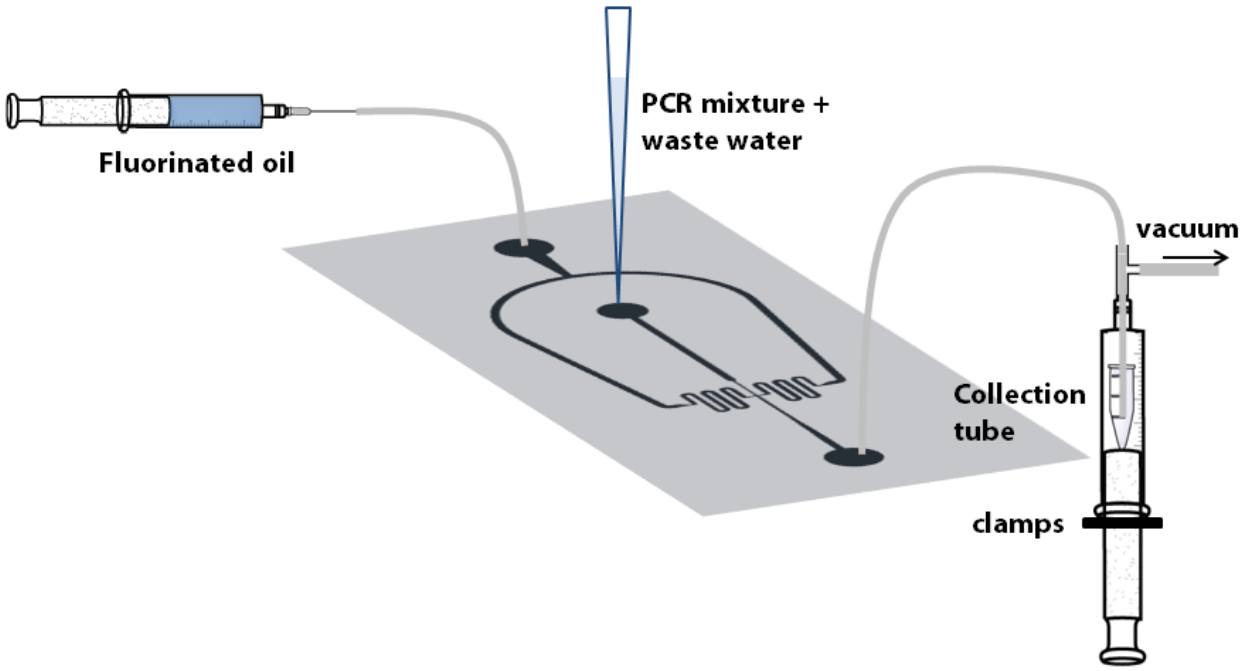
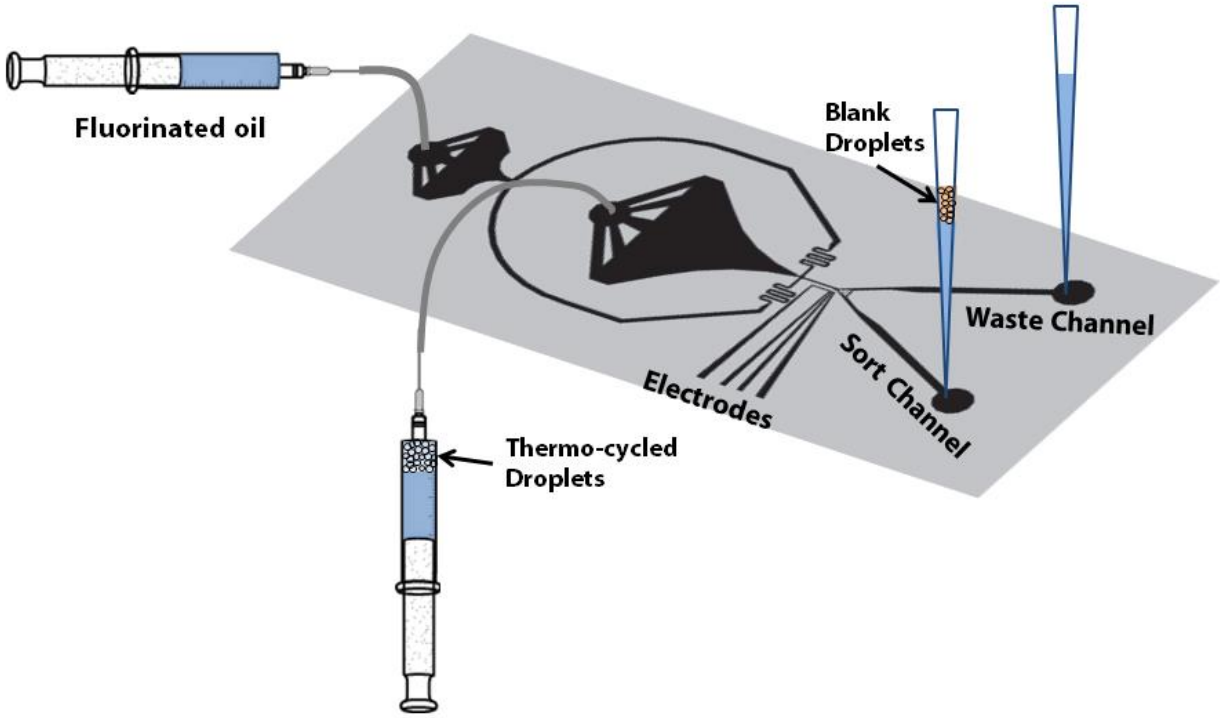
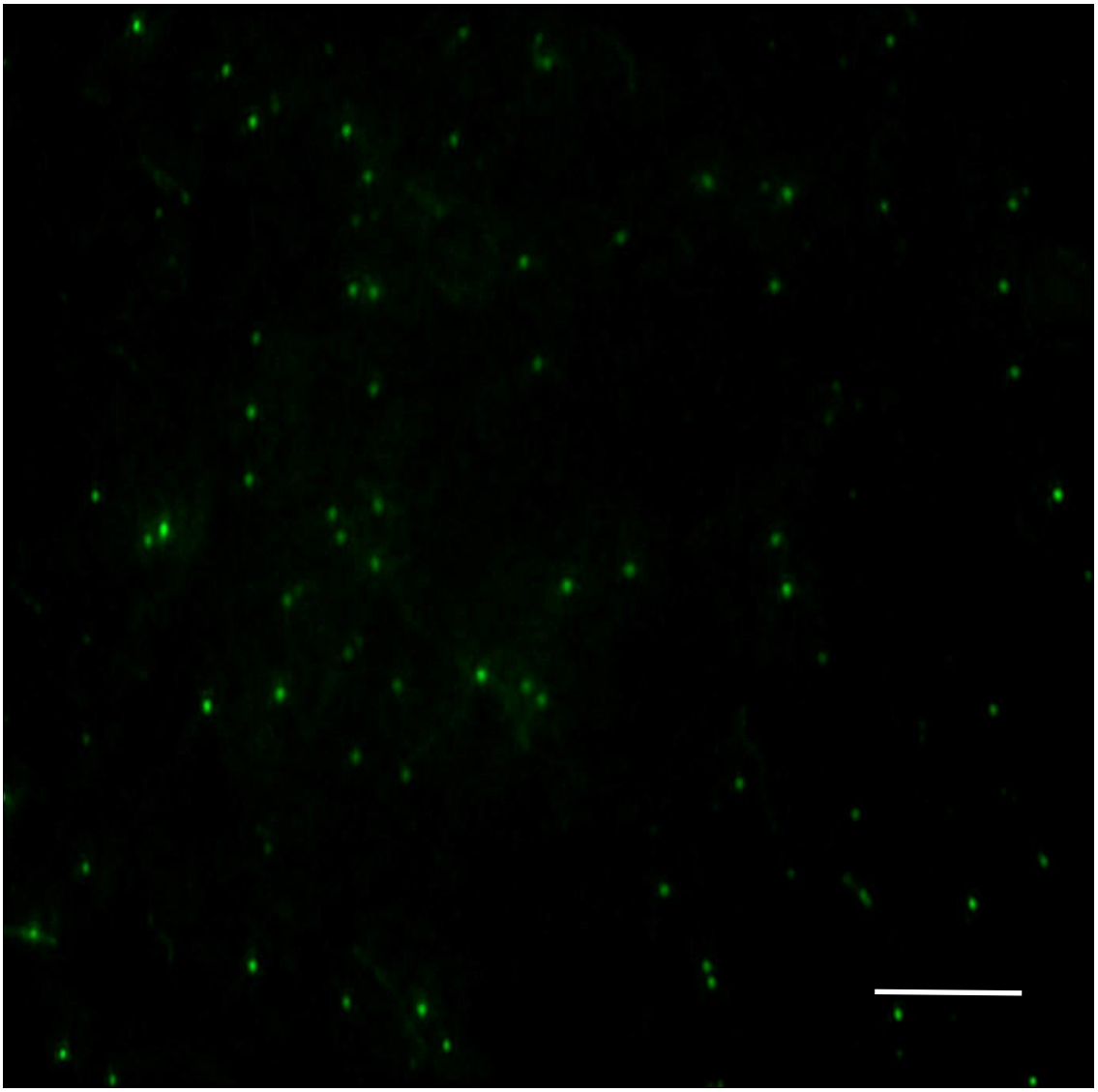
Supplementary Figure S1. Illustration of the microfluidic droplet maker



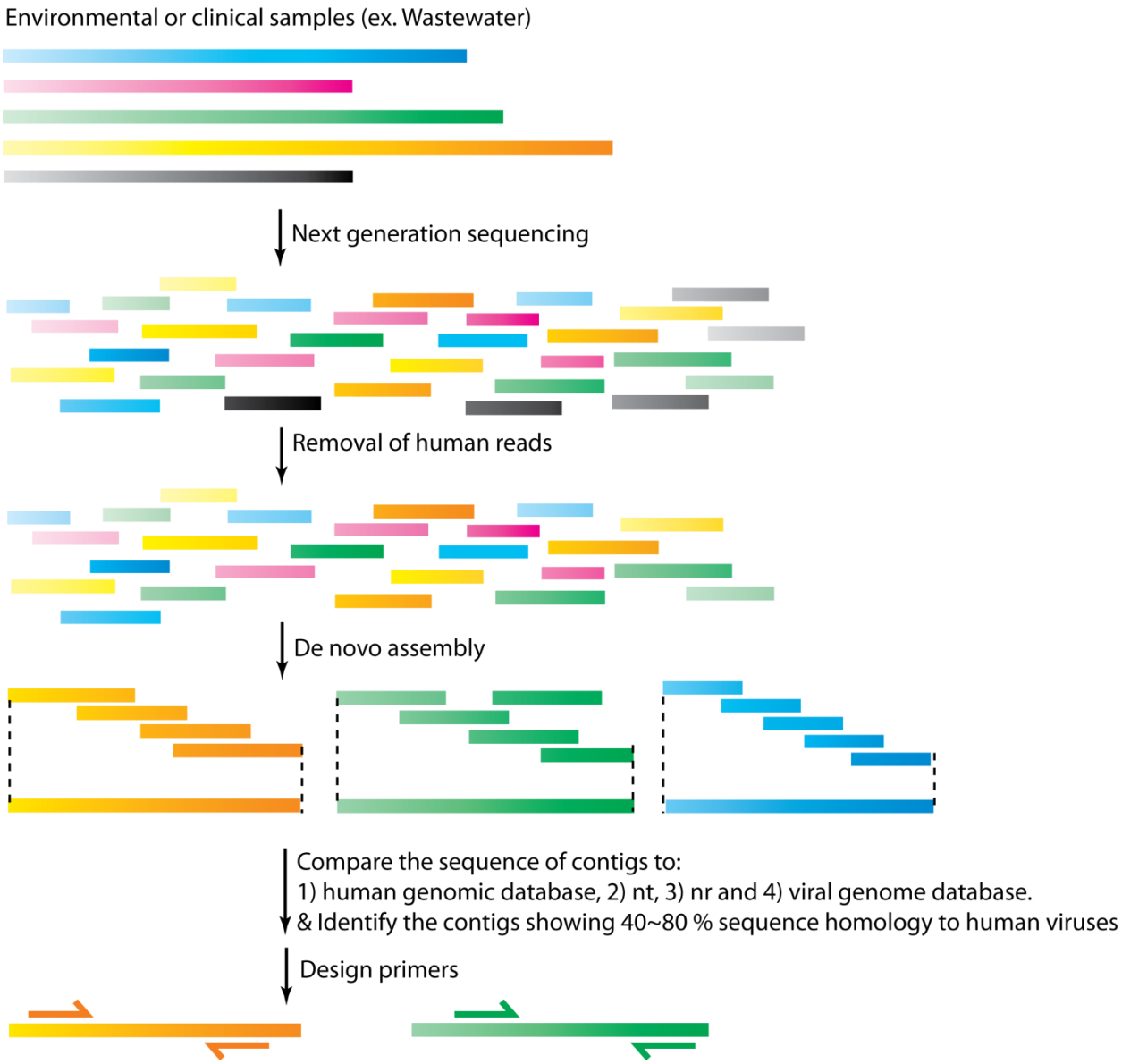
Supplementary Figure S2. Illustration of the microfluidic sorter



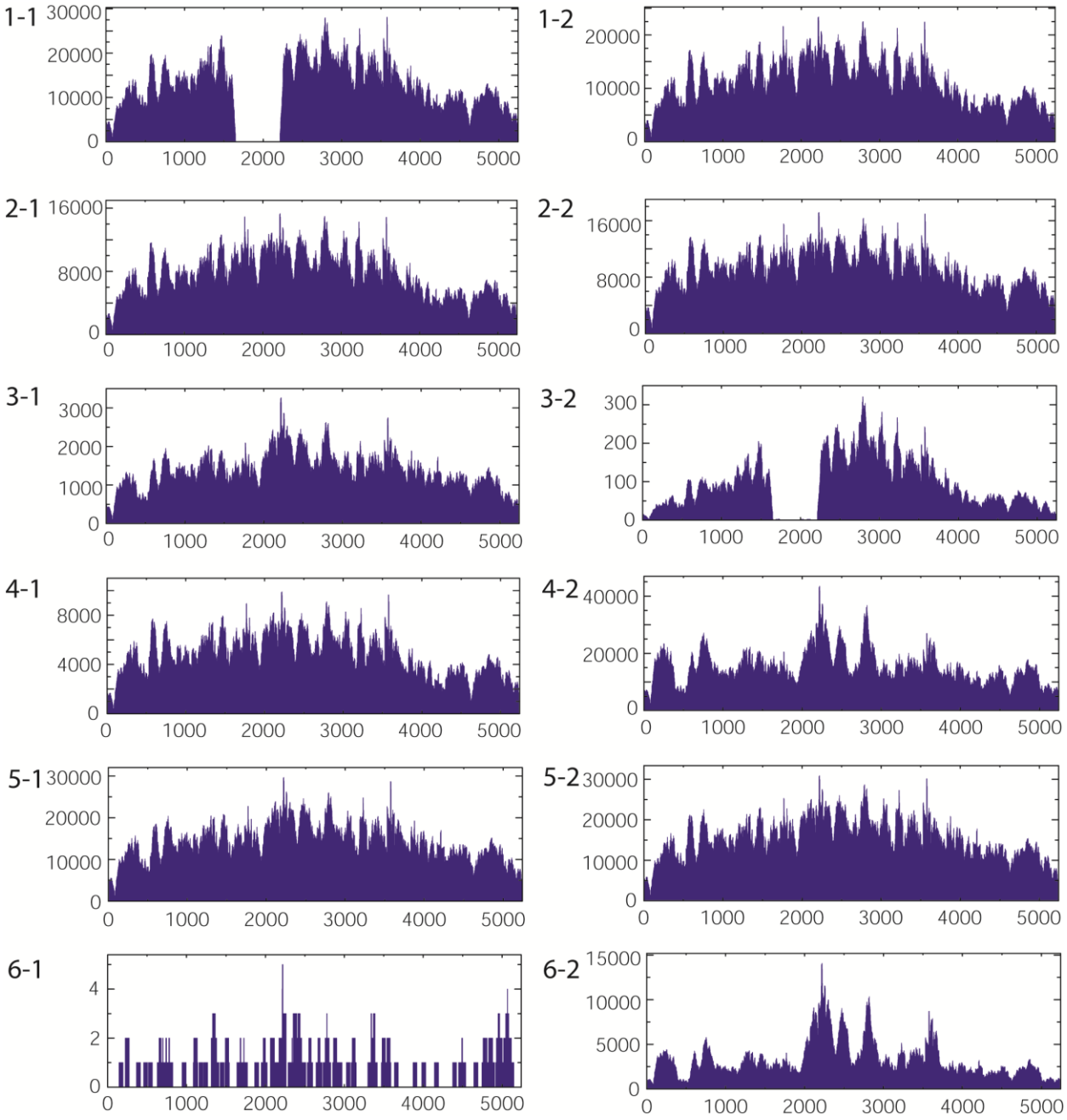
Supplementary Figure S3. Fluorescence image of wastewater after staining virus particles using SYBR Gold. Wastewater was clarified by low-speed centrifugation and the supernatant was affixed to a 0.02 micron aluminum oxide filter. The filter was stained with SYBR Gold after clarified wastewater was filtered through a 0.22 micron filter. Scale bar = 10 µm.



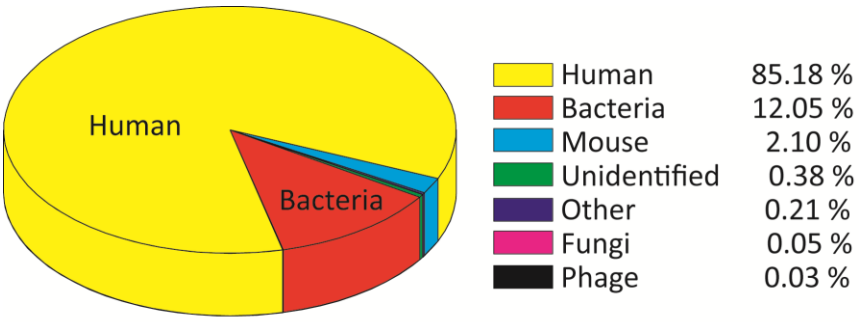
Supplementary Figure S4. Scheme for primer design for novel viruses[1]. First, next generation sequencing is performed on environmental samples. High quality, non-human reads are assembled to yield longer contigs. The contigs are then sequentially compared against (i) the human genome using BLASTN, (ii) GenBank nt database using BLASTN and TBLASTX and (iii) the NCBI viral genome database using TBLASTX. Contigs showing 40~80% sequence homology to human viruses are identified as "contig-of-interest". Based on the sequence of these contigs, primers specific to each contig-of-interest can be designed.



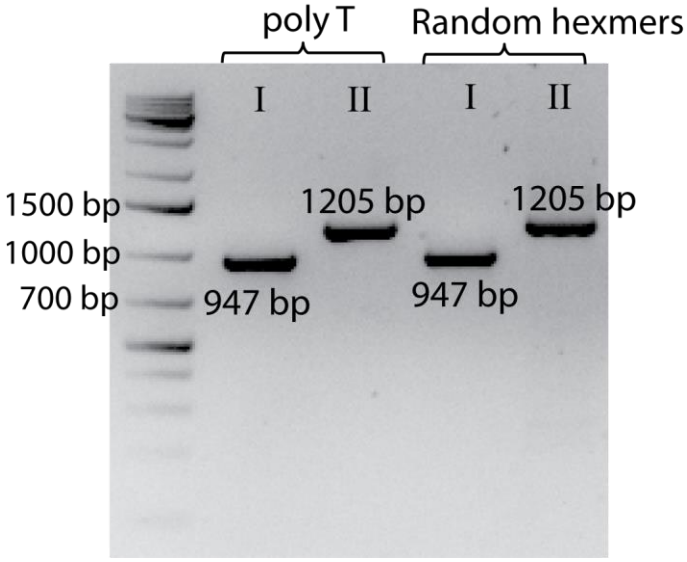
Supplementary Figure S5. Reference mapping results on the enriched SV40 samples from different sets of sorting experiments.



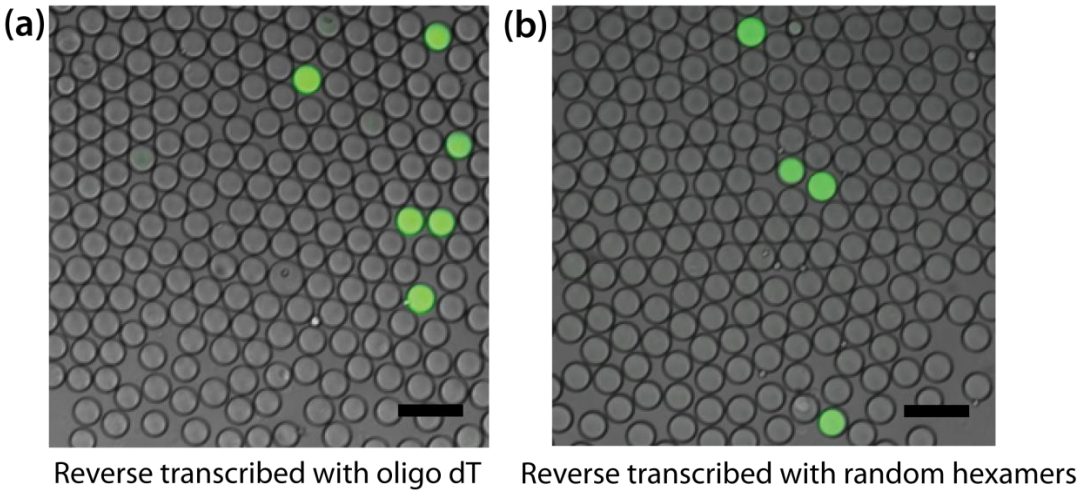
Supplementary Figure S6. Summary of the origin of contaminant sequences. To analyse the origin of contaminant sequences, a BLAST search is performed on non-SV40 contigs against human, nt, nr, and viral ref-seq databases for annotation. Sequences are taxonomically binned based on the top of BLAST hit. A pie chart is created from a combined dataset of non-SV40 contigs and the sequence reads which align to human databases and are omitted for assembly. This chart shows the percentage of these non-SV40 reads into the various bins.



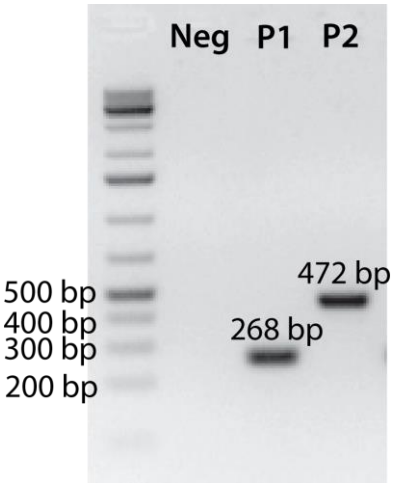
Supplementary Figure S7. The cDNA copies of the genome of norovirus are synthesized by extracting the RNA genome from virions and performing reverse transcription. Both oligo dT and random hexamer primers can be used for reverse-transcription. The reverse-transcribed genomes are then mixed with PCR mixtures and two sets of primers targeting different regions of the genome (Table S4) and undergo thermo-cycling. Primer set I is expected to generate amplicons of 947 bp and Primer II is expected to generate amplicons of 1205 bp. As expected, in all cases, amplicons of the predicted size are generated after PCR. This result indicates that the genome of norovirus is successfully reverse-transcribed into the cDNA copies.



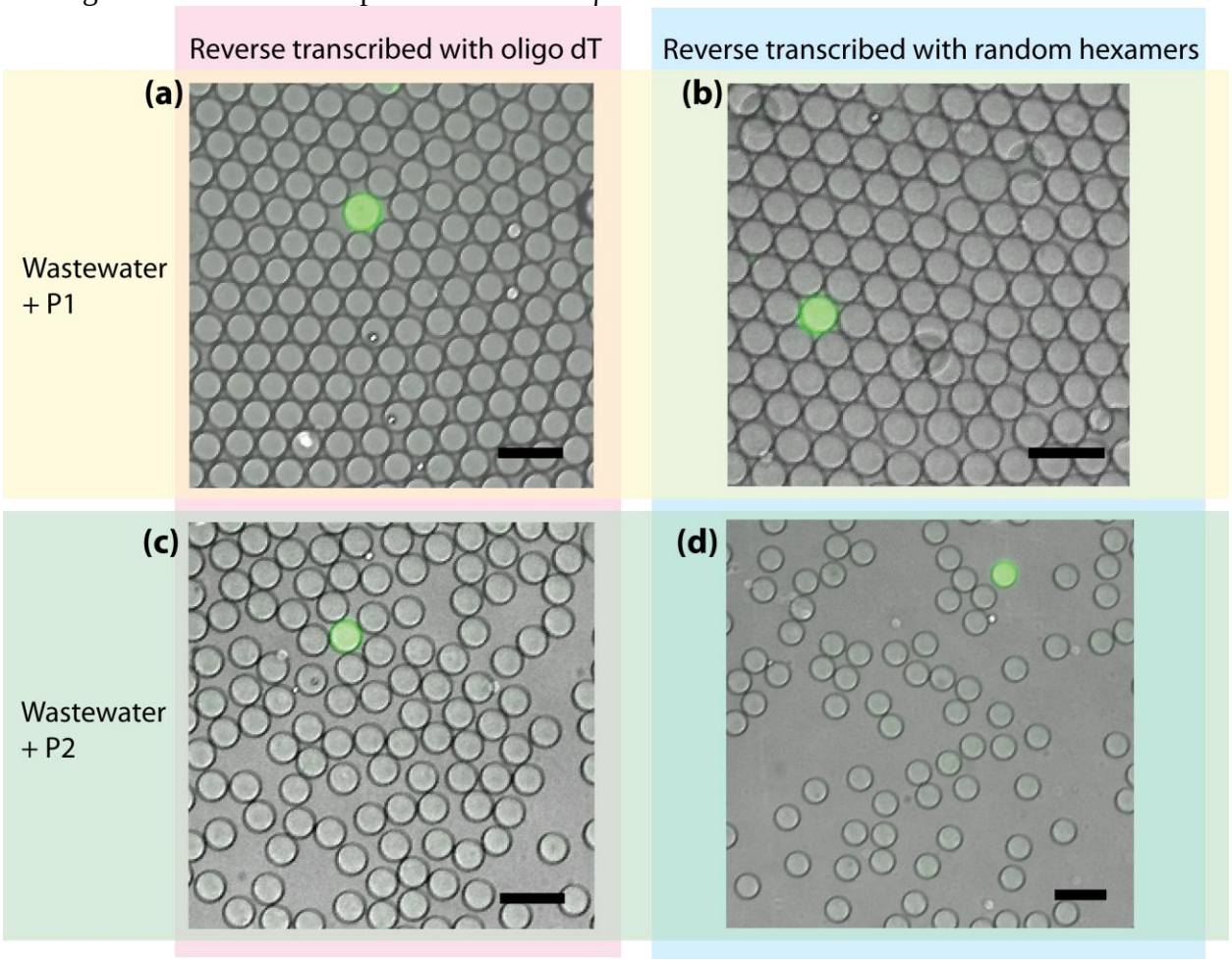
Supplementary Figure S8. Fluorescence images of drops after performing droplet digital PCR with the cDNA copies of norovirus. The genomic cDNA copies are encapsulated into 25 µm drops with Primer I (Table S4), and the drops undergo thermo-cycling. After PCR, a small fraction of the drops display a high level of fluorescence under excitation at 470 nm indicating the presence of the genomic cDNA copies of norovirus in those drops. (a) Reverse transcription is performed with oligo dT primers. (b) Reverse transcription is performed with random hexamers. Scale bar = 50 µm.



Supplementary Figure S9. The reverse-transcribed wastewater sample is mixed with either P1, primers targeting novel virus related to human rhinovirus A, or P2, primers targeting novel virus related to human rhinovirus C, and undergo thermo-cycling (Table S5). The predicted size of P1 and P2 amplicons is 268 bp and 472 bp, respectively. Gel electrophoresis analysis shows that our strategy for designing primers for novel viruses is valid.



Supplementary Figure S10. Fluorescence images of the drops after performing droplet digital PCR with the cDNA copies of novel viruses related to human rhinovirus. Reverse transcription is performed with either (a, c) oligo dT primers or (b, d) random hexamers. (a, b) The reverse-transcribed wastewater sample is mixed with P1, which is specific to the novel virus related to human rhinovirus A, and undergoes thermo-cycling (Table S5). (c, d) A reverse-transcribed wastewater sample is mixed with P2, which is specific to the novel virus related to human rhinovirus C, and undergoes thermo-cycling (Table S5). In all cases, a small fraction of drops display a high level of fluorescence signals under excitation at 470 nm, indicating the presence of the target virion in those drops. Scale bar = 50 µm.



Supplementary Table S1. Percentage of SV40 reads in total reads from different sets of sorting experiments.

Sample	Population of SV40	Number of droplets collected	MDA	% of SV40 reads in total reads	Average depth (×)
1-0	3%	No sorting	O	0.004%	N.A.
1-1		100	O	94.0%	12060
1-2		1000	O	98.6%	11085
1-3		waste	O	0.01%	N.A.
3-1*		100	O	8.9%	49
3-2*		1000	O	0.74%	93
6-1	30%	100	O	0.01%	N.A.
6-2		1000	O	93.7%	3295
4-1		1000	O	57.3%	4727
4-2		5000	O	93.5%	15398
5-1		1000	O	96.8%	14185
5-2		5000	O	96.9%	15326
2-1		1000	O	95.8%	7478
2-2		5000	O	98.8%	9118

Supplementary Table S2. The attributes of the contigs generated from various sets of sorting experiments

Sample	Contig	Length of contig (bp)	Organism	Sequence Identity (%)	Reads
1-1	1	5066	SV40 (Virus)	100	1264585
	2	585	Mouse	99.69	40818
1-2	1	5162	SV40 (Virus)	100	1162340
	2	714	Mouse	100	3808
2-1	1	5134	SV40 (Virus)	100	784159
	2	4322	Hypothetical (Plant)	46.21	2928
	3	2614	Mouse	100	2020
	4	1623	Mouse	100	1940
	5	2074	Mouse	100	1940
2-2	1	5187	SV40 (Virus)	99.98	956067
	2	835	Human	92.18	5318
	3	761	Human	92.35	3522
	4	614	Human	90.49	2954
	5	600	Human	88.73	1547
3-1	1	5159	SV40 (Virus)	99.98	5159
	2	728	Mouse	100	728
	3	630	Mouse	100	630
3-2	1	5016	SV40 (Virus)	100	9793
	2	5369	Bacteria	99.66	6108
	3	2186	N.A.	N.A.	4759
	4	3605	Mouse	99.64	3602
	5	4199	Bacteria	59.29	2844
4-1	1	5135	SV40 (Virus)	100	495682
	2	2299	Bacteria	67.41	358853
4-2	1	5121	SV40 (Virus)	100	1625104
5-1	1	5163	SV40 (Virus)	100	1487457
5-2	1	5159	SV40 (Virus)	100	1607112
6-2	1	5159	SV40 (Virus)	99.98	345563
	2	3174	Bacteria	70.38	664

Supplementary Table S3. Sequence of primers used for SV40 enrichment and sequencing

Virus-gene	Locus (bp)	Primer name	Genome bp	Sequence
SV40-VP1*	1499-2593	SV40-VP1-F	1741-1764	5'- CTCTCCAGACAAAGAACAACTGCC
		SV40-VP1-R	2259-2237	5'- CCAACACCCTGCTCATCAAGAAG
SV40-gp2**	344-522	SV40-gp2-F	237-261	5'- GGACTTTCCACACCCTAACTGACAC
		SV40-gp2-R	495-472	5'- TTCTGGTTTTTTGCGTTTCCCGTC

* This primer set is used for estimation of the concentration of SV40 virus particles in SV40 solution and enrichment of SV40 from SV40/wastewater mixture followed by USER treatment and MDA

** This primer set is used for the enrichment of SV40 from SV40/wastewater mixture followed by MDA without USER treatment

Supplementary Table S4. Sequence of primers used for norovirus

Virus	Primer name	Genome bp	Sequence
Murine norovirus	MNV1-I-F	3778	5'- CCCCTCAAGAAATCCTTGATGCAGTC
	MNV1-I-R	4724	5'- ACGGCGCAGGAAGGAGATGCC
	MNV1-II-F	4700	5'- GCAGGGCATCTCCTTCCTGCGC
	MNV1-II-R	5904	5'- ACCGGAGATTGGGGTGGTACCAAGC

Supplementary Table S5. Sequence of the primers designed for novel viruses related to human rhinovirus. The predicted sequence of amplicons is also presented.

<table>
<tr><th rowspan="2">Primer set</th><th rowspan="2">Virus</th><th colspan="3">BLASTX results</th><th rowspan="2">Primer name</th><th rowspan="2">Sequence</th></tr>
<tr><th>Accession</th><th>Sequence homology</th><th>E-value</th></tr>
<tr><td rowspan="3">P1</td><td rowspan="3">Novel virus related to human rhinovirus A</td><td rowspan="2">gb|AER92559.1| Human rhinovirus A</td><td rowspan="2">50.8%</td><td rowspan="2">6e-04</td><td>P1-F</td><td>5’- TTGTCTGCGATCAACTCACC</td></tr>
<tr><td>P1-R</td><td>5’- ACGATTTCTTTGCCACTTCG</td></tr>
<tr><td colspan="6">Predicted amplicon sequence: 472 bp
TTGTCTGCGATCAACTCACCTAAAACTTATAACGCTGTTTATAGTTATAATATTAATGGAATTTTTAGGTATACAAACTCGT
TTTCGGAGAATTCGTTCTTTTATATGAAGCGAGGACCTGGTATGTGTGGTTTACCTATAGTAGACCAAGAATGTGGTGTTTT
AGGAATACATGTTTCGGGTAGTGCCGAAGTCGGGGTTGGAGTTTCAATGCGCTGGAGTGATCTGGTAAGAGATAAGATTGGT
GAAATTTTAAGAGAAGACCAGCAATTTCTGATGCAGGTCGATATAAATAACAAAGTAAGGGATGGGGAAAGTGTAATGCGTT
TAGAGGCAGATGTACACCAAACTATTCCAAAAAAGAGTAACTTTCAACCATCACCTCTATATGATATATATCCTATTGAAAG
AGAACCAGCCAACTTGAGTAAATTTGGAGTTGGTACTGTTTGCGAAGTGGCAAAGAAATCGT</td></tr>
<tr><td rowspan="3">P2</td><td rowspan="3">Novel virus related to human rhinovirus C</td><td rowspan="2">gb|AET72458.1| Human rhinovirus C</td><td rowspan="2">65.5%</td><td rowspan="2">4e-08</td><td>P2-F</td><td>5’- GACCCGAACCTGCTTGTTTA</td></tr>
<tr><td>P2-R</td><td>5’- AGATCAGCTGACGCACAATC</td></tr>
<tr><td colspan="6">Predicted amplicon sequence: 268 bp
GACCCGAACCTGCTTGTTTAATACTTGATGGACCTCCTGGAACTCAAAAATCTATTTTTGTCACACAGTTACTACAATATTT
AGGCGAATCATCTTATATACATGCAACTAAGTCTATAGATGATGGAAAAGATTTCTATGACACTTATAATAATGAAACGATT
TTTGTAATGGATGATGTAGGCCAACAAAGCACATCCCAGTGGAGATCTTTAATAAATATGGTGTCTGTAGTTAAAATGCCTT
TAGATTGTGCGTCAGCTGATCT</td></tr>
</table>

Methods and Materials

Wastewater collection, removal of bacteria, and nucleic acid extraction. Wastewater (3 L) was transported in a carboy on ice from the wastewater treatment facility (McKeesport, PA) to the laboratory. Wastewater was loaded into 500 ml polycarbonate tubes and centrifuged at 10,000 rpm for 20 minutes in an SLA-1500 rotor (Sorvall) to pellet bacteria and other debris. The supernatant was then passed through a 0.22 µm filter. Total nucleic acids were purified with a QIAgen DNeasy Blood and Tissue kit. Nucleic acids were eluted with 30-60 µl dH_2O. The concentration of total virus particles in the wastewater solution was estimated by staining virus particles with SYBR-gold and counting them using an epi-fluorescence microscopy[2]. However, the staining efficiencies for different types of viruses are not well characterized, and therefore, the actual viral concentration in a wastewater solution may be higher than the calculated value.

Design of the microfluidic platform

The microfluidic platform consists of two PDMS devices, a drop maker and a sorter. The drop maker is designed to accommodate the small volume of aqueous samples. To minimize the loss of reagents, a pipette tip loaded with aqueous reagents is directly inserted into a channel inlet. Vacuum is then applied to the channel outlet to draw the reagents and oil through a flow-focusing junction where drops are generated (Figure S1). This setup allows us to work with an aqueous solution whose volume is less than 10 µL. The sorter is designed to maximize the collection efficiency of the sorted drops, which is critical when collecting a small number of drops from a rare population. To minimize the loss of drops during collection, the sorted drops are collected into a pipette tip that is directly inserted into the sort outlet (Figure S2). The collected drops accumulate on top of the collected oil as their density is lower than that of fluorocarbon oil, which is used as a continuous phase. Evaporation of the collected drops is prevented by loading drops containing pure PCR buffer on top of the solution.

Fabrication of microfluidic devices

The microfluidic devices are fabricated using standard soft lithographic technique in poly-dimethylsiloxane(PDMS). Photoresist (SU8-3025) is spun at 3000 rpm to yield a coating thickness of 25 µm. After photomask exposure, baking and development, a 25 µm tall positive master of the devices are formed on the silicon wafer. The silicon surface is then treated with 2% trichloro(1H,1H,2H,2H-perfluorooctyl)silane (Sigma Aldrich) in Novec HFE-7500 for easy peeling of PDMS. PDMS elastomer (Sylgard 184) mixed with crosslinker at the ratio of 10:1 is poured onto the master and baked for 4 h at 65 °C. The cured PDMS replicates are peeled from the master and inlet/outlet ports are punched using 1.00 mm biopsy punch. The peeled PDMS replicates are then bonded to glass slides using oxygen plasma treatment. For a sorter device, eight-pin terminal blocks (digikey) are inserted to the outlets of the electrodes and a low melting point solder (Indally 19 (52 In, 32.5 Bi, 16.5 Sn, Indium Corp.)) is introduced into the electrode

channels which are heated to 85 ℃. The pins are then glued (Loctite UV cured) to the surface of the device. Electrical contacts are made with alligator clips connected to a high voltage amplifier (Trek) and function generator (LabWIEW FPGA). After fabrication, both drop makers and sorters are flushed with Aquapel (PPG Industries). The channels are immediately dried with pressurized air and dried at room temperature overnight to permanently make the devices hydrophobic.

Compartmentalize single virions into droplets

Fluorocarbon oil is used as a continuous phase because these oils, being both hydrophobic and lipophobic, effectively inhibit molecular diffusion between droplets and do not swell PDMS devices. To prevent coalescence of droplets during droplet formation and sorting, we use wt2% poly(ethylene glycol)-Di-(krytox-FSH amide) (Ran Biotechnologies, Inc.) in Novec HFE-7500. Droplets are generated by applying negative pressure to the outlet of the drop-maker using a vacuum pump. A syringe loaded with the fluorocarbon oil is connected to the one of the inlet of the device with polyethylene tubing (Scientific commodities Inc.) while PCR solution containing wastewater, Taq buffer, Taq polymerase (Phusion U Hot start DNA polymerase, Thermo Scientific), dUTP/dATP/dCTP/dGTP, specific primers, 0.25X SYBR Green I is loaded in a 200 µL plastic pipette tip and directly inserted into the other channel inlet. For droplet generation, negative pressure is applied to the channel outlet using a vacuum pump. The pressure was kept at -40 ~ -50kPa for stable droplet generation.

Droplet digital PCR on SV40 solution and preparation of SV40/wastewater mixtures

The concentration of SV40 virus particles in cell lysates is calculated by performing ddPCR using primers specific for VP1 and SYBR Green I[3]. A dilution series of SV40 solution was encapsulated into 25µm droplets with primers annealed to the virion glycoprotein VP1, dNTP, SYBR Green I and Taq polymerase, and the droplets were thermo-cycled off-chip using the following protocol: 95 ℃ for 4 m, 35 cycles of 95 ℃ for 15 s/52 ℃ for 15 s/72 ℃ for 30 s, 72 ℃ for 5 m. The primer sequences are listed in Supplementary Table S3. After PCR, droplets containing SV40 virions exhibited increased level of fluorescence due to intercalation of SYBR Green I into the amplicons. The number of bright droplets was counted to calculate the concentration of SV40 virions in the stock solution. An aliquot of the SV40 solution is added to wastewater to prepare 3% and 30% SV40/wastewater mixtures.

Purification of a single viral species using a microfluidic sorter

The collected droplets are thermo-cycled off chip using the following protocol: 95 ℃ for 4 m, 35 cycles of 95 ℃ for 15 s/52 ℃ for 15 s/72 ℃ for 30 s, 72 ℃ for 5 m. After PCR, droplets were visually inspected under an epifluorescent microscope (Leica) equipped with a mercury arc lamp, 480/40 nm excitation filter, and 527/30 emission filter. Droplets are then loaded into a syringe and injected into a microfluidic sorter using a syringe pump. Protocol for sorting droplets based

on fluorescence was slightly modified from the previous studies[4]. The geometry of the sorter forces the close-packed droplets into a constriction where they flow in a single-file line. Fluorinated oil containing 2% (w/v) surfactant was co-flowed with droplets to evenly space close-packed droplets. The flow rate of droplets and fluorinated oil was kept at 10 µL/hr and 200 µL/hr. When a droplet passes the laser slit, its fluorescence is read using a photomultiplier tube (Hammamatsu). When the level of fluorescence is higher than a manually set threshold, the custom built Labview program activates the sorting electrodes. For our experiment, droplets were flowed at ~200 Hz and we used 20 cycles of 20kHz single-ended square wave at 3kV to minimize false positive sorting. 200 µL plastic pipette tips (gel-loading tips, VWR) are attached to the two outlets, sort channel and waste channel, for droplet collection. Dummy droplets containing 1× PCR buffer was added on top of the pipette tips inserted into the sort outlet to prevent vaporization of reagents in the sorted droplets.

Whole genome amplification

For whole genome amplification on the sorted droplets, the sorted droplets are transferred to a PCR tube and broken by adding 30% (v/v) 1H,1H,2H,2H-perfluoro-1-octanol (Oakwood Chemical) to the droplet emulsion. The tube is then agitated to mix droplet emulsion with the destabilizing reagent and briefly centrifuged to yield separate aqueous phase on top. The aqueous phase is pipetted into a new PCR tube followed by addition of USER (New England Biolabs Inc.) to a final volume of 2.5 µL. The mixture is then incubated at 37 ℃ for 1 hr followed by heat treatment at 85 ℃ for 10 mins and 95 ℃ for 1 min. This heat treatment step irreversibly deactivates USER enzyme, which inhibits MDA reaction when remaining active. After selective digestion of amplicons using USER enzyme, 1.5 µL 2.6× MDA mixture is added to the solution to yield 4 µL of 1× MDA mixture. 1× MDA mixture constitutes of 37 mM Tris HCl (pH7.5), 50 mM KCl, 10 mM MgCl2, 5 mM $(NH_4)_2SO_4$, 1 mM dNTP, 3 mM DTT, 1X BSA, 0.2% Tween 20, 1 unit/ml yeast pyrophosphatase, 50 µM exonuclease-resistant hexamer primers, and 540 unit/ml φ29 DNA polymerase (RepliPHI, Epicentre).[5] To remove the contaminant DNAs, we treated MDA buffer except φ29 polymerase with UV (Stratalinker 2400, 30 mins). Φ29 is not treated with UV since UV treatment drastically lowers its activity. The MDA mixture is then incubated at 30 ℃ for 22 hours to yield microgram quantities of DNA. Lastly, the mixture was heat treated at 65℃ for 10 mins to deactivate φ29 polymerase. The resulting mixture was directly used for next-generation library preparation without purification.

Next-generation sequencing and reference mapping

Next-generation library is prepared following the protocol provided by Illumina (Nextera® XT DNA sample prep kit and Nextera® XT index kit). The barcoded library is then deep sequenced using Illumina platforms, either MiSeq or HiSeq (single end 50 bp reads). Raw sequencing reads are mapped against the SV40 reference genome (NCBI reference sequence: NC_001668.1) using Bowtie-0.12.7. Bowtie assumes a linear genome. Sequencing coverage graphs (Figure 3 and

Supplementary Fig. S3) are created using a MATLAB code by extracting the genomic address of each aligned sequence from .SAM files generated from Bowtie.

De novo assembly of the selected genome

Raw sequencing reads were processed with Prinseq (http://prinseq.sourceforge.net/) to remove low quality or low entropy reads and to trim poor quality ends. Prinseq was run with the following command line options: "-lc_method entropy -lc_threshold 60 -min_qual_mean 15 -ns_max_p 5 -trim_qual_right 10 -trim_qual_left 10 -min_len 30". Human reads were computationally subtracted by aligning the high-quality reads against the hg19 human reference genome and two Ensembl databases (http://useast.ensembl.org/info/data/ftp/index.html), cDNA and ncRNA using Bowtie 2 with default parameters. Remaining reads were assembled into contigs with clc_assembler (http://www.clcbio.com/) using default options. Then, sequences are masked with RepeatMasker to identify contigs containing interspersed repeats that are commonly found in the human genome and masked sequences were filtered as described[1]. Then, remaining contigs were further filtered by removing any sequence < 500bp or had low entropy (<65). Finally, high quality contigs were annotated by sequentially aligning them against (i) the NCBI human genome database using BLASTN; (ii) GenBank nt database using BLASTN; (iii) GenBank nr database using BLASTX; and (iv) the NCBI viral refseq genome database using TBLASTX. A minimal E-value cutoff of 1e-5 for all steps was applied. Additionally, a minimal query coverage of 50% and minimal percent identity of 80% were applied to the BLASTN steps. Contigs were taxonomically binned as human, fungi, bacteria, archaea, phage, virus, or other based on the identity of the top BLAST hit sequence. Any contig that did not have a significant hit in the BLAST pipeline were considered unannotated. The relative abundance of contigs in each sample is determined by the number of reads assembled into the contig normalized by contig length (no. reads / contig length).

Design of primers for novel viruses

To design primers for novel viruses, viral sequences present in a virus sample are first determined using metagenomic techniques[6]. After performing deep sequencing on a complex virus mixture such as an environmental sample, the sequence reads are processed following the protocol described in the previous section.[1]

Virus sequences that align to a database sequence at high% identity (>90%) across most of the query sequence are considered known viral sequences. Most of the viral hits in the BLASTN step to Genbank nt fall into this category. Most sequences with BLASTX or TBLASTX hits to a virus are considered novel viral sequences since the% identity is lower (< 80%). Putative novel human-virus sequences are identified by their having a BLAST hit to a human virus or to a virus family that is known to infect humans. Primers are then designed to viral sequences using Primer

3 (http://bioinfo.ut.ee/primer3-0.4.0/) with default parameters to generate amplicons of approximately 100-500bp in length.

The length requirements for the pre-known sequence of novel viruses

Our platform requires a small fraction of the target viral genome to be known so that the target virions can be selectively enriched using the microfluidic sorter; however, the length of this pre-known sequence can be as short as 100 bp. In our platform, genome sequencing of a target virus, either known or novel, is achieved by selective enrichment of the target virions, whole genome amplification, sequencing and de novo assembly. The enrichment step is essential as MDA, a whole genome amplification method used in our platform, is a sequence independent amplification method. Target virions are enriched by encapsulating individual virions into drops with primers specific to the target virions, thermo-cycling the drops, and selecting PCR positive drops using a microfluidic sorter. To design primers specific to the target virion, a small fraction of the target viral genome, a marker sequence, must be known. In case of novel viruses, a marker sequence can be identified using metagenomic analysis. Because the purpose of marker sequence is to design PCR primers, the length of the marker sequence can be as short as 100 bp, the minimum length of DNA that can be amplified using PCR.

References

1. Cantalupo, P. G.; Calgua, B.; Zhao, G.; Hundesa, A.; Wier, A. D.; Katz, J. P.; Grabe, M.; Hendrix, R. W.; Girones, R.; Wang, D.; Pipas, J. M., Raw Sewage Harbors Diverse Viral Populations. *mBio* **2011,** *2* (5), e00180-11.
2. Patel, A.; Noble, R. T.; Steele, J. A.; Schwalbach, M. S.; Hewson, I.; Fuhrman, J. A., Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. *Nat. Protocols* **2007,** *2* (2), 269-276.
3. Hindson, B. J.; Ness, K. D.; Masquelier, D. A.; Belgrader, P.; Heredia, N. J.; Makarewicz, A. J.; Bright, I. J.; Lucero, M. Y.; Hiddessen, A. L.; Legler, T. C.; Kitano, T. K.; Hodel, M. R.; Petersen, J. F.; Wyatt, P. W.; Steenblock, E. R.; Shah, P. H.; Bousse, L. J.; Troup, C. B.; Mellen, J. C.; Wittmann, D. K.; Erndt, N. G.; Cauley, T. H.; Koehler, R. T.; So, A. P.; Dube, S.; Rose, K. A.; Montesclaros, L.; Wang, S.; Stumbo, D. P.; Hodges, S. P.; Romine, S.; Milanovich, F. P.; White, H. E.; Regan, J. F.; Karlin-Neumann, G. A.; Hindson, C. M.; Saxonov, S.; Colston, B. W., High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Analytical Chemistry* **2011,** *83* (22), 8604-8610.
4. (a) Mazutis, L.; Gilbert, J.; Ung, W. L.; Weitz, D. A.; Griffiths, A. D.; Heyman, J. A., Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* **2013,** *8* (5), 870-891; (b) Abate, A. R.; Hung, T.; Mary, P.; Agresti, J. J.; Weitz, D. A., High-throughput injection with microfluidics using picoinjectors. *Proc. Natl. Acad. Sci.* **2010,** *107* (45), 19163-19166.
5. Hutchison, C. A.; Smith, H. O.; Pfannkoch, C.; Venter, J. C., Cell-free cloning using φ29 DNA polymerase. *Proc. Natl. Acad. Sci.* **2005,** *102* (48), 17332-17336.
6. (a) Edwards, R. A.; Rohwer, F., Viral metagenomics. *Nat. Rev. Microbiol.* **2005,** *3* (6), 504-510; (b) Delwart, E. L., Viral metagenomics. *Rev. Med. Virol.* **2007,** *17* (2), 115-131; (c) Rosario, K.; Breitbart, M., Exploring the viral world through metagenomics. *Curr. Opin. Virol.* **2011,** *1* (4), 289-297.