Unbiased metagenomic sequencing complements specific routine diagnostic methods and increases chances to detect rare viral strains

Dagmara W. Lewandowska [a], Osvaldo Zagordi [a,1], Andrea Zbinden [a,1], Macé M. Schuurmans [b,1], Peter Schreiber [c,1], Fabienne-Desirée Geissberger [a], Jon B. Huder [a], Jürg Böni [a], Christian Benden [b], Nicolas J. Mueller [c,*,1], Alexandra Trkola [a,*,1], Michael Huber [a,*,1]

[a] Institute of Medical Virology, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[b] Division of Pulmonary Medicine, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland
[c] Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland

ARTICLE INFO

ABSTRACT

Multiplex PCR assays for respiratory viruses are widely used in routine diagnostics, as they are highly sensitive, rapid, and cost effective. However, depending on the assay system, cross-reactivity between viruses that share a high sequence homology as well as detection of rare virus isolates with sequence variations can be problematic. Virus sequence-independent metagenomic high-throughput sequencing allows for accurate detection of all virus species in a given sample, as we demonstrate here for human Enterovirus and Rhinovirus in a lung transplant patient. While early in infection a commercial PCR assay recorded Rhinovirus, high-throughput sequencing correctly identified human Enterovirus C104 as the source of infection, highlighting the potential of the technology and the benefit of applying open assay formats in complex diagnostic situations.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The clinical outcome and associated risks of viral infections can differ substantially between immunocompetent and immunocompromised individuals. In the latter, viral infections including those that commonly are only mildly pathogenic or cause self-limited diseases can lead to protracted infections with chronic viral shedding, to disseminated disease with infection of organs rarely affected in immunocompetent hosts and to severe disease courses with increased mortality (Quan et al., 2010; Wunderli et al., 2011). For instance, human Enterovirus (HEV) and human Rhinovirus (HRV) cause asymptomatic or mild infections in immunocompetent individuals (common cold and rash), while causing more severe diseases in immunocompromised individuals (e.g., acute hemorrhagic conjunctivitis, aseptic meningitis, encephalitis, acute flaccid paralysis, and acute febrile illness in young children) (Tapparel et al., 2013). Although often no specific treatment for virus infections is available, accurate discrimination of viral pathogens is nevertheless important at least to limit nosocomial spread.

To date, sequence-specific PCR is the most common approach for direct virus identification and allows precise quantification of noncultivatable or low replicating viruses. A drawback of the method is that it requires prior knowledge of the viral sequence and specific assays need to be designed for individual viruses. Using broadly reactive PCR detection systems for the genus Enterovirus, which comprises HEV and HRV, accurate discrimination of HEV and HRV species can be challenging as these viruses share a high homology and identical genome organization (Tapparel et al., 2009a). A further significant limitation of sequence-specific molecular methods is the detection of newly arising virus strains, as their sequences may deviate from known members and may not be detected accurately by the applied molecular probes.

Metagenomic-based methods for the identification of viruses comprise comprehensive sequencing of random nucleic acid fragments from all genomes present in an environmental or patient sample (Delwart, 2007). Several studies have used this technology in recent years to explore the breadth of the virome in diverse biological samples including human and animal feces (Breitbart et al., 2003; Cann et al., 2005; Finkbeiner et al., 2008, 2009; Kapoor et al., 2009; Li et al., 2010; Victoria et al., 2009), blood (Breitbart & Rohwer, 2005; Jones et al., 2005), tissue (Briese et al., 2009; Kistler et al., 2008; Palacios et al., 2010; Victoria et al., 2008), or respiratory tract secretions (Allander et al., 2005, 2007; Willner et al., 2009) and highlighted the validity of the approach to detect rare and novel viruses (Bibby, 2013; Fancello et al., 2012; Relman, 2013).

In this study, we demonstrate the limitations of a common, commercially available multiplex assay system in detecting and discriminating

between HEV and HRV infection. Following a sample early in infection that was recorded HRV positive, the subsequent samples collected during the course of infection proved highly positive for HRV and additionally also borderline for HEV, suggesting either a dual infection with both viruses or a HEV infection that was not properly detected by the multiplex PCR assay and cross-reacted with HRV. Therefore, to determine the actual status of infection, we performed whole nucleic acid high-throughput sequencing of throat swabs and a stool sample from several time points, which revealed an HEV infection and clearly ruled out HRV.

2. Material and methods

2.1. Ethics statement

Noninvasive samples were obtained from a patient in the frame of the Viral Metagenome Study of the Clinical Research Priority Program "Viral Infectious Diseases" of the University of Zurich. The ethics committee of the canton of Zurich approved the study, and written informed consent was obtained.

2.2. Nucleic acid extraction for routine diagnostics

Two hundred microliters of throat swab (in 3-mL virus transport medium) stored at −20 °C was used for nucleic acid extraction with the automated NucliSENS EasyMAG system (BioMérieux, Craponne, France) and eluted in 110 μL of elution buffer.

2.3. FTD Respiratory pathogens 21 multiplex PCR assay

Respiratory viruses were tested by routine diagnostic procedures using the multiplex PCR FTD Respiratory pathogens 21 (FTD21) assay (Fast-track Diagnostics, Junglinster, Luxembourg) according to the manufacturer's protocol on a ViiA7 Real-Time PCR System (Life Technologies Europe, Zug, Switzerland), which detects 18 different respiratory viruses and virus types: influenza A, influenza A H1N1 pdm 09, and influenza B viruses; coronaviruses NL63, 229E, OC43, and HKU1; parainfluenzavirus 1, 2, 3, and 4; and human metapneumovirus A/B, *Rhinovirus*, respiratory syncytial viruses A/B, adenovirus, *Enterovirus*, parechovirus, and bocavirus. The FTD21 assay is IVD and CE marked (www.fast-trackdiagnostics.com/products/ftd-respiratory-pathogens-21/).

2.4. In-house HEV PCR

Extracted nucleic acids (10 μL) were used for *Enterovirus*-specific PCR (Tapparel et al., 2009b). Primers and probes were synthesized by Microsynth AG (Balgach, Switzerland). Primer concentrations were 1 μmol/L each, and probe concentrations was 0.4 μmol/L in a reaction volume of 50 μL. Amplification and detection were performed with the QuantiTect Probe RT-PCR Kit (Qiagen AG, Hombrechtikon, Switzerland) on an ABI 7300 (Applied Biosystems/Life Technologies, Zug, Switzerland) with cycling conditions 30 min 48 °C, 10 min 95 °C followed by 50 cycles 15 s at 95 °C, 1 min at 60 °C. The assay was validated using external Quality Control for Molecular Diagnostics controls.

2.5. Virus enrichment and nucleic acid extraction for metagenomic sequencing

Of the throat swabs samples (in 3-mL virus transport medium), 300–1000 μL was centrifuged (1200 rpm for 5 min, Heraeus Multifuge X3 R; Thermo Fisher Scientific) and filtered (0.45 μm, TPP, Trasadingen, Switzerland). Stool samples were suspended in phosphate buffered saline (1:6), and viral particles were enriched by centrifugation (3000 rpm for 30 min) and filtration (0.45 μm). Total nucleic acids were extracted from 1000-μL enriched material using the automated

NucliSENS EasyMAG system (BioMérieux) and eluted in 25 μL of elution buffer.

2.6. Metagenomic sequencing

Anchored, random PCR amplification of total nucleic acids was performed for all samples in duplicates or quadruplicates. Complementary DNA (cDNA) was generated by reverse transcription with a primer containing a random octamer linked to an anchor sequence ATCGTCGTCGTAGGCTGCTCNNNNNNNNN (Froussard, 1992; Victoria et al., 2008; Wang et al., 2003). In particular, 5 μL of extraction eluate was used as a template in a total volume of 20 μL, together with 4 μmol/L of random primer, 1 mmol/L of deoxyribose triphosphates, 1× first-strand buffer, 20 mmol/L dithiothreitol and 20 U/μL of SuperScript III (Invitrogen/Life Technologies, Zug, Switzerland). The template and random primers were heated at 65 °C for 5 min, before added to the rest of the reaction components. Reverse transcription was performed at 42 °C for 60 min followed by inactivation at 96 ° C for 5 min. Before second-strand synthesis, cDNA was denatured at 94 °C for 2 min and cooled down to 10 ° C for 5 min. The second strand for both RNA and DNA templates was synthesized with 0.39 U/μL T7 DNA polymerase (New England Biolabs, Ipswich, MA, USA) while ramping up to 37 °C at 3.37 °C/min for 8 min followed by a denaturation step at 94 °C for 2 min and cooling at 10 °C for 5 min. Additional T7 polymerase was added (0.39 U/μL), and the synthesizing step was repeated. Subsequently, DNA was amplified with primers to the anchor sequences only using AmpliTaq Gold (Applied Biosystems/Life Technologies), as previously described (Wang et al., 2003). Briefly, the PCR was performed in a total volume of 50 μL with the anchor-specific primer (1 μmol/L), dNTPs (0.25 mmol/L each), AmpliTaq Gold buffer, AmpliTaq Gold enzyme (0.05 U/μL), and the 3 μL of cDNA/DNA input from the previous step. The temperature protocol involved initial denaturation at 94 °C for 15 min, 40 cycles of 94 °C 30 s, 40 °C 30 s, 50 °C 30 s, and 72 °C 1 min and final extension at 72 °C for 5 min. The quality and size of the generated library were checked by capillary-based gel electrophoresis (Fragment Analyzer; Advanced Analytical, Ames, IA, USA). DNA was quantified with PicoGreen (Invitrogen/Life Technologies) and diluted to 0.2 ng/μL as required for the NexteraXT protocol (Illumina, San Diego, CA, USA). Individual samples were dual indexed during the library preparation and pooled for sequencing. Libraries were sequenced on a MiSeq (Illumina) for 1 × 150 cycles with version 3 reagents and the "FASTQ only" workflow. Samples were demultiplexed using MiSeq Reporter version 2.4.60. Raw reads of biological samples were pooled together from 2 different runs for analysis. The raw sequencing data have been uploaded to Zenodo (doi:10.5281/zenodo.14611).

2.7. Bioinformatic data analysis

Sequencing reads were analyzed with a dedicated bioinformatics pipeline (https://github.com/ozagordi/VirMet). In the filtering step, reads are trimmed by removing low-quality bases and removed if shorter than 75 bp, if their average PHRED score is below 20 or if they have low entropy (i.e., they consist mainly of repeats). Reads passing quality filters are cleaned from nonviral reads by aligning with STAR (Dobin et al., 2013) against the following genomes in this order: human, bacterial, bovine, canine. The latter 2 genomes are included to eliminate contaminating reads originating from fetal calf serum and from Madin-Darby canine kidney epithelial cells. Reads not matching any of the above genomes are blasted (Camacho et al., 2009) against a viral database that contains approximately 35,000 different complete viral genomes. This database was created from sequences downloaded from GenBank nuccore that 1) have the expression "complete genome" in the title, 2) are of viral origin (organism taxonomy id is child of id 10239, denoting all viruses), and 3) are not of cellular origin (excluding all organisms children of id 131567, denoting cellular organism) in order to avoid chimeric sequences. For each sequencing read that

passed the filter step, we report the blast hit with minimum e-value, provided that coverage and identity of the hit are higher than 75%. Reads matching neither genome nor any viral sequences were reported as of unknown origin.

2.8. Full-genome reconstruction

Viral reads reported by the pipeline were used for constructing a consensus. Reads were aligned to HEV-C104 strain AK11 genome (GenBank accession number AB686524.1) with SMALT (version 0.7.6, word length 7, step size 2, identity threshold 0.80, coverage threshold 0.80, www.sanger.ac.uk/resources/software/smalt). *FreeBayes* (Garrison & Marth, 2012) was used to detect variants with respect to the AK11 strain (–pooled-continuous option was used). Variants were then filtered to exclude artifacts due to low quality or strand bias, and for positions where a minor variant was detected at a frequency greater than 20%, the corresponding ambiguous code was inserted.

2.9. 5′ rapid amplification of cDNA ends

Rapid amplification of cDNA ends (RACE) kit version 2.0 (Invitrogen) was used according to the manufacturer's instructions to amplify the missing 5′-region. Ten microliters of eluate as extracted for metagenomic analysis (described above) was used. The gene-specific primer GAGCAATGGCAACAACTG was used for reverse transcription. The tailed product was amplified with primer AAP and the specific primer GAAACACGGACACCCAAAGTAGT. A PCR product of the expected size (about 600 bp) was cloned into the pcDNA 3.1/V5-His TOPO TA Expression vector (Invitrogen/Life Technologies), and DNA from minipreps was sequenced on an ABI 3130 XL Genetic Analyzer (Applied Biosystems/Life Technologies).

2.10. Phylogenetic analysis

Phylogenetic analysis of HEV isolates was performed for capsid protein VP1 (corresponding to nucleotide positions 2461-3348 in HEV-C104 strain AK11 [AB686524.1], which was used as a reference). HEV-C104 strains Pavia259-7712 (JX982253.1) and CL-12310945 (EU840733.1) were used to illustrate relationship of strains within the HEV-C104 cluster. HEV-C117 strain LIT22 (JX262382.1) was used as an out-group. Alignment of sequences and calculations of divergence and percent identity were performed with MegAlign 10.1.2 (DNASTAR, Madison, WI, USA) using the Clustal W method with default settings.

2.11. Sequencing of FTD21 amplicons

One hundred nanograms cleaned-up PCR amplicons of FTD21 reaction RsEPA (which includes HEV) of both internal control and patient kvv585 were diluted into 60 μL of Resuspension Buffer (Illumina). Sequencing libraries were prepared with the TruSeq Nano DNA LS Preparation kit (Illumina) according to the manufacturer's instructions, except that the workflow was started with end repair and that size selection for both small and large fragments was omitted. Instead of size selection, the end repair mix was purified (NucleoSpin Gel and PCR Clean-up kit; Macherey-Nagel, Düren, Germany) and eluted in 20-μL resuspension buffer. Libraries were sequenced on an Illumina MiSeq system for 2×75 cycles with v3 reagents. Samples were demultiplexed using MiSeq Reporter 2.4.60. Reads were aligned with SMALT (version 0.7.5, word length 7, step size 2, identity and coverage threshold 0.75, www.sanger.ac.uk/resources/software/smalt/) to reference strain HEV-C104 AK11 (AB686524.1), and a consensus was generated.

3. Results

3.1. Performance of multiplex assay and in-house HEV PCR

A 51-year-old man enrolled in the Viral Metagenome Study at the University of Zurich (patient kvv585) underwent bilateral lung transplantation due to end-stage cystic fibrosis-related obstructive and suppurative lung disease. Bacterial colonization (detection of pathogen, but not causative for symptoms) with *Pseudomonas aeruginosa* was known prior to transplantation and was documented also after transplantation. Three and a half months after transplantation, the patient presented with a deterioration of lung function and signs of inflammation, compatible with a viral infection.

Serial throat swab specimens of the patient were analyzed for presence of respiratory viruses by multiplex PCR using the commercial FTD21 kit. At week 15 posttransplantation, a positive result was recorded for HRV. However, 1 week later, the multiplex assay recorded positive results for both HRV and HEV (Fig. 1). No other respiratory viruses were detected in the throat swabs by the FTD21 kit, nor was cytomegalovirus detected in the patient's blood. No fungi or bacteria other than *P. aeruginosa* were detected in clinical specimens.

Since HRV is known to cross-react with HEV in this assay, when both HRV and HEV are detected, the patient is reported only *Enterovirus* positive according to the assay's interpretation rules. The results obtained at consecutive time points were inconsistent, and a dual infection with HEV and HRV or an infection with HEV that lead to cross-reaction with HRV was suspected. Accurate definition of which virus caused the infection was of interest as the patient developed a protracted significant lung function loss.

To verify the results, the samples were analyzed by an HEV specific in-house PCR detecting the 5′ untranslated region (Tapparel et al., 2009b), which gave a clear positive HEV result (Fig. 1). There was a striking difference in the cycle threshold (Ct) values of the multiplex (Ct = 33.0) and the HEV-specific PCR (Ct = 24.0), which would translate to an approximately 5×10^5 higher copy number. HEV load detected by the multiplex assay stayed low in the following weeks, and HEV was not detected after week 20. In contrast, the signal reported for HEV by the in-house test in respective time points was consistently much higher than in the multiplex assay (Fig. 1). Noteworthy, during the course of infection, detection levels of HEV and HRV by the FTD21 kit followed the same pattern: an increase in copy number between weeks 16 and 19 followed by a distinct drop in copy number between weeks 19 and 20 (Fig. 1).
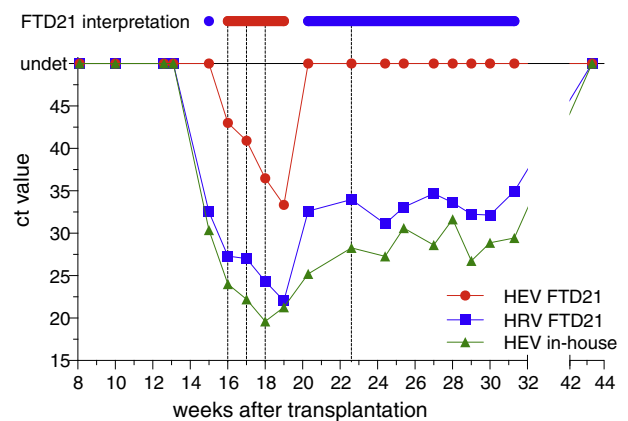
Fig. 1. Comparison of FTD21 and in-house HEV PCR shows higher sensitivity by the in-house PCR. Viral load (Ct values) for HEV (red) and HRV (blue) obtained with the commercial multiplex real-time PCR FTD21 assay and results obtained with an in-house HEV PCR (green) are shown over time. Vertical lines represent time points that were analyzed by high-throughput sequencing. The result interpretation according to the rules of the FTD21 kit is shown on top (red bar = HEV positive; blue bar = HRV positive).

3.2. Detection of Enterovirus by metagenomic sequencing

To confirm the status of the infection, we performed whole nucleic acid high-throughput sequencing of throat swabs and a stool sample from several time points. The steps involved in sample processing were virus enrichment, extraction of total RNA and DNA, amplification of nucleic acids, and library preparation. Sequencing on Illumina MiSeq platform in several runs produced between 2 and 9.5 million reads per sample. Our bioinformatic analysis reported very high number of reads matching to HEV-C104 strains (~99% of viral reads) in the swab samples (Table 1). Significantly lower numbers of HEV reads were obtained from the stool sample. Of note, no HRV reads were found in any of the samples. In addition, other viruses most notably Karolinska Institute polyomaviruses (KIPyV), Torque teno virus (TTV), human endogenous retroviruses (HERV), pepino mosaic virus, and bacteriophages were identified in some samples (Table 1). The detection of HERV is most likely due to an incomplete filtering of reads of human origin.

3.3. Full-length genome reconstruction

To obtain the full-length genome, reads reported by the pipeline as viral were merged within biological samples per time point and aligned to HEV-C104 AK11 with SMALT. We were able to construct near full-length genomes of throat swab samples from weeks 16, 17, and 18 and the nasal swab sample from week 22 (4 out of 6 samples), which had a high coverage and uniform distribution of reads throughout the whole genome. Compared to HEV-C104 AK11, around 20 nucleotides were missing at the 5′-end, and those were retrieved by 5′-RACE for the throat swab sample of week 16. Surprisingly, the 3′-end was covered until the end of the reference genome. The sequence of the derived HEV-C isolate was deposited in GenBank (accession number KR815824). In contrast to the swab samples, viral reads obtained from a stool sample were insufficient for full-length genome reconstruction. Phylogenetic analysis showed that the patient's isolates segregate into a distinct cluster within the HEV-C104 group with the closest identity to HEV-C104 AK11 (Fig. 2).

3.4. Prediction of FTD21 primer binding sites

The Ct values for HEV observed with the commercial multiplex FTD21 and the in-house PCR assays differed significantly and indicated a low analytical sensitivity of the FTD21 kit for this patients virus isolate. As details on primer and probe sequences of the FTD21 kit are not publicly available, we sequenced amplicons obtained from positive reactions of the FTD21 kit to define the cause for the lack in detection. The alignment of the obtained multiplex PCR amplicon with the reference genome of HEV-C104 AK11 and HEV-A, B, C, and D consensus sequences (Bailey & Tapprich, 2006) revealed several differences in the nucleotide sequences. The sequences of the retrieved amplicon of the FTD21 HEV

reaction suggest that the kit utilizes a forward primer that extends to position 433 and a probe up to position 488, where the sequences in HEV-C104 strains differ at several locations compared to the HEV consensus (Fig. 3). This would explain the decreased sensitivity of the FTD21 assay in detecting HEV-C104 as a result of binding sites of the forward primer and the probe outside conserved regions.

4. Discussion

In this study, we demonstrate the potential of metagenomic sequencing in complementing clinical routine diagnostics by successfully resolving an ambiguous case of an *Enterovirus* infection that could not be resolved by a commercial multiplex PCR assay.

Our metagenomic analysis reported high numbers of HEV reads across longitudinal samples (Table 1). The lower amount of viral reads identified during the later time points of the study might be a consequence of the onset of virus clearing as HEV levels recorded by PCR dropped from week 19 onwards (Fig. 1). As the patient suffered mainly from respiratory symptoms, a spread of HEV in the gut was not suspected. Consistently, only a comparatively low number of HEV reads were recorded in the stool sample. Importantly, for most of the samples analyzed, the number of viral reads was sufficient to reconstruct a full-length genome. Phylogenetic analysis of capsid protein VP1 showed that the isolates from different time points in this patient segregated into a distinct cluster with the closest identity to HEV-C104 AK11 (Fig. 2).

It is known that HRV is restricted to the respiratory tract, whereas HEV, besides infecting the respiratory tract both in adults and children causing upper and lower respiratory tract diseases (Kaida et al., 2012; Piralla et al., 2013; Victoria et al., 2009; Xiang et al., 2013), may also infect various organs other than the gastrointestinal tract. Two distinct groups within the HEV-C serotype have been described causing either respiratory or gastrointestinal disease. Notably, HEV-C104 is classified within the group associated with respiratory disease and has been detected in nasal or nasopharyngeal samples exclusively, but not in stool, which is consistent with our results (Tokarz et al., 2013).

The metagenomic sequencing approach that we applied here clearly ruled out HRV as the source of the high positive signal initially reported by the commercial multiplex PCR test, as no HRV sequences were found in any of the samples. Other viruses reported by our metagenomic approach and of potential clinical relevance were KIPyV and TTV. KIPyV is frequently identified in patients with respiratory symptoms, especially in immunocompromised individuals (Mourez et al., 2009). However, a clear link between detection of KIPyV and respiratory symptoms is yet to be established (Bialasiewicz et al., 2007; Hormozdi et al., 2010). TTV viremia has been suggested to correlate with the intensity of immunosuppression, as several recent studies reported a higher level of Anelloviridae (with a predominance of TTV) in transplant recipients with more intense immunosuppression (Burra et al., 2008; De Vlaminck et al., 2013; Focosi et al., 2014; Görzer et al., 2014). Our analysis also

Table 1
Metagenomic sequencing identified HEV-C104 in each tested sample.

	Throat week 16	Throat week 17	Throat week 18	Throat week 22	Nasal week 22	Stool week 22
Total reads	9,560,297	9,292,252	6,183,486	5,778,571	2,315,153	2,032,088
Virus reads	3,609,488	5,006,695	3,225,654	768,747	1,088,595	123,937
Virus reads (%)	37.8	53.9	52.2	13.3	47.0	6.1
HEV-C104 reads	3,576,982	4,998,690	3,219,925	768,092	1,087,656	181
HEV-C104 within virus reads (%)	99.1	99.8	99.8	99.9	99.9	0.1
TTV reads	28,125	16	774	0	495	5
KIPyV reads	5	6054	2002	1	9	0
HERV[a] reads	215	16	65	0	28	0
Bacteriophage reads	0	0	9	584[b]	22[c]	38[b]
Pepino mosaic virus reads	0	0	0	55	71	123,688

[a] Strains K113, HCML-ARV, and K115.
[b] *Enterococcus* phages, *Lactobacillus* phages, and *Staphylococcus* phages.
[c] *Staphylococcus* phages.
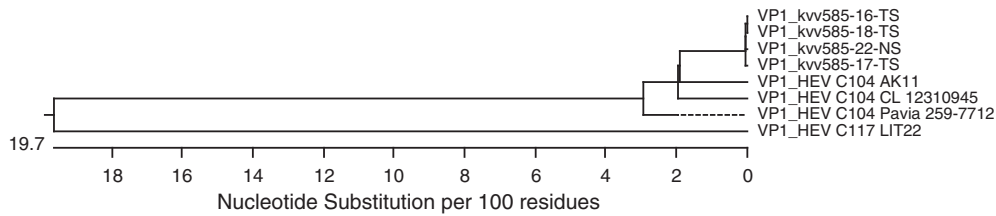
Fig. 2. Phylogenetic analysis reveals that the patient's isolates cluster within HEV-C104 strains. Phylogenetic analysis of the VP1 gene of 4 isolates from patient kvv585 and selected HEV-C104 strains. Alignment of strains and bootstrapping was performed with MegAlign (clustal W). Isolates detected in this study contain the code of the patient, the weeks after transplantation, and the sample type. TS = throat swab; NS = nasal swab; ST = stool.

reported the presence of some plant viruses such as pepino mosaic virus, which may be due to food uptake, and bacteriophage sequences, which are expected to occur in abundance due to the presence of commensal bacterial species (Table 1).

The interpretation of the multiplex PCR results first suggested an HRV infection, as the HEV signal was below the threshold of detection and the HRV signal was clearly positive. Due to cross-reactivity issues of this assay, however, if both HEV and HRV are positive, samples are recorded *Enterovirus* positive. Most detection systems described for enteroviruses target the 5′ noncoding part of the genome. High sequence homology in this region between HRV and HEV is responsible for the observed cross-reactivity within multiplex assays (Brown et al., 2009; van Doornum et al., 2007). In addition, for HEV-C104, it was previously

reported that its genome could neither be amplified with standard *Enterovirus*-specific PCR nor with VP1-specific primers that are routinely used to type enteroviruses (Tapparel et al., 2009b). Alignment of the predicted amplicon sequences of the FTD21 kit identified several nucleotide differences compared to HEV-C104 strains. Although the exact sequences of primers and probe(s) used in the FTD21 kit are not available, these mutations are the most likely explanation for the observed low efficiency of amplification of the patient's HEV-C104 strain (Fig. 3). Of course, in a routine diagnostic setting, it is not feasible even with multiplex assays to have type-specific primers and probes available for the detection of all rare virus isolate that might be clinically relevant (Whiley et al., 2008). This is where metagenomic diagnostics step in as it is more comprehensive in detecting virus isolates with sequence
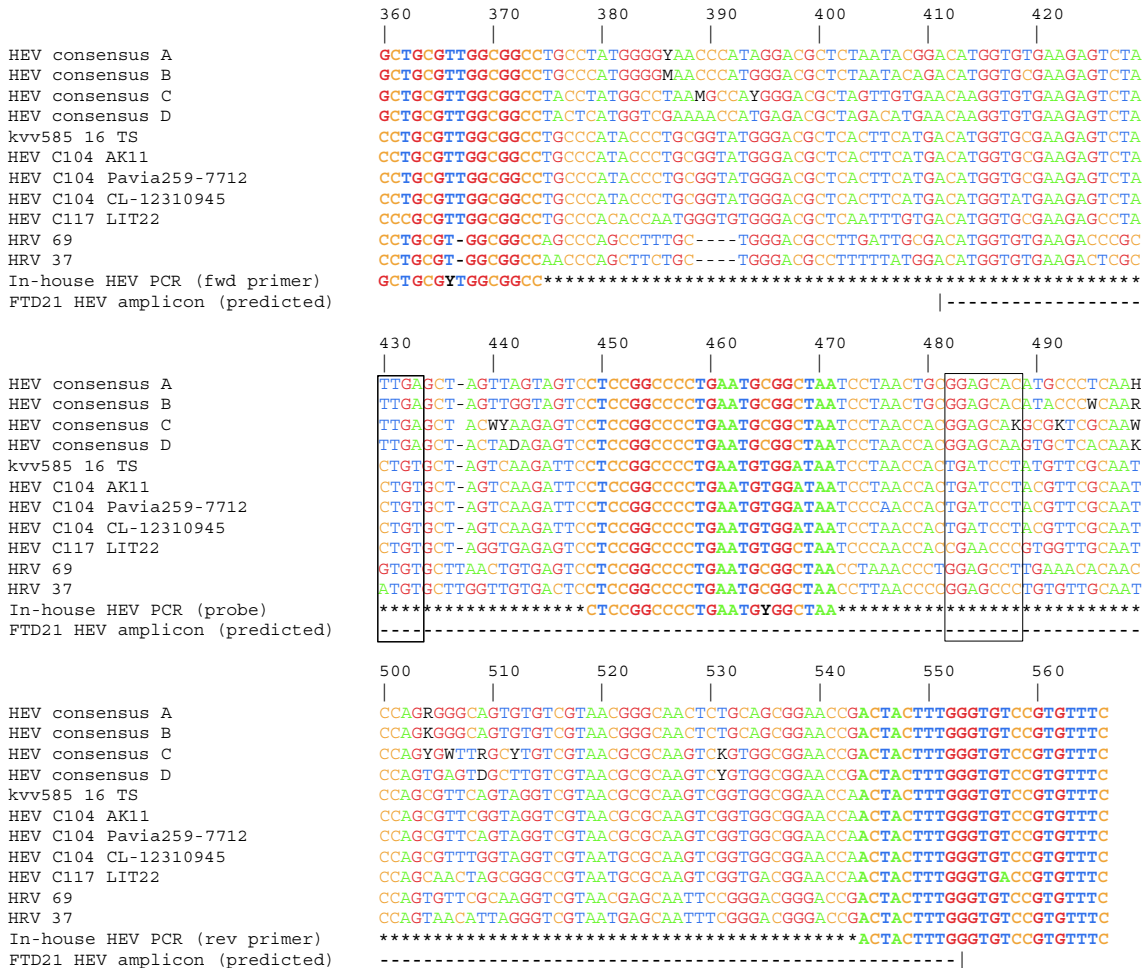
Fig. 3. Alignment of HEV-C104 sequences reveal several nucleotide differences in the region of the putative FTD21 HEV amplicon. The week 16 throat swab sequence (kvv585-16-TS) was aligned to other strains of HEV-C104; enteroviruses group A, B, C, and D consensus sequences; and 2 HRV strains in the region targeted by the in-house PCR system (nucleotide positions 360-556 in HEV-C104 AK11, represented by *). A dashed line denotes the position of the predicted amplicon of the FTD21 kit. Regions where forward primer, reverse primer, and the probe from the in-house HEV PCR system anneal are color coded (regions 360-374, 544-566, 449-471, respectively) as well as putative regions of mismatches responsible for impaired detection by the FTD21 kit (regions 430-433 and 482-488).

variation that otherwise could be insufficiently detected or even missed by specific routine tests. As our analysis highlights, it is, therefore, important to include periodically broader detection methods to screen for novel or emerging virus strains and to verify the specificity of the routine (multiplex) assays.

In summary, the metagenomic approach used in our study proved to be successful in both facilitating the HEV-C104 diagnosis and clarifying potential viral coinfections. In contrast to routine diagnostics, which is only able to detect viruses for which diagnostic tools are established, metagenomic analysis offers unprecedented possibilities for characterization, diagnostics, and management of rare and novel viral infections.

Acknowledgments

References

Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B, et al. Cloning of a human parvovirus by molecular screening of respiratory tract samples. PNAS 2005;102:12891–6.

Allander T, Andreasson K, Gupta S, Bjerkner A, Bogdanovic G, Persson MAA, et al. Identification of a third human polyomavirus. J Virol 2007;81:4130–6.

Bailey JM, Tapprich WE. Structure of the 5′ nontranslated region of the coxsackievirus B3 genome: chemical modification and comparative sequence analysis. J Virol 2006;81:650–68.

Bialasiewicz S, Whiley DM, Lambert SB, Wang D, Nissen MD, Sloots TP, et al. A newly reported human polyomavirus, KI virus, is present in the respiratory tract of Australian children. J Clin Virol 2007;40:15–8.

Bibby K. Metagenomic identification of viral pathogens. Trends Biotechnol 2013;31:275–9.

Breitbart M, Rohwer F. Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. Biotech 2005;39:729–36.

Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, et al. Metagenomic analyses of an uncultured viral community from human feces. J Bacteriol 2003;185:6220–3.

Briese T, Paweska J, McMullan L, Hutchison S, Street C, Palacios G, et al. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. PLos Pathog 2009;5:e1000455.

Brown BA, Maher K, Flemister MR, Naraghi-Arani P, Uddin M, Oberste MS, et al. Resolving ambiguities in genetic typing of human enterovirus species C clinical isolates and identification of enterovirus 96, 99 and 102. J Gen Virol 2009;90:1713–23.

Burra P, Masier A, Boldrin C, Calistri A, Andreoli E, Senzolo M, et al. Torque teno virus: any pathological role in liver transplanted patients? Transpl Int 2008;21:972–9.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Biochem 2009;10:421.

Cann AJ, Fandrich SE, Heaphy S. Analysis of the virus population present in equine faeces indicates the presence of hundreds of uncharacterized virus genomes. Virus Genes 2005;30:151–6.

De Vlaminck I, Khush KK, Strehl C, Kohli B, Luikart H, Neff NF, et al. Temporal response of the human virome to immunosuppression and antiviral therapy. Cell 2013;155:1178–87.

Delwart EL. Viral metagenomics. Rev Med Virol 2007;17:115–31.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.

Fancello L, Raoult D, Desnues C. Computational tools for viral metagenomics and their application in clinical research. Virology 2012;434:162–74.

Finkbeiner SR, Allred AF, Tarr PI, Klein EJ, Kirkwood CD, Wang D, et al. Metagenomic analysis of human diarrhea: viral detection and discovery. PLos Pathog 2008;4:e1000011.

Finkbeiner SR, Holtz LR, Jiang Y, Rajendran P, Franz CJ, Zhao G, et al. Human stool contains a previously unrecognized diversity of novel astroviruses. Virol J 2009;6:1–5.

Focosi D, Macera L, Pistello M, Maggi F. Torque Teno virus viremia correlates with intensity of maintenance immunosuppression in adult orthotopic liver transplant. J Infect Dis 2014;210:667–8.

Froussard P. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. Nucleic Acids Res 1992;20. (2900–2900).

Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing; 2012 (arXiv).

Görzer I, Haloschan M, Jaksch P, Klepetko W, Puchhammer-Stöckl E. Plasma DNA levels of Torque teno virus and immunosuppression after lung transplantation. J Heart Lung Transplant 2014;33:320–3.

Hormozdi DJ, Arens MQ, Le B-M, Buller RS, Agapov E, Storch GA. KI polyomavirus detected in respiratory tract specimens from patients in St. Louis, Missouri. Pediatr Infect Dis J 2010;29:329–33.

Jones MS, Kapoor A, Lukashov VV, Simmonds P, Hecht F, Delwart EL. New DNA viruses identified in patients with acute viral infection syndrome. J Virol 2005;79:8230–6.

Kaida A, Kubo H, Sekiguchi J-i, Hase A, Iritani N. Enterovirus 104 infection in adult, Japan, 2011. Emerg Infect Dis 2012;18:882–3.

Kapoor A, Slikas E, Simmonds P, Chieochansin T, Naeem A, Shaukat S, et al. A newly identified bocavirus species in human stool. J Infect Dis 2009;199:196–200.

Kistler AL, Gancz A, Clubb S, Skewes-Cox P, Fischer K, Sorber K, et al. Recovery of divergent avian bornaviruses from cases of proventricular dilatation disease: identification of a candidate etiologic agent. Virol J 2008;5:88.

Li L, Victoria JG, Wang C, Jones M, Fellers GM, Kunz TH, et al. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. J Virol 2010;84:6955–65.

Mourez T, Bergeron A, Ribaud P, Scieux C, de Latour RP, Tazi A, et al. Polyomaviruses KI and WU in immunocompromised patients with respiratory disease. Emerg Infect Dis 2009;15:107–9.

Palacios G, Lovoll M, Tengs T, Hornig M, Hutchison S, Hui J, et al. Heart and skeletal muscle inflammation of farmed salmon is associated with infection with a novel reovirus. PLoS One 2010;5:e11487.

Piralla A, Fiorina L, Daleno C, Esposito S, Baldanti F. Complete genome characterization of enterovirus 104 circulating in Northern Italy shows recombinant origin of the P3 region. Infect Genet Evol 2013;20:111–7.

Quan PL, Wagner TA, Briese T, Torgerson TR, Hornig M, Tashmukhamedova A, et al. Astrovirus encephalitis in boy with X-linked agammaglobulinemia. Emerg Infect Dis 2010;16:918–25.

Relman DA. Metagenomics, infectious disease diagnostics, and outbreak investigations: sequence first, ask questions later? JAMA 2013;309:1531–2.

Tapparel C, Junier T, Gerlach D, Van-Belle S, Turin L, Cordey S, et al. New respiratory enterovirus and recombinant rhinoviruses among circulating picornaviruses. Emerg Infect Dis 2009a;15:719–26.

Tapparel C, Cordey S, Van Belle S, Turin L, Lee WM, Regamey N, et al. New molecular detection tools adapted to emerging rhinoviruses and enteroviruses. J Clin Microbiol 2009b;47:1742–9.

Tapparel C, Siegrist F, Petty TJ, Kaiser L. Picornavirus and enterovirus diversity with associated human diseases. Infect Genet Evol 2013:282–93.

Tokarz R, Haq S, Sameroff S, Howie SR, Lipkin WI. Genomic analysis of coxsackieviruses A1, A19, A22, enteroviruses 113 and 104: viruses representing two clades with distinct tropism within enterovirus C. J Gen Virol 2013;94:1995–2004.

van Doornum GJJ, Schutten M, Voermans J, Guldemeester GJJ, Niesters HGM. Development and implementation of real-time nucleic acid amplification for the detection of enterovirus infections in comparison to rapid culture of various clinical specimens. J Med Virol 2007;79:1868–76.

Victoria JG, Kapoor A, Dupuis K, Schnurr DP, Delwart EL. Rapid identification of known and new RNA viruses from animal tissues. PLos Pathog 2008;4:e1000163.

Victoria JG, Kapoor A, Li L, Blinkova O, Slikas B, Wang C, et al. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. J Virol 2009;83:4642–51.

Wang D, Urisman A, Liu Y-T, Springer M, Ksiazek TG, Erdman DD, et al. Viral discovery and sequence recovery using DNA microarrays. PLoS Biol 2003;1:E2.

Whiley DM, Lambert SB, Bialasiewicz S, Goire N, Nissen MD, Sloots TP, et al. False-negative results in nucleic acid amplification tests–do we need to routinely use two genetic targets in all assays to overcome problems caused by sequence variation? Crit Rev Microbiol 2008;34:71–6.

Willner D, Furlan M, Haynes M, Schmieder R, Angly FE, Silva J, et al. Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. PLoS One 2009;4:e7370.

Wunderli W, Meerbach A, Gungor T, Berger C, Greiner O, Caduff R, et al. Astrovirus infection in hospitalized infants with severe combined immunodeficiency after allogeneic hematopoietic stem cell transplantation. PLoS One 2011;6:e27483.

Xiang Z, Xie Z, Wang Z, Ren L, Xiao Y, Li L, et al. Human enterovirus genotype C104, China. Emerg Infect Dis 2013;19:689–91.