Two Years of Viral Metagenomics in a Tertiary Diagnostics Unit: Evaluation of the First 105 Cases

Verena Kufner [1,†] (ORCID), **Andreas Plate** [2,†], **Stefan Schmutz** [1] (ORCID), **Dominique L. Braun** [1,2], **Huldrych F. Günthard** [1,2] (ORCID), **Riccarda Capaul** [1], **Andrea Zbinden** [1], **Nicolas J. Mueller** [2,*] (ORCID), **Alexandra Trkola** [1,*] and **Michael Huber** [1,*] (ORCID)

1   Institute of Medical Virology, University of Zurich, 8057 Zurich, Switzerland
2   Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, 8091 Zurich, Switzerland
*   Correspondence: nicolas.mueller@usz.ch (N.J.M.); trkola.alexandra@virology.uzh.ch (A.T.); huber.michael@virology.uzh.ch (M.H.); Tel.: +41-446-342-639 (M.H.)
†   These authors contributed equally to this work.

Abstract: Metagenomic next-generation sequencing (mNGS) can capture the full spectrum of viral pathogens in a specimen and has the potential to become an all-in-one solution for virus diagnostics. To date, clinical application is still in an early phase and limitations remain. Here, we evaluated the impact of viral mNGS for cases analyzed over two years in a tertiary diagnostics unit. High throughput mNGS was performed upon request by the treating clinician in cases where the etiology of infection remained unknown or the initial differential diagnosis was very broad. The results were compared to conventional routine testing regarding outcome and workload. In total, 163 specimens from 105 patients were sequenced. The main sample types were cerebrospinal fluid (34%), blood (33%) and throat swabs (10%). In the majority of the cases, viral encephalitis/meningitis or respiratory infection was suspected. In parallel, conventional virus diagnostic tests were performed (mean 18.5 individually probed targets/patients). mNGS detected viruses in 34 cases (32%). While often confirmatory, in multiple cases, the identified viruses were not included in the selected routine diagnostic tests. Two years of mNGS in a tertiary diagnostics unit demonstrated the advantages of a single, untargeted approach for comprehensive, rapid and efficient virus diagnostics, confirming the utility of mNGS in complementing current routine tests.

Keywords: metagenomic sequencing; virus infection; virus; clinical impact; diagnostics

1. Introduction

Viruses are one of the major causes of human infectious diseases. The identification of the causative pathogen remains challenging in several disease settings [1–4]. Many clinical syndromes, such as viral meningitis/encephalitis, respiratory syndromes or febrile illness often call for a broad differential diagnosis. This results in extensive diagnostic testing most commonly in an iterative manner starting with the most likely pathogens. This approach is time consuming and associated with considerable costs and can, in the worst case, require weeks to deliver a final result, if at all. As rapid diagnosis is essential for initiating tailored treatment, a diagnostic test with the capacity to capture all potential pathogens in a single assessment would be a tremendous improvement [5].

Viral metagenomic next-generation sequencing (mNGS) fulfills the requirement for comprehensive testing as it enables the detection of the full spectrum of viral pathogens in nearly any clinical specimen by a single assay [6]. As a hypothesis-free, unbiased test, it offers numerous advantages compared

to conventional diagnostic approaches providing the possibility to detect unexpected [7,8] or even unknown pathogens [9–11].

The application of mNGS as a diagnostic tool is intensely discussed [5,12–17]. Elaborate workflows for mNGS have been developed over the years [18–21] and it is increasingly successfully applied in the clinic [15,22]. As any other test, it also has drawbacks [23,24]: reference standards are often missing [25], costs are still high, assay turnover times are considerably longer than for sequence specific DNA/RNA detection by polymerase chain reaction (PCR) and sensitivity may remain lower than for specific PCR assays. These facts have to be counterbalanced once the usefulness of mNGS in correctly diagnosing diverse infectious syndromes has been formally evaluated.

In 2017, we introduced virus sequence-independent metagenomic high-throughput sequencing at our tertiary diagnostic unit and established protocols for sample preparation, sequencing and bioinformatics analysis [26]. This analysis workflow has proven highly successful enabling us to detect disease causing viral pathogens in complex clinical cases [27–29]. Since 2017, viral mNGS was applied in 105 primarily hard-to-diagnose cases upon request by the treating clinician. In this study, we systematically evaluated the impact of mNGS on diagnosis, comparing outcome and workload to results obtained by conventional routine testing in these cases.

2. Materials and Methods

2.1. Ethical Statement

Samples were obtained in the frame of the ongoing diagnostic workup from patients of the University Hospital Zurich and other regional hospitals. The overall analysis was performed on all anonymized samples obtained. For a more detailed clinical analysis, written informed general consent of the University Hospital Zurich was sought and obtained from 67 participants, referred to as the study subgroup.

2.2. Study Design

In this study, all mNGS assays performed at the Institute of Medical Virology on clinical cases analyzed over a period of two years (May 2017–June 2019) were evaluated and compared to conventional routine testing regarding workload and outcome. Criteria for performing metagenomic analysis were either: (i) an unknown etiology of infection even after extensive conventional testing; or (ii) a very broadly formulated initial differential diagnosis. The analysis was either requested by the treating physician or the infectious disease (ID) consultant service of the University Hospital Zurich.

2.3. Viral Metagenomic Sequencing of Clinical Samples

Clinical samples were pre-processed upon arrival and nucleic acid extracts were collected and batched for weekly sequencing runs. mNGS and data analysis were performed as described previously in a subset of patients ($n = 11$, protocol v1.0.0, [26,28]). The remaining patients ($n = 94$) were analyzed using an adapted metagenomic workflow, which was established during the past years (https://github.com/medvir/virome-protocols). Detailed description is provided in Appendix A. Briefly, samples were centrifuged and filtered (0.45 µm). Total nucleic acids were extracted followed by reverse transcription with random hexamers and second strand synthesis in separate reactions for RNA and DNA genomes. Next, sequencing libraries were constructed using the NexteraXT protocol (Illumina, San Diego, CA, USA) and sequenced on a MiSeq for 1×151 cycles using version 3 chemistry. A maximum of five samples (plus a negative control) were sequenced per run, resulting in on average 6.8 million reads per sample. Reads were analyzed with a dedicated bioinformatic pipeline called "VirMet" (github.com/medvir/VirMet/releases/tag/v1.1.1) [27]. The raw viral sequencing reads have been uploaded to zenodo (doi: 10.5281/zenodo.3355228).

2.4. Criteria for Positive Virus Hits

Sequencing was considered positive for a specific virus species if: (i) at least three reads of that species were detected; (ii) the reads were distributed over the whole genome with a high coverage score (coverage score = number of bases covered/number of bases expected to be covered with the amount of virus reads); (iii) the virus was not detected more than 100 times more often in the negative control or in other samples of the same run (carry-over, index hopping); and (iv) the virus reads were detected in the corresponding workflow (RNA/DNA). While for research purposes all detected viruses were listed ("virus found" in Table S1), only viruses that met all defined detection criteria and those with a potential clinical significance for humans were reported to the physician ("virus reported" in Table S1; e.g., Brome mosaic virus was clearly detected according to our criteria, but not reported). Viruses were reported without any quantitative information (e.g., read numbers).

2.5. Evaluation of the Utility of Viral Metagenomic Sequencing Compared to Conventional Testing

The utility of mNGS was evaluated by comparing outcome and workload to conventional routine testing. To evaluate the outcome, all mNGS tests (or separately by sample type) were compared to the respective conventional test (same sample type, direct detection method, time point ±2 days apart) and positive and negative percent agreement (PPA/NPA) were calculated. Overall percent agreement (OPA) was additionally calculated as a direct comparison metric. For these analyses, co-infections and multiple tested specimens per patient were treated as individual datasets.

The evaluation of the workload for conventional testing was done in two ways. As a first measure, all viral targets were considered individually and a mean of performed tests over all patients was calculated. As a second measure, viral targets were grouped according to multiplex panels (ePlex Respiratory Pathogen Panel, GeneXpert Flu/RSV, Fast-Track Diagnostics (FTD) Respiratory pathogens 21, RIDAGENE Gastrointestinal panel). The number of panels and additional individually performed tests was named "diagnostic requests" and its mean over all patients was calculated. In both ways, all virus detection methods were considered, including specific PCR, immunofluorescence, serology, culture and intrathecal antibody synthesis testing.

2.6. Evaluation of the Clinical Impact

A subgroup of patients (*n* = 67) had an extended informed consent (IC). In these patients, we retrospectively analyzed clinical charts to determine the clinical impact of the performed test.

3. Results

3.1. Case Statistics of Samples Analyzed by Viral Metagenomics

Since May 2017, 163 specimens from 105 patients (one case per patient) were analyzed by mNGS with increasing numbers over the years (2017: 11 patients; 2018: 53 patients; 2019 until 12 June: 41 patients). In some cases, multiple specimens were sequenced. The most frequently analyzed sample type was cerebrospinal fluid (CSF) (34.4%), followed by blood samples (including ethylenediaminetetraacetic acid (EDTA) and native blood, 32.5%), throat swabs (9.8%), stool samples and biopsies (5.5%) (Figure 1A). Diverse departments and units requested mNGS. The test was most frequently ordered by the subspecialties neurology (30.7%), general internal medicine (including intensive care units) (19.6%), infectious diseases (11.7%) and cardiology (6.7%) as well as the emergency department (8%) (Figure 1B). CSF was the major sample type analyzed for the neurology department (56%, 28 specimens), but also blood samples and brain biopsies were processed (30% and 10%, 15 and 5 specimens, respectively). General internal medicine and infectious diseases primarily sent blood samples for analysis (37.5% and 57.9%, 12 and 11 specimens, respectively). For pulmonary diseases, either throat swabs (66.7%, 4 specimens) or broncho-alveolar lavages (BAL; 33.3%, 2 specimens) were analyzed. The remaining clinics sent in more diverse specimens, including punctures (pericardium, pleura), stool, urine, and swabs from different locations. Some rare analyzed

specimens included a DNA extract of a myocardial biopsy prepared by the ordering center, a sonicate of an infected knee implant and cell culture supernatant (Figure 1B).
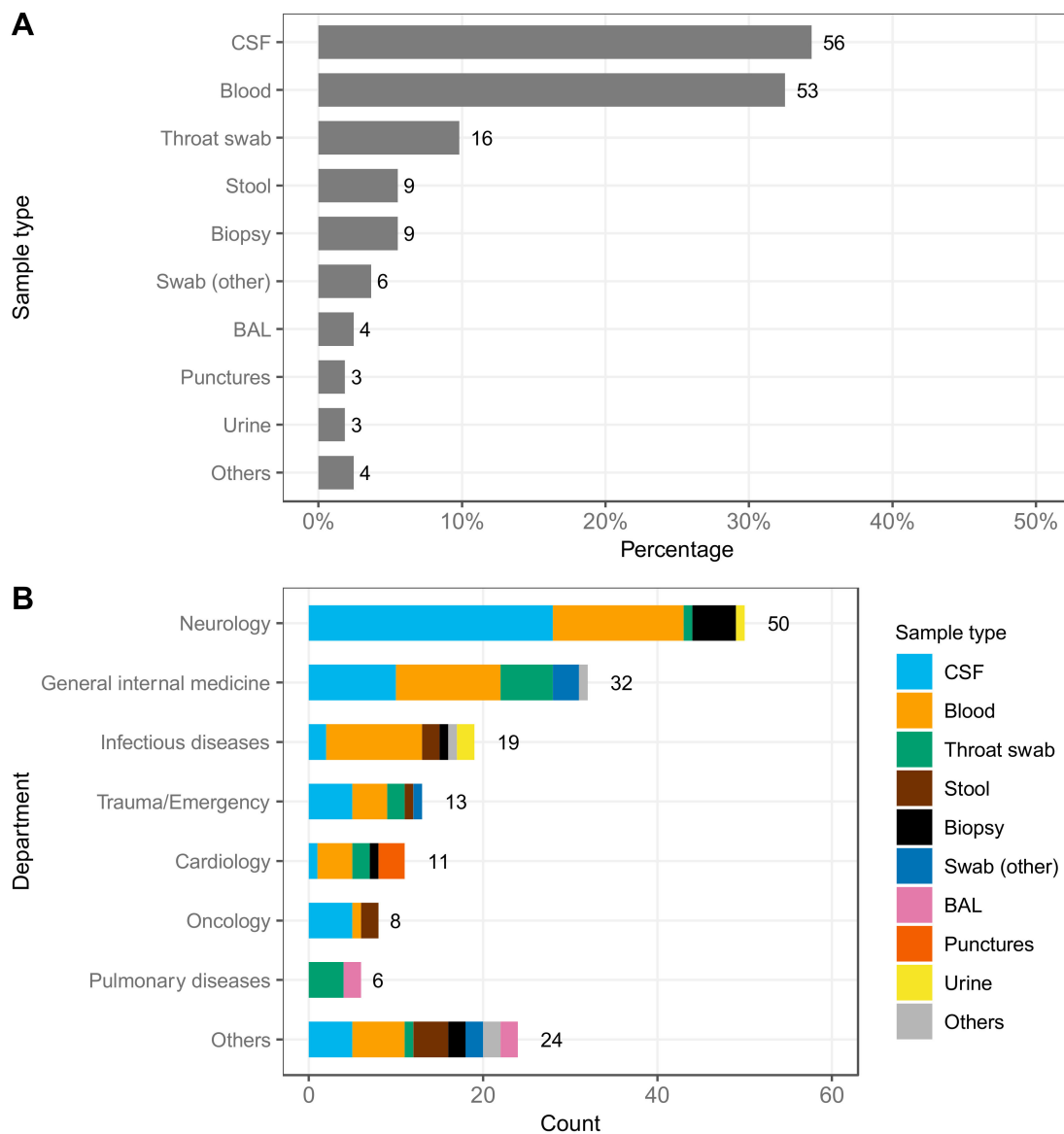

Figure 1. Case statistics of samples analyzed by viral metagenomic next-generation sequencing (mNGS). (**A**) Sample types analyzed in 105 cases over two years shown in percentage. Swab (other) includes enoral swab, skin swab, nasopharyngeal secretion and unknown swab. Others includes DNA extract of myocardial biopsy, sonicate of a knee implant and cell culture supernatant. (**B**) Number of samples and sample types analyzed per department. Others includes otorhinolaryngology, dermatology, rheumatology and external sources. BAL: Broncho-alveolar lavage, CSF: Cerebrospinal fluid. Bar labels indicate the total count.

3.2. Viruses Detected by Viral Metagenomic Sequencing

Viral mNGS was performed for all 105 cases as requested by the treating physician. In 43 cases (41%), mNGS found at least one virus. Among these, mNGS detected multiple viruses in 13 cases (30.2%) resulting in a mean of 1.6 detected viruses over all positive cases. In one case, four viruses were detected in one sample (Patient 39, Table S1). In all others, no relevant viral reads could be detected (excluding bacteriophages, plant viruses and endogenous retrovirus sequences).

In the 43 positive cases, mNGS detected 32 distinct virus species belonging to 14 virus families. Anelloviruses, which are considered body flora [30], were found most frequently (21%), followed by Herpes- and Flaviviruses (18% and 15%, respectively, Figure S1). Of those, 27 species and 13 families were reported to the physicians (Figure 2). Here, we included one case, where the detection of bacteriophage Aeromonas virus phiO18P supported microbiological routine workup which was known to be positive for *Aeromonas*.

Overall, more DNA than RNA viruses were found (17 DNA and 15 RNA virus species, Figure S1), whereas less DNA than RNA viruses were reported (13 DNA and 14 RNA virus species, Figure 2). An interactive version of the krona charts (Figure 2 and Figure S1) including found and reported viruses is provided in Figure S2.
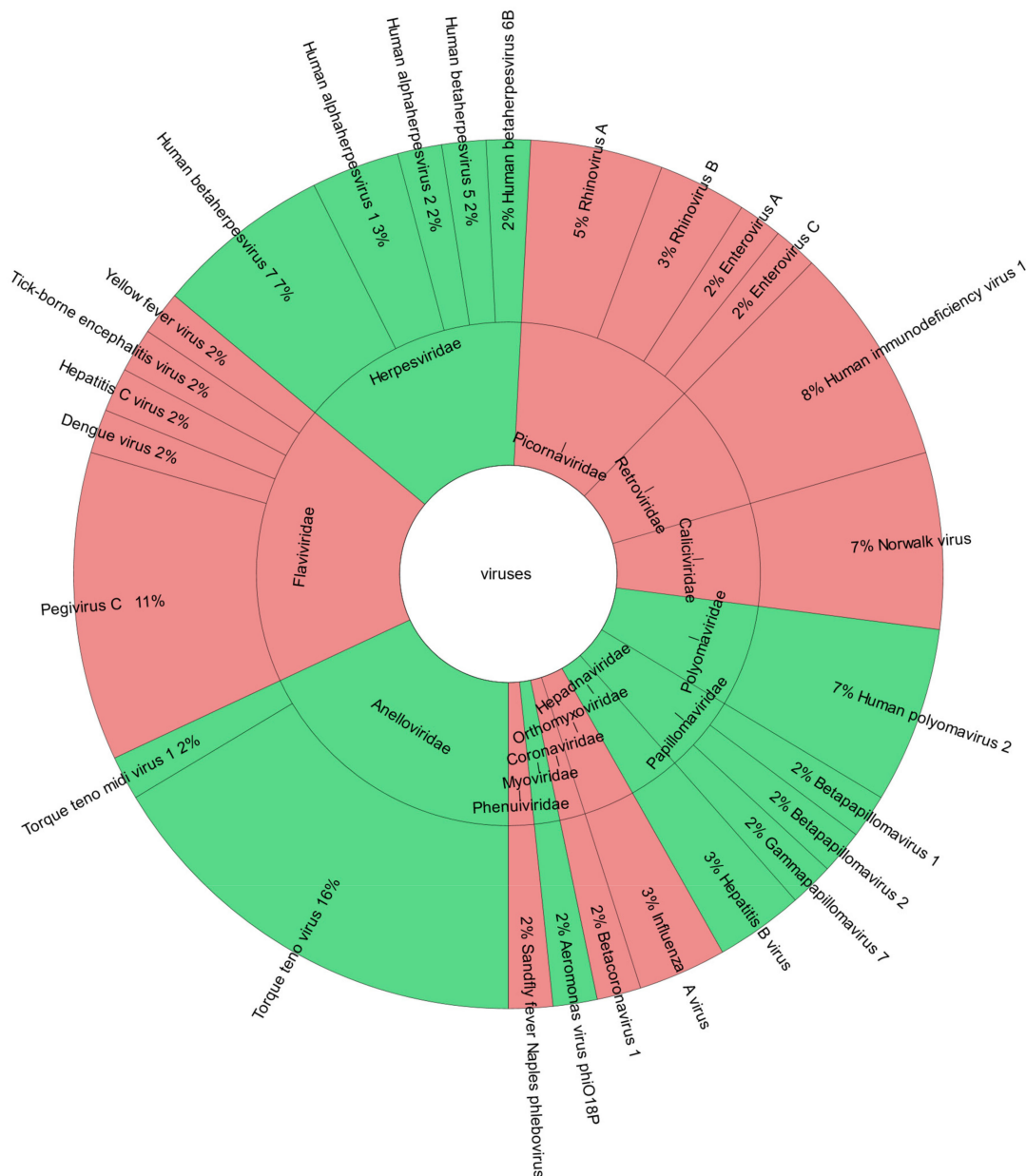

Figure 2. Viruses reported by viral mNGS. Krona chart of detected and reported viruses on family and species level with DNA viruses shown in green and RNA viruses shown in red. Reference krona chart: [31].

3.3. Outcome of Viral Metagenomic Assay Versus Conventional Testing

We analyzed individual results of mNGS compared to routine diagnostic testing. In cases where mNGS yielded a positive result, often no respective conventional test (same viral target/direct detection method/similar time point) was done for specimens analyzed by metagenomics. Therefore, these samples were excluded for the following analyses. For the remaining 75 datasets, positive percent agreement (PPA, sensitivity) was 65% and negative percent agreement (NPA, specificity) was 95%. Twelve samples tested positive by specific PCR but not by mNGS; 10 of these tested low positive and are probably out of the detection range of mNGS. Excluding these, the sensitivity would be 92% (Table 1). The low positive conventional tests missed by mNGS were different species of Human herpesviruses and one case of Dengue virus. The two positive conventional tests detected Human gammaherpesvirus 4 (EBV) in BAL and CSF, respectively (Patients 8 and 96, Table 1).

To disentangle differences in the outcome of mNGS compared to conventional testing regarding different sample types, we performed an additional analysis for the three most frequently analyzed specimen types. Including tests reported as low positive by respective conventional testing, overall percent agreement (OPA) is 81% in CSF, 68% in blood samples and 100% in the tested throat swabs (Table 1).

Table 1. Comparison of the outcome of viral mNGS testing to respective conventional tests in all samples and the most frequently analyzed sample types.

					Respective Conventional Testing	
					+	−
All Samples	OPA = 81/94 [1]	PPA = 65/92% [1]	mNGS	+	22	2
		NPA = 95%		−	2 pos 10 low pos [1]	39
CSF	OPA = 81/91% [2]	PPA = 64/88% [2]	mNGS	+	7	1
		NPA = 93%		−	1 pos 3 low pos [2]	14
Blood	OPA = 68/100% [3]	PPA = 46/100% [3]	mNGS	+	5	0
		NPA = 100%		−	0 6 low pos [3]	8
Throat swab	OPA = 91%	PPA = 100%	mNGS	+	4	1
		NPA = 86%		−	0	6

[1] Ten out of the 12 samples that tested negative by mNGS and positive by conventional testing were actually reported as "low positive" (at the detection threshold). [2] Three out of the four CSF samples that tested negative by mNGS and positive by conventional testing were actually reported as "low positive" (at the detection threshold). [3] Six blood samples tested negative by mNGS and "low positive" by conventional testing (at the detection threshold). NPA: Negative percent agreement, OPA: Overall percent agreement, PPA: Positive percent agreement.

We next looked at the cases with a result reported positive by mNGS that were excluded from the analysis above because no respective conventional test was performed. If we excluded viruses that are considered body flora (Anelloviruses [30]) or common skin contamination (Papillomaviruses [21,32]), mNGS detected 24 "infections" of 11 different virus species which were not tested for by a respective conventional test: Pegivirus C (7), Human betaherpesvirus 7 (4), Norwalk virus (3), Human immunodeficiency virus 1 (2), Hepatitis B virus (2), Influenza A virus (1), Human alphaherpesvirus 1 (1), Human alphaherpesvirus 2 (1), Hepatitis C virus (1), Betacoronavirus 1 (1), and Aeromonas virus phiO18P (1).

3.4. Workload of Viral Metagenomic Assay Versus Conventional Testing

mNGS is an open and unbiased approach with the potential to replace multiple specific diagnostic tests. We evaluated the number and workload of conventional tests that were performed in the

cases analyzed in this study. We distinguished between: (i) the individually performed tests; and (ii) "diagnostic requests", which included syndromic panels (multiplex assays, block analyses) and additional individually performed tests. For example, a single request for a multiplex panel can target multiple different virus species or subtypes. The rationale behind the two categories was the simplified differential diagnostic algorithm and the lower workload (hands-on time) when using multiplex panels. In our 105 cases, the physicians placed on average 10.8 diagnostic requests that covered 18.5 different viral targets, which shows the inherent difficulty in obtaining a correct and rapid diagnosis in these patients and the large number of viruses taken into consideration.

On the other hand, mNGS was performed once per patient, except in seven cases where an additional test was ordered at a later timepoint always including a different specimen. 70% of these additional metagenomic tests were performed on biopsies. Overall, this resulted in a mean of 1.07 mNGS analyses over all patients.

3.5. Timing Viral Metagenomic Assay Versus Conventional Testing

We were further interested in temporal aspects of mNGS in clinical use. We looked at the delay until mNGS was considered by the treating physician or ID consult service. Most often, mNGS was requested the same day as the sample was taken, i.e., ID specialist recognized the cases as hard-to-diagnose and necessitating an open diagnostics approach (Figure 3A). In other cases, it took several days until the test was requested, probably because in the meantime routine diagnostics did not provide a positive result. Tests requested after a prolonged time period correspond to retrospective analyses of archived specimens. In the majority of cases (65.4%), the results were validated and reported to the clinicians within seven days (Figure 3B). For urgent cases and with optimal timing (sample available early in the morning), the results could even be provided the next day. In some cases, however, it took several weeks until a final result was reported, due to delays in sample shipment or batching of less urgent cases into one sequencing run.
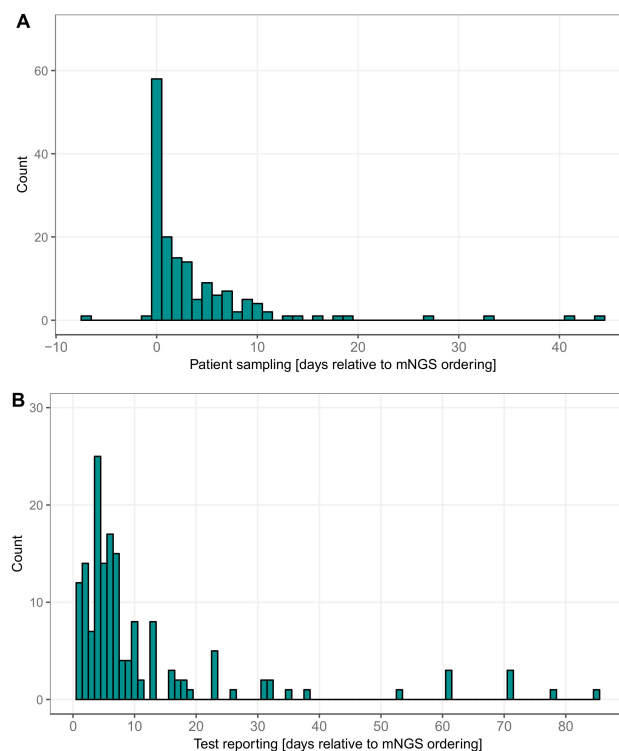

Figure 3. Timing of viral mNGS: (**A**) the delay (in days) between the time point when the clinical specimen was collected and the metagenomic test was requested; and (**B**) the delay (in days) between metagenomic test request and validation, corresponding to reporting of the results to the physician.

3.6. Patient Characteristics in the Study Subgroup

For 67 patients, an extended collection of clinical data was available. Basic demographic and clinical characteristics are shown in Table 2: 43 (64.2%) patients were male and the median age was 53 years. In total, 24 patients (35.8%) were immunocompromised. Most of the tests were ordered for patients treated at the department of neurology ($n = 26$ (38.8%)), and the department of general internal medicine ($n = 15$ (22.4%)). ID consultant service was involved in 45 (67.2%) cases. CSF, blood and throat swabs were the most common clinical specimens (Table 3). mNGS was most often performed in patients with a suspected meningitis/meningoencephalitis, peri-/myocarditis and febrile syndromes (Table 4).

Table 2. Basic characteristics in a study subgroup of 67 patients.

Age: Median (Range)	53 (17–88 Years)
Male gender	43 (64.2%)
Patients immunocompromised	24 (35.8%)
Post SOT	7 (29.2%)
Malignancy	5 (20.8%)
HIV	5 (20.8%)
Autoimmune disorder	7 (29.2%)
Patients by department	
Internal medicine and subspecialties	35 (52.2%)
General internal medicine	15 (22.4%)
Cardiology	7 (10.4%)
Infectious diseases	7 (10.4%)
Pulmonolgy	3 (4.5%)
Rheumatology	2 (3%)
Hematology/Oncology	1 (1.5%)
Neurology/Neurosurgery	28 (41.8%)
Neurology	26 (38.8%)
Neurosurgery	2 (3%)
Other	4 (6%)
Emergency department	1 (1.5%)
Otorhinolaryngology	1 (1.5%)
Dermatology	2 (3%)

Data shown as absolute number and percentage (in parenthesis) if not stated otherwise. HIV: Human immunodeficiency virus, SOT: Solid organ transplant.

Table 3. Sample types analyzed in the study subgroup.

Clinical Samples:	**101**
CSF	35 (34.7%)
Blood	32 (31.7%)
Throat swab	11 (11%)
Biopsy	6 (6%)
Stool	4 (4%)
Urine	3 (3%)
Punctures	3 (3%)
BAL	2 (2%)
Others	5 (5%)

Data shown as absolute number and percentage (in parenthesis). Others includes nasopharyngeal swab, cell culture, skin swab, unknown swab, and nasopharyngeal secretion.

Table 4. Tentative diagnosis when test was ordered (study subgroup).

Disease	Number of Cases
Neurological disorders	
Meningitis and/or encephalitis	17
Other central nervous system disorders [1]	11
Cerebral lesion/abscess	3
Peripheral nervous system disorders	2
PML	1
Other diseases, disorders & syndromes	
Pericarditis and/or myocarditis	8
Febrile syndromes (including FUO)	8
Respiratory tract infections	4
Allograft dysfunction after lung transplantation	3
Diarrhea	3
Sepsis in neutropenia	1
Cytokine-Release-Syndrome	1
Unspecific polyarthritis and lymphadenopathy	1
Constitutional symptoms unknown etiology	1
Unspecific myalgia syndrome	1
Unspecific cutaneous lesions	1
Chronic sinusitis	1

[1] Consists of patients with syndrome of transient headache and neurologic deficits with cerebrospinal fluid lymphocytosis (1), aphasia (1), multiple sclerosis (1), ocular flutter and vertigo (1), orofacial myoclonus and impairment of the oculomotor nerve (1), optic neuritis (1), cervical myelopathy (2), Parry-Romberg Syndrome (1), multiple sclerosis (1), stroke (1) and suspected cerebral vasculitis (1). PML: Progressive multifocal leukoencephalopathy, FUO: Fever unknown origin.

3.7. Clinical Impact of Viral Metagenomics Results in the Study Subgroup

While specific treatment is often not available against viruses, a rapid confirmation or exclusion of a viral infection is important to limit nosocomial spread and avoid unnecessary antibiotic treatment. Reviewing patient charts in the study subgroup, we were interested if the positive or negative result of mNGS had a direct impact on patient treatment on top of routine diagnostics. In three cases, mNGS provided the final diagnosis for viruses not considered or available by routine diagnostic tests (see Case Vignettes 1–3 in Appendix B). Further diagnostic tests were therefore not necessary and empirical antibiotic treatment could be stopped [29]. In other cases, mNGS essentially contributed to exclude an actively replicating virus infection. In one heart transplant recipient, treatment was adjusted after confirmation of a positive routine test (Norwalk virus PCR) and exclusion of further viruses in a stool sample. One patient with an inflammatory central nervous system disorder was treated with high dose immunosuppression after exclusion of a viral infection by mNGS.

4. Discussion

In the present study, we evaluated the clinical utility of mNGS by assessing all cases ($n = 105$) submitted for mNGS to our tertiary diagnostics unit since May 2017. The 105 cases comprise a study set with highly diverse clinical syndromes. All were referred for mNGS by the treating physician or the ID consult service because prior routine testing had failed to define an etiology of infection or because the initial differential diagnosis was very broad. Clinical chart review revealed that, in three cases, mNGS detected a virus infection that fitted the observed clinical picture but was not considered by

the original differential diagnosis and hence initially not tested by routine diagnostics. mNGS thus provided the final diagnosis in these cases and proved to be a helpful or even decisive diagnostic tool with a relevant impact on treatment decisions and case management. As determined in the outcome analysis, for 22 datasets (29.3%), mNGS matched with routine diagnostic results. In most of the samples (52%), neither mNGS nor routine diagnostic testing could detect a viral pathogen. In such cases, it remained open whether a non-viral pathogen was the cause or whether no pathogen was involved at all. Negative results are nevertheless important for antibiotic treatment decisions and lifting of precautionary isolation measures. However, it is not possible to rule out viral infections entirely as sampling may have been post peak virus production and remaining virus loads may have been below the sensitivity threshold of mNGS.

Given the large number of different viruses and virus families detected in the positive cases, it might be difficult to assign whether a detected virus is really relevant for the observed symptoms, in particular for viruses not commonly detected in humans. On the other hand, this is an advantage of the unbiased approach as potentially clinically relevant pathogens might not be considered by a conventional approach. For reporting test results to clinicians, we decided to report positive virus hits as binary results without quantitative information as for any other qualitative specific test.

Our current analysis strategy focused solely on viral pathogens. As parallel assessment of bacterial or fungal pathogens by metagenomic sequencing would provide a comprehensive analysis of infectious agents, we are currently adjusting our procedures to implement this.

On average, in the analyzed cases, physicians requested 10.8 diagnostic tests which covered 18.5 different viral targets. This highlights the inherent difficulty in obtaining a correct and rapid diagnosis in these patients and the large number of viruses taken into consideration. A single, unbiased metagenomic assay could not only limit the workload for diagnostic labs, but also be overall more cost effective, less work intensive and faster, given the high number of individual diagnostic requests on average over all patients.

The amount of nucleic acid background in clinical samples will influence the performance of metagenomics, which together with the variable sensitivity of mNGS for different virus classes and strains noted in the literature [33], renders the establishment of a reference standard very challenging. A direct comparison between standard diagnostics and metagenomics always depends on the definition of the gold standard, adding to the difficulty in validation of such a new method. Internal validation by specific PCR is key to determine sensitivity and specificity. As a quality control assessment, inter-laboratory ring trials are necessary to test any new metagenomics pipeline, to disentangle sources of variability, to formulate best practices and to gain the trust of the clinician for successful application in the clinical setting.

Despite the indisputable innovation and potential, the exact role for metagenomics in the diagnostic algorithm still has to be defined. In our study, mNGS was often ordered as a first-in-line diagnostic tool, in parallel with routine diagnostic tests. As mNGS is currently used for difficult-to-diagnose cases, there is a strong bias towards patients where no diagnosis can be made—which should not be held against metagenomics but rather seen in a clinical context, where non-infectious causes may contribute to some of the clinical syndromes. The lower sensitivity of mNGS compared to specific PCR will, at least in the initial phase, make it necessary to do specific tests in parallel, importantly for pathogens where a rapid treatment initiation and diagnosis is crucial. Given that often more than one virus is detected, a very close interdisciplinary collaboration with the clinician is key to put the findings into a clinical context.

Author Contributions: Conceptualization, N.J.M., M.H., and A.T.; methodology, V.K., S.S., and M.H.; investigation, V.K., A.P. and S.S.; resources, D.L.B., H.F.G., R.C. and A.Z.; data curation, V.K., A.P. and S.S.; writing—original draft

Appendix A (Supplemental Methods)

Appendix A.1 Metagenomic Sequencing v2.1.0

Appendix A.1.1 Sample Pre-Processing, Nucleic Acid Extraction and DNase Treatment

Samples were centrifuged at 2000 rpm for 10–30 min depending on sample type, the supernatant was filtered using a 0.45 μm PES filter (TPP, Trasadingen, Switzerland) and aliquoted for optional storage at –80 °C. The NucliSENS EasyMAG system (BioMérieux; Craponne, France) was used for extraction of total nucleic acids according to manufacturer's instructions with an input volume of 1000 μL that was eluted in 25 μL. Next, DNase treatment was performed in a separate reaction for RNA viruses prior to reverse transcription using the TURBO DNA-*free* Kit (TURBO DNase Treatment and Removal Reagents; Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA) as per manufacturer's instructions.

Appendix A.1.2 Random Reverse Transcription and Second Strand Synthesis

Random reverse transcription of total nucleic acids was performed in separated reactions for RNA and DNA viruses. For the RNA workflow, complementary DNA was generated with a random hexamer primer. Then, 5 μL of DNase-treated eluate were used as template in a total volume of 20 μL, with 2.5 μM of random primer, 0.5 mM of dNTPs (Thermo Scientific), 1× SuperScript IV buffer, 5 mM DTT, 2 U/μL RNaseOUT (Invitrogen/Life Technologies, Waltham, MA, USA) and 10 U/μL of SuperScript IV Reverse Transcriptase (Invitrogen/Life Technologies). The template and random primers were heated at 65 °C for 5 min, followed by reverse transcription. First, the reaction mixture was incubated at 23 °C for 10 min before reverse transcription at 50 °C for 10 min and inactivation at 80 °C for 10 min. A treatment with 2U of RNase H (New England Biolabs, Ipswich, MA, USA) was performed for 20 min at 37 °C followed by an inactivation step at 94 °C for 2 min. The second strand was synthesized with 5 U DNA Polymerase I Large Klenow Fragment (NEB Biolabs) in 1× NEB buffer 2 in a final volume of 30 μL, at 37 °C for 30 min followed by an enzyme inactivation step at 75 °C for 20 min.

Similarly, the DNA workflow started at the denaturation step at 94 °C for 2 min. Subsequent second strand synthesis was performed using random hexamer primers with the same conditions as described for the RNA workflow and a final reaction volume of 20 μL.

After second strand synthesis, the products were purified using 3× volumes of Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA), eluted in 20 μL of UltraPure DNase/RNase-free distilled water (Invitrogen/Life Technologies) each.

Appendix A.1.3 Viral Metagenomic High-Throughput Sequencing

PCR products were quantified with QuantiFluor ONE dsDNA system (Promega, Madison, WI, USA) on the Quantus Fluorometer (Promega) and diluted to 0.2 ng/μL if necessary. One nanogram of DNA was used as input for library construction using the NexteraXT protocol (Illumina). For samples with a DNA concentration below 0.2 ng/μL, 5 μL of undiluted sample were used for constructing sequencing libraries. DNA and RNA preparations were kept separate for constructing sequencing libraries and individual samples were dual indexed during the library preparation and pooled for sequencing. The quality and size of the libraries were assessed after limited-cycle Nextera PCR by

capillary gel electrophoresis (Fragment Analyzer, Advanced Analytical, Ames, IA). Libraries were sequenced on a MiSeq (Illumina) for 1×151 cycles with version 3 reagents and the "FASTQ only" workflow. Samples were demultiplexed using MiSeq Reporter v2.6.2.

Appendix A.1.4 Bioinformatic Analysis

The obtained reads were processed with a dedicated bioinformatic pipeline "VirMet" version 1.1.1 developed in our laboratory (https://github.com/medvir/VirMet/releases/tag/v1.1.1) [27]. Briefly, reads were quality-filtered by removing low quality bases (average PHRED score below 20), reads shorter than 75 bp and reads with low entropy (i.e., consisting mainly of repeats). Reads passing quality filters were cleaned from non-viral reads by aligning with Burrows-Wheeler Aligner (BWA, algorithm BWA-MEM) [34] against, in this order, human, bacterial, bovine and canine genomes. Reads not matching any of the above genomes were aligned with BLAST [35] against an in-house viral database that contains approximately 69,000 different virus sequences. For each sequencing read that passed the quality filter, the BLAST hit with lowest e-value was reported, given the coverage and identity were higher than 75%. Reads which did not match genomes used in the cleaning step and did not match viral genomes included in the database were reported as of unknown origin.

Appendix A.1.5 Metagenomic Sequencing v2.2.1

In this version of the protocol, the ratio of AMPure XP beads (Beckman Coulter) was reduced from 3× to 2× volumes for purification of the DNA products after second strand synthesis. The rest of the workflow was carried out according to protocol v2.1.0.

Appendix B (Case vignettes)

Appendix B.1 Sandfly Fever Naples phlebovirus Associated Meningitis: Direct Impact of mNGS on Diagnosis and Antibiotic Treatment

The case was already reported [29]: A 23-year-old man presented with meningitis and epididymitis. Routine diagnostic workup yielded no diagnosis, but viral metagenomics in CSF revealed a Toscana virus infection. Due to the diagnosis, empirical antibiotic therapy for meningitis was suspended.

Appendix B.2 Tick-Borne Encephalitis: Rapid Diagnosis of an Unexpected Virus in Encephalitis

A 58-year-old female patient with the history of a kidney transplantation was transferred from a regional hospital to our hospital because of somnolence, and an acute kidney injury. During the hospital stay, the patient showed a prolonged comatose state. At the time of referral, C-reactive protein (CRP) was elevated (31 mg/L [normal range: <10 mg/L]) and leucocytes were normal (7.69 G/L [normal range: 3.0–9.6 G/L]). Lumbar puncture revealed 5 cells/µL. Cranial magnetic resonance imaging (MRI) revealed a leptomeningeal and pachymeningeal enhancement as well as a cerebellar swelling. Multiple serologic tests could not identify an infectious cause. Noticeable was a serum borderline positive IgM for tick-borne encephalitis virus. Serum IgG was negative as well as IgM in the CSF. In contrast, a subsequently performed viral metagenomics test in the CSF was positive for tick-borne encephalitis virus. The result was confirmed by specific PCR. Despite intensified symptomatic treatment, the patient remained comatose, showed absence of brainstem reflexes, and flaccid paresis. Due to the unfavorable prognosis, treatment was finally stopped. Viral metagenomics revealed the final diagnosis in this case, whereas serology in the CSF was still negative and a specific PCR was pending.

Appendix B.3 Pegivirus C Associated Meningoencephalitis: A Possible New Encephalitis Pathogen

A 43-year-old male patient was evaluated for a limbic encephalitis. Initial systemic infectious parameters were normal (leucocytes: 7.6 G/L (normal range: 3.0–9.6 G/L); CRP: <1 mg/L (normal range: <10 mg/L)). CSF puncture revealed a lymphocytic pleocytosis (57 cells/µL) with normal glucose

(3.9 mmol/L (normal range: 2.3–3.9 mmol/L)) and elevated protein (0.9 g/L (normal range 0.15–0.45 g/L)). Cranial computed tomography scan and angiography were unremarkable. A cranial magnetic resonance imaging revealed a bilateral enhancement in the entorhinal cortex. An empirical treatment with ceftriaxone and acyclovir was stopped after exclusion of a Lyme neuroborreliosis and negative test results for Cytomegalovirus (CMV) and Herpes simplex virus (HSV). Because of worsening symptoms, the patient was transferred to our hospital. Multiple diagnostic testing (serology, CT scan and positron-emission tomography (PET) scan) could not identify an infectious, auto-immunological, or malignant cause of the encephalitis. In contrast, viral metagenomics detected Pegivirus C in both blood and CSF samples. Pegivirus C is a member of the family Flaviviridae and was recently described as the putative pathogen in several cases of encephalitis [36]. As Pegivirus C is not part of the routine workup in patients with a meningo-/encephalitis, viral metagenomics was able to establish a putative diagnosis, and further diagnostic tests were postponed. The recently described role of Pegivirus as a potential cause of an infection of the central nervous system has to be validated further.

References

1. Glaser, C.A.; Honarmand, S.; Anderson, L.J.; Schnurr, D.P.; Forghani, B.; Cossen, C.K.; Schuster, F.L.; Christie, L.J.; Tureen, J.H. Beyond viruses: clinical profiles and etiologies associated with encephalitis. *Clin. Infect. Dis.* **2006**, *43*, 1565–1577. [CrossRef] [PubMed]

2. Ewig, S.; Torres, A.; Angeles Marcos, M.; Angrill, J.; Rano, A.; de Roux, A.; Mensa, J.; Martinez, J.A.; de la Bellacasa, J.P.; Bauer, T. Factors associated with unknown aetiology in patients with community-acquired pneumonia. *Eur. Respir. J.* **2002**, *20*, 1254–1262. [CrossRef] [PubMed]

3. Bleeker-Rovers, C.P.; Vos, F.J.; de Kleijn, E.M.; Mudde, A.H.; Dofferhoff, T.S.; Richter, C.; Smilde, T.J.; Krabbe, P.F.; Oyen, W.J.; van der Meer, J.W. A prospective multicenter study on fever of unknown origin: The yield of a structured diagnostic protocol. *Medicine (Baltimore)* **2007**, *86*, 26–38. [CrossRef] [PubMed]

4. Jain, S.; Self, W.H.; Wunderink, R.G.; Team, C.E.S. Community-acquired pneumonia requiring hospitalization. *N. Engl. J. Med.* **2015**, *373*, 2382. [CrossRef] [PubMed]

5. Greninger, A.L. The challenge of diagnostic metagenomics. *Expert Rev. Mol. Diagn.* **2018**, *18*, 605–615. [CrossRef]

6. Delwart, E.L. Viral metagenomics. *Rev. Med. Virol.* **2007**, *17*, 115–131. [CrossRef] [PubMed]

7. Murkey, J.A.; Chew, K.W.; Carlson, M.; Shannon, C.L.; Sirohi, D.; Sample, H.A.; Wilson, M.R.; Vespa, P.; Humphries, R.M.; Miller, S.; et al. Hepatitis E virus-associated meningoencephalitis in a lung transplant recipient diagnosed by clinical metagenomic sequencing. *Open Forum Infect. Dis.* **2017**, *4*, ofx121. [CrossRef]

8. Wilson, M.R.; Suan, D.; Duggins, A.; Schubert, R.D.; Khan, L.M.; Sample, H.A.; Zorn, K.C.; Rodrigues Hoffman, A.; Blick, A.; Shingde, M.; et al. A novel cause of chronic viral meningoencephalitis: Cache Valley virus. *Ann. Neurol.* **2017**, *82*, 105–114. [CrossRef]

9. Pastrana, D.V.; Peretti, A.; Welch, N.L.; Borgogna, C.; Olivero, C.; Badolato, R.; Notarangelo, L.D.; Gariglio, M.; FitzGerald, P.C.; McIntosh, C.E.; et al. Metagenomic discovery of 83 new human papillomavirus types in patients with immunodeficiency. *mSphere* **2018**, *3*. [CrossRef]

10. Eibach, D.; Hogan, B.; Sarpong, N.; Winter, D.; Struck, N.S.; Adu-Sarkodie, Y.; Owusu-Dabo, E.; Schmidt-Chanasit, J.; May, J.; Cadar, D. Viral metagenomics revealed novel betatorquevirus species in pediatric inpatients with encephalitis/meningoencephalitis from Ghana. *Sci. Rep.* **2019**, *9*, 2360. [CrossRef]

11. McMullan, L.K.; Folk, S.M.; Kelly, A.J.; MacNeil, A.; Goldsmith, C.S.; Metcalfe, M.G.; Batten, B.C.; Albarino, C.G.; Zaki, S.R.; Rollin, P.E.; et al. A new phlebovirus associated with severe febrile illness in Missouri. *N. Engl. J. Med.* **2012**, *367*, 834–841. [CrossRef]

12. Brown, J.R.; Bharucha, T.; Breuer, J. Encephalitis diagnosis using metagenomics: application of next generation sequencing for undiagnosed cases. *J. Infect.* **2018**, *76*, 225–240. [CrossRef] [PubMed]

13. Deurenberg, R.H.; Bathoorn, E.; Chlebowicz, M.A.; Couto, N.; Ferdous, M.; García-Cobos, S.; Kooistra-Smid, A.M.D.; Raangs, E.C.; Rosema, S.; Veloo, A.C.M.; et al. Application of next generation sequencing in clinical microbiology and infection prevention. *J. Biotechnol.* **2017**, *243*, 16–24. [CrossRef] [PubMed]

14. Gu, W.; Miller, S.; Chiu, C.Y. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. *Annu. Rev. Pathol.* **2019**, *14*, 319–338. [CrossRef] [PubMed]

15. Wilson, M.R.; Sample, H.A.; Zorn, K.C.; Arevalo, S.; Yu, G.; Neuhaus, J.; Federman, S.; Stryke, D.; Briggs, B.; Langelier, C.; et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N. Engl. J. Med.* **2019**, *380*, 2327–2340. [CrossRef] [PubMed]

16. Zanella, M.C.; Lenggenhager, L.; Schrenzel, J.; Cordey, S.; Kaiser, L. High-throughput sequencing for the aetiologic identification of viral encephalitis, meningoencephalitis, and meningitis. A narrative review and clinical appraisal. *Clin. Microbiol. Infect.* **2019**, *25*, 422–430. [CrossRef] [PubMed]

17. Ruppé, E.; Schrenzel, J. Messages from the third International Conference on Clinical Metagenomics (ICCMg3). *Microbes Infect.* **2019**, 1–18. [CrossRef] [PubMed]

18. Victoria, J.G.; Kapoor, A.; Li, L.; Blinkova, O.; Slikas, B.; Wang, C.; Naeem, A.; Zaidi, S.; Delwart, E. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. *J. Virol.* **2009**, *83*, 4642–4651. [CrossRef]

19. Greninger, A.L.; Naccache, S.N.; Federman, S.; Yu, G.; Mbala, P.; Bres, V.; Stryke, D.; Bouquet, J.; Somasekar, S.; Linnen, J.M.; et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med.* **2015**, *13*, 1–13. [CrossRef] [PubMed]

20. Briese, T.; Kapoor, A.; Mishra, N.; Jain, K.; Kumar, A.; Jabado, O.J.; Lipkin, W.I. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *mBio* **2015**, *6*, e01491-15. [CrossRef]

21. Miller, S.; Naccache, S.N.; Samayoa, E.; Messacar, K.; Arevalo, S.; Federman, S.; Stryke, D.; Pham, E.; Fung, B.; Bolosky, W.J.; et al. Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res.* **2019**, *29*, 831–842. [CrossRef]

22. Luk, K.-C.; Berg, M.G.; Naccache, S.N.; Kabre, B.; Federman, S.; Mbanya, D.; Kaptué, L.; Chiu, C.Y.; Brennan, C.A.; Hackett, J. Utility of metagenomic next-generation sequencing for characterization of HIV and human Pegivirus diversity. *PLoS ONE* **2015**, *10*, e0141723. [CrossRef] [PubMed]

23. Simner, P.J.; Miller, S.; Carroll, K.C. Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin. Infect. Dis* **2017**, *380*, 2095. [CrossRef] [PubMed]

24. Boers, S.A.; Jansen, R.; Hays, J.P. 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*, 1059–1070. [CrossRef] [PubMed]

25. Hardwick, S.A.; Deveson, I.W.; Mercer, T.R. Reference standards for next-generation sequencing. *Nat. Rev. Genet.* **2017**, *18*, 473–484. [CrossRef] [PubMed]

26. Lewandowska, D.W.; Zagordi, O.; Geissberger, F.-D.; Kufner, V.; Schmutz, S.; Böni, J.; Metzner, K.J.; Trkola, A.; Huber, M. Optimization and validation of sample preparation for metagenomic sequencing of viruses in clinical samples. *Microbiome* **2017**, *5*, 1–13. [CrossRef]

27. Lewandowska, D.W.; Zagordi, O.; Zbinden, A.; Schuurmans, M.M.; Schreiber, P.; Geissberger, F.-D.; Huder, J.B.; Böni, J.; Benden, C.; Mueller, N.J.; et al. Unbiased metagenomic sequencing complements specific routine diagnostic methods and increases chances to detect rare viral strains. *Diagn. Microbiol. Infect. Dis.* **2015**, *83*, 133–138. [CrossRef] [PubMed]

28. Lewandowska, D.W.; Capaul, R.; Prader, S.; Zagordi, O.; Geissberger, F.-D.; Kügler, M.; Knorr, M.; Berger, C.; Güngör, T.; Reichenbach, J.; et al. Persistent mammalian orthoreovirus, coxsackievirus and adenovirus co-infection in a child with a primary immunodeficiency detected by metagenomic sequencing: A case report. *BMC Infect. Dis.* **2018**, *18*, 1–5. [CrossRef] [PubMed]

29. Tschumi, F.; Schmutz, S.; Kufner, V.; Heider, M.; Pigny, F.; Schreiner, B.; Capaul, R.; Achermann, Y.; Huber, M. Meningitis and epididymitis caused by Toscana virus infection imported to Switzerland diagnosed by metagenomic sequencing: A case report. *BMC Infect. Dis.* **2019**, *19*, 1–4. [CrossRef]

30. Moustafa, A.; Xie, C.; Kirkness, E.; Biggs, W.; Wong, E.; Turpaz, Y.; Bloom, K.; Delwart, E.; Nelson, K.E.; Venter, J.C.; et al. The blood DNA virome in 8,000 humans. *PLoS Pathog.* **2017**, *13*, e1006292. [CrossRef]

31. Ondov, B.D.; Bergman, N.H.; Phillippy, A.M. Interactive metagenomic visualization in a Web browser. *BMC Bioinform.* **2011**, *12*, 385. [CrossRef]

32. Kaiser, L. Virome and Transplantation. 23 April 2018. Available online: https://www.eccmidlive.org/#resources/virome-and-transplantation. (accessed on 31 July 2019).

33. Parras-Molto, M.; Rodriguez-Galet, A.; Suarez-Rodriguez, P.; Lopez-Bueno, A. Evaluation of bias induced by viral enrichment and random amplification protocols in metagenomic surveys of saliva DNA viruses. *Microbiome* **2018**, *6*, 119. [CrossRef] [PubMed]

34. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]

35. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [CrossRef] [PubMed]

36. Bukowska-Osko, I.; Perlejewski, K.; Pawelczyk, A.; Rydzanicz, M.; Pollak, A.; Popiel, M.; Cortes, K.C.; Paciorek, M.; Horban, A.; Dzieciatkowski, T.; et al. Human Pegivirus in patients with encephalitis of unclear etiology, Poland. *Emerg. Infect. Dis.* **2018**, *24*, 1785–1794. [CrossRef]