HYBRID-CAPTURE NEXT-GENERATION SEQUENCING FOR VIRAL DETECTION IN ACUTE FEBRILE ILLNESS OF UNKNOWN ETIOLOGY PATIENTS IN NAKHON PHANOM AND TAK PROVINCES, THAILAND

Sumonmal Uttayamakul[1], Kornkanok Prasertsom[1], Don Changsom[1], Naphak Moonnamarng[1], Peera Hemarajata[2] and Pilailuk Akkapaiboon Okada[1]

[1]National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi Province, Thailand; [2]Association of Public Health Laboratories, Thailand Office, Nonthaburi Province, Thailand

Abstract. An acute febrile illness (AFI) with an unknown etiology is a major diagnostic challenge in tropical settings. In this study we aimed to evaluate hybrid-capture next-generation sequencing (NGS) for detecting viral pathogens in hospitalized AFI patients with negative routine diagnostic results in order to determine the kinds of viruses causing an AFI in the study locations that were missed by routine testing and the proportion of those cases that could be detected by hybrid-capture NGS testing. We retrospectively examined previously collected serum samples obtained from patients hospitalized with AFI during 2017-2020 in Nakhon Phanom and Tak Provinces, Thailand in whom no etiology was determined and in which an adequate sample was available to conduct the specified testing. In each sample, total nucleic acids were extracted and analyzed using a hybrid-capture NGS test with the Twist Comprehensive Viral Research Panel and interpreted using a metagenomics pipeline with predefined background-subtraction and positivity criteria. The median age of the study subjects in whom the samples were obtained was 36 (range: 2-79) years, 60.0% females. Viral pathogens were detected by hybrid-capture NGS in 65 (56.5%) of the 115 samples. Cytomegalovirus (n = 29, 25.2%) and Epstein-Barr virus (n = 27, 23.5%) were the most frequently identified pathogens, followed by enterovirus A (n = 2, 1.7%), enterovirus B (n = 1, 0.9%), hepatitis B virus (n = 2, 1.7%), human herpesvirus 7 (n = 2, 1.7%), influenza A virus (n = 1, 0.9%), and Kaposi's sarcoma-associated herpesvirus (n = 1, 0.9%). It is important to remember easily tested for viruses, such as influenza viruses, were already tested for and found to be negative in order to be included in this study. In summary, hybrid-capture NGS was able

to determine more than half of samples previously testing negative with routine testing and the most common viruses detected were CMV and EBV. We conclude the hybrid-capture NGS may be beneficial for evaluating viral AFI in which the etiology is unclear with routine testing. Further studies in other populations are needed to determine when and in what situations the hybrid-capture NGS would be of most benefit in guiding management of patients with AFI of unknown etiology.

Keywords: acute febrile illness, viral pathogens, next-generation sequencing, Twist Comprehensive Viral Panel, hybrid capture, metagenomics, Chan Zuckerberg ID, Thailand

Correspondence: Sumonmal Uttayamakul, National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Tiwanon Road, Nonthaburi 11000, Thailand
Tel: +66 (0) 2951 0000 Ext 98384; Fax: +66 (0) 2951 5974
E-mail: sumonmal.u@dmsc.mail.go.th

INTRODUCTION

An acute febrile illness (AFI) is defined as a sudden onset of fever ≥38°C lasting 2-14 days without an identified etiology (Saravanan *et al*, 2022). It is frequently associated with non-specific symptoms, such as headache, myalgia, malaise or gastrointestinal discomfort (Saravanan *et al*, 2022, Wodniak *et al*, 2024). AFIs are especially found in resource-limited settings where the disease burden is high and infectious diseases are common (Bhaskaran *et al*, 2019). The etiologies of AFIs vary and may be caused by viral, bacterial, parasitic or fungal pathogens. More than 70 different pathogens in these 4 groups have been reported to cause AFI (Grundy and Houpt, 2022) making diagnosis challenging. In Thailand surveillance studies of AFI, 61.3% of cases after routine laboratory testing had no identified etiology (Leelarasamee *et al*, 2004; Wodniak *et al*, 2024) highlighting ongoing diagnostic challenges.

Viral infections are a common cause of AFI and diagnosis can be difficult since many viruses have overlapping presentations and there is little access to comprehensive diagnostic testing (Yozwiak *et al*, 2012). Rapid diagnostic tests, serology and quantitative real-time polymerase chain reaction (qRT-PCR) tests are restricted to predefined target viruses and may fail to detect unexpected or emerging pathogens. These limitations show the need for sensitive, comprehensive testing that identifies a variety of the most common viral pathogens.

Next-generation sequencing (NGS) may meet this need since this testing allows for identification of a variety of viral agents which are difficult to cultivate (Boers *et al*, 2019). Amplicon-based metagenomic tests, such as 16S ribosomal RNA gene sequencing, are available to identify bacterial pathogens (Janda and Abbott, 2007). Viral metagenomics attempts to detect a variety of viruses which are not detected by routine diagnostic testing (Gauthier *et al*, 2023). However, untargeted metagenomic sequencing has difficulty in identifying viruses in samples with a low viral load. Hybrid capture-based enrichment strategies can improve viral detection by amplifying viral nucleic acids in clinical specimens improving the sensitivity of the test (Metsky *et al*, 2019). One study reported a hybrid capture panel targeting vertebrate viruses in cerebrospinal fluid from patients with meningoencephalitis yielded a 100- to 10,000-fold increase in viral detection compared to unenriched metagenomic sequencing, with a sensitivity comparable to PCR testing (Mourik *et al*, 2024).

The Twist Comprehensive Viral Research Panel targets thousands of viral genomes and provides a broad, sensitive platform for viral detection (Twist Bioscience, 2020). This method allows detection of a wide range of viral pathogens, including emerging and re-emerging viruses, with improved sensitivity compared to untargeted

metagenomics (Wylie *et al*, 2015). Combining hybrid-capture based enrichment with NGS improves detection of viruses with a low viral load and co-infections often missed by previous diagnostic testing (Metsky *et al*, 2019).

In this study we aimed to evaluate hybrid-capture next-generation sequencing (NGS) for detecting viral pathogens in hospitalized AFI patients with negative routine diagnostic results in order to determine the kinds of viruses causing an AFI in the study locations that were missed by routine testing and the proportion of those cases that could be detected by hybrid-capture NGS testing.

MATERIALS AND METHODS

Study subjects

Study subjects were those hospitalized for acute febrile illness (AFI) in Nakhon Phanom and Tak Provinces, Thailand, during 2017-2020 in whom routine testing, including routine testing, blood cultures and polymerase chain reaction (PCR) testing did not show an etiology. Common real-time PCR testing checks for dengue virus, chikungunya virus, Zika virus, Japanese encephalitis virus, pathogenic *Leptospira*, *Salmonella* spp, *Burkholderia pseudomallei*, *Plasmodium* spp and *Orientia tsutsugamushi*.

Study samples

Serum samples had previously been obtained from each study subject and we retrospectively examined these samples for our study. The samples had been stored at -80° until use. Samples with insufficient volume or in which subject data was missing were not examined in our study and the subjects were excluded.

Sample evaluation

A total of 115 serum samples, from patients with a viral infection but no identified type of virus, were retrospectively analyzed using the Twist Comprehensive Viral Research Panel. The performance of the panel for viral genome detection was evaluated through comparison of positive samples obtained by

real-time PCR and hybridization based next-generation sequencing.

Nucleic acid extraction

Total nucleic acid was extracted from 200 µl of each serum sample using the ZymoBIOMICS MagBead DNA/RNA kit (Zymo Research, Irvine, CA), according to the manufacturer's instructions. The extracted nucleic acids were eluted in DNase/RNase-free water (Zymo Research, Irvine, CA) to give a final volume of 35 µl and then stored at -80°C until used.

Library preparation

Library preparation was performed using the Twist Comprehensive Viral Research Panel workflow (Twist Bioscience, South San Francisco, CA) following the manufacturer's instructions. Briefly, 5 µl of extracted nucleic acid was synthesized into first-stranded complementary DNA (cDNA) using the ProtoScript II First Strand cDNA Synthesis Kit (New England Biolabs, Ipswich, MA). Second-strand synthesis was subsequently performed using the NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module (New England Biolabs, Ipswich, MA).

After cDNA synthesis, DNA fragmentation, end repair and dA-tailing were performed using the Twist Library Preparation EF Kit 1.2 (Twist Bioscience, South San Francisco, CA) following the manufacturer's thermal cycler program consisting of: an initial hold at 4°C, incubation at 37°C for 20 minutes and 65°C for 30 minutes, followed by a final hold at 4°C. Adapter ligation was carried out using 2.5 µl of the Twist Universal Adapter, mixing and then adding 20 µl of Ligation Master Mix. The reaction was incubated at 20°C for 15 minutes in a thermal cycler. Then, 60 µl of homogenized (0.8x) DNA Purification Beads were added and incubated for 5 minutes. The tube was placed on a magnetic stand for 1 minute and washed twice with 200 µl 80% ethanol. The beads were air-dried and 17 µl nuclease-free water was added to resuspend the beads. The tube

was returned to the magnetic stand until a clear bead pellet formed. An aliquot of fifteen microliters of the eluate was transferred to a PCR tube for library amplification.

PCR amplification was performed by adding 10 µl Twist UDI primers and 25 µl Equinox Library Amp Mix. We performed thermocycling following denaturation at 98°C for 45 seconds, 8 cycles of 98°C for 15 seconds, 60°C for 30 seconds, and 72°C for two minutes, with extension at 72°C for one minute and a hold at 4°C. Following amplification, purification was conducted by adding 50 µl homogenized DNA Purification Beads, incubating for 5 minutes and placing the tube on a magnetic stand for 1 minute. The resultant pellet was washed twice with 200 µl 80% ethanol, air-dried and then eluted in 22 µl nuclease-free water. After a two-minute incubation, the tube was placed back on the magnetic stand for three minutes and 20 µl of the final eluate was transferred to a new PCR tube. Library quantity was assessed using the Qubit dsDNA High-Sensitivity Assay Kit (Thermo Fisher Scientific, Waltham, MA).

Target enrichment and sequencing

Enriched viral libraries were generated by hybridization with biotinylated probes specific for the Twist Comprehensive Viral Research Panel. The required volume of each indexed library was calculated as the volume of each library needed for hybridization, aiming for the final quantity of 1,500 ng pooled library per hybridization. The Twist Viral Panel reagents were then added and the hybridization reaction was performed at 70°C for 16 hours in a thermal cycler with the lid at 85°C. Next, the hybridized samples were bound to the prepared streptavidin binding beads and incubated for 5 minutes at 60°C. After incubation, the tubes were placed on a magnetic stand for 1 minute, and the supernatant was discarded. The beads were then washed with Wash Buffer 1 at 68°C and incubated for 5 minutes at 68°C. After incubation, the mixture was transferred to a microcentrifuge tube, placed on a magnetic stand,

and the supernatant was removed. The beads were washed again with Wash Buffer 2 at 48°C and incubated for 5 minutes at 48°C, followed by two additional wash steps with the same buffer. Finally, 45 µl of purified water was added to elute the captured targets.

The PCR mixture containing the streptavidin binding bead slurry was prepared and amplified under the following cycling conditions: initial denaturation at 98°C for 45 seconds, followed by 17 cycles of 98°C for 15 seconds, 60°C for 30 seconds, and 72°C for 30 seconds. A final extension was performed at 72°C for 1 minute, and the reaction was then held at 4°C. The PCR product was purified using 90 µl of 1.8x DNA Purification Beads. The tube was then placed on a magnetic plate for 1 minute to separate the supernatant, which was removed. The beads were washed twice with 200 µl of 80% ethanol and were air-dried. Following, 32 µl of water was added to elute the DNA. The mixture was incubated for 2 minutes and then placed on the magnetic plate for 3 minutes. Finally, 30 µl of the clear supernatant containing the enriched library was transferred to a PCR tube. The quality of the final library was assessed by measuring the DNA concentration with a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA) and evaluating fragment size distribution using a TapeStation system (Agilent Technologies, Santa Clara, CA). The enriched pooled library was diluted per Illumina's instructions for MiSeq Reagent and sequenced on Illumina MiSeq (v3 600-cycle).

Bioinformatic analysis

Raw sequence reads were processed through the CZ ID Illumina mNGS pipeline v8.3 (Chan Zuckerberg Biohub, San Francisco, CA). Quality filtering, subtraction of host-derived reads and taxonomic classification were carried out using default pipeline parameters. Each sequencing run included one non-template control (NTC), which was processed through the complete library preparation

and sequencing workflow. Reads detected in the NTC across runs were combined to construct the background model. Taxa appearing in the NTC or falling below the NTC-derived abundance thresholds were classified as background signal and removed during analysis. Only taxa that remained after background subtraction were considered for downstream interpretation. We applied the "lightbulb" criteria from the CZ ID sample report: taxa with nucleotide alignment (NT) reads per million (rPM) >1, non-redundant protein alignment (NR) rPM >1, NT Z-score >1 and NR Z-score >1 and ranked among the top three scoring species were highlighted (blue lightbulb icon) and considered candidates for positivity. NT and NR rPM reflected the quantity and genomic dispersion of aligned reads, while NT and NR Z-scores indicate statistical enrichment above the run-specific background model. Samples in which microbial taxa were detected, but none met the predefined CZ ID lightbulb criteria, were classified as indeterminate. For our study, we further required a minimum threshold of 10 aligned reads or non-redundant contigs to call a taxon positive. When viral taxa met both the lightbulb and minimum-read criteria, we then performed reference-based genome mapping and coverage visualization to confirm taxon identity and assess genome completeness.

Statistical analysis

Descriptive statistics were used to summarize demographic data. Medians and interquartile ranges (IQR) were calculated and used to describe subject ages and genders. Continuous variables were analyzed using the Kruskal-Wallis test and categorical variables were analyzed using the chi-square test. A p-value <0.05 was considered statistically significant.

Ethical considerations

The study protocol was approved by the Research Ethics Committee, Department of Medical Sciences, Ministry of Public Health, Thailand (approval ID: DMSc-EC038).

RESULTS

Subject characteristics

A total of 115 subjects were included in the study. The median age of subjects was 36 (range 2-79; IQR = 37) years. Sixty-nine subjects (60.0%) were females. Most subjects were Thai (95.7%, 110/115) and the remaining 5 subjects were Myanmar, Karen and stateless. The most commonly reported clinical symptoms among subjects were fatigue (91.3%), loss of appetite (73.9%), cough (61.7%) and chills (60.0%).

Validation of Twist Comprehensive Viral Panel

To validate the performance of the Twist Comprehensive Viral Panel for viral genome detection, 7 serum samples from the original AFI surveillance project were selected. These samples had all been previously confirmed positive by real-time qPCR for their respective pathogens. The Twist panel successfully detected all targeted pathogens in these samples, demonstrating 100% concordance with the prior qPCR results (Table 1). Based on these findings, the described wet-lab and bioinformatics workflows were used to identify viral pathogens in the 115 from the 115 study subjects.

Detection profile of viral pathogen

The Twist panel was applied according to the criteria defined in the Methods (taxa meeting the background-model subtraction, lightbulb threshold, and minimum read/contig counts). Viral pathogens were detected in 65 of the 115 samples (56.5 %). Although multiple taxa were occasionally detected at low levels, no samples met the predefined positivity criteria for more than one pathogen. Therefore, no coinfections were identified. Of the 65 positive samples, the most frequently detected viruses were CMV (n = 29, 25.2%) and EBV (n = 27, 23.5%) followed by HBV (n = 2, 1.7%), HHV-7 (n = 2, 1.7%), enterovirus A (n = 2, 1.7%), enterovirus B (n = 1, 0.9%), influenza A virus (n = 1, 0.9%) and Kaposi's sarcoma-associated herpesvirus (KSHV) (n = 1, 0.9%) (Fig 1).

Table 1

Results of Twist Comprehensive Viral Panel for the validation of 7 PCR-positive control samples

No.	Sample ID	PCR results/ expected results	CT values (multiplex PCR)	CZ ID	Interpretation (concordant/ discordant)	NT rPM values
1.	11-9-00810	Chikungunya virus	21.794	Chikungunya virus	Concordant	931,413.60
2.	11-9-00626	Chikungunya virus	21.344	Chikungunya virus	Concordant	937,860.20
3.	11-2-00757	Dengue virus 4	21.988	*Orthoflavivirus denguei*	Concordant	82,416.50
4.	11-1-00654	Dengue virus 4	17.192	*Orthoflavivirus denguei*	Concordant	868,719.00
				Chikungunya virus	-	234.90
5.	10-2-00395	Dengue virus 4	22.471	*Orthoflavivirus denguei*	Concordant	10,602.60
				Lymphocryptovirus humangamma4	-	138.50
6.	09-9-03014	Chikungunya virus	23.668	Chikungunya virus	Concordant	844,885.30
				Simplexvirus humanalpha1	-	107.10
				Orthoflavivirus denguei	-	48.90
7.	09-5-02672	Dengue virus 2	18.096	*Orthoflavivirus denguei*	Concordant	927,356.40

Table 1 (cont)

No.	Sample ID	PCR results/ expected results	CT values (multiplex PCR)	CZ ID	Interpretation (concordant/ discordant)	NT rPM values
8.	Vendor positive control	Influenza A virus H1N1	-	Influenza A virus	Concordant	138
		Influenza A virus H3N2		Influenza A virus	Concordant	138
		Influenza B virus		Betacoronavirus 1	Concordant	17.1
		SARS-CoV-2		Severe acute respiratory syndrome-related virus	Concordant	19.4
		HCoV-OC43		Betacoronavirus 1	Concordant	17.1
		HCoV-229E		Betacoronavirus 1	Concordant	17.1
		HCoV-NL63		Human coronavirus NL63	Concordant	9.1
		Measles virus		*Morbillivirus hominis*	Concordant	32.5
		Mumps virus		*Orthorubulavirus parotitidis*	Concordant	38.8
		Human bocavirus		Human bocavirus	Concordant	28.5
		Human enterovirus 68		Rhinovirus A	Concordant	9.1
		Parainfluenza virus type 1		*Respirovirus laryngotracheitidis*	Concordant	11.4
		Parainfluenza virus type 4		*Respirovirus laryngotracheitidis*	Concordant	11.4
		Rhinovirus type 89		Rhinovirus A	Concordant	9.1

Ct: cycle threshold; CZ ID: Chan Zuckerberg ID; HCoV: human coronavirus; NT rPM: nucleotide reads per million; PCR: polymerase chain reaction; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

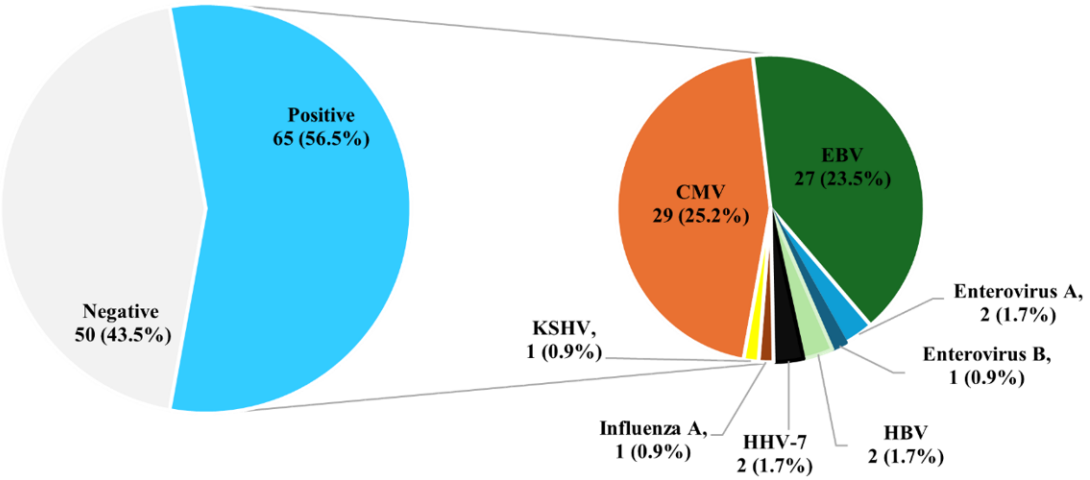

Fig 1 - Types, numbers, and percentages of viruses detected among study samples using the Twist Comprehensive Viral Panel (N = 115)

CMV: cytomegalovirus; EBV: Epstein-Barr virus; HBV: hepatitis B virus; HHV-7: human herpesvirus-7; KSHV: Kaposi's sarcoma-associated herpesvirus

We also summarized representative coverage metrics for CMV and EBV, the two most frequently detected viruses in our dataset. Because CZ ID does not generate traditional genome-wide breadth or depth, we used the platform's abundance metrics (NT and NR rPM) and significance metrics (NT and NR Z-scores) as surrogates for sequence coverage. Among CMV-positive samples, the median NT rPM was 17.5 (interquartile range (IQR for the middle 50%): 36.7) and the median NR rPM was 14.5 (IQR for the middle 50%: 35.5). EBV-positive samples had a median NT rPM of 22.8 (IQR for the middle 50%: 78.0) and a median NR rPM of 21.5 (IQR for the middle 50%: 52.1). For both CMV and EBV viruses, the NT and NR Z-scores consistently reached the CZ ID maximum of 100, indicating a strong signal above background. Although

hybrid-capture sequencing often yielded partial genome recovery, the combination of measurable rPM values, maximal Z-scores, and reads mapping to multiple genomic regions (data not shown) support the validity of these detections.

The median age of EBV-positive patients (47 years) was older than the median age of CMV-positive patients (25 years) and the median age of EBV-negative patients (40 years). Females comprised 55.6% of EBV-positive and 62.1% of CMV-positive patients, with a similar proportion among EBV- and CMV-negative patients. Headaches and chills were common in all groups. Headaches were reported in 66.7% of EBV-positive and 65.5% of CMV-positive patients. Sore throat was reported in 25.9% of EBV-positive and 24.1% of CMV-positive patients. Rash was reported in 7.4% of EBV-positive patients and 0% of CMV-positive patients. Musculoskeletal complaints were frequent across groups. Gastrointestinal symptoms, including nausea, vomiting, diarrhea, and abdominal pain, were reported in 48.1% of EBV-positive, 75.9% of CMV-positive, and 78.0% of negative patients. Respiratory symptoms were reported in 63.0% of EBV-positive, 69.0% of CMV-positive, and 36.0% of negative patients. Gastrointestinal and respiratory symptoms were significantly associated with viral detection (p-value <0.05) (Table 2).

DISCUSSION

In our study, hybrid-capture NGS substantially increased viral detection among hospitalized AFI patients whose illnesses remained unexplained after routine diagnostic testing. More than half of the selected serum samples yielded a viral pathogen using the Twist Comprehensive Viral Panel, demonstrating an important diagnostic gap between conventional culture, rapid diagnostic assays and targeted molecular testing and the Twist Comprehensive Viral Panel. Our findings are consistent with previous studies showing metagenomic and target-enriched

Table 2

Demographics and clinical characteristics among viral detection status

Characteristics	Frequency, *n* (%)*			*p*-value
	EBV (N = 27)	CMV (N = 29)	Negative (N = 50)	
Age in years, median (IQR)	47.0 (28.0)	25.0 (37.0)	40.0 (35.8)	0.022
Female sex	15 (55.6)	18 (62.1)	31 (62.0)	0.839
Headache	18 (66.7)	19 (65.5)	38 (76.0)	0.523
Chills	23 (85.2)	26 (89.7)	38 (76.0)	0.329
Sore throat	7 (25.9)	7 (24.1)	5 (10.0)	0.131
Rash	2 (7.4)	0 (0.0)	2 (4.0)	0.450
Myalgia/arthralgia	22 (81.5)	18 (62.1)	39 (78.0)	0.196
GI symptoms[a]	13 (48.1)	22 (75.9)	39 (78.0)	0.019
Respiratory symptoms[b]	17 (63.0)	20 (69.0)	18 (36.0)	0.009

*Unless otherwise stated

Note: p-values derived from the Kruskal-Wallis test (continuous) and chi-square test (categorical). Statistically significant when *p*-value <0.05.

[a]GI (gastrointestinal) symptoms included nausea, vomiting, diarrhea, and abdominal pain; [b]Respiratory symptoms included cough, rhinitis, and dyspnea

CMV: cytomegalovirus; EBV: Epstein-Barr virus; IQR: interquartile range

sequencing can identify additional pathogens in undiagnosed acute febrile illness, sepsis, respiratory infection and central nervous system infection cases when standard diagnostic panels are negative or incomplete (Yozwiak *et al*, 2012; Grundy *et al*, 2023; Castellot *et al*, 2023; Gauthier *et al*, 2023). Hybrid-capture enrichment is particularly useful in this setting because it improves detection of low-abundance viral genomes while maintaining broad coverage

for both expected and unexpected viruses (Wylie *et al*, 2015; Metsky *et al*, 2019; Mourik *et al*, 2024). These findings support the use of hybrid-capture NGS as a complementary second-line diagnostic approach for AFI cases without a clearly identified etiology after routine testing.

CMV was the most frequently detected virus in this study. Similar metagenomic studies in patients with undiagnosed acute febrile illness have demonstrated a variety of viruses not usually tested for on routine laboratory testing (Ashraf *et al*, 2025). The frequent detection of CMV may reflect the high background exposure to CMV in Thailand, since previous serologic studies among Thai blood donors showed it was common to have a prior CMV infection (Urwijitaroon *et al*, 1993). CMV detection may represent active CMV infection, CMV virus reactivation or the finding of CMV DNA in the blood but the patient is asymptomatic as described previously (Oliveira *et al*, 2026). The finding of large numbers of sequencing signals in CMV-positive samples suggests actual viral infection rather than a nonspecific background signal, but clinical interpretation needs to take into consideration the host's immune status, the timing of the illness and whether the detected CMV represents primary infection, reactivation or clinically relevant viremia (Martelius *et al*, 2010; Chiu and Miller, 2019).

EBV was the second most frequently detected virus. Previous studies have also reported EBV detection in patients with acute undifferentiated febrile illness (Dhodapkar *et al*, 2021). The high prevalence of EBV detection in our cohort may be associated with the widespread exposure to EBV in Thailand. EBV test result interpretation needs to take into consideration the same factors as those of CMV discussed in the previous paragraph (Martelius *et al*, 2010; Chiu and Miller, 2019).

Other viruses were detected less frequently, with enteroviruses representing the next most common

group after CMV and EBV. Enteroviruses have previously been reported as causes of acute illness in tropical settings (Le-Viet *et al*, 2019). Because these viruses can present with non-specific febrile syndromes, their detection in our cohort of AFI patients is clinically plausible and underscores the value of a metagenomic approach.

The clinical characteristics of patients with detected viruses were generally non-specific, reinforcing the difficulty of identifying viral AFI etiologies based on symptoms alone. Headache, chills and musculoskeletal symptoms were common among subjects who were EBV-positive, CMV-positive and virus-negative. Sore throat was reported in 25.9% of EBV-positive patients and 24.1% of CMV-positive patients, whereas rash was uncommon in both groups, occurring in 7.4% of EBV-positive patients and 0% of CMV-positive patients. Gastrointestinal and respiratory symptoms were significantly associated with subjects in whom a viral etiology was detected, suggesting these symptoms suggest patients may be more likely to have a positive viral test. However, this may vary type of pathogen and study population (Goumballa *et al*, 2020). Other studies have also reported relying on symptoms to diagnose viruses is not clinically useful (Grundy and Houpt, 2022; Saravanan *et al*, 2022; Wodniak *et al*, 2024).

Our results show the benefit of the hybrid-capture workflow. The Twist Comprehensive Viral Research Panel detected all the viruses in the qRT-PCR-positive validation samples, similar to the findings of previous studies showing that targeted or enriched NGS approaches can effectively identify viral nucleic acid that is present at detectable levels (Wylie *et al*, 2015; Metsky *et al*, 2019; Mourik *et al*, 2024). It is important to note our validation set was small and was not designed to determine sensitivity or specificity.

In our study, CZ ID background subtraction, rPM metrics, Z-score thresholds and minimum read

criteria strengthened the confidence of the reported detections. Similar approaches using abundance metrics, enrichment scores and background-subtraction models have been applied to metagenomic studies to distinguish true pathogen signals from background or contaminant reads (Mostafa *et al*, 2020; Castellot *et al*, 2023). For CMV and EBV, the presence of elevated abundance metrics, elevated enrichment scores and reads mapping to multiple genomic regions supports the validity of the findings. However, these metrics do not replace standardized viral-load quantification or complete genome coverage assessment (Chiu and Miller, 2019; Gauthier *et al*, 2023).

The main limitations of our study were the retrospective selection of residual serum samples, possible selection bias due to sample-volume availability, lack of paired specimen types and lack of viral-load, serologic and longitudinal clinical data. These limitations are particularly relevant for CMV and EBV, because both viruses are common latent infections and may be detected during primary infection, reactivation or be an incidental finding of viremia not representing an infection (Martelius *et al*, 2010; Chiu and Miller, 2019).

In summary, hybrid-capture NGS was able to determine more than half of samples previously testing negative with routine testing and the most common viruses detected were CMV and EBV. We conclude the hybrid-capture NGS may be beneficial for evaluating viral AFI in which the etiology is unclear with routine testing. Further studies in other populations are needed to determine when and in what situations the hybrid-capture NGS would be of most benefit in guiding management of patients with AFI of unknown etiology.

ACKNOWLEDGEMENTS

We would like to extend our gratitude to Drs Suzanne Beard, Saithip Bhengsri, Pongpun Sawatwong and Ms Duangkamon Siludjai from the Thailand

MOPH-US CDC Collaboration, and to Dr Beth Skaggs for their contributions. We also thank Dr Navin Horthongkham from the Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, and Dr Wittawin Worakitchanon, who contributed technical support while affiliated with the Association of Public Health Laboratories, for their assistance.

This research was funded under the CDC cooperative agreement with the Thailand Ministry of Public Health entitled 'Technical Collaboration to Strengthen and Expand Capacity to Detect, Prevent, Respond and Mitigate the Effects of Public Health Threats of National and Global Importance with the Ministry of Public Health (MOPH) in the Kingdom of Thailand' (Cooperative Agreement NU2GGH002341) and was also supported by Cooperative Agreement NU2HGH000080 funded by the Centers for Disease Control and Prevention through the Association of Public Health Laboratories.

CONFLICT OF INTEREST DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

Ashraf S, Jerome H, Bugembe DL, *et al*. Uncovering the viral etiology of undiagnosed acute febrile illness in Uganda using metagenomic sequencing. *Nat Commun* 2025; 16(1): 2844.

Bhaskaran D, Chadha SS, Sarin S, Sen R, Arafah S, Dittrich S. Diagnostic tools used in the evaluation of acute febrile illness in South India: a scoping review. *BMC Infect Dis* 2019; 19(1): 970.

Boers SA, Jansen R, Hays JP. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *Eur J Clin Microbiol Infect Dis* 2019; 38(6): 1059-70.

Castellot A, Camacho J, Fernández-García MD, Tarragó D. Shotgun metagenomics to investigate unknown viral etiologies of pediatric meningoencephalitis. *PLoS One* 2023; 18(12): e0296036.

Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet* 2019; 20(6): 341-55.

Dhodapkar R, Mugunthan M, Thangavelu K, Sivaradjy M, Veerappan K, Gunalan A. Epstein-Barr virus: an infrequent pathogen of acute undifferentiated febrile illness from a tertiary care hospital in southern India. *Cureus* 2021; 13(9): e18207.

Gauthier NPG, Chorlton SD, Krajden M, Manges AR. Agnostic sequencing for detection of viral pathogens. *Clin Microbiol Rev* 2023; 36(1): e0011922.

Goumballa N, Diop A, Hoang VT, *et al*. Pathogens associated with respiratory, gastrointestinal and febrile illness in patients consulting at Mbacke healthcare centre during the 2018 Grand Magal of Touba: a preliminary study. *Travel Med Infect Dis* 2020; 37: 101820.

Grundy BS, Houpt ER. Opportunities and challenges to accurate diagnosis and management of acute febrile illness in adults and adolescents: a review. *Acta Trop* 2022; 227: 106286.

Grundy BS, Parikh H, Jacob S, *et al*. Pathogen detection using metagenomic next-generation sequencing of plasma samples from patients with sepsis in Uganda. *Microbiol Spectr* 2023; 11(1): e0431222.

Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 2007; 45(9): 2761-4.

Leelarasamee A, Chupaprawan C, Chenchittikul M, Udompanthurat S. Etiologies of acute undifferentiated febrile illness in Thailand. *J Med Assoc Thai* 2004; 87(5): 464-72.

Le-Viet N, Le VH, Chung H, *et al*. Prospective case-control analysis of the aetiologies of acute undifferentiated fever in Vietnam. *Emerg Microbes Infect* 2019; 8(1): 339-52.

Martelius T, Lappalainen M, Aalto SM, Nihtinen A, Hedman K, Anttila VJ. Clinical characteristics, outcome and the role of viral load in nontransplant patients with Epstein-Barr viraemia. *Clin Microbiol Infect* 2010; 16(6): 657-62.

Metsky HC, Siddle KJ, Gladden-Young A, *et al*. Capturing sequence diversity in metagenomes with

comprehensive and scalable probe design. *Nat Biotechnol* 2019; 37(2): 160-8.

Mostafa HH, Fissel JA, Fanelli B, *et al*. Metagenomic next-generation sequencing of nasopharyngeal specimens collected from confirmed and suspect COVID-19 patients. *mBio* 2020; 11(6): e01969-20.

Mourik K, Sidorov I, Carbo EC, *et al*. Comparison of the performance of two targeted metagenomic virus capture probe-based methods using reference control materials and clinical samples. *J Clin Microbiol* 2024; 62(6): e0034524.

Oliveira LP, Paton EJA, Giovanardi MF, *et al*. Assessing the significance of cytomegalovirus reactivation in recipients of allogeneic hematopoietic stem cell transplantation: a cohort study. *Braz J Med Biol Res* 2026; 59: e15073.

Saravanan N, Rajendiran P, Nandagopal B, Ramamurthy MB, Vadivel K. A review of acute febrile illness. *Indian J Microbiol Res* 2022; 9(4): 232-40.

Twist Bioscience. Twist Comprehensive Viral Research Panel, 2020 [cited 2026 Mar 13]. Available from: URL: https://www.twistbioscience.com/content/dam/twistbioscience/resources/2020-11/ProductSheet_NGS_ComprehensiveViralResearchPanel_11Nov20_Rev1.0.pdf

Urwijitaroon Y, Teawpatanataworn S, Kitjareontarm A. Prevalence of cytomegalovirus antibody in Thai-northeastern blood donors. *Southeast Asian J Trop Med Public Health* 1993; 24(Suppl 1): 180-2.

Wodniak NR, Bhengsri S, Skaggs B, *et al*. Demographic and clinical factors associated with bacterial or nonbacterial etiologies of acute undifferentiated febrile illness: findings from a 3-year observational study in Thailand, 2017-2020. *Am J Trop Med Hyg* 2024; 111(3): 650-60.

Wylie TN, Wylie KM, Herter BN, Storch GA. Enhanced virome sequencing using targeted sequence capture. *Genome Res* 2015; 25(12): 1910-20.

Yozwiak NL, Skewes-Cox P, Stenglein MD, Balmaseda A, Harris E, DeRisi JL. Virus identification in unknown tropical febrile illness cases using deep sequencing. *PLoS Negl Trop Dis* 2012; 6(2): e1485.