

METHODOLOGY

Open Access



# Development and evaluation of a non-ribosomal random PCR and next-generation sequencing based assay for detection and sequencing of hand, foot and mouth disease pathogens

Anh To Nguyen<sup>1\*</sup>, Thanh Tan Tran<sup>1</sup>, Van Minh Tu Hoang<sup>2</sup>, Ngoc My Nghiem<sup>3</sup>, Nhu Nguyen Truc Le<sup>1</sup>, Thanh Thi My Le<sup>3</sup>, Qui Tu Phan<sup>3</sup>, Khanh Huu Truong<sup>4</sup>, Nhan Nguyen Thanh Le<sup>4</sup>, Viet Lu Ho<sup>2</sup>, Viet Chau Do<sup>2</sup>, Tuan Manh Ha<sup>2</sup>, Hung Thanh Nguyen<sup>4</sup>, Chau Van Vinh Nguyen<sup>3</sup>, Guy Thwaites<sup>1,5</sup>, H. Rogier van Doorn<sup>1,5</sup> and Tan Van Le<sup>1</sup>

## Abstract

**Background:** Hand, foot and mouth disease (HFMD) has become a major public health problem across the Asia-Pacific region, and is commonly caused by enterovirus A71 (EV-A71) and coxsackievirus A6 (CV-A6), CV-A10 and CV-A16. Generating pathogen whole-genome sequences is essential for understanding their evolutionary biology. The frequent replacements among EV serotypes and a limited numbers of available whole-genome sequences hinder the development of overlapping PCRs for whole-genome sequencing. We developed and evaluated a non-ribosomal random PCR (rPCR) and next-generation sequencing based assay for sequence-independent whole-genome amplification and sequencing of HFMD pathogens. A total of 16 EV-A71/CV-A6/CV-A10/CV-A16 PCR positive rectal/throat swabs (Cp values: 20.9–33.3) were used for assay evaluation.

**Results:** Our assay evidently outperformed the conventional rPCR in terms of the total number of EV-A71 reads and the percentage of EV-A71 reads: 2.6 % (1275/50,000 reads) vs. 0.1 % (31/50,000) and 6 % (3008/50,000) vs. 0.9 % (433/50,000) for two samples with Cp values of 30 and 26, respectively. Additionally the assay could generate genome sequences with the percentages of coverage of 94–100 % of 4 different enterovirus serotypes in 73 % of the tested samples, representing the first whole-genome sequences of CV-A6/10/16 from Vietnam, and could assign correctly serotyping results in 100 % of 24 tested specimens. In all but three the obtained consensus of two replicates from the same sample were 100 % identical, suggesting that our assay is highly reproducible.

**Conclusions:** In conclusion, we have successfully developed a non-ribosomal rPCR and next-generation sequencing based assay for sensitive detection and direct whole-genome sequencing of HFMD pathogens from clinical samples.

**Keywords:** Hand, foot and mouth disease, Enterovirus A, Random PCR, FR26RV-Endoh primer, Next-generation sequencing

\* Correspondence: anhnt@oucru.org

<sup>1</sup>Oxford University Clinical Research Unit, 764 Vo Van Kiet Street, Ward 1, District 5, Ho Chi Minh City, Vietnam

Full list of author information is available at the end of the article



## Background

Hand, foot and mouth disease (HFMD) is a common and usually mild disease of children worldwide. The disease is caused by different genotypes of the species Enterovirus A, genus *Enterovirus*, family *Picornaviridae* (including coxsackievirus A (CV-A) 6, 10 and 16 and particularly EV-A71). However, EV-A71 has emerged and caused large and sometimes severe/fatal HFMD outbreaks [1] across the Asia-Pacific region since 1997. Of note, the frequent replacements between EV-As have been observed over the last decade in the regions where HFMD is endemic [2–6]. In recent years CV-A6 has emerged and replaced CV-A16 to become the dominant EV-A detected in HFMD patients [7, 8]. While the underlying mechanism of this phenomenon remains unknown, the data highlight the importance of continued effort to monitor the evolution of the causative agents of HFMD.

Currently, there is no clinically proven antiviral drug available to treat severe disease. Likewise, although phase III trials of three monovalent inactivated EV-A71 vaccines have been completed in China with an efficacy of over 95 %, routine use is still far away. Moreover, to what degree the implementation of a monovalent vaccine for EV-A71 may influence the epidemic patterns of HFMD and the evolution of the causative agents in endemic countries is a subject that merits follow-up research.

Collectively, the ability to generate viral whole-genome sequences is essential for understanding the evolutionary biology and epidemiology of HFMD. It is also important for the development of intervention strategies, especially vaccines. While the availability of relatively large numbers of EV-A71 whole-genome sequences ( $n = \sim 524$ ) deposited in GenBank has facilitated the development of a sensitive overlapping PCR based whole-genome sequencing assay [9], smaller numbers of whole-genome sequences of other EV-As are available (CV-A16;  $n = 61$ , A6; 35, A10; 11) from limited localities. This is problematic for the selection of specific PCR primers that can amplify diverse EV-As. Additionally, one of the major drawbacks of specific-PCR based sequencing assays is that due to the nature of quick evolution rates of RNA viruses, selected primers may need to be adjusted regularly to be able to amplify newly emerging viral variants or genotypes. As a consequence, a sequence-independent approach is thus attractive to overcome such obstacles.

Developed by Froussard in 1992 [10], random PCR (rPCR) primer (FR26RV-N6: 5'-GCCGGAGCTCTGCA GATATCNNNNNN-3') consists of a fixed 20 nucleotides (FR20RV: GCCGGAGCTCTGCAGATATC) at the 5'-end and a random hexanucleotides at 3' end (N6: NNNNNN). In 2005 Endoh and his colleagues designed a set of 96 hexanucleotides for specific amplification of

viral sequences called non-ribosomal hexanucleotides [11]. For sequence-independent whole-genome amplification and sequencing of HFMD pathogens, herein we describe the development and evaluation of a non-ribosomal random amplification assay utilizing the 96 non-ribosomal hexanucleotide oligos designed by Endoh [11] and the 5'-end fixed oligo of the conventional random PCR primers (FR20RV) [10]. When combined with next-generation sequencing, our assay showed that it could generate full-genome sequences of HFMD pathogens directly from clinical specimens.

## Methods

### Samples

The clinical samples used included two residual throat swabs from anonymous HFMD patients with EV-A71 infection admitted to the Hospital for Tropical Diseases in Ho Chi Minh City in 2012. Additionally, 13 throat/rectal swabs of diverse viral load (including CV-A6;  $n = 4$ , CV-A10;  $n = 4$ , CV-A16;  $n = 3$  and EV-A71;  $n = 2$ ) derived from patients enrolled into an on-going prospective observational HFMD study of all severities in three referral hospitals in Ho Chi Minh City, Vietnam since 2013 were also used [9]. The clinical samples were collected in viral transport medium, divided into three aliquots and stored at  $-80^{\circ}\text{C}$  until use. Viral detection and serotype identification were done as per the study protocol using previous described assays [12, 13].

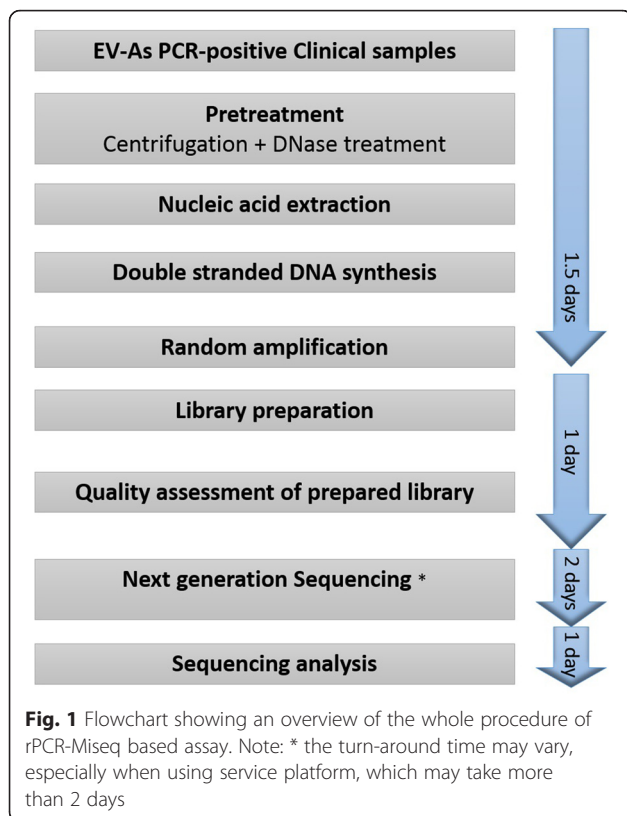
### Development and preparation of non-ribosomal random PCR primers

For selective amplification of viral sequences, we replaced the random hexanucleotide motif at the 3'-end of the primer FR26RV-N6 by those 96 hexanucleotides designed by Endoh. This resulted in a set of 96 separate primers consisting of an FR20RV sequence at 5'-end plus one of the 96 Endoh's hexanucleotides at the 3'-end (Additional file 1: Table S1).

Each individual primer was synthesized at a concentration of  $100\ \mu\text{M}$ , and an equal amount of each synthesized oligo was pooled together to make working solution ( $\sim 1\ \mu\text{M}$ ). This primer mixture was named FR26RV-Endoh.

### Sample pretreatment and nucleic acid extraction

An overview of the whole procedure is described in Fig. 1. Sample pretreatment was carried out as previously described [14]. In short, prior to nucleic acid isolation  $110\ \mu\text{l}$  of clinical samples was centrifuged at  $10,000\ \text{g}$  for 10 min. The resulting  $100\ \mu\text{l}$  of supernatants were collected and treated with  $2\ \text{U}/\mu\text{l}$  of turbo DNase (Ambion, Life Technology, Carlsbad, CA, USA) at  $37^{\circ}\text{C}$  for 30 min. Viral RNA was then extracted from the treated material using QIAamp viral RNA kit



(QIAGEN GmbH, Hilden, Germany), following the manufacturer’s instructions, and finally eluted in 50 µl of elution buffer (provided with the extraction kit).

**cDNA and double stranded DNA synthesis**

Double stranded (ds) DNA was synthesized from the extracted RNA using either FR26RV-N6, FR26RV-Endoh, random hexanucleotides or non-ribosomal hexanucleotides primer. Firstly, 10 µl of extracted RNA was mixed with 0.1 µM of the primer and 0.5nM of dNTPs (Roche Diagnostics GmbH, Mannheim, Germany). The mixture was incubated at 65 °C for 5 min, and was then immediately chilled on ice for 1 min. Secondly, 7 µl of a reaction mix containing 200U of Super Script III, 40 U of RNase OUT, 0.1 M DTT and 1X first strand buffer (Invitrogen, Carlsbad, CA, USA) was added into the first reaction mixture. The reaction was continued at 25 °C for 10 min, 37 °C for 1 min and 94 °C for 2 min, and then immediately chilled on ice for 2 min. Next, 5U of exo-Klenow fragment (Ambion) and 10U of Ribonuclease H (Ambion) were added into the reaction mixture, which was finally subjected to a double-stranded (ds) DNA synthesis step consisting of 25 °C for 5 min, 37 °C for 1 h and 75 °C for 10 min.

**Random amplification**

The resulting dsDNA products generated by FR26RV-N6 and FR26RV-Endoh primers were amplified using

FR20RV primer (5’-GCCGGAGCTCTGCAGATATC-3’). PCR amplification was carried out in a total reaction volume of 50 µl consisting of 3 µl of dsDNA, 0.4 µM of primer FR20RV and 45 µl of Platinum PCR supermix high fidelity (Invitrogen). The thermal cycling condition consisted of 94 °C for 2 min and followed by 40 cycles of 94 °C for 30s, 55 °C for 30s and 72 °C for 3 min and 1 cycle of 72 °C for 2 min.

**Next generation sequencing library preparation and sequencing**

The resulting dsDNA generated by hexanucleotides or non-ribosomal hexanucleotides and rPCR products were purified with use of QIAquick PCR purification kit (QIAGEN GmbH, Hilden, Germany). DNA concentration of the purified products was measured by Qubit dsDNA HS kit (Invitrogen). One nanogram of the purified DNA was then subjected to library preparation steps by using Nextera XT DNA library preparation kit (Illumina, San Diego, CA, USA), according to manufacturer’s instructions. Prior to sequencing, the quantity of the prepared library was measured by using KAPA Library Quant Kit (Kapa Biosystems, Wilmington, MA, USA), following manufacturer’s instructions.

The prepared library was sequenced using MiSeq reagent kit V2 in an Illumina MiSeq platform (Illumina). For each run, tested samples were multiplexed and differentiated by double indexes using Nextera XT Index Kit (Illumina).

**Sequence analysis**

The sequences generated by Illumina MiSeq were analyzed using Geneious 8.1.5 (Biomatters, San Francisco, CA, USA). The obtained sequences were processed to remove primer sequences. Sequence assembly was carried out by using a reference-based mapping strategy available in Geneious (CV-A10, HQ728262; CV-A6, JN582001; CV-A16, JX481738; EV-A71 B5, DQ341363; EV-A71 C4, AB550338), followed by manual editing of the obtained consensus.

Representatives of viral protein 1 (VP1) sequences of CV-A16 (n = 39), A6 (38), A10 (29) and EV-A71 (36) of different subgenotypes and from various localities worldwide were used for phylogenetic inference. Pairwise alignment was performed using Geneious alignment tool. Phylogenetic reconstructions were performed using maximum likelihood method (ML) with general time reversible (GTR) nucleotide substitution model available in Geneious package, and support for individual nodes was assessed using a bootstrap procedure (1000 replicates).

The sequences obtained in this study were submitted to NCBI (GenBank) and assigned accession numbers KX430795-KX430824.

## Results

### Non-ribosomal rPCR vs. conventional rPCR

To test whether our modified rPCR, which we named non-ribosomal rPCR, can selectively amplify viral sequences in clinical specimen as compared to the conventional rPCR, two EV-A71 positive swabs with Cp values of 26 (ID.13) and 30 (ID.14) (i.e. high and low viral load) were selected and subjected to random amplification procedures utilizing either FR26RV-N6 or FR26RV-Endoh, and followed by Illumina Miseq sequencing. The total- and percentage of EV-A71 reads, genome coverage and sequencing depth/coverage (i.e. the number of times a single nucleotide was sequenced) were taken into account for comparison.

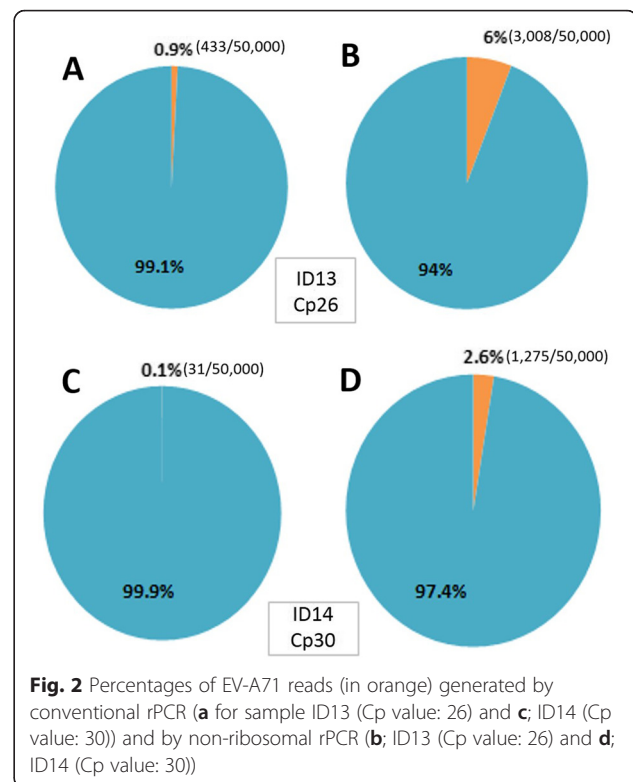
In order to avoid the potential biases introduced by variable number of reads between barcodes, a total of 50,000 reads were randomly taken from each index for the analysis. In both tested EV-A71 positive samples, the total number of EV-A71 reads and the percentage of EV-A71 reads generated by non-ribosomal rPCR based assay was higher than the corresponding outputs generated by the conventional rPCR-based assay; 2.6 % (1275/50,000 reads) vs. 0.1 % (31/50,000 reads) for the sample ID14 with Cp value of 30 and 6 % (3008/50,000 reads) vs. 0.9 % (433/50,000 reads) for the sample ID13 with Cp value of 26 (Fig. 2). Additionally, a higher EV-A71 genome coverage and sequencing depth were also observed in both samples sequenced by non-ribosomal rPCR-based assay (Fig. 3). Taken together, the data indicated that our non-ribosomal rPCR is more viral specific and efficient than the conventional rPCR.

### Non-ribosomal rPCR vs. direct sequencing

Previous studies shown that viral load enrichment by random amplification step resulted in biases in genome coverage [15, 16]. We therefore further evaluated our non-ribosomal rPCR by comparing its performance against that of direct sequencing of dsDNA library generated by hexanucleotide or non-ribosomal hexanucleotide primers. An EV-A71 positive throat swab (sample ID15) with a Cp value of 31 was used. After normalization, the obtained reads of each DNA library were map to an EV-A71 genome (DQ341363.1). Despite biases in terms of sequencing depth across the genome, non-ribosomal rPCR based workflow could generate nearly complete EV-A71 genome sequence (KX430823), while dsDNA library produced by hexanucleotide and non-ribosomal hexanucleotide primers could not (Additional file 1: Figure S1).

### Detection and sequencing of HFMD pathogens: assessment of assay sensitivity and reproducibility

To further evaluate the performance of our non-ribosomal rPCR assay in terms of sensitivity and reproducibility a series of 12 swabs that were EVs real time PCR



**Fig. 2** Percentages of EV-A71 reads (in orange) generated by conventional rPCR (a for sample ID13 (Cp value: 26) and c; ID14 (Cp value: 30)) and by non-ribosomal rPCR (b; ID13 (Cp value: 26) and d; ID14 (Cp value: 30))

positive with different common HFMD pathogens (including CV-A6, CV-A10, CV-A16 and EV-A71) and with a wide range of Cp values from 20.8 to 33.3 [12] (i.e. from high to low viral load) (described in Methods section) were included for testing (Table 1). The included samples were tested in duplicate from sample pretreatment to nucleic acid isolation, random amplification by FR26RV-Endoh primers and sequencing by Illumina Miseq, resulting a total of 24 MiSeq datasets (Table 1).

### Assay sensitivity

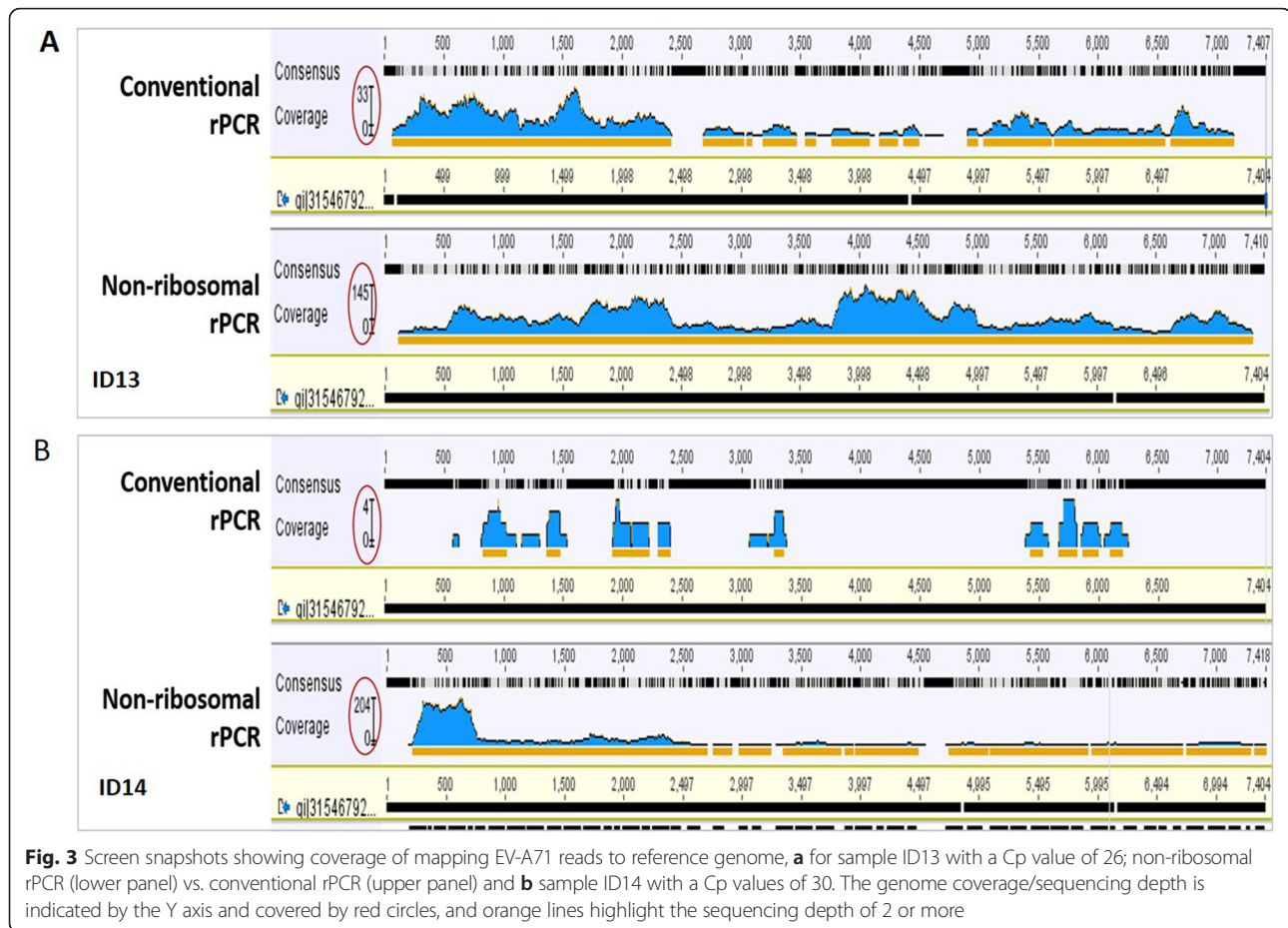
Illumina Miseq sequencing results showed that in addition to successfully providing correct serotype information (i.e. diagnostic results) in 100 % (24/24) of the tested samples, the assay could generate 17/24 (71 %) genome sequences of HFMD pathogens with the percentages of coverage of between 94 and 100 % (Table 1).

Collectively, of 24 tested samples, whole-genome sequencing success rates of 100 % (8/8), 93 % (13/14) and 71 % (17/24) with genome coverage of 94-100 % without internal gap were achieved among samples with Cp values of  $\leq 25$ ,  $\leq 30$  and  $\leq 33.3$ , respectively (Table 1).

### Assay reproducibility

To investigate the reproducibility of the assay, we compared the level of sequence identity between the obtained consensus of the tested sample and its replicate. In 9/12 tested samples the consensus of both





replicates were 100 % identical (Table 1). In the remaining 3 samples, the differences of between 0.01 - 0.04 % were recorded (Additional file 1: Table S2). Additionally, the level of genome coverage, mean coverage (i.e. the numbers of times that a single nucleotide was sequenced) and the percentage of viral reads were comparable between two replicates (Table 1).

**Phylogenetic analysis**

Currently there are relatively few whole-genome sequences of CV-A6, CV-A10 and CV-A16 from limited geographical localities available in GenBank. To make more meaningful phylogenetic inference, we therefore first focused our analysis on representative VP1 sequences collected from different geographic locations worldwide.

Phylogenetic analysis of VP1 sequences suggested that the EV-A71 strains obtained in the present study sampled in 2012 belonged to subgenogroup C4, whereas the viruses collected in 2013 belonged to subgenogroup B5 (Additional file 1: Figure S1), which reconfirmed our previous finding about the replacement between these two subgenogroups occurring in Vietnam around 2012 [17]. All CV-A16 sequences belonged to genogroup B1a. In Vietnam, this B1a genogroup was first detected in the

2005 outbreak [18] and showed a close relatedness to the viruses circulating in the Asia-Pacific region (e.g. China, Japan, Thailand and Malaysia) (Additional file 1: Figure S2). In contrast, the analysis of CV-A6 sequences indicated that our CV-A6 belonged to genogroup A, which consists of CV-A6 strains sampled from United Kingdom and others viruses from China and Taiwan (Fig. 4b). Likewise, the CV-A10 strains sequenced in the present study belonged to genogroup C consisting of viral trains originating from various parts of the world and associated with HFMD outbreaks in Europe and Asia including in Spain, France and China (Fig. 4a). Similar results in terms of phylogenetic clustering of the sequences were obtained when whole-genome sequences were analyzed separately (data not shown).

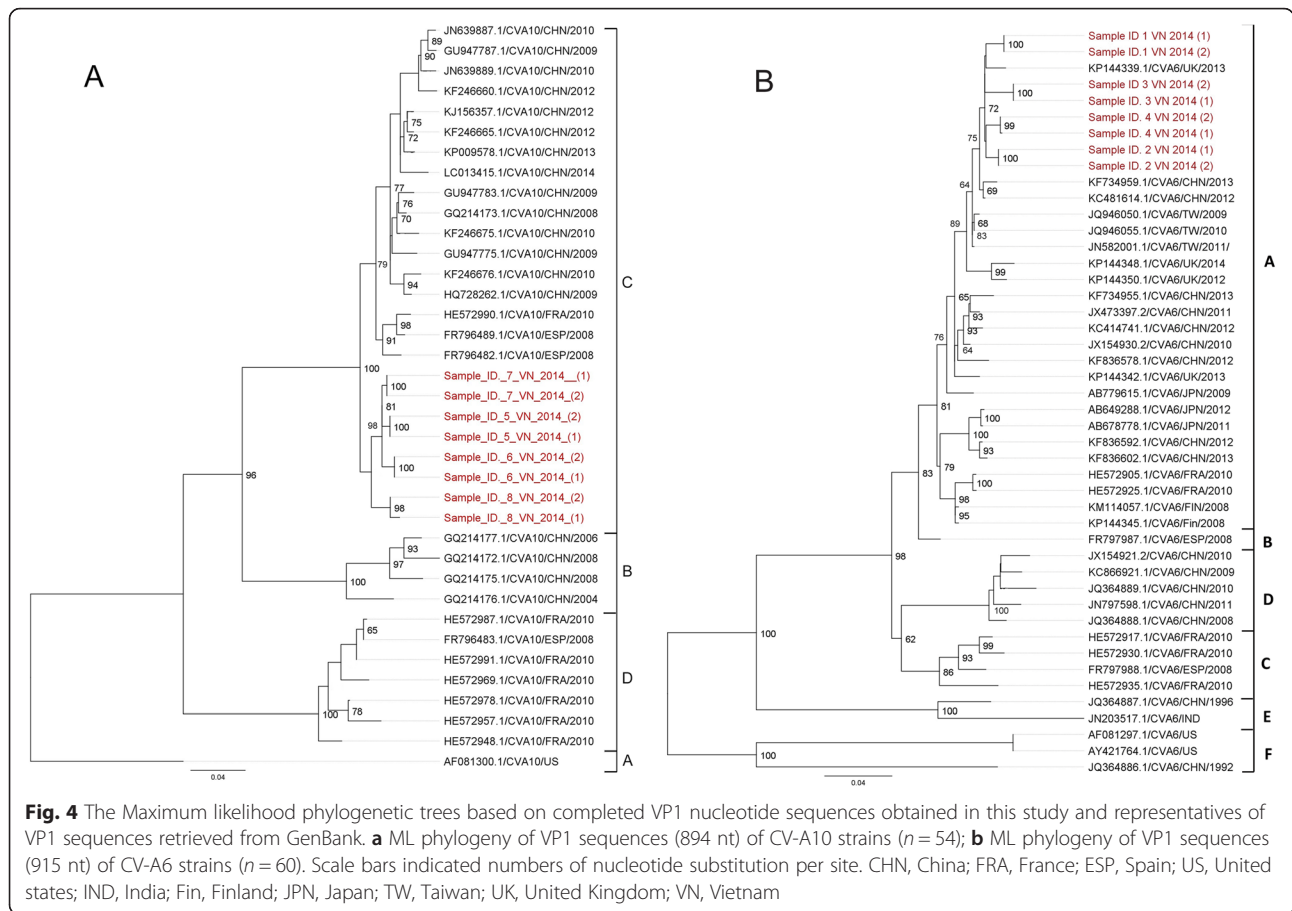
**Discussion**

Traditionally, obtaining whole-genome sequence of a pathogen requires the design of several overlapping specific PCR primers based on the basis of sequence alignment of the published genome sequences. Although such strategies have been successfully applied for sequencing of HFMD pathogens including EV-A71 and other EV-As [9, 19–21], except for EV-A71, these

**Table 1** Result summary of non-ribosomal rPCR and Miseq run

Virus <sup>a</sup>	Sample ID	Sample type	Cp values	% of enteroviral read	% Genome coverage	Internal gap length (bp)	Mean coverage	Accession numbers	Pairwise identity (%)
CV-A6	1	RS	22.69	90.2	99.5	0	26542	KX430795	100
				85.1	100	0	22630	KX430796	
	2	TS	28.34	11.2	97.5	0	1173	KX430797	100
				10.9	95.3	0	1119	KX430798	
CV-A10	3	RS	30.5	7.9	97.7	0	1822	KX430799	100
				8.3	97.3	57	2244	KX430800	
	4	TS	32.06	7.1	75	1625	1328	KX430801	99.96
				13.3	96.8	41	3061	KX430802	
CV-A16	5	RS	20.92	40.7	99.2	0	11189	KX430803	100
				53.8	98	0	12725	KX430804	
	6	RS	23.59	53.1	97.5	0	17439	KX430805	100
				51.9	97.6	0	14086	KX430806	
CV-A71	7	RS	26.71	18.2	98	0	4299	KX430807	99.99
				14.5	97	0	3820	KX430808	
	8	TS	33.2	42.4	94	0	12216	KX430809	99.99
				30.5	97.2	26	8412	KX430810	
EV-A71	9	TS	24.97	83.8	99.7	0	3161	KX430811	100
				80.1	99	0	20597	KX430812	
	10	TS	26.72	2	91	71	475	KX430813	100
				2.1	96	0	509	KX430814	
EV-A71	11	TS	33.26	4.5	96	0	971	KX430815	100
				4.2	86	712	1101	KX430816	
	12	TS	31.1	0.2	72.5	1447	5.3	KX430817	100
				0.3	94	0	52.6	KX430818	

<sup>a</sup>Miseq run was multiplexed. Only run output of relevant samples were shown here; CV-A6: coxsackievirus A6, CV-A10: coxsackievirus A10, CV-A16: coxsackievirus A16 and EV-A71 B5: enterovirus A71 subgenogroup B5; TS: Throat swab; RS: rectal swab



overlapping primers were designed based on a limited numbers of sequences of EV-As and therefore may not function properly on diverse circulating viral strains whose complete genomes are yet to be sequenced. In addition, to be able to amplify emerging outbreak/novel strain, such viral specific PCR primers often need to be updated regularly, which is always challenging.

There have been several reports regarding the use of random primers, e.g. FR26RV-N6 primer, to generate whole-genome sequence of viral pathogens [22, 23]. However, as FR26RV-N6 primer contains a random hexamer motif at the 3' end, which is not viral specific, assays may therefore lack specificity when used on materials such as rectal/throat swabs, which contain high amounts of host genetic materials and low concentrations of targeted virus. Meanwhile, Endoh's non-ribosomal hexanucleotide oligos have recently been successfully used as an alternative to random hexamers for selective amplification of viral RNA in the field of viral pathogen discovery [24–26]. For specific amplification and sequencing of viral pathogens in particular HFMD viruses (which were the focus of the present study) in clinical specimens, we adapted the fixed 5' end oligo of the normal random PCR and Endoh's non-

ribosomal hexanucleotides to create a novel 96 viral specific rPCR primer set (Additional file 1: Table S1).

When compared back-to-back using EV-A71 positive swabs, our non-ribosomal rPCR evidently outperformed the normal rPCR utilizing FR26RV-N6 primers and direct sequencing of dsDNA libraries generated by either hexanucleotides or non-ribosomal hexanucleotides. In subsequent testing we showed that without the requirement of viral specific PCR, our assay could generate whole-genome sequences of 4 different common HFMD pathogens (including CV-A6, CV-A10, CV-A16 and EV-A71) in either rectal or throat swabs with diverse viral load. Of 24 tested samples with Cp values between 20.9 and 33.2, (nearly) complete genomes were obtained in 17/24 (71 %) samples, representing the first whole-genome sequences of CV-A6, CV-A10 and CV-A16 from Vietnam. In three tested swabs and their replicates, the obtained consensus occupied between 0.01–0.04 % of differences. This is however below the reported error rate of next generation sequencing (0.1 %). Of note, 2 out of the 4 EV-A71 genomes sequenced in the present study (sample IDs: 13 and 15) were previously recovered (KJ686266 and KX430824) using an overlapping PCRs and deep sequencing based workflow [9, 17]. And

pairwise comparisons of the obtained consensus sequences generated by both workflows revealed only 0.03 % and 0.04 % of variations without amino acid substitution observed (data not shown). Collectively, the data points to the fact that potential biases (if any) introduced by enrichment steps as 40-cycle PCR amplification by FR20RV primer of the present workflow is negligible and that our non-ribosomal rPCR and next-generation based assay is reproducible and sensitive.

Despite the use of non-ribosomal primers and the employment of a sample pretreatment step incorporating centrifugation and DNase treatment to enrich for enteroviral content in the swabs, the percentage of enteroviral reads in the obtained MiSeq libraries ranged between 0.2 and 90.2 %. This might have been attributed to the difference in terms of the compositions of non-enteroviral contents between the samples and/or the viral load of the tested viruses. Meanwhile there have been other reports about alternative sequence-independent whole-genome next-generation sequencing based assays including those incorporating sample pretreatment steps as physical virion enrichment and RNase digestion [27–29]. It is therefore of interest to evaluate the usefulness of those sample pretreatment steps when combined with our non-ribosomal rPCR. Likewise, comparing the performance of our non-ribosomal rPCR with those existing sequence-independent assays warrants further research, which is however beyond the scope of the present study.

For clinical diagnostics, obtaining partial viral genome sequence is sufficient for establishment of the diagnostic result. Exploring the use of next-generation sequencing based assay as a diagnostic tool was an objective in many recent reports [30–32]. In addition, next-generation sequencing has been shown to be able to establish the diagnostics in swabs from HFMD patients that were enterovirus specific PCR negative [33]. Similarly, our assay could sequence and provide correct serotype information of the targeted enteroviruses in all tested samples with Cp values between 20.9 and 33.2, although we did not test our assay on samples with lower viral load (i.e. Cp value of >33.2). Assuming that a Cp value of 33.2 is the assay limit of detection, and a Cp value of <30 is required for the purpose of whole-genome sequencing; among a sample collection from over 1300 HFMD patients enrolled in our ongoing HFMD study in Ho Chi Minh City, Vietnam (data not shown), we can conservatively extrapolate that our assay can detect enterovirus in 97 % and generate complete or nearly complete genome sequence of enteroviruses in 62 % of the RT-PCR positive clinical samples, respectively.

The advantages of random amplification and NGS based assay include: i) there is no requirement for several pathogen specific assays to diagnose diseases caused by multiple pathogens as HFMD, and ii) in addition to providing

diagnostic information, the obtained sequencing result is informative for study of viral evolution and identification of the source of an outbreak. Indeed, by analyzing the obtained sequences we were able to reveal interesting insights into the evolution and origin of the CV-A6, CV-A10, CV-A16 and EV-A71 in Vietnam, albeit the sample size was small. As a consequence, further effort to obtain full genome sequences of HFMD causing pathogens is currently ongoing as part of our HFMD research program, which ultimately would lay the foundation for future research focusing on genetic diversity and evolutionary dynamics of HFMD in Vietnam and beyond, and can now be facilitated by our viral specific rPCR and next-generation sequencing based assay.

Our study has some limitations: i) we only evaluated our assay performance on rectal/throat swabs, whereas in HFMD, viral detection in vesicle swab, blood, CSF and urine has been reported, albeit at a lower frequency in the latter 3 sample types. Evaluation of the assay on these sample types is therefore needed, ii) similar to other reports [15, 16], we observed that the level of sequencing depth varied across the genomes of the tested viruses generated in the present study. While *in silico* investigation did not reveal any biases in terms of binding preference of the non-ribosomal hexanucleotides to specific genomic regions of the targeted viruses, it is attempting to speculate that such biases were attributed to the transposome-based library workflow as previously reported [34], iii) the current high cost (~\$USD 75 per sample as compared to \$USD 5–8 per one multiplex PCR reaction), low throughput (total operation time is about 5 days to complete) and bioinformatics requirements remain major barriers for next-generation sequencing-based assays to be widely applied in a diagnostic setting, in particular in less developed countries in Asia where HFMD is endemic, iv) the capacity of rPCR and next generation sequencing based assay to detect mixed infection and to identify novel/new viral variants [27, 35, 36] was not explored as it is beyond the scope of this study. For the latter, *de novo* assembly approach followed by metagenomic analysis using appropriate bioinformatics tool is recommended. Likewise, evaluating the viability of the 96 non-ribosomal hexanucleotides on new viral species discovered from 2005 onward is needed.

## Conclusion

We have successfully developed a non-ribosomal rPCR and next-generation sequencing based assay for sensitive detection and whole-genome sequencing of HFMD pathogens in clinical samples. Our assay can be used to study the genetic diversity and evolutionary biology of HFMD pathogens, which may aid the development of intervention strategies (including vaccines), and guide



public health plans in response to future HFMD outbreaks. As next-generation sequencing associated cost has been going down quickly, and once the bioinformatics challenge becomes less burden, one would expect the expanding use of next-generation sequencing based methodologies in clinical research and routine care, both in developed and less developed countries.

## Additional file

**Additional file 1: Table S1.** List of 96 FR26RV-Endoh and FR20RV primer sequences. **Table S2.** Result summary of consensus sequence variations recorded between 2 replicates of 3 tested swabs. Note: NA: not applicable. **Figure S1.** Screen snapshots showing the mapping results of EV-A71 MiSeq reads to an EV-A71 reference genome of sample ID15; non-ribosomal rPCR assay (bottom panel), non-ribosomal hexanucleotide primers assay (middle panel) and hexanucleotide assay (top panel); the genome sequencing depth is indicated by the Y axis and covered by red circles. **Figure S2.** Maximum likelihood phylogenetic tree based on completed VP1 nucleotide sequences (891 nt) of EV-A71 strains obtained from this study (in bold red) and representatives retrieved from GenBank. Scale bars indicated numbers of nucleotide substitution per site. CHN, China; USA, United states; TW, Taiwan; NL, Netherlands; MY, Malaysia; KOR, Korean; VN, Vietnam. **Figure S3.** Maximum likelihood phylogenetic tree based on completed VP1 nucleotide sequences (891 nt) of CV-A16 strains obtained from this study (in bold red) and representatives retrieved from GenBank. Scale bars indicated numbers of nucleotide substitution per site. CHN, China; US, United states; TL, Thailand; JPN, Japan; AUS, Australia; MY, Malaysia; KOR, Korean; VN, Vietnam. (PDF 783 kb)

## Acknowledgements

We thank Mrs. Le Kim Thanh from Oxford University Clinical Research Unit in Ho Chi Minh City, Vietnam for her logistic assistance. We are indebted to patients and their parents for their participation in this study, and all the nursing and medical staff at the Children's Hospital 1 and 2, and the Hospital for Tropical Diseases who provided care for the patients and helped collect clinical data.

## Funding

The research leading to these results has received funding from the Wellcome Trust (101104/Z/13/Z and 089276/Z/09/Z).

## Authors' contributions

NTA and LVT: designed the study, did laboratory testing, analysed the test results, performed statistical analysis and drafted the manuscript. TTT, HMTV, NMN, LNTN, LTMT, PTQ, THK, LNTN, HLV, DCV, HMT, NTH, NWC, GT: enrolled patients, took samples and did laboratory testing. RHvD: designed the study and involved in drafting the manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Ethics approval and consent to participate

The clinical samples used in this study were derived from an on-going HFMD study in three referral hospitals in Ho Chi Minh city, Vietnam. The study was reviewed and approved by the local Institutional Review Boards and the Oxford Tropical Research Ethics Committee (OxTREC), University of Oxford, Oxford, United Kingdom. The institutional review board of HTD in HCMC, Vietnam approved the whole-genome sequencing of residual swabs of anonymous HFMD patients. Written informed consent was obtained from a parent or guardian of each enrolled patients.

## Author details

<sup>1</sup>Oxford University Clinical Research Unit, 764 Vo Van Kiet Street, Ward 1, District 5, Ho Chi Minh City, Vietnam. <sup>2</sup>Children's Hospital 2, Ho Chi Minh City, Vietnam. <sup>3</sup>Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam.

<sup>4</sup>Children's Hospital 1, Ho Chi Minh City, Vietnam. <sup>5</sup>Centre for Tropical Medicine, Nuffield Department of Medicine, University of Oxford, Oxford, UK.

Received: 24 November 2015 Accepted: 29 June 2016

Published online: 07 July 2016

## References

- Brown BA, Oberste MS, Alexander JP, Kennett ML, Pallansch MA. Molecular epidemiology and evolution of enterovirus 71 strains isolated from 1970 to 1998. *J Virol*. 1999;73:9969–75.
- Li L, He Y, Yang H, Zhu J, Xu X, Dong J, Zhu Y. Genetic Characteristics of Human Enterovirus 71 and Coxsackievirus A16 Circulating from 1999 to 2004 in Shenzhen, People's Republic of China Genetic Characteristics of Human Enterovirus 71 and Coxsackievirus A16 Circulating from 1999 to 2004 in Shenzhen. *Society*. 2005;43:3835–9.
- Bin LQ, Zhang XA, Wo Y, Xu HM, Li XJ, Wang XJ, Ding SJ, Chen XD, He C, Liu LJ, Li H, Yang H, Li TY, Liu W, Cao WC. Circulation of Coxsackievirus A10 and A6 in Hand-Foot-Mouth Disease in China, 2009-2011. *PLoS One*. 2012;7:e52073.
- Gopalkrishna V, Patil PR, Patil GP, Chitambar SD. Circulation of multiple enterovirus serotypes causing hand, foot and mouth disease in India. *J Med Microbiol*. 2012;61:420–5.
- Puenja J, Chieochansin T, Linsuwanon P, Korkong S, Thongkomplew S, Vichaiwattana P, Theamboonlers A, Poovorawan Y. Hand, foot, and mouth disease caused by Coxsackievirus A6, Thailand, 2012. *Emerg Infect Dis*. 2013; 19:641–3.
- Hu Y-Q, Xie G-C, Li D-D, Pang L-L, Xie J, Wang P, Chen Y, Yang J, Cheng W-X, Zhang Q, Jin Y, Duan Z-J. Prevalence of Coxsackievirus A6 and Enterovirus 71 in Hand, Foot, and Mouth Disease in Nanjing, China in 2013. *Pediatr Infect Dis J*. 2015;34:951–7.
- Bian L, Wang Y, Yao X, Mao Q, Xu M, Liang Z. Coxsackievirus A6: a new emerging pathogen causing hand, foot and mouth disease outbreaks worldwide. *Expert Rev Anti Infect Ther*. 2015;13:1061–71.
- Lu J, Zeng H, Zheng H, Yi L, Guo X, Liu L. Hand, foot and mouth disease in Guangdong, China, in 2013: new trends in the continuing epidemic. *Clin Microbiol Infect*. 2014;20:7.
- Van Tan L, Tuyen NTK, Thanh TT, Ngan TT, Van HMT, Sabanathan S, Van TTM, Thanh LTM, Nguyet LA, Geoghegan JL, Ong KC, Perera D, Hang VTT, Ny NTH, Anh NT, Ha DQ, Qui PT, Viet DC, Tuan HM, Wong KT, Holmes EC, Chau NVV, Thwaites G, van Doorn HR. A generic assay for whole-genome amplification and deep sequencing of enterovirus A71. *J Virol Methods*. 2015;215–216:30–6.
- Froussard P. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Res*. 1992;20:2900.
- Endoh D, Mizutani T, Kirisawa R, Maki Y, Saito H, Kon Y, Morikawa S, Hayashi M. Species-independent detection of RNA virus by representational difference analysis using non-ribosomal hexanucleotides for reverse transcription. *Nucleic Acids Res*. 2005;33:1–11.
- Thanh T, Anh N, Tham N, Van H, Sabanathan S, Qui P, Ngan T, Van T, Nguyet L, Ny N, Thanh L, Chai O, Perera D, Viet D, Khanh T, Ha D, Tuan H, Wong K, Hung N, Chau N, Thwaites G, van Doorn H, Van Tan L. 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.
- Allan Nix W, Oberste MS, Pallansch MA. Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. *J Clin Microbiol*. 2006;44:2698–704.
- Van Tan L, van Doorn HR, van der Hoek L, Hien VM, Jebbink MF, Ha DQ, Farrar J, van Vinh Chau N, de Jong MD. Random PCR and ultracentrifugation increases sensitivity and throughput of VIDISCA for screening of pathogens in clinical specimens. *J Infect Dev Ctries*. 2011;5:142–8.
- Rossee T, Van Borm S, Vandenbussche F, Hoffmann B, van den Berg T, Beer M, Höper D. The Origin of Biased Sequence Depth in Sequence-Independent Nucleic Acid Amplification and Optimization for Efficient Massive Parallel Sequencing. *PLoS One*. 2013;8:1–9.
- Karlsson OE, Belák S, Granberg F. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. *Biosecur Bioterror*. 2013;11:S227–34.
- Geoghegan JL, Van Tan L, Kühnert D, Halpin RA, Lin X, Simenauer A, Akopov A, Das SR, Stockwell TB, Shrivastava S, Ngoc NM, Uyen LTT, Tuyen NTK, Thanh TT, Hang VTT, Qui PT, Hung NT, Khanh TH, Think LQ, Nhan LNT, Van HMT, Viet DC, Tuan HM, Viet HL, Hien TT, Chau NVV, Thwaites G,

- Grenfell BT, Stadler T, Wentworth DE, Holmes EC, Van Doorn HR. Phylodynamics of Enterovirus A71-Associated Hand, Foot and Mouth Disease in Viet Nam. *J Virol*. 2015;89:8871–9.
18. Perera D, Yusof MA, Podin Y, Ooi MH, Thao NTT, Wong KK, Zaki A, Chua KB, Malik YA, Tu PV, Tien NTK, Puthavathana P, McMinn PC, Cardoso MJ. Molecular phylogeny of modern coxsackievirus A16. *Arch Virol*. 2007;152:1201–8.
  19. Hu YF, Yang F, Du J, Dong J, Zhang T, Wu ZQ, Xue Y, Jin Q. Complete genome analysis of coxsackievirus A2, A4, A5, and A10 strains isolated from hand, foot, and mouth disease patients in China revealing frequent recombination of human enterovirus A. *J Clin Microbiol*. 2011;49:2426–34.
  20. Liu W, Wu S, Xiong Y, Li T, Wen Z, Yan M, Qin K, Liu Y, Wu J. Co-circulation and genomic recombination of coxsackievirus A16 and enterovirus 71 during a large outbreak of hand, foot, and mouth disease in Central China. *PLoS One*. 2014;9:e96051.
  21. Feng X, Guan W, Guo Y, Yu H, Zhang X, Cheng R, Wang Z, Zhang Z, Zhang J, Li H, Zhuang Y, Zhang H, Lu Z, Li M, Yu H, Bao Y, Hu Y YZ. Genome Sequence of a Novel Recombinant Coxsackievirus A6 Strain. *Genome Announc*. 2015;3:e01347–14.
  22. Djikeng A, Halpin R, Kuznickas R, Depasse J, Feldblyum J, Sengamalay N, Afonso C, Zhang X, Anderson NG, Ghedin E, Spiro DJ. Viral genome sequencing by random priming methods. *BMC Genomics*. 2008;9:5.
  23. Rosseel T, Lambrecht B, Vandenbussche F, van den Berg T, Van Borm S. Identification and complete genome sequencing of paramyxoviruses in mallard ducks (*Anas platyrhynchos*) using random access amplification and next generation sequencing technologies. *Virology*. 2011;8:463.
  24. De Vries M, Deijs M, Canuti M, van Schaik BDC, Faria NR, van de Garde MDB, Jachimowski LCM, Jebbink MF, Jakobs M, Luyf ACM, Coenjaerts FEJ, Claas ECJ, Molenkamp R, Koekkoek SM, Lammens C, Leus F, Goossens H, Ieven M, Baas F, van der Hoek L. A sensitive assay for virus discovery in respiratory clinical samples. *PLoS One*. 2011;6:e16118.
  25. 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.
  26. De Groof A, Guelen L, Deijs M, van der Wal Y, Miyata M, Ng KS, van Grinsven L, Simmelink B, Biermann Y, Grisez L, van Lent J, de Ronde A, Chang SF, Schrier C, van der Hoek L. A Novel Virus Causes Scale Drop Disease in *Lates calcarifer*. *PLoS Pathog*. 2015;11:e1005074.
  27. Li L, Deng X, Mee ET, Collot-Teixeira S, Anderson R, Schepelmann S, Minor PD, Delwart E. Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. *J Virol Methods*. 2014;213C:139–46.
  28. Rosseel T, Ozhelvaci O, Freimanis G, Van Borm S. Evaluation of convenient pretreatment protocols for RNA virus metagenomics in serum and tissue samples. *J Virol Methods*. 2015;222:72–80.
  29. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, Yinda CK, Lavigne R, Maes P, Van Ranst M, Heylen E, Matthijnsens J. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep*. 2015;5:16532.
  30. Raman R, Thomas RG, Weiner MW, Jack CR, Ernstrom K, Aisen PS, Tariot PN, Quinn JF. Actionable Diagnosis of Neuroleptospirosis by Next-Generation Sequencing. *N Engl J Med*. 2014;23:333–6.
  31. Prachayangprecha S, Schapendonk CME, Koopmans MP, Osterhaus ADME, Schurch AC, Pas SD, van der Eijk AA, Poovorawan Y, Haagmans BL, Smits SL. Exploring the Potential of Next-Generation Sequencing in Detection of Respiratory Viruses. *J Clin Microbiol*. 2014;52:3722–30.
  32. Thorburn F, Bennett S, Modha S, Murdoch D, Gunson R, Murcia PR. The use of next generation sequencing in the diagnosis and typing of respiratory infections. *J Clin Virol*. 2015;69:96–100.
  33. Linsuwanon P, Poovorawan Y, Li L, Deng X, Vongpunsawad S, Delwart E. The Fecal Virome of Children with Hand, Foot, and Mouth Disease that Tested PCR Negative for Pathogenic Enteroviruses. *PLoS One*. 2015;10:e0135573.
  34. Aird D, Ross MG, Chen W-S, Danielsson M, Fennell T, Russ C, Jaffe DB, Nusbaum C, Gnirke A. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol*. 2011;12:R18.
  35. Blomström AL, Widén F, Hammer AS, Belák S, Berg M. Detection of a novel astrovirus in brain tissue of mink suffering from shaking mink syndrome by use of viral metagenomics. *J Clin Microbiol*. 2010;48:4392–6.
  36. Allander T, Emerson SU, Engle RE, Purcell RH, Bukh J. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proc Natl Acad Sci U S A*. 2001;98:11609–14.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

