify our mNGS by performing specific PCR on original materials. Second, the DNase treatment step in our assay meant to reduce cellular DNA concentration in CSF

might reduce the sensitivity of mNGS for the detection of DNA viruses such as herpes simplex virus (25,26). Third, some of the non-PCR-confirmed viral sequences likely originated from contamination of reagents, which is a lingering problem for mNGS (27,28).

In summary, our results emphasize that mNGS could detect a broad range of viral nucleic acids in CSF. In spite of extensive investigation, establishing the etiology in many patients with CNS infections remains a challenge. However, our findings indicate that enteroviruses are important causes of viral CNS infections in Vietnam and thus should be considered in the differential diagnosis among young patients with CNS infections.

VIZIONS Consortium members: from the Oxford University Clinical Research Unit, Bach Tuan Kiet, Stephen Baker, Alessandra Berto, Maciej F. Boni, Juliet E. Bryant, Bui Duc Phu, James I. Campbell, Juan Carrique-Mas, Dang Manh Hung, Dang Thao Huong, Dang Tram Oanh, Jeremy N. Day, Dinh Van Tan, H. Rogier van Doorn, Duong An Han, Jeremy J. Farrar, Hau Thi Thu Trang, Ho Dang Trung Nghia, Hoang Bao Long, Hoang Van Duong, Huynh Thi Kim Thu, Lam Chi Cuong, Le Manh Hung, Le Thanh Phuong, Le Thi Phuc, Le Thi Phuong, Le Xuan Luat, Luu Thi Thu Ha, Ly Van Chuong, Mai Thi Phuoc Loan, Behzad Nadjm, Ngo Thanh Bao, Ngo Thi Hoa, Ngo Tri Tue, Nguyen Canh Tu, Nguyen Dac Thuan, Nguyen Dong, Nguyen Khac Chuyen, Nguyen Ngoc An, Nguyen Ngoc Vinh, Nguyen Quoc Hung, Nguyen Thanh Dung, Nguyen Thanh Minh, Nguyen Thi Binh, Nguyen Thi Hong Tham, Nguyen Thi Hong Tien, Nguyen Thi Kim Chuc, Nguyen Thi Le Ngoc, Nguyen Thi Lien Ha, Nguyen Thi Nam Lien, Nguyen Thi Ngoc Diep, Nguyen Thi Nhung, Nguyen Thi Song Chau, Nguyen Thi Yen Chi, Nguyen Thieu Trinh, Nguyen Thu Van, Nguyen Van Cuong, Nguyen Van Hung, Nguyen Van Kinh, Nguyen Van Minh Hoang, Nguyen Van My, Nguyen Van Thang, Nguyen Van Thanh, Nguyen Van Vinh Chau, Nguyen Van Xang, Pham Ha My, Pham Hong Anh, Pham Thi Minh Khoa, Pham Thi Thanh Tam, Pham Van Lao, Pham Van Minh, Phan Van Be Bay, Maia A. Rabaa, Motiur Rahman, Corinne Thompson, Guy Thwaites, Ta Thi Dieu Ngan, Tran Do Hoang Nhu, Tran Hoang Minh Chau, Tran Khanh Toan, Tran My Phuc, Tran Thi Kim Hong, Tran Thi Ngoc Dung, Tran Thi Thanh Thanh, Tran Thi Thuy Minh, Tran Thua Nguyen, Tran Tinh Hien, Trinh Quang Tri, Vo Be Hien, Vo Nhut Tai, Vo Quoc Cuong, Voong Vinh Phat, Vu Thi Lan Huong, Vu Thi Ty Hang, and Heiman Wertheim; from the Centre for Immunity, Infection, and Evolution, University Of Edinburgh, Edinburgh, Scotland, UK, Carlijn Bogaardt, Margo Chase-Topping, Al Ivens, Lu Lu,

Dung Nyugen, Andrew Rambaut, Peter Simmonds, and Mark Woolhouse; from The Wellcome Trust Sanger Institute, Hinxton, UK, Matthew Cotten, Bas B. Oude Munnink, Paul Kellam, and My Vu Tra Phan; from the Laboratory of Experimental Virology, Department of Medical Microbiology, Center for Infection and Immunity Amsterdam (CINIMA), Academic Medical Center of the University of Amsterdam, Amsterdam, the Netherlands, Martin Deijs, Lia van der Hoek, Maarten F. Jebbink, and Seyed Mohammad Jazaeri Farsani; and from Metabiota, California, USA, Karen Saylor and Nathan Wolfe.

### Acknowledgments

We would like to thank Le Thi Kim Thanh for her logistic support and the patients for their participation in this study.

The study was funded by the Wellcome Trust of Great Britain (awards nos. WT/093724 and 106680/B/14/Z [to the Oxford University Clinical Research Unit in Vietnam], 100087/Z/12/Z [to S.B.] and 204904/Z/16/Z [to L.V.T.]), and the Royal Society (grant no. 098511/Z/12/Z [to S.B.]). X.D. and E.D. were supported by the Blood Systems Research Institute and the National Heart, Lung, and Blood Institute (grant no. R01 HL105770).

### About the Author

Ms. Nguyen is a PhD student in life science at Open University, Milton Keynes, UK. Her research interests are virus discovery and evolution of emerging pathogens such as enteroviruses.

### References

1. Granerod J, Tam CC, Crowcroft NS, Davies NWS, Borchert M, Thomas SL. Challenge of the unknown. A systematic review of acute encephalitis in non-outbreak situations. *Neurology*. 2010;75:924–32. <https://doi.org/10.1212/WNL.0b013e3181f11d65>
2. Feigin VL, Krishnamurthi RV, Theadom AM, Abajobir AA, Mishra SR, Ahmed MB, et al.; GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol*. 2017;16:877–97. [https://doi.org/10.1016/S1474-4422\(17\)30299-5](https://doi.org/10.1016/S1474-4422(17)30299-5)
3. World Health Organization. Neurological disorders: public health challenges. 2006 [cited 2017 Nov 11]. [https://www.who.int/mental\\_health/publications/neurological\\_disorders\\_ph\\_challenges](https://www.who.int/mental_health/publications/neurological_disorders_ph_challenges)
4. Sejvar JJ. West Nile virus: an historical overview. *Ochsner J*. 2003;5:6–10.
5. Rabaa MA, Tue NT, Phuc TM, Carrique-Mas J, Saylor K, Cotten M, et al. The Vietnam Initiative on Zoonotic Infections (VIZIONS): a strategic approach to studying emerging zoonotic infectious diseases. *EcoHealth*. 2015;12:726–35. <https://doi.org/10.1007/s10393-015-1061-0>



6. Ho Dang Trung N, Le Thi Phuong T, Wolbers M, Nguyen Van Minh H, Nguyen Thanh V, Van MP, et al.; VIZIONS CNS Infection Network. Aetiologies of central nervous system infection in Viet Nam: a prospective provincial hospital-based descriptive surveillance study. *PLoS One*. 2012;7:e37825. <https://doi.org/10.1371/journal.pone.0037825>
7. Brown JR, Bharucha T, Breuer J. Encephalitis diagnosis using metagenomics: application of next generation sequencing for undiagnosed cases. *J Infect*. 2018;76:225–40. <https://doi.org/10.1016/j.jinf.2017.12.014>
8. Anh NT, Hong NTT, Nhu LNT, Thanh TT, Lau C-Y, Limmathurotsakul D, et al. Viruses in Vietnamese patients presenting with community-acquired sepsis of unknown cause. *J Clin Microbiol*. 2019;57:e00386–19. <https://doi.org/10.1128/JCM.00386-19>
9. Aiemjoy K, Altan E, Aragie S, Fry DM, Phan TG, Deng X, et al. Viral species richness and composition in young children with loose or watery stool in Ethiopia. *BMC Infect Dis*. 2019;19:53. <https://doi.org/10.1186/s12879-019-3674-3>
10. Kroneman A, Vennema H, Deforche K, v d Avoort H, Peñaranda S, Oberste MS, et al. An automated genotyping tool for enteroviruses and noroviruses. *J Clin Virol*. 2011;51:121–5. <https://doi.org/10.1016/j.jcv.2011.03.006>
11. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32:268–74. <https://doi.org/10.1093/molbev/msu300>
12. Phan TG, Mori D, Deng X, Rajidrajith S, Ranawaka U, Fan Ng TF, et al. Small viral genomes in unexplained cases of human encephalitis, diarrhea, and in untreated sewage. *Virology*. 2015;482:98–104. <https://doi.org/10.1016/j.virol.2015.03.011>
13. Sun Y, Miao Z, Yan J, Gong L, Chen Y, Chen Y, et al. Sero-molecular epidemiology of enterovirus-associated encephalitis in Zhejiang Province, China, from 2014 to 2017. *Int J Infect Dis*. 2019;79:58–64. <https://doi.org/10.1016/j.ijid.2018.11.002>
14. Richter J, Tryfonos C, Christodoulou C. Molecular epidemiology of enteroviruses in Cyprus 2008–2017. *PLoS One*. 2019;14:e0220938. <https://doi.org/10.1371/journal.pone.0220938>
15. Taylor WR, Nguyen K, Nguyen D, Nguyen H, Horby P, Nguyen HL, et al. The spectrum of central nervous system infections in an adult referral hospital in Hanoi, Vietnam. *PLoS One*. 2012;7:e42099. <https://doi.org/10.1371/journal.pone.0042099>
16. Tan V, Thai H, Phu NH, Nghia HDT, Chuong LV, Sinh DX, et al. Viral aetiology of central nervous system infections in adults admitted to a tertiary referral hospital in southern Vietnam over 12 years. *PLoS Negl Trop Dis*. 2014;8:e3127. <https://doi.org/10.1371/journal.pntd.0003127>
17. Lema C, Torres C, Van der Sanden S, Cisterna D, Freire MC, Gómez RM. Global phylodynamics of Echovirus 30 revealed differential behavior among viral lineages. *Virology*. 2019;531:79–92. <https://doi.org/10.1016/j.virol.2019.02.012>
18. Schibler M, Brito F, Zanella MC, Zdobnov EM, Laubscher F, L'Huillier AG, et al. Viral sequences detection by high-throughput sequencing in cerebrospinal fluid of individuals with and without central nervous system disease. *Genes (Basel)*. 2019;10:1–12. <https://doi.org/10.3390/genes10080625>
19. Zayyad Z, Spudich S. Neuropathogenesis of HIV: from initial neuroinvasion to HIV-associated neurocognitive disorder (HAND). *Curr HIV/AIDS Rep*. 2015;12:16–24. <https://doi.org/10.1007/s11904-014-0255-3>
20. Zhou C, Zhang S, Gong Q, Hao A. A novel gemycircular-virus in an unexplained case of child encephalitis. *Virol J*. 2015;12:197. <https://doi.org/10.1186/s12985-015-0431-0>
21. Wilson MR, Sample HA, Zorn KC, Arevalo S, Yu G, Neuhaus J, et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N Engl J Med*. 2019;380:2327–40. <https://doi.org/10.1056/NEJMoa1803396>
22. Thanh TT, Anh NT, Tham NT, Van HM, Sabanathan S, Qui PT, et al. Validation and utilization of an internally controlled multiplex Real-time RT-PCR assay for simultaneous detection of enteroviruses and enterovirus A71 associated with hand foot and mouth disease. *Virol J*. 2015;12:85. <https://doi.org/10.1186/s12985-015-0316-2>
23. Wilson MR, Fedewa G, Stenglein MD, Olejnik J, Rennick LJ, Nambulli S, et al. Multiplexed metagenomic deep sequencing to analyze the composition of high-priority pathogen reagents. *mSystems*. 2016;1:1–9. <https://doi.org/10.1128/mSystems.00058-16>
24. Yang J, Yang F, Ren L, Xiong Z, Wu Z, Dong J, et al. Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach. *J Clin Microbiol*. 2011;49:3463–9. <https://doi.org/10.1128/JCM.00273-11>
25. Nguyen TTH, Nguyen TA, Nguyen Thi Hoang M, Ho DTN, Le NNT, Tran tan T, et al. Performance of metagenomic next-generation sequencing for the diagnosis of viral meningoencephalitis in a resource limited setting. *Open Forum Infect Dis*. 2020;7:ofaa046. <https://doi.org/10.1093/ofid/ofaa046>
26. Perlejewski K, Popiel M, Laskus T, Nakamura S, Motooka D, Stokowy T, et al. Next-generation sequencing (NGS) in the identification of encephalitis-causing viruses: Unexpected detection of human herpesvirus 1 while searching for RNA pathogens. *J Virol Methods*. 2015;226:1–6. <https://doi.org/10.1016/j.jviromet.2015.09.010>
27. Asplund M, Kjartansdóttir KR, Møllerup S, Vinner L, Fridholm H, Herrera JAR, et al. Contaminating viral sequences in high-throughput sequencing viromics: a linkage study of 700 sequencing libraries. *Clin Microbiol Infect*. 2019;25:1277–85. <https://doi.org/10.1016/j.cmi.2019.04.028>
28. Holmes EC. Reagent contamination in viromics: all that glitters is not gold. *Clin Microbiol Infect*. 2019;25:1167–8. <https://doi.org/10.1016/j.cmi.2019.06.019>

---

Address for correspondence: Nguyen To Anh or Le Van Tan, Oxford University Clinical Research Unit, 764 VoVan Kiet, District 5, Ho Chi Minh City, Vietnam; email: anhnt@oucru.org or email: tanlv@oucru.org

# Viral Metagenomic Analysis of Cerebrospinal Fluid from Patients with Acute Central Nervous System Infections of Unknown Origin, Vietnam

## Appendix

**Appendix Table 1.** List of diagnostic assays conducted by the original study

Pathogens*	Assay	References
Dengue virus	ELISA	(1)
Japanese encephalitis virus	ELISA	(1)
Herpes simplex virus 1 and 2	PCR	(2)
Varicella-zoster virus	PCR	(3)
Enterovirus	PCR	(4)
<i>Neisseria meningitidis</i>	PCR	(5)
<i>Streptococcus pneumoniae</i>	PCR	(5)
<i>Hemophilus influenza</i> type b	PCR	(5)
<i>Streptococcus suis</i>	PCR	(6)

\*Diagnostic tests performed in every case as part of standard of care at participating hospitals including complete blood count, blood culture, gram/ZN smears, CSF culture and CSF examination if patients have neurologic symptoms and CNS infection is suspected.

**Appendix Table 2.** List of primers and probes used for PCR confirmatory

Viruses	Oligo sequence (5'→3')			Sources
	Forward	Reverse	Probe	
HBV	GGACCCCTGCTCGTGTTACA	GAGAGAAGTCCACCMCGAGTCTAGA	FAM-TGTTGACAARAATCCTCACAAATACCRGAGA-TAMRA	Newly designed
Rotavirus	ACC ATC TWC ACR TRA CCC TC	GGT CAC ATA ACG CCC CTA TA	FAM-ATGAGCACAATAGTTAAAAGCTAACACTGT CAA-BHQ1	(7)
Enterovirus	CCCTGAATGCGGCTAAT	ATTGTCACCATAAGCAGCC	CY5-ACCCAAAGTAGTCGGTTCCG -BHQ3	(8)
HIV1	GGTGCAGAGCGTC	ATGCTRTCATCATYTTCTTC		(9)
Gemycircularvirus	ATGGGTRAARGTARTAGAAGAAAAGGG	CTGCCTGRTGYCCYCCCACTA		
	GTGGTAATGGTCGTCGGTATTC	CCTCATCATTCGTAGTAAGCAATCTCA		(10)
	AGTCCTGAATGTTTCCACTCG	CAAGCGTTCCTCGAAAATGAC		Newly designed
Cyclovirus VN	GAGCGCACATTGAAAGAGCTAAA	TCTCCTCCTTCAATGACAGAAACAAC	FAM-CGADAATAAGGMATACTGCTCTAAAGSTGGCG-BHQ1	(11)
Molluscum contagiosum virus	AACCTACGCTACCTGAAGMTGGA	CAGGCTCTTGATGGTCGARATGGA		(12)

**Appendix Table 3.** Baseline characteristics, and CSF white cell count of patients with viruses detected by mNGS

Characteristic	Hepatitis B virus (n = 3)	Human immunodeficiency virus (n = 2)	Molluscum contagiosum virus (n = 1)	Gemycircularvirus (n = 1)
Male, no. (%)	3 (100)	2 (100)	1 (100)	1 (100)
Age (median; min-max)	42 (16–63)	57.5 (49–66)	7	12
Location				
Hue, no. (%)	0	0	0	0
Dak lak, no. (%)	0	0	1 (100)	0
Khanh Hoa, no. (%)	2 (75)	0	0	0
Dong Thap, no. (%)	1 (25)	2 (100)	0	1 (100)
3-d fever (at enrollment or last 3 days)				
Fever, no. (%); median (min-max)	3 (100); 38 (38–39)	1 (50); 39	1 (100); 38.5	0
Fever with unknown temp, no. (%)	0	0	0	1 (100)
No fever, no. (%)	0	1 (50)	0	0
Unknown, no. (%)	0	0	0	0
Outcome				
Death or discharged to die, no. (%)	0	0		
Discharge with complete recovery, no. (%)	2 (75)	0	1 (100)	1 (100)
Discharged with incomplete recovery, no. (%)	1 (25)	1 (50)	0	0
Transferred to another hospital, no. (%)	0	1 (50)	0	0
Other (patient request), no. (%)	0	0	0	0
Unknown, no. (%)	0	0	0	0
CSF White Cells (cell/mm <sup>3</sup> )	1530 (7–2590)	835 (340–1330)	70	60
Median (min-max)				

## References

- Cardosa MJ, Wang SM, Sum MSH, Tio PH. Antibodies against prM protein distinguish between previous infection with dengue and Japanese encephalitis viruses. *BMC Microbiol.* 2002;2:9. [PubMed https://doi.org/10.1186/1471-2180-2-9](https://doi.org/10.1186/1471-2180-2-9)
- van Doornum GJJ, Guldemeester J, Osterhaus ADME, Niesters HGM. Diagnosing herpesvirus infections by real-time amplification and rapid culture. *J Clin Microbiol.* 2003;41:576–80. [PubMed https://doi.org/10.1128/JCM.41.2.576-580.2003](https://doi.org/10.1128/JCM.41.2.576-580.2003)
- de Jong MD, Weel JF, Schuurman T, Wertheim-van Dillen PM, Boom R. Quantitation of varicella-zoster virus DNA in whole blood, plasma, and serum by PCR and electrochemiluminescence. *J Clin Microbiol.* 2000;38:2568–73. [PubMed https://doi.org/10.1128/JCM.38.7.2568-2573.2000](https://doi.org/10.1128/JCM.38.7.2568-2573.2000)
- Beld M, Minnaar R, Weel J, Sol C, Damen M, van der Avoort H, et al. Highly sensitive assay for detection of enterovirus in clinical specimens by reverse transcription-PCR with an armored RNA internal control. *J Clin Microbiol.* 2004;42:3059–64. [PubMed https://doi.org/10.1128/JCM.42.7.3059-3064.2004](https://doi.org/10.1128/JCM.42.7.3059-3064.2004)
- Corless CE, Guiver M, Borrow R, Edwards-Jones V, Fox AJ, Kaczmarek EB. Simultaneous detection of *Neisseria meningitidis*, *Haemophilus influenzae*, and *Streptococcus pneumoniae* in suspected cases of meningitis and septicemia using real-time PCR. *J Clin Microbiol.* 2001;39:1553–8. [PubMed https://doi.org/10.1128/JCM.39.4.1553-1558.2001](https://doi.org/10.1128/JCM.39.4.1553-1558.2001)

6. Nga TVT, Nghia HDT, Tu TP, Diep TS, Mai NTH, Chau TTH, et al. Real-time PCR for detection of *Streptococcus suis* serotype 2 in cerebrospinal fluid of human patients with meningitis. *Diagn Microbiol Infect Dis*. 2011;70:461–7. [PubMed](#) <https://doi.org/10.1016/j.diagmicrobio.2010.12.015>
7. Dung TTN, Phat VV, Nga TVT, My PVT, Duy PT, Campbell JI, et al. The validation and utility of a quantitative one-step multiplex RT real-time PCR targeting rotavirus A and norovirus. *J Virol Methods*. 2013;187:138–43. [PubMed](#) <https://doi.org/10.1016/j.jviromet.2012.09.021>
8. Thanh TT, Anh NT, Tham NT, Van HM, Sabanathan S, Qui PT, et al. Validation and utilization of an internally controlled multiplex Real-time RT-PCR assay for simultaneous detection of enteroviruses and enterovirus A71 associated with hand foot and mouth disease. *Virology*. 2015;12:85. [PubMed](#) <https://doi.org/10.1186/s12985-015-0316-2>
9. Chook JB, Ong LY, Takebe Y, Chan KG, Choo M, Kamarulzaman A, et al. Molecular detection of HIV-1 subtype B, CRF01\_AE, CRF33\_01B, and newly emerging recombinant lineages in Malaysia. *Am J Trop Med Hyg*. 2015;92:507–12. [PubMed](#) <https://doi.org/10.4269/ajtmh.14-0681>
10. Phan TG, Mori D, Deng X, Rajidrajith S, Ranawaka U, Fan Ng TF, et al. Small viral genomes in unexplained cases of human encephalitis, diarrhea, and in untreated sewage. *Virology*. 2015;482:98–104. [PubMed](#) <https://doi.org/10.1016/j.virol.2015.03.011>
11. Van Tan L, Van Doorn HR, Trung D, Hong T, Phuong T, De Vries M. Identification of a New Cyclovirus in Cerebrospinal Fluid of Patients with Acute Central Nervous System Infections. *MBio*. 2013;4:1–10.
12. Hošnjak L, Kocjan BJ, Kušar B, Seme K, Poljak M. Rapid detection and typing of *Molluscum contagiosum* virus by FRET-based real-time PCR. *J Virol Methods*. 2013;187:431–4. [PubMed](#) <https://doi.org/10.1016/j.jviromet.2012.11.008>