An alternative method to multiple displacement amplification for preparing virome DNA in a way adapted for the sequencing of both double-strand and single-strand DNA viruses

Maud Billaud[1], Ilias Theodorou[1,3], Shiraz Shah[4], Luisa de Sordi[2,3], Marianne De Paepe[1] and Marie-Agnès Petit[1]

1. Université Paris-Saclay, INRAE, Micalis unit, Jouy en Josas France
2. Sorbonne Université, INSERM, Centre de Recherche St Antoine, Paris, France
3. Paris Center for Microbiome Medicine (PaCeMM) FHU, AP-HP, Paris, France.
4. Copenhagen Prospective Studies on Asthma in Childhood, Copenhagen University Hospital, Herlev-Gentofte, Ledreborg Allé 34, DK-2820 Gentofte, Denmark

Summary

Unless some dedicated treatments are applied, virome shotgun sequencing will only reveal the double stranded DNA (dsDNA) content of a given sample. However bacterial virus genomes often consist of a circular single stranded DNA (ssDNA) molecule. Pre-treatment and amplification of DNA with the multiple displacement amplification (MDA) method allows to convert the ssDNA to ds DNA, but it introduces an over-representation of these circular ssDNA genomes. A more recent alternative consists in using an xGen kit, which permits the ligation of adaptators directly to sheared DNA. However, this kit is expensive, and the sonication step might shear ssDNA more efficiently than dsDNA, therefore introducing another bias in virome sequencing. This prompted us to search for an alternative method of DNA preparation for the sequencing of mixed ssDNA and dsDNA viromes. We present here this new method based on the use of the T7 DNA polymerase (T7pol), which is easy to implement and cheap, and compare it to MDA with 30 minutes incubation time (sMDA) and to the direct shearing of DNA without prior conversion of ssDNA to dsDNA. All samples were then prepared with the xGen kit. We found that sMDA overamplifies with 2 to 8.6-fold excess compared to input phage particles, which was less than reported for the full time, 90 minutes MDA treatment. In contrast, the T7pol method led to the under-representation (1.3 to 8.2-fold) of these genomes, and was overall a better reflection of input phage cocktails, compared to the other methods. We conclude a T7pol pre-treatment is a good alternative to MDA for the shotgun sequencing of viromes.

Introduction

In the last ten years, many methodological improvements have been put forward, in order to generate informative metagenomics datasets for the viral fraction (or virome) of any type of environmental or biological sample (1-5). A particular attention has to be paid to the fact that viromes are mixtures of single-stranded (ss) and double-stranded (ds) DNA viruses. Having a simple method to include the sequencing of ssDNA phages in any type of environmental shotgun sequencing would also be relevant in the future, as more groups are now analysing the viral fraction of environmental or biological samples directly from whole metagenomics shotgun sequences, rather than first preparing a virome filtrate (6, 7). To do so, they rely on the sorting of viral contigs among all assembled contigs, based on their characteristics and gene contents (8, 9). While this method is certainly efficient to detect the phage DNA replicating inside bacteria, as well as prophages, its ability to efficiently extract and sequence phage particle content is an open question. In particular, ssDNA phage genomes will certainly escape analysis, as no pre-treatment step ensuring their conversion into dsDNA is done at present (10).

Initially, most viromes were sequenced after a multiple displacement amplification (MDA) step, which was a convenient way to increase the amounts of DNA for library preparation. It became rapidly obvious that MDA introduced a bias favouring small (~2000 to 7000 nt long) circular ssDNA viral genomes (11, 12). This is due to the fact that the Phi29 DNA polymerase used for MDA has a potent strand displacement activity (13). Therefore, once these small ssDNA genomes are converted into ds DNA, the polymerase proceeds for further rounds, generating long concatemers (Figure 1A). Tests with synthetic mixes of viral genomes demonstrated unambiguously this caveat of MDA for mixes of ss- and ds-DNA genomes (14).

Two main solutions were then used to circumvent this problem. The first, the “short MDA (sMDA)“ method consisted in reducing the incubation time of the MDA treatment from the normally recommended 90-120 minutes to 30 minutes (2). The second solution consisted in replacing the MDA

kit by an Accel-NGS kit from Swift Biosciences (recently acquired by Integrated DNA Technologies – IDT- and rebranded as xGen kit, below xGen) developed for the sequencing of ancient DNA (14). Here, after an initial shearing step, all DNA molecule extremities, both ss- and ds-DNA, receive a short extension using an "adaptase", the nature of which is kept secret by the manufacturer. Then an additional linker is added, allowing PCR amplification and sample labelling. For this second method, tests with synthetic mixes showed that the bias in favour of ssDNA viruses was significantly reduced(14).

For some reason that will be explained below, the DNA shearing step applied prior library preparations with the xGen kit failed to produce satisfying results for ssDNA viral DNA in our hands. Further, the cost of this kit is high (~40 euros per reaction). This prompted us to search for an alternative method of DNA preparation for the sequencing of mixed ssDNA and dsDNA viromes. We present here this new method based on the use of the T7 DNA polymerase (T7pol), which is easy to implement and cheap, and compare it to sMDA as well as to the direct shearing of DNA without any pre-treatment. Then all samples were treated with the xGen kit. We found that sMDA overamplifies with 2 to 8.6-fold excess compared to input phage particles, which was less than reported for the full time, 90 minutes MDA treatment. In contrast, the T7pol method led to an under-representation (1.3 to 8.2-fold) of these genomes, and was overall more congruent to input cocktails.

Methods

Purchased ssDNA, enzymes, kits and oligonucleotides

ssDNA of M13mp18 and PhiX174 were purchased from New England Biolabs (references N4040S and N3023L). The T7pol was purchased from NEB (ref. M0274), and used following the manufacturer's instructions. Random oligonucleotides, 20 nt long, were purchased from Eurofins Genomics. For MDA treatments, the Genomiphi2 kit was purchased from Fisher Scientific (10628505). The "xGen ssDNA Low-Input DNA Lib Prep 16rxn" (previously, Accel NGS S2 kit) was purchased from IDT (reference 10009859).

Phage stock preparations

Two *Escherichia coli* infecting phages with ssDNA genomes were selected for the study: PhiX174 is a Microviridae forming a tail-less icosaedric capsid enclosing its circular 5386 nt ssDNA genome. It was grown on *E. coli* strain C, which codes for its appropriate core LPS receptor. M13-yfp (15) is a filamentous phage harbouring a 7916 nt long ssDNA circular genome. This M13mp18 derivative makes less turbid "plaques" than M13, allowing better estimates of its phage titer. It was grown on *E. coli* strain MG1655 F'kanR, a strain coding for the F pilus, which is the M13 receptor. Three dsDNA Caudovirales were selected as well for the study: T4 and lambda are *Escherichia coli* infecting phages, and were grown on *E. coli* strain MG1655. Inside the virions, the T4 genome is a 169 kb-long, linear dsDNA molecule, while the lambda linear genome has a shorter, 48 kb-long linear dsDNA genome. Finally, the *Bacillus subtilis* infecting phage SPP1 was propagated on strain YP886, generating virions containing a linear, 46 kb long genome.

Phage lysates of T4, SPP1, lambda and phiX174 were prepared as described previously (16), and bacterial remains were removed by centrifugation (5000 g, 7 minutes, 4°C), followed by filtration through polyestersulfone (PES) membranes (0.2 µm pore size). A lysate of M13-yfp was obtained by propagating the infected strain in 50 mL of LB medium with 50 µg/mL kanamycin up to OD 1, removing bacteria by centrifugation (5000g, 7 minutes), and filtering the supernatant (0.45µm PES filter). Stocks were kept at 4°C in SM buffer (200 mM NaCl, 50 mM Tris pH 7, 10 mM MgSO4). Just prior qPCR verifications and preparing phage mixes, lysates were treated with DNaseI (Turbo DNAse I,

ThermoFisherScientific, reference AM2222), using 1 µL of enzyme/ml of phage sample (1h at 37°C) to eliminate any residual bacterial DNA. Phage concentration of the stocks was then estimated by qPCR or epifluorescence (see below).

Phage DNA extractions

Total DNA of the two phage mixes (500 µL) was extracted by double phenol-chloroform extraction followed by ethanol precipitation, as described earlier (16), resuspended into 25 µL Tris 10 mM (pH 8), and dispatched for the three treatments as follows: 1 µL for MDA pre-treatment, 5 µL for T7pol pre-treatment, and 19 µL for the AcceI library without pre-treatment.

To prepare the qPCR DNA standards, highly purified phage stocks of T4, SPP1 and lambda were prepared first. For this, phage lysates were concentrated by over-night centrifugation (20 000 g, 16h), resuspended in 0.5 mL SM buffer, and loaded on SM-iodixanol gradients as described (16). The phage-containing fraction was dialysed against SM buffer and stored at 4°C. Prior phage DNA extraction, purified particles were treated for 1h at 37°C with Turbo DNaseI (ThermoFisherScientific, reference AM2222), using 1 µL of enzyme/ml of phage sample to eliminate any residual bacterial DNA. Next, DNA extraction was performed by the same phenol-chloroform method (16). Pure PhiX174 ssDNA was purchased from NEB. Phage dsDNA concentration was estimated using a Qubit device, and ssDNA with a Nanodrop, and dilutions of the pure DNA stocks (6µL aliquots corresponding to 5 x 10^6 phage genomes) were kept at -20°C.

Phage quantification for the synthetic phage mixes

To prepare mixes containing phage in known concentrations, phage genome quantities of T4, lambda, SPP1 and PhiX174 were determined by qPCR directly on the phage stocks (after a 95°C treatment) using the primers indicated in Table 1. Three dilutions of each phage stock were compared with a calibration curve made with highly purified DNA of each phage (see above). For this, each stock of purified DNA was diluted in 10 mM Tris (pH 8), to obtain a concentration of 5 x 10^6 phage genomes in 6 µL. The samples were further diluted four times, by 3-fold dilution steps, to obtain the calibration range (1.8×10^5 to 5×10^6 genomes). Primers were used at a concentration of 200 nM. qPCR was performed in a total volume of 15 µL (6 µL purified DNA or phage particle dilution, and 9 µL of amplification mix) in MicroAmp Fast Optical 96-well plates sealed with MicroAmp Optical Adhesive Film using the Takyon ROX SYBR Mastermix blue dTTP kit. Amplifications were run in duplicate on a StepOnePlus real-time PCR system with the following cycling conditions: 95°C for 5 min, (95°C for 15 s, 58°C for 45 s, 72°C for 30 s) for 45 cycles, 72°C for 5 min, 95°C for 15s, 60°C for 15 s, 95°C for 15 s. The analysis of the melting curves confirmed the specificity of the primers. Data analysis were done with the StepOne software 2.3.

Table 1: primers used for qPCR

Phage	Forward	Reverse	Reference
T4	ACTGGCCAGGTATTCGCA	ATGCTTCTTTAGCACCGGCA	(17)
lambda	GAGTGCGGAAGATGCAAAGG	TTAACAGTGCGTGACCAGG	(18)
SPP1	CGGGCTGAAATACCTGTGGA	TAGCCCCTCCTCCGATTGTT	(19)
PhiX174	GGTTCGTCAAGGACTGGTT	TTGAACAGCATCGGACTCAG	this work

M13 titer could not be quantified by qPCR due to coat resistance to heat, so it was quantified by plaque formation on the MG165 F'kan strain (M13mp18 derivatives make much clearer plaques than the parental M13).

Each phage stock quantified by one of the two indicated methods was then used to prepare the mixes. Initial phage genome concentrations in each stock, as well as volumes added to compose each mix, are presented in table 2.

Table 2: Mix preparations

Phage	initial stock	Dilution of the initial stock		volume added for each mix (µL)		ratio of each phage	
	phage /mL	Mix high ssDNA	Mix low ssDNA	Mix high ssDNA	Mix low ssDNA	Mix high ssDNA	Mix low ssDNA
T4	3.29E+09	no	no	370.2	370.2	24.9%	30.1%
lambda	7.47E+11	1E-01	1E-01	16.5	32.9	25.2%	60.8%
SPP1	5.39E+10	1E-01	1E-01	13.2	23.7	1.5%	3.2%
phiX174	4.67E+10	1E-01	1E-02	32.1	31.1	3.1%	0.4%
M13	3.00E+12	1E-02	1E-03	73.8	73.8	45.3%	5.5%

T7pol pre-treatment

To test whether degenerated primers could be used with the T7pol, to convert any phage ssDNA to dsDNA, we ordered 20-nt long degenerated primers, and tested them on purchased ssDNA from PhiX174 and M13mp18. For the annealing step, samples containing 10 ng/µL of M13mp18 and 10 ng/µL PhiX174 ssDNA, in a final volume of 50 µL of water were incubated with varying concentrations of oligonucleotides for 5 minutes at 65°C. The samples were then allowed to cool down slowly by disposing them on floaters over a beaker containing 25 mL of 65°C water, placed at room temperature for 10 minutes (final temperature 46°C). Replication was then initiated by adding 300 µM of dNTP, 1X of T7pol buffer, 50 µg/mL of BSA and 5 units of T7pol in each tube completed with water to a final volume of 100µL, followed by incubation for 5 minutes at 37°C. Reactions were stopped by placing the samples at 75°C for 20 minutes. Products (20 µL) were separated on a 0.7% agarose gel containing 40 µg/mL ethidium bromide. For the T7pol pre-treatment of phage DNA mixes, the same conditions were applied, except that degenerated primers were used at a 25 µM final concentration, together with 5 µL of DNA mix complemented with water to a final volume of 25 µL. The next replication step was done in a final volume of 50 µL.

Library preparations, virome sequencing and analysis

DNA preparations from the synthetic phage mixes were first pre-treated (or not) for the conversion of ssDNA to dsDNA using either the Phi29 polymerase of the Genomiphi2 kit according to supplier's instructions with a 30 minutes incubation, or the T7pol (see Result section). Next all samples were sonicated (100 µL in 1.5 mL tubes) in a cooled water bath, during 21 sec with 90 sec breaks for 3 cycles, then centrifuged and sonicated again for 21 sec with 90 sec breaks for 2 cycles with a Diagenode Bioruptor Pico device, to generate DNA segments of an average 300 bp size. DNA samples were then concentrated to 20 µL using SPRIselect magnetic beads from Beckman Coulter (ratio beads/DNA 1.2, also for next steps). All samples were then processed for library preparation using the xGen kit. These library steps included: (1) adaptase (37°C, 15 min then 95°C, 2 min); (2) Extension (98°C, 30sec ; 63°C, 15 sec ; 68°C, 5 min); (3) Magnetic beads cleaning; (4) linker ligation (25°C, 15 min); (5) Magnetic beads cleaning; (6) DNA concentration estimate by Qubit (high sensitivity kit); (7) Indexing by PCR (98°C, 30 sec followed by 4 to 14 cycles of (98°C, 10 sec ; 60°C, 30 sec ; 68°C, 60 sec) depending on initial DNA concentration.

Paired-end sequencing (2x150 nt) was performed on an Illumina ISeq100 system, at a depth of 800 000 reads per sample. Read cleaning was then computed with fastp (v 0.23.1, (20)), using the following parameters: --trimfront2 15 --trimtail1 2 --trimtail2 2 –trim_poly_g –trim_poly_x –r –W 4 –M 20 –u 30 –e 20 –l 100 –p –w 12. These parameters were used to remove the adapters present at read extremities with xGen libraries. Read mapping on the phage genomes was then conducted with bwa (22), and the

following parameters: bwa mem INDEX –t 6. Next, data were treated and analysed with msamtools and the following parameters: msamtools filter –b –u –l 80 –p 95 –z 80 –besthit –S and then msamtools profile –multi=proportional. The number of reads mapped per genome were normalized in 'transcripts per million reads' TPM, at the msamtools profile step, using the default parameter. TPM are computed as follows: an "abundance" value of each genome was estimated by the number of reads mapped to that genome, divided by its length and normalized by the sum of abundances across all five genomes. The proportion of each phage within the sequenced sample was then estimated by its TPM ratio.

Results

Setting up an ssDNA to dsDNA conversion protocol using the T7 DNA polymerase.

A necessary step of library preparations involves the shearing of DNA with ultrasounds. It is explicitly mentioned in the xGen protocol that "fragmentation of ssDNA may result in shorter fragments and lower library yields". Indeed, when comparing T4 dsDNA and M13 ssDNA sonication, we observed an increased fragmentation (Figure 1A). In addition, our sonication protocol was done on 100 µL aliquots, which needed then to be concentrated to 20 µL for the adaptase step of the xGen kit. For this, magnetic beads were used (see methods), a step where probably most of the small ssDNA fragments were lost. We designate this as 'in-house xGen' protocol, to underline that protocols in which sonicated DNA can directly be used for the adaptase treatment will probably no suffer the same drastic ssDNA loss.

We therefore searched for a quantitative method to convert ssDNA molecules to dsDNA, prior DNA shearing. Contrary to the Phi29 DNA polymerase, the DNA polymerase encoded by phage T7 has no strand displacement activity, required to open up the DNA strands ahead of the polymerase (13, 23, 24) (Figure 1B) . Instead, a separate, replicative DNA helicase is responsible for this step, as in most living organisms (24). As a consequence, once a first round of replication has been completed on a circular ssDNA molecule, the T7 DNA polymerase (T7pol) should stop and unload, therefore preventing DNA over-amplification. We reasoned this should help preparing virome libraries that are less biased than the one generated after an MDA step, which uses the Phi29 DNA polymerase.

We found that 25 µM of 20 nt long degenerated oligonucleotides were sufficient to achieve maximal conversion of M13mp18 and phiX174 ssDNA molecules to dsDNA (Figure 1C). In addition, and as expected, no smear indicative of further rounds of replication beyond the first one was detected. The marked intensity of the dsDNA bands, relative to ssDNA, is due to the higher affinity of ethidium bromide for dsDNA.

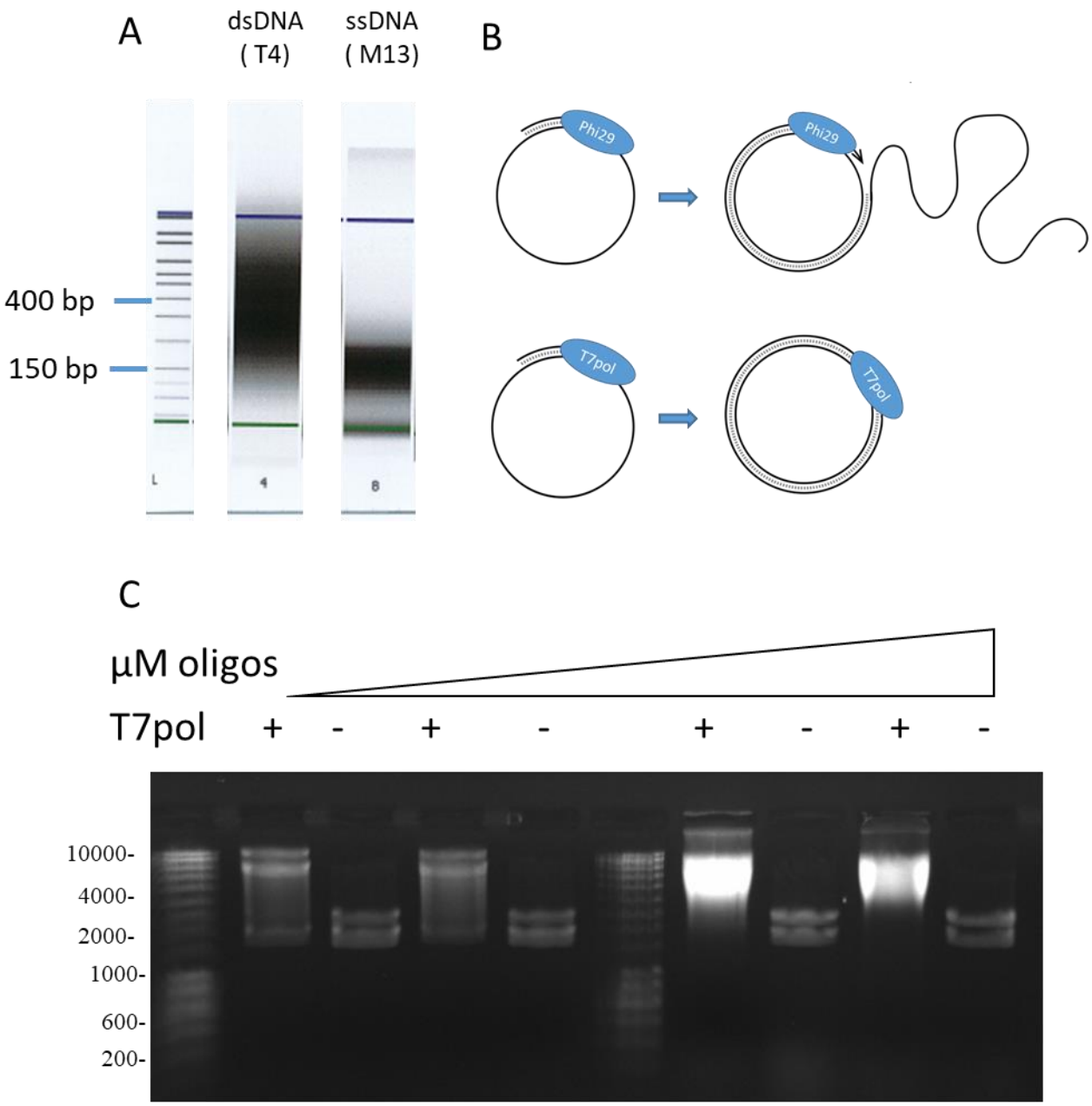

Figure 1: A: Bioanalyzer runs of sheared DNA of T4 dsDNA (9 ng/µL, 100 µL) and M13ssDNA (12ng/µL, 100 µL) sheared using 7 cycles of 20 seconds separated by 90 seconds pauses. B: The different strand displacement activities of Phi29 DNA polymerase (up, rolling circle), and T7 DNA polymerase (down, stop after one round). C. Adjusting primer concentrations for T7 DNA polymerase conversion of ssDNA to dsDNA of M13 and PhiX174. Degenerated oligonucleotides (20 nt) were used at increasing concentrations of 6µM, 12.5µM, 25µM and 50µM on a mix of M13 and PhiX174 ssDNA. Products were separated on 0,7% agarose gel. Sizes of linear molecular dsDNA standards (lanes 1 and 6) are shown on the left.

Comparing T7pol pre-treatment of synthetic viromes with short MDA pre-treatment, and viromes with standard XGen library preparation

From stocks of two ssDNA phages (M13-yfp and PhiX174) and three dsDNA phages (T4, lambda and SPP1) two synthetic mixes were prepared. One mix contained 48.4% of ssDNA phage genomes (45.3% of M13-yfp, 3.1% of PhiX174), while the other one contained only 5.9% of ssDNA phage genomes (5.5% M13-yfp and 0.4% PhiX174, see dsDNA genome proportions in Table 3). The final and global phage concentration in each mix was around 10^9 genomes/mL.

Next, DNA was extracted from the two mixes, and each was subjected to one of the three following treatments (Figure 2A): (i) short MDA (sMDA) with an incubation time at 30°C of 30 minutes only, (ii) T7pol-dependent DNA replication using conditions described above and 25 µM of oligonucleotides, (iii) no pre-treatment. Then all samples were sonicated in a final volume of 100 µL, and concentrated using magnetic beads. Finally, all were treated for xGen library preparation (see Methods for details). Each treatment was done in duplicate, and the 12 samples were then sequenced on an Illumina platform. Read mapping on the 5 phage genomes were converted into relative abundance, or TPM, ie reads per nt of the mapped genome, divided by the sum of all such read densities per sample. This allowed to estimate the proportion of each phage in the mix at the end of all treatments, and to compare them to their initial ratio (Figure 2B, and Table 3).

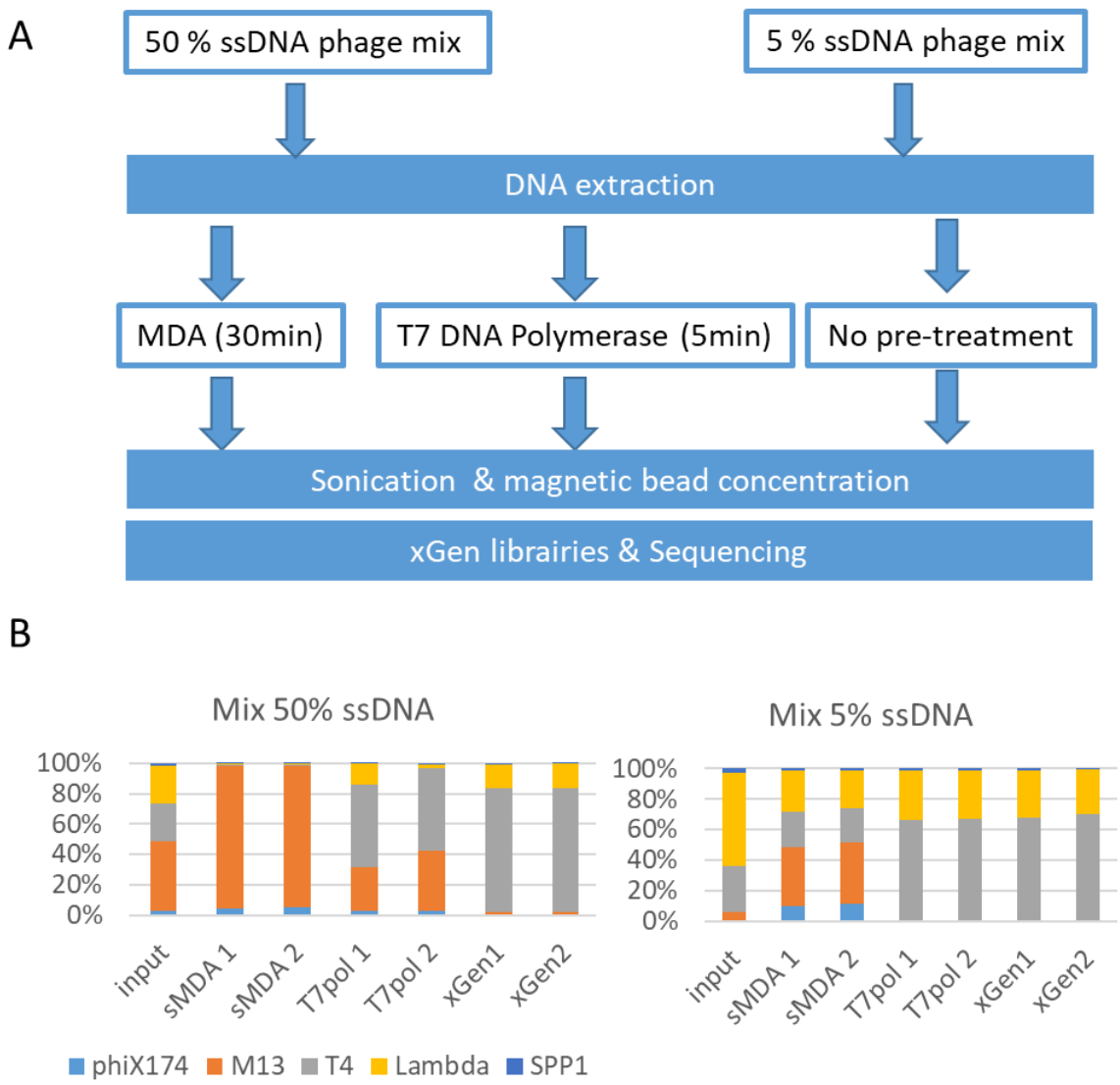

Figure 2. A: Comparison protocols for sequencing ss and dsDNA mixes. B. Proportions of the 5 phages in input mixes, and after deep sequencing of the various samples.

Table 3: proportion of the five genomes in input mixes, and after sequencing using the different protocols.A, high ssDNA mix, B, low ssDNA mix.

A. high ssDNA mix	input	short MDA 1	short MDA 2	T7pol 1	T7pol 2	xGen without pretreatment 1	xGen without pretreatment 2
T4	24.9%	1.2%	1.0%	54.8%	53.6%	81.4%	81.4%
lambda	25.2%	0.8%	0.9%	13.4%	2.9%	16.2%	16.2%
SPP1	1.5%	0.0%	0.0%	0.5%	0.7%	0.6%	0.6%
PhiX174	3.1%	4.7%	5.1%	2.6%	2.5%	0.2%	0.1%
M13-yfp	45.3%	93.3%	92.9%	28.6%	40.2%	1.7%	1.7%
%ssDNA, in total	48.4%	98.0%	98.0%	31.3%	42.7%	1.9%	1.8%

B. low ssDNA mix	input	short MDA 1	short MDA 2	T7pol 1	T7pol 2	xGen without pretreatment 1	xGen without pretreatment 2
T4	30.1%	23.3%	22.2%	65.3%	66.5%	68.1%	69.7%
lambda	60.8%	27.5%	24.9%	33.1%	31.9%	31.0%	29.5%
SPP1	3.2%	1.0%	1.0%	1.0%	1.0%	0.9%	0.7%
PhiX174	0.4%	9.8%	11.4%	0.2%	0.1%	0.00%	0.00%
M13-yfp	5.5%	38.4%	40.3%	0.5%	0.6%	0.00%	0.09%
%ssDNA, in total	5.8%	48.2%	51.8%	0.7%	0.7%	0.0%	0.1%

We found that the DNA samples treated directly with xGen libraries (absence of pre-treatment) led to a major depletion of the ssDNA phages: starting with the 48.4% input ssDNA mix, ssDNA phages only contributed 1.8-1.9% of the final sequencing result. The result with the 5.8% input ssDNA mix was even more drastic, as no reads matching the two ssDNA phages were obtained in one of the repeats, and M13 contributed only 0.09% of the sequenced virome in the other repeat.

With the short MDA pre-treatment, starting with the 48.4% input ssDNA mix, 98% of ssDNA phage genomes were recovered, and 48-52% when using the 5.8% input ssDNA mix. We conclude that over-amplification occurred in both cases, by a factor of 2 and 8.6, respectively. It should be noted that for the high ssDNA input mix, the sMDA effect is much higher than implied by the factor of 2, as almost all sequenced DNA ended up being ssDNA.

Finally, the T7pol pre-treatment gave a result that was more congruent to input DNA: for the 48.4% ssDNA input mix, similar proportions (31-43%) of ssDNA phage were recovered, and 8.2-fold lower ssDNA (0.7 %) was found for the 5.8% input mix.

In order to quantify the congruence between the initial mix compositions and the different pre-treatments, while accounting for the compositionality of the data, we computed Aitchison-distances (21) between each input DNA mix and the various outputs, using the abundances from Table 3. Judging from the distances (Table 4) the short MDA pre-treatment was more quantitative than direct xGen libraries at low ssDNA concentrations, while the opposite was true at higher ssDNA concentrations. The T7pol pre-treatment was more quantitative than both other methods regardless of the input

ssDNA/dsDNA ratio. The distance between the two input mixes themselves was also computed for comparison, revealing that the T7pol pre-treatment gave results closer to its initial mix (a distance around 2) than the distance between the two input stocks, which was around 3.

A.	Mix low ssDNA
xGen 1	8.17
xGen 2	5.97
sMDA 1	3.86
sMDA2	4.02
T7pol 1	2.29
T7pol 2	2.23
Mix high ssDNA	3.05

B.	Mix high ssDNA
xGen 1	3.72
xGen 2	3.93
sMDA 1	4.39
sMDA 2	4.34
T7pol 1	2.15
T7pol 2	2.31
Mix low ssDNA	3.05

Table 4: A. Aitchison distances between the compositions obtained from the different pre-treatments and the input composition with low ssDNA amounts (5.8%). The distance to the high ssDNA mix (48.4% ssDNA) was also calculated for comparison. B. Same calculations for the high ssDNA input mix.

Discussion

In this benchmarking experiment, xGen libraries performed on untreated sheared DNA were the most detrimental to ssDNA sequencing: ssDNA was recovered at levels 40 to 100-fold less than expected. However, a previous study which also used synthetic mixes to compare MDA and xGen methods, had come to the conclusion that ssDNA sequencing was quantitative with xGen libraries (14). Some important details in the shearing steps might explain these different conclusions. ssDNA might be more sensitive to shearing during the sonication step, which parameters are likely to vary between laboratories. We also suspect some ssDNA loss in our protocol at the step where magnetic beads were used to concentrate the sheared DNA prior the first step of the xGen kit (adaptase treatment). Clearly, all steps of DNA treatments should be carefully tested to ensure proper recovery of ssDNA. Nevertheless, this observation confirms the notion that pre-treatments permitting to convert all ssDNA molecules to dsDNA prior to sonication should prevent downstream bias, such as those we observed here.

MDA was the first method used to convert ssDNA to dsDNA, however this method has been known for a long time to be detrimental for quantitative assessment of viromes containing a mix of ss- and dsDNA genomes (11, 14). The standard, 90-120 minutes incubation time used for MDA was reported to amplify ssDNA by an excess of 33 to 44-fold when input ssDNA was 2% (14). We found, however, that a shortening of the incubation time to 30 minutes (called here “sMDA”) reduces somewhat the bias favouring ssDNA genomes: ssDNA genomes were only in 8.6-fold excess when starting from 5.8% input ssDNA and 2-fold (that is, ssDNA represented 98% of all phage DNA), when starting with 48.4% ssDNA.

A T7pol pre-treatment appears as a better method compared to sMDA for sequencing DNA viromes. The overall distance between the final sequencing product and the input DNA was the least in this case as well as being the most consistent between both tested compositions. In particular, with respect to ssDNA, the final ssDNA genome proportions were only slightly lower than input ssDNA when the input phage mix had a high proportion of ssDNA (48.4%). However, for the low input ssDNA mix, the final proportion was 8.2-fold lower than expected. It may be that the annealing step required for T7pol-

mediated DNA replication is less efficient when the ssDNA molecules are a minority within a complex sample.

The T7pol pre-treatment adds only a marginal cost (1 euro per sample) to the library preparations, compared to a Genomiphi MDA pre-treatment (20 euros per sample). Moreover, it can be applied to samples planned for any kind of downstream library preparation kit for low DNA amounts, such as the Nextera kit from Illumina, making it independent of the costly xGen kit used here for comparison purposes. Furthermore, we think the T7pol method described here might be a useful addition in the future, for projects aiming at describing simultaneously the bacterial and viral components of their microbial ecosystems, without missing ssDNA viruses.

1. Castro-Mejia JL, Muhammed MK, Kot W, Neve H, Franz CM, Hansen LH, Vogensen FK, Nielsen DS. 2015. Optimizing protocols for extraction of bacteriophages prior to metagenomic analyses of phage communities in the human gut. Microbiome 3:64.
2. Deng L, Silins R, Castro-Mejia JL, Kot W, Jessen L, Thorsen J, Shah S, Stokholm J, Bisgaard H, Moineau S, Nielsen DS. 2019. A Protocol for Extraction of Infective Viromes Suitable for Metagenomics Sequencing from Low Volume Fecal Samples. Viruses 11.
3. Hsieh SY, Tariq MA, Telatin A, Ansorge R, Adriaenssens EM, Savva GM, Booth C, Wileman T, Hoyles L, Carding SR. 2021. Comparison of PCR versus PCR-Free DNA Library Preparation for Characterising the Human Faecal Virome. Viruses 13.
4. Shkoporov AN, Ryan FJ, Draper LA, Forde A, Stockdale SR, Daly KM, McDonnell SA, Nolan JA, Sutton TDS, Dalmasso M, McCann A, Ross RP, Hill C. 2018. Reproducible protocols for metagenomic analysis of human faecal phageomes. Microbiome 6:68.
5. Wang G, Li S, Yan Q, Guo R, Zhang Y, Chen F, Tian X, Lv Q, Jin H, Ma X, Ma Y. 2022. Optimization and evaluation of viral metagenomic amplification and sequencing procedures toward a genome-level resolution of the human fecal DNA virome. J Adv Res doi:10.1016/j.jare.2022.08.011.
6. Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. 2020. The Gut Virome Database Reveals Age-Dependent Patterns of Virome Diversity in the Human Gut. Cell Host Microbe 28:724-740 e8.
7. Tisza MJ, Belford AK, Dominguez-Huerta G, Bolduc B, Buck CB. 2021. Cenote-Taker 2 democratizes virus discovery and sequence annotation. Virus Evol 7:veaa100.
8. Kieft K, Zhou Z, Anantharaman K. 2020. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome 8:90.
9. Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, Pratama AA, Gazitua MC, Vik D, Sullivan MB, Roux S. 2021. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome 9:37.
10. Billaud M, Lamy-Besnier Q, Lossouarn J, Moncaut E, Dion M, Moineau S, Traoré F, Le Chatelier E, Denis C, Estelle J, Achard C, Zemb O, Petit M-A. 2021. Analysis of viromes and microbiomes from pig fecal samples reveals that phages and prophages rarely carry antibiotic resistance genes. ISME Communication 1.
11. Kim KH, Bae JW. 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. Appl Environ Microbiol 77:7663-8.
12. Yilmaz S, Allgaier M, Hugenholtz P. 2010. Multiple displacement amplification compromises quantitative analysis of metagenomes. Nat Methods 7:943-4.
13. Blanco L, Bernad A, Lazaro JM, Martin G, Garmendia C, Salas M. 1989. Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. J Biol Chem 264:8935-40.

14. Roux S, Solonenko NE, Dang VT, Poulos BT, Schwenck SM, Goldsmith DB, Coleman ML, Breitbart M, Sullivan MB. 2016. Towards quantitative viromics for both double-stranded and single-stranded DNA viruses. PeerJ 4:e2777.
15. De Paepe M, De Monte S, Robert L, Lindner AB, Taddei F. 2010. Emergence of variability in isogenic Escherichia coli populations infected by a filamentous virus. PLoS One 5:e11823.
16. Sausset R, Krupova Z, Guédon E, Peron S, Grangier A, Petit M-A, De Sordi L, De Paepe M. 2022. Comparison of interferometric light microscopy with nanoparticle tracking analysis for the study of extracellular vesicles and bacteriophages. bioRxiv.
17. Klopot A, Zakrzewska A, Lecion D, Majewska JM, Harhala MA, Lahutta K, Kazmierczak Z, Laczmanski L, Klak M, Dabrowska K. 2017. Real-Time qPCR as a Method for Detection of Antibody-Neutralized Phage Particles. Front Microbiol 8:2170.
18. Fogg PC, Rigden DJ, Saunders JR, McCarthy AJ, Allison HE. 2011. Characterization of the relationship between integrase, excisionase and antirepressor activities associated with a superinfecting Shiga toxin encoding bacteriophage. Nucleic Acids Res 39:2116-29.
19. Labarde A, Jakutyte L, Billaudeau C, Fauler B, Lopez-Sanz M, Ponien P, Jacquet E, Mielke T, Ayora S, Carballido-Lopez R, Tavares P. 2021. Temporal compartmentalization of viral infection in bacterial cells. Proc Natl Acad Sci U S A 118.
20. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884-i890.
21. Aitchison, John (1982). "The Statistical Analysis of Compositional Data". Journal of the Royal Statistical Society. Series B (Methodological). 44 (2): 139–177
22. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-60.
23. Canceill D, Viguera E, Ehrlich SD. 1999. Replication slippage of different DNA polymerases is inversely related to their strand displacement efficiency. J Biol Chem 274:27481-90.
24. Lechner RL, Richardson CC. 1983. A preformed, topologically stable replication fork. Characterization of leading strand DNA synthesis catalyzed by T7 DNA polymerase and T7 gene 4 protein. J Biol Chem 258:11185-96.