Isolation of vaccine-like poliovirus strains in sewage samples from the UK

Manasi Majumdar[1], Dimitra Klapsa[1], Thomas Wilton[1], Joyce Akello[2], Catherine Anscombe[2], David Allen[2], Edward T. Mee[1], Philip D. Minor[1] and Javier Martin[1*]

[1]Division of Virology, National Institute for Biological Standards and Control (NIBSC), South Mimms, Potters Bar, Herts EN6 3QG, UK

[2]Enterovirus Unit, Public Health England, London NW9 5EQ, UK

[*]To whom correspondence should be addressed. E-mail: javier.martin@nibsc.org

Abstract word count: 192

Main text word count: 3485

Key words: Global Polio Eradication, environmental surveillance, human enterovirus, vaccine-like poliovirus, intertypic recombination

Brief Summary: We describe the isolation of vaccine-like poliovirus in sewage samples using concentration methods followed by cell culture infection and next generation sequencing. Using this approach, we rapidly obtained whole-genome sequences of polio and non-polio enterovirus strains present in mixtures.

ABSTRACT

Background:

Environmental surveillance (ES) is a sensitive method for detecting human enterovirus (HEV) circulation and it is used worldwide to support global polio eradication. We describe a novel ES approach using next generation sequencing (NGS) to identify HEVs in sewage samples taken in London, UK, from June-2016 to May-2017.

Methods and Results:

Two different sewage concentration methods were used: a two-phase aqueous separation system and size-exclusion by filtration and centrifugation, in combination with virus isolation in cell cultures and NGS. Type 1 and 3 vaccine-like poliovirus (PV) strains were detected in samples from September 2016 and January 2017. NGS analysis allowed us to rapidly obtain whole-genome sequences of polio and non-polio HEV strains. As many as six virus strains from different HEV serotypes were identified in a single cell culture flask. PV isolates contained only a small number of mutations from vaccine strains commonly seen in early isolates from vaccinees.

Conclusions:

Our ES setup has high sensitivity for polio and non-polio HEV detection generating nearly whole-genome sequence information. Such ES systems provide critical information to assist the polio eradication endgame and contribute to improve our understanding of HEV circulation patterns in humans.

INTRODUCTION

The Global Polio Eradication Initiative (GPEI) has been very successful in reducing PV circulation in humans to the brink of global extinction [1]. However, some areas in Afghanistan and Pakistan remain where PV transmission has never been eliminated and type 1 wild PV (WPV1) and type 2 circulating vaccine-derived PV (cVDPV2) are still being transmitted from person to person. In addition, as of July 2017, PV transmission still occurs in some areas of the Middle East and Africa where there are severe difficulties in accessing children for vaccination. WPV1 was last detected in Nigeria in September 2016 and recent cVDPV2 outbreaks have been reported in Syria and DR Congo [2, 3].

As only a small proportion of infections by PV cause paralytic disease, the establishment of efficient and sensitive surveillance systems to guide public health interventions has been essential in ensuring the GPEI success. The GPEI mostly relies on acute flaccid paralysis (AFP) surveillance which is monitored by strict quality performance indicators [4, 5]. A WHO global polio laboratory network of more than 140 laboratories exists where stool samples from AFP cases are processed and analysed for the presence of PV. In addition, several countries have successfully conducted ES for PV for many years. Virus genotypes of isolates found in sewage samples from Egypt, Nigeria, India, Afghanistan and Pakistan closely matched those found in stool samples from AFP cases which has helped tracing the elimination of wild PV in some areas [6-9]. This approach has also been used to monitor the disappearance of vaccine virus after vaccination campaigns in countries such as Cuba, New Zealand and Japan [10-12]. Furthermore, widespread circulation of WPV1 imported from Pakistan, in the absence of reported paralytic cases, was identified in Israel in 2013 as the virus was found in sewage samples across the country clearly suggesting that PV can circulate for long periods of time undetected in countries using IPV, even in those with high vaccination coverage such as Israel [13, 14]. Thus, ES is seen as a powerful tool to

support the GPEI endgame helping to identify any remaining PV transmission. With this in mind, a pilot study was set up to analyse sewage samples from London for the presence of PV. The last polio case due to wild PV in the UK was reported in 1982, but the country switched to the exclusive use of IPV in 2004 and has frequent population exchanges with areas where PV is still circulating and/or OPV is being used routinely. We concentrated sewage samples using two different methods and added them to cell cultures susceptible for HEV infection. HEV isolates were then characterised by nucleotide sequencing. Nucleotide sequence information of HEV clinical and environmental isolates is limited, often restricted to HEV species-specific real-time PCR positive/negative results and/or short genomic sequences that can help identifying the HEV serotype but have limited use for detailed phylogenetic analysis. NGS metagenomics and target-specific techniques have recently been described by us and others to obtain nucleotide sequences of HEV strains present in stool, sewage and cell culture samples [15-21]. A novel approach using NGS analysis is described here to quickly obtain nearly whole-genome sequences of polio and non-polio HEVs present in cell cultures infected with sewage concentrates. This approach has the power to detect known HEVs as well as divergent strains and novel serotypes providing information that should assist with tracing the source and transmission of HEVs, including PV, in human populations.

MATERIALS AND METHODS

Sample collection

One litre composite sewage samples were collected during a 24-hour period once monthly at Beckton Sewage Treatment Works in London (UK), currently treating the waste of 3.5 million people. Samples were transported to the laboratory on the same day and processed

within one day of arrival. A total of twelve samples were collected from June 2016 to May 2017.

Sample processing

Raw sewage samples were processed by two different methods previously described: two-phase (TP) aqueous separation system [22] and concentration by filtration and centrifugation (FC) using Centriprep® YM-50 centrifugal concentration devices (Merck) [23].

Cell lines and Virus Isolation

Virus isolation in cell cultures was performed according to WHO recommendations [24]. Rhabdomyosarcoma (RD) and mouse L20B cells expressing the human PV receptor were used as detailed in the Supplementary Data section

Intratypic differentiation (ITD) of PV isolates by real-time RT-PCR (rRT-PCR) assays.

Conventional PV ITD rRT-PCR was performed using a PV diagnostic rRT-PCR kit provided by the US CDC using a Rotorgene Q (Qiagen) platform and following kit instructions [25].

Nucleotide sequence analysis of the VP1 coding region of PV isolates.

RT-PCR fragments containing the VP1 coding region were generated from purified viral RNAs by one-step RT-PCR using a SuperScript III (Invitrogen) system and primers Y7 (5′-GGGTTTGTGTCAGCCTGTAATGA-3′) and Q8 (5′-AAGAGGTCTCTRTTCCACAT-3′) [25]. Amplified products were purified using QIAquick Gel Extraction Kit (Qiagen,) and sequenced by Sanger using an ABI Prism 3130 genetic analyser (Applied Biosystems).

Preparation of RT-PCR templates for NGS analysis

Whole-genome PV RT-PCR fragments were amplified from purified RNAs from infected cells by one-step RT-PCR as described [20]. In addition, RT-PCR products were also generated by Sequence-Independent Single-Primer amplification (SISPA) of purified RNAs as described elsewhere [18, 19, 26]. Two primer sets were used to generate two different dsDNA templates from each sample. Details of primers used and amplification conditions are provided in the Supplementary Data section.

Generation of sequencing libraries and quality trimming of NGS reads

Sequencing libraries were prepared using Nextera XT reagents and sequenced on a MiSeq using a 2 x 301 paired-end v3 Flow Cell and manufacturer's protocols (Illumina). Raw sequence data were imported into Geneious R10 software (Biomatters) and sequence files processed using a custom workflow. Parameters used for quality trimming of NGS reads are available in the Supplementary Data section. Raw fastq files are available from NCBI's Sequence Read Archive (SRA) under project code PRJNA417977.

Generation of polio and non-polio HEV sequence contigs from NGS data by reference-guided or *de novo* assembly of filtered NGS reads

The filtered reads were mapped to a set of HEV sequences using a curated HEV sequence database and contig sequences were generated. Filtered reads were then iteratively reassembled to consensus sequences from the longest contigs with an aim to build whole-genome contig sequences. Final consensus sequences were obtained by assigning the most common nucleotide sequence at each nucleotide position. The filtered reads were also independently assembled *de novo* using similar assembly conditions. Whole-genome contig sequences were generated following the same analysis workflow as above. As a result, we obtained nearly whole-genome nucleotide sequences for various HEV strains in each sample,

including various PV strains. Results using the two different assembly approaches were almost identical. Manual analyses for visualizing and quantifying assembly results were performed throughout the process. Full details regarding settings used for genome assembly are provided in the Supplementary Data section. Consensus nucleotide sequences are available from DDBJ/EMBL/GenBank with accession numbers MG451802 to MG451811

Phylogenetic analysis of polio and non-polio HEV isolates

The closest virus relatives to the London sewage HEV isolates were identified using the RIVM and BLAST online sequence analysis tools [27, 28] and HEV serotypes were assigned on the basis of their VP1 sequence. Whole genome PV sequences were aligned to type 1 Sabin PV AY184219 or type 3 Sabin PV AY184221 reference genome sequences using the program ClustalW (within Geneious) to identify mutations and/or recombination events. Any PV3/PV1 recombinant genomic structure was confirmed by independently mapping filtered reads to the PV1 and PV3 Sabin reference sequences mentioned above, with a minimum 50 base overlap, minimum overlap identity of 95%, maximum 5% mismatches per read and both end pair reads mapping. Percentages of maximum coverage were calculated. Mapped reads against each reference were combined in a graph for each NGS product.

RESULTS

Concentration of sewage samples and virus isolation in cell cultures

Twelve samples from London sewage were analysed. 500 ml and 120 ml of raw sewage from each sample were used for the TP and FC methods, respectively**.** Typically, 10 ml and 4 ml were obtained using each of the concentration procedures, which means approximate concentration factors of 50 and 30 times, respectively. Aliquot samples of sewage concentrates were used to infect RD and L20B cells. Sewage concentrates from all 12

175 samples produced cytopathic effect (CPE) in RD cells after 3-4 days. Only two samples, from
176 September 2016 and January 2017, produced CPE in L20B cells. As shown in Table 1,
177 concentrates from both sewage samples, obtained with both concentration methods produced
178 CPE in L20B cells but only in a proportion of flasks, 5 out of 10 flasks with concentrates
179 from September 2016 and 2 out of 12 flasks with concentrates from January 2017.
180
181 **Typing of PV isolates by ITD rRT-PCR and VP1 sequencing**
182 PV strains present in L20B cell cultures showing CPE were initially characterized by ITD
183 rRT-PCR. As shown in Table 1, all PV isolates from September 2016 were identified as PV3
184 and Sabin (vaccine)-Like (SL). The sewage sample from January 2017 produced both PV1
185 and PV3 isolates. The TP-RD isolate from January 2017 was characterised as a PV3-SL
186 strain while the FC-L20B$_1$ isolate was found to be a PV1-SL strain. The PV serotype and SL
187 classification were confirmed by nucleotide sequencing of the VP1 coding region. PV
188 isolates showed very few VP1 mutations from Sabin references (Table 1).
189
190 **Genetic characterization of PV isolates by NGS**
191 The genomes of the seven selected PV isolates were further characterized by high resolution
192 sequencing. Whole-genome PV RT-PCR products were obtained and analysed by NGS.
193 Consensus sequences for each of the PV isolates were generated by *de novo* assembly. As
194 shown in Table 2, all five isolates from September 2016 were confirmed as PV3-SL by NGS
195 analysis, very closely related to the Sabin 3 vaccine strain and containing reversions at known
196 Sabin 3 attenuation sites, nucleotide 472 in the 5'NCR and nucleotide 2493 coding for capsid
197 amino acid VP1-6. Few additional nucleotide differences from Sabin 3 were found in the PV
198 isolates, all showing unique sequences. The TP-RD PV isolate from the 2017 sample was
199 also a PV3-SL strain and also contained reversions at nucleotides 472 and 2493. However,

the whole-genome consensus sequence of this 2017 isolate revealed a type 3/type 1 PV recombinant structure with a crossover point between nucleotides 4904 and 4914. As shown in Fig. 1, NGS reads from this isolate mapped to both Sabin 1 and Sabin 3 reference sequences in different regions of the genome. This finding was reproducible using the random PCR sequencing approach discussed in next section (Fig. 1). Finally, the 2017 L20B isolate from the FC concentrate was confirmed as PV1, very closely related to the Sabin 1 vaccine strain, with only one nucleotide change from the vaccine strain. The VP1 sequences of all PV isolates determined by NGS analysis were identical to those obtained by the Sanger method

Sequence analysis of HEV mixtures found in RD cells

RT-PCR products generated using random primers were also analysed by NGS with an aim to sequence any other HEV strain that might be present in the infected RD cell cultures. We first analysed virus control samples (Reference strains Enterovirus 20 JV-1, Echovirus 7 Wallace, Coxsackievirus B5 Faulkner and PV Sabin 1), to test and optimize our ability to detect and identify HEV mixtures. NGS reads were filtered and analysed as described in Materials and Methods. Nucleotide sequences obtained from RA01 and M13 random RT-PCR products from these reference strains were almost identical except in the extreme ends where sequence coverage was low. These extreme regions were discarded from the final consensus sequence assigned to the virus. As shown in Fig. 2A, the results showed excellent coverage across most of the genome for all viruses and nearly whole-genome sequences (≥95% of the genome) were obtained for all four reference virus strains in both single and mixed samples. Sequences in these single and mixed samples were identical and highly similar (>99.8%) to the corresponding Genebank Sanger sequences (Table 3).

The same analytical process was followed for TP-RD isolates from the two London sewage samples. Six and four different HEV strains were identified in the September 2016 and January 2017 samples, respectively, including PV3 strains found in both samples. Results are shown in Fig. 2 and Table 3. Again, nearly whole-genome sequences (>90% of genome) were obtained for all virus strains identified in both samples and nucleotide sequences determined by either RA01 or M13 primers were almost identical. The relative proportions of sequence reads mapping to each of the different HEV strains identified in each sample are shown in Fig. 3. There were some differences in the results for the September 2016 sample in that Echovirus 3 sequence reads were the most prevalent in sequences from the M13 RT-PCR while Coxsackievirus B3 reads showed the highest proportion when the RA01 RT-PCR product was analysed. These differences likely reflect some degree of bias in the amplification of viral genomes from some strains using different primer sets. Echovirus 7 was the most prevalent strain in the 2017 RD culture with > 92% of reads mapping to this strain. The proportion of sequence reads mapping to PV sequences was relative low, with only 6.28% and 0.76% of the total number of reads mapping to PV3 sequences for the September 2016 and January 2017 samples, respectively. The closest virus relatives to the London sewage non-polio HEV isolates were identified by BLAST analysis of VP1 sequences. VP1 genetic similarities to related non-polio HEV strains ranged between 87.4% to 96.9 % for nucleotides and 97.9% to 100% for amino acid sequences, confirming the HEV serotype assignment (Table 3). Two different Echovirus 7 strains were found in the September 2016 sewage sample with 83.4% and 98.2% nucleotide and amino acid sequence identity between them across the whole genome, respectively.

DISCUSSION

All twelve sewage samples from London (UK), tested as part of a pilot ES study, were positive for HEVs and two of them contained PV. Type 1 and 3 vaccine-like PV isolates were found in samples taken in September 2016 and January 2017. Both concentration methods used resulted in PV isolation and we found the FC method to be simpler, quicker, less technically demanding and free from bacterial contamination as compared to the TP separation system used in most WHO laboratories. Finding PV in these samples was rather unexpected as the UK has exclusively used IPV for polio immunisation since 2004 [29]. However, vaccine-like PVs have occasionally been found in environmental samples collected in countries using only IPV, presumably imported by people coming from countries where OPV is still used [30]. The London sewage sampling site serves areas that include large migrant groups from countries where there is still OPV use.

Nearly whole-genome viral sequences were rapidly generated by NGS analysis of RT-PCR products obtained using random or PV-specific primers. Five different vaccine-like PV3 strains were identified in the 2016 sewage sample, all containing few mutations from PV3 Sabin vaccine strain. These included reversions at attenuation mutations at nucleotides 472 in domain V of the 5'NCR, with a role in protein translation, and 2493 coding for amino acid VP1-6 and possibly involved in virus particle stability [31, 32]. A different PV3 strain was found in the 2017 sewage sample also containing few mutations from Sabin 3 and the two reversion mutations mentioned above. This 2017 PV3 isolate had a PV3/PV1 recombinant genomic structure with a crossover point locating in the region coding for non-structural protein 2C. In addition, a PV1 vaccine-like isolate with a single mutation from the Sabin 1 vaccine strain was also found in the sewage sample from 2017. The fact that all PV isolates from sewage contained a very low number of mutations from the Sabin vaccine strains indicates a very short period of replication/transmission in humans, from several days

to few weeks after vaccination, with these PV strains possibly having been excreted by just one or few recent vaccinees and/or their immediate contacts.

It is striking that significant sequence differences were found between virus isolates found in the same sewage sample, including the presence of unique PV3 genetic variants in the sample from September 2016 and a PV1 strain together with a PV3/PV1 recombinant virus in the sample from January 2017. This is likely due to PV being in very low concentration in sewage leading to a strong sampling effect that is reflected when using different aliquots of concentrate to infect different cell culture flasks. This is in agreement with the fact that only a proportion of L20B cell culture flasks incubated with sewage concentrates showed CPE. Indeed, complex virus mixtures are commonly found in sewage samples, with parallel cell culture flasks infected with aliquots of the same sewage concentrate producing very different results. This might include PV in different homotypic and/or heterotypic mixtures often in combination with non-polio HEVs [6]. Furthermore, it is not at all unexpected that vaccinees excrete virus mixtures containing mixed serotype and recombinant variants [20, 33]. PV3/PV1 recombinant strains similar to the 2017 isolate found in this study are commonly found in stool samples from vaccinees taken soon after vaccination in combination with other non-recombinant and recombinant variants from all serotypes present in the OPV vaccine [33, 34].

Considering that there is no OPV use in the UK and this sewage site covers a population of more than 3.5 million people, we conclude that our ES set up is sensitive enough to detect low concentrations of PV. In this context, it is expected that, should transmission of WPV1 or cVDPV2 occur in the UK following importation, it would be readily detected using our system. Importantly, no such PV isolates were detected in the London samples. It is also reassuring that no PV2 vaccine sequences were detected in any of the PV isolates identified since the type 2 component was removed from OPV in August

2016. The results shown here are compatible with the viruses found being derived from vaccinees that received type 1 and 3 bOPV.

It is also expected that several non-polio HEV strains would be present in sewage concentrates, and hence in RD cultures from them, reflecting their circulation in human populations. Using NGS analysis we found 5 and 3 non-polio HEV strains of different serotypes, in addition to PV, in the London sewage samples from September 2016 and January 2017, respectively. A recent report, also using NGS, described the presence of Coxsackievirus B1, B3 and B5 strains in single RD cultures from sewage samples in Pakistan [17]. All non-polio HEV strains found in the 2016 and 2017 London samples correspond to species B HEV serotypes, a common finding that, rather than reflecting the actual prevalence of HEV serotypes in human populations, might be a consequence of the high sensitivity of RD cells for infection with species B HEVs [35]. Virus strains from all four HEV species can infect RD cells as shown by the analysis of stool extracts from AFP cases [36]. However, the complex HEV composition in sewage means that species B HEVs would likely outcompete viruses from other species when growing on RD cells. Indeed, some studies have shown that HEV strains from all four species A, B, C and D are frequently found in stool and sewage samples [7, 30, 37, 38]. Identifying them has required the use of several cell culture systems and/or sequencing RT-PCR products from multiple PCR reactions or from a large number of cDNA clones. Our NGS approach can reveal the presence of several non-polio HEV strains in a single cell culture system providing nearly whole-genome nucleotide sequence information of each of them.

PV strains are known to replicate efficiently on RD cells but RD infected cultures from the two London samples were found to contain only a low proportion of PV relative to other species B non-polio HEV strains. This observation highlights the relevance of using L20B cells to increase the sensitivity for PV detection in clinical and environmental samples.

Several studies have described the microbiome in stool and sewage samples but information on HEV content is very limited. Our results show the great value of using NGS technology for HEV surveillance, particularly for PV, as it can detect low concentrations of PV possibly excreted by one or few individuals and can quickly provide whole-genome genetic information including evidence for recombination events. Identifying genetic features that link PV isolates is essential to establish temporal and geographical relationships between them that help tracing virus transmission. Previous work in our laboratory using NGS for the analysis of both vaccine products and isolates from vaccinees has shown that these methods can also accurately identify PV strains in homotypic and heterotypic mixtures [18, 20], a critical feature that will help identifying wild PV and cVDPV strains in a background of OPV. The use of NGS methods for HEV identification and characterisation represents a major step forward in HEV molecular diagnosis and will greatly contribute to improve our knowledge on HEV circulation patterns in human populations and their association with human disease.

Notes

This report is independent research which, in part, was funded by the Department of Health Policy Research Programme (NIBSC Regulatory Science Research Unit, 044/0069)
All authors declare no conflict of interests.

ACKNOWLEDGEMENTS

We thank Jim McLauchlin and Chris Roberts for organising the regular composite sampling of sewage from Becton between June 2016 and May 2017 to enable this project to take place. We would also like to thank the NGS team at NIBSC for their technical support.

Figure legends

Figure 1. Coverage of Sabin PV 1 and 3 genomes following RT-PCR and sequencing by NGS of the TP-RD PV isolate from the January 2017 sewage sample. Filtered reads from (A) PV-specific, (B) M13-random or (C) RA01-random RT-PCR products obtained from RNA purified from infected cells were independently mapped against type 1 Sabin PV AY184219 (red line) or type 3 Sabin PV AY184221 (blue line) reference genome sequences. Mapped reads against each reference were combined in the graph for each NGS product. Percentages of maximum coverage are reported. The results were identical for all RT-PCR products and identified a type 3/type 1 PV recombinant with a crossover point between nucleotides 4904 and 4914.

Figure 2. Genome coverage of HEV genomes following RT-PCR and sequencing by NGS of TP-RD isolates from sewage samples. Filtered reads from the virus control mixture (A, B), sewage sample from September 2016 (C, D) and sewage sample from January 2017 (E, F) were mapped to the respective final HEV consensus sequences identified by *de novo* assembly. Results obtained with NGS reads generated with M13-random (A, C, E) and RA01-random (B, D, F) RT-PCR products are shown. The number of sequence reads at each nucleotide position is shown for each HEV strain. The results show excellent coverage across most of the genome for all viruses.

Figure 3. Percentage of filtered reads mapping to VP1 sequences. Filtered reads from the virus control mixture (A, B), sewage sample from September 2016 (C, D) and sewage sample from January 2017 (E, F) were mapped to VP1 consensus sequences of each of the HEV strains identified by *de novo* assembly. Results for M13-random (A, C, E) and RA01-random (B, D, F) RT-PCR products are shown. Percentages of total reads mapping to each of the

HEV VP1 sequences are indicated. No HEV isolates were identified in any of the negative controls analysed. The results show that the proportion of PV sequences found in infected RD cells was low, particularly in the sample from January 2017.

REFERENCES

1. Poliomyelitis Fact Sheet. World Health Organization. Available at: http://www.who.int/mediacentre/factsheets/fs114/en/. 2017.

2. Wild polio and vaccine derived polio in Nigeria. Available at: http://www.who.int/csr/don/06-october-2016-polio-nigeria/en/. 2017.

3. Circulating vaccine-derived poliovirus type 2 – Syrian Arab Republic. Available at: http://www.who.int/csr/don/13-June-2017-polio-syrian-arab-republic/en/. 2017.

4. WHO-recommended surveillance standard of poliomyelitis. Available at: http://www.who.int/immunization/monitoring_surveillance/burden/vpd/surveillance_type/active/poliomyelitis_standards/en/. 2017.

5. Diop OM, Kew OM, de Gourville EM, Pallansch MA. The Global Polio Laboratory Network as a Platform for the Viral Vaccine-Preventable and Emerging Diseases Laboratory Networks. Journal of Infectious Diseases **2017**; 216:S299-S307.

6. Hovi T, Blomqvist S, Nasr E, et al. Environmental surveillance of wild poliovirus circulation in Egypt--balancing between detection sensitivity and workload. J Virol Methods **2005**; 126:127-34.

7. Ndiaye AK, Diop PA, Diop OM. Environmental surveillance of poliovirus and non-polio enterovirus in urban sewage in Dakar, Senegal (2007-2013). Pan Afr Med J **2014**; 19:243.

8. Johnson Muluh T, Hamisu AW, Craig K, et al. Contribution of Environmental Surveillance Toward Interruption of Poliovirus Transmission in Nigeria, 2012-2015. J Infect Dis **2016**; 213 Suppl 3:S131-5.

9. Deshpande JM, Shetty SJ, Siddiqui ZA. Environmental surveillance system to track wild poliovirus transmission. Appl Environ Microb **2003**; 69:2919-27.

10. Mas Lago P, Caceres VM, Galindo MA, et al. Persistence of vaccine-derived poliovirus following a mass vaccination campaign in Cuba: implications for stopping polio vaccination after global eradication. Int J Epidemiol **2001**; 30:1029-34.

11. Huang QS, Greening G, Baker MG, et al. Persistence of oral polio vaccine virus after its removal from the immunisation schedule in New Zealand. Lancet **2005**; 366:394-6.

12. Nakamura T, Hamasaki M, Yoshitomi H, et al. Environmental surveillance of poliovirus in sewage water around the introduction period for inactivated polio vaccine in Japan. Appl Environ Microbiol **2015**; 81:1859-64.

13. Shulman LM, Gavrilin E, Jorba J, et al. Molecular epidemiology of silent introduction and sustained transmission of wild poliovirus type 1, Israel, 2013. Euro Surveill **2014**; 19:20709.

14. Shulman LM, Martin J, Sofer D, et al. Genetic analysis and characterization of wild poliovirus type 1 during sustained transmission in a population with >95% vaccine coverage, Israel 2013. Clin Infect Dis **2015**; 60:1057-64.

15. Victoria JG, Kapoor A, Li LL, et al. Metagenomic Analyses of Viruses in Stool Samples from Children with Acute Flaccid Paralysis. Journal of Virology **2009**; 83:4642-51.

16. Bessaud M, Sadeuh SA, Joffret ML, et al. Whole Genome Sequencing of Enterovirus species C Isolates by High-Throughput Sequencing: Development of Generic Primers. Front Microbiol **2016**; 7.

17. Furtak V, Roivainen M, Mirochnichenko O, et al. Environmental surveillance of viruses by tangential flow filtration and metagenomic reconstruction. Euro Surveill **2016**; 21.

18. Mee ET, Minor PD, Martin J. High resolution identity testing of inactivated poliovirus vaccines. Vaccine **2015**; 33:3533-41.

19. Montmayeur AM, Ng TFF, Schmidt A, et al. High-Throughput Next-Generation Sequencing of Polioviruses. J Clin Microbiol **2017**; 55:606-15.

20. Stern A, Yeh MT, Zinger T, et al. The Evolutionary Pathway to Virulence of an RNA Virus. Cell **2017**; 169:35-46 e19.

21. Terry FFN, Marine R, Wang CL, et al. High Variety of Known and New RNA and DNA Viruses of Diverse Origins in Untreated Sewage. Journal of Virology **2012**; 86:12161-75.

22. Poyry T, Stenvik M, Hovi T. Viruses in Sewage Waters during and after a Poliomyelitis Outbreak and Subsequent Nationwide Oral Poliovirus Vaccination Campaign in Finland. Appl Environ Microb **1988**; 54:371-4.

23. Harvala H, Calvert J, Van Nguyen D, et al. Comparison of diagnostic clinical samples and environmental sampling for enterovirus and parechovirus surveillance in Scotland, 2010 to 2012. Euro Surveill **2014**; 19.

24. Guidelines for environmental surveillance of poliovirus circulation. Available at:http://archives.who.int/vaccines-documents/DocsPDF03/www737.pdf. Accessed 25th September 2017.

25. Burns CC, Kilpatrick DR, Iber JC, Chen Q, Kew OM. Molecular Properties of Poliovirus Isolates: Nucleotide Sequence Analysis, Typing by PCR and Real-Time RT-PCR. Methods Mol Biol **2016**; 1387:177-212.

26. 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.

27. Madden T. The BLAST Sequence Analysis Tool. 2002 Oct 9 [Updated 2003 Aug 13]. The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2002. In: McEntyre J OJ, ed. The NCBI Handbook. The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2002-. .

446 28. Kroneman A, Vennema H, Deforche K, et al. An automated genotyping tool for
447 enteroviruses and noroviruses. J Clin Virol **2011**; 51:121-5.
448 29. Minor PD. The polio-eradication programme and issues of the end game. J Gen Virol
449 **2012**; 93:457-74.
450 30. Pellegrinelli L, Binda S, Chiaramonte I, et al. Detection and distribution of culturable
451 Human Enteroviruses through environmental surveillance in Milan, Italy. J Appl Microbiol
452 **2013**; 115:1231-9.
453 31. Minor PD. Poliovirus biology. Structure **1996**; 4:775-8.
454 32. Weeks-Levy C, Tatem JM, DiMichele SJ, Waterfield W, Georgiu AF, Mento SJ.
455 Identification and characterization of a new base substitution in the vaccine strain of Sabin 3
456 poliovirus. Virology **1991**; 185:934-7.
457 33. Minor PD. The molecular biology of poliovaccines. J Gen Virol **1992**; 73 (Pt 12):3065-
458 77.
459 34. Cuervo NS, Guillot S, Romanenkova N, et al. Genomic features of intertypic recombinant
460 sabin poliovirus strains excreted by primary vaccinees. J Virol **2001**; 75:5740-51.
461 35. Faleye TO, Adeniji JA. Enterovirus Species B Bias of RD Cell Line and Its Influence on
462 Enterovirus Diversity Landscape. Food Environ Virol **2015**; 7:390-402.
463 36. Fernandez-Garcia MD, Kebe O, Fall AD, Ndiaye K. Identification and molecular
464 characterization of non-polio enteroviruses from children with acute flaccid paralysis in West
465 Africa, 2013-2014. Sci Rep **2017**; 7:3808.
466 37. Apostol LN, Imagawa T, Suzuki A, et al. Genetic diversity and molecular
467 characterization of enteroviruses from sewage-polluted urban and rural rivers in the
468 Philippines. Virus Genes **2012**; 45:207-17.

469 38. Wang H, Tao Z, Li Y, et al. Environmental surveillance of human enteroviruses in
470 Shandong Province, China, 2008 to 2012: serotypes, temporal fluctuation, and molecular
471 epidemiology. Appl Environ Microbiol **2014**; 80:4683-91.

472