# Herpes Simplex Virus Type 1 Co-Infection Leads to the Formation of Rolling Circle Amplification-Like Adeno-Associated Virus DNA Replication Products

Anita F. Meier[1, †], Kurt Tobler[1, †], Remo Leisi[2], Anouk Lkharrazi[1], Carlos Ros[2] and Cornel Fraefel[1, *]

1 Institute of Virology, University of Zurich, Zurich, 8057, Switzerland
2 Department for Chemistry and Biochemistry, University of Bern, Bern, 3012, Switzerland
† These authors contributed equally to this work.
* To whom correspondence should be addressed. Tel: +41 44 635 87 01; Fax: +41 44 635 89 07; Email: cornel.fraefel@uzh.ch

## ABSTRACT

Adeno-associated virus (AAV) genome replication only occurs in the presence of a co-infecting helper virus such as adenovirus type 5 (AdV5) or herpes simplex virus type 1 (HSV-1). AdV5-supported replication of the AAV genome has been described to occur in a strand-displacement rolling hairpin mechanism initiated at the AAV 3' inverted terminal repeat (ITR) end. It has been assumed that the same mechanism applies to HSV-1-supported AAV genome replication. We demonstrate the formation of double-stranded head-to-tail concatemers of AAV genomes in presence of HSV-1, and thus provide evidence for an unequivocal rolling circle amplification (RCA) mechanism. This study reveals the ability of AAV to modify the canonical rolling hairpin replication mechanism and to mimic the replication strategy of a co-infecting herpesvirus. This stands in contrast to the textbook model of AAV genome replication when HSV-1 is the helper virus. Furthermore, we introduce nanopore sequencing as a novel, high-throughput approach to study viral genome replication in unprecedented detail.

## INTRODUCTION

Adeno-associated virus (AAV) is a small, non-pathogenic *Dependoparvovirus* and is predominantly known for its application in gene therapy (1, 2). The wildtype AAV genome consists of a single-stranded DNA (ss DNA) with a length of 4.7 kb containing a *rep* and a *cap* coding region flanked by 145 nt-long inverted terminal repeat (ITR) sequences (3, 4). AAV can only replicate in presence of helper factors, which can be provided by co-infecting helper viruses such as adenoviruses (AdV), herpesviruses and papillomaviruses (5–7). Helper factors are required to transactivate transcription of AAV genes and support AAV genome replication. It was suggested that concatemeric AAV DNA isolated from AAV2/ AdV5 co-infected cells is organized as alternating plus and minus strands (8). However, that study did not provide definite proof of such concatemers. Nevertheless, that observation together with preceding studies on the replication of autonomous parvoviruses and the proposal of a novel replication model for linear double-stranded DNA (ds DNA) (9) led to the currently widely accepted model of the rolling hairpin amplification for AAV genome replication (9–14). According to that model the rolling hairpin amplification of the AAV genome is initiated at the palindromic terminal ITR sequences, which are capable of self-

annealing into double hairpin structures (3) and thus providing the 3'-OH primer. Extension of the annealed 3' end leads to the generation of a duplex structure, which is covalently closed at one end. To resolve the closed end, a nick is introduced into one strand of the covalently closed ITR sequence at the terminal resolution site (*trs*) by AAV Rep-proteins (15). The newly generated 3'-OH primer is then used to fill in the remaining gap to form a full-length linear open duplex configuration. Following denaturation of the ends and the subsequent re-annealing into double hairpin structures, a new 3'-OH primer is formed, and, upon strand-displacement, the progeny genome is released.

Several lines of evidence suggest that the described model is valid for AdV-supported AAV genome replication: (i) The predicted covalently closed double-stranded (ds) monomer structures have been isolated from AAV-infected cells (8, 13). (ii) It was shown that the palindromic ITR-sequences behave like origins of DNA replication (13, 16, 17). (iii) Two different ITR orientations (flip and flop) have been observed and their relative orientation to each other was found to be independent (3, 18). The observation of two ITR orientations is in accordance with the model as it predicts that the ITRs are inverted relative to each other during each round of replication. (iv) Site-specific nicking of the *trs* was observed (15, 19, 20). (v) It has been suggested that concatemers isolated from AAV/AdV co-infected cells consisted of alternating plus and minus strand AAV genomes (8). Such concatemers are believed to arise in the absence of *trs*-nicking. It has been generally accepted that herpes simplex virus type 1 (HSV-1)-supported AAV genome replication occurs in the same manner as described for AdV-supported replication.

Although never conclusively demonstrated, the widely accepted model is that HSV-1 DNA circularizes upon entry, followed by an origin-dependent bidirectional replication, which then switches to a rolling circle amplification (RCA) mechanism (21, 22). It has further been shown that addition of HSV-1 proteins essential for genome replication to non-viral plasmid DNA mediates amplification of the plasmid and leads to the formation of head-to-tail concatemers (23, 24). Since the AAV genome was found to circularize within the host cell (25, 26), which is a prerequisite for RCA, we assessed if HSV-1 mediated AAV genome replication occurs in an RCA mechanism. We explored AAV genome replication intermediates using nanopore sequencing. This technology provides the unique feature of extremely long reads and therefore enables the analysis of single replication intermediate molecules. As hypothesized, we found head-to-tail concatemers, which cannot be explained by the current model of AAV genome replication. Therefore, we suggest RCA as a novel replication model for the AAV genome amplification in presence of HSV-1 helper factors. In addition, we present nanopore sequencing as a novel, high-throughput approach to investigate the mechanism of viral genome replication. Nanopore sequencing has previously been applied to analyze the integrity of genomes of recombinant AAV vector stocks (27). However, we are unaware of any studies using nanopore sequencing to assess the mechanism of viral genome replication.

## MATERIAL AND METHODS

### Cells and viruses

BJ (foreskin fibroblasts, human), Vero (kidney, African green monkey) and A9 (fibroblasts, mouse) cells were obtained from ATCC (Manassas, Virginia, USA). All cell lines were maintained in Dulbecco's Modified Eagle's Medium (DMEM) with high glucose, supplemented with 10% fetal bovine serum (FBS), 2 mM glutamine, 100 units/ml penicillin, 100 µg/ml streptomycin, in a humidified incubator at 37°C and 5% $CO_2$. Wildtype HSV-1 (strain F) was grown as previously described (28, 29). A confluent layer of Vero cells was infected at a low multiplicity of infection (MOI) and harvested when cells showed complete cytopathic effect (CPE). Cells were scraped into media, centrifuged, and the cell pellet was frozen and thawed three times. The supernatant and cell pellet were combined and centrifuged again. Aliquots were made from supernatant and stored at -70°C. Plaque-forming units (pfu) were determined by plaque assay.

Purified AAV2 was produced by the Viral Vector Facility (University of Zurich, Switzerland) as previously described (30, 31). Wildtype AAV2 stocks were produced using the pAV2 plasmid, whereas the AAV2 variant AAV201 was produced using the pSub201 plasmid (32, 33). Stocks of MVM (rodent protoparvovirus 1; clone VR-1346; ATCC) were produced by infection of semi-confluent A9 cell culture as previously described (34). Cell culture supernatant was collected four days post-infection (dpi), cleared by low-speed centrifugation (3000 x g) and stored in aliquots at -70°C. Viral load was determined by quantitative polymerase chain reaction (qPCR) and median tissue culture infectious dose (TCID50) assay.

### Preparation of AAV2 stock sample for sequencing

The AAV2 stock sample ($1.5*10^{11}$ genome-containing particles) was pre-treated with 10U DNase I (04716728001, Roche, Switzerland) in 1x DNase buffer for 15 min at 37°C to remove any DNA outside the particles. DNase was then inactivated, and capsids were denatured at 95°C for 5 min. The sample was used directly for nanopore sequencing.

### Infection protocol

Cells were seeded 24 h before infection. BJ cells were seeded at $8*10^5$ cells/ tissue culture plate (10cm in diameter; 172958, Nunc, Thermo Fisher Scientific, Waltham, MA, USA) with 8 tissue culture plates per condition. Virus (wtAAV in the presence or absence of HSV-1) was diluted in an appropriate volume of DMEM (0% FBS). Low infection volumes were chosen to enhance infection but ensuring the complete submersion of the cell layer. Cells were placed in a humidified incubator at 37°C and 5% $CO_2$. One to two hours post infection (hpi), the supernatant was removed and sufficient DMEM containing 2% FBS was added. Cells were incubated in a humidified incubator at 37°C and 5% $CO_2$ for the indicated time. For MVM infection, A9 cells were inoculated at a confluency of 20-30% to ensure exponential cell growth during replication (35). One T-150 flask per condition was infected with either 10 ml undiluted or 1/10 diluted MVM stock to obtain 10'000 and 1'000 genome equivalents per cell, respectively. Inoculum was removed 1 hpi and replaced with 15 ml DMEM, 10% FBS. Infected cells were incubated in a humidified incubator at 37°C and 5% $CO_2$ for the indicated time.

**Hirt extraction**

Extraction of extrachromosomal DNA was performed according to the Hirt protocol (36). Cells were washed with PBS and detached using 0.05% Trypsin-EDTA. The cell pellet was resuspended in 50µl TBS (50mM Tris-HCl, 150mM NaCl, pH7.5). Then 500µl Hirt buffer (0.6% SDS, 10mM Tris-HCl, 10mM EDTA, pH7.5) was added, and the suspension was incubated at room temperature for 1h. After adding 120µl of a 5M NaCl solution, the sample was incubated at 4°C for at least 12h. For phenol/ chloroform extraction of DNA, the sample was centrifuged at 15'500 g for 10 min at 4°C. The supernatant was transferred into a fresh tube and 1 volume of phenol: chloroform: isoamyl alcohol (25:24:1, v/v) was added. The sample was centrifuged at 15'500 g for 5 min at 4°C. The supernatant was transferred into a fresh tube and 1 volume of chloroform was added. The sample was centrifuged for 1min at 15'500 g at 4°C. The supernatant was transferred into a fresh tube and 2.5 volumes of EtOH (pure) and 0.1 volume of 3M NaAc pH5.5 was added, and the suspension was incubated at -80°C for at least 20 min to precipitate the DNA. The sample was centrifuged at 18'000 g for 10min at 4°C and the supernatant was discarded. The pellet was washed with 70% EtOH. After centrifugation at 18'000 g for 10min at 4°C, the supernatant was removed and the pellet was air-dried before being resuspended in 10mM Tris-HCl, pH8.5.

**Rolling circle amplification of pUC19**

Plasmid DNA (pUC19, 2 ng) was amplified using the circular DNA amplification kit TempliPhi (25-6400-10, Cytiva, Marlborough, MA, USA) according to the manufacturer's protocol. Then, phenol/ chloroform DNA purification was performed as described in the Hirt extraction protocol above.

**Nanopore sequencing**

Nanopore sequencing was performed by the research group of Dr. A. Ramette (University of Bern, Switzerland). Hirt-extracted DNA samples were prepared for sequencing using the ligation sequencing kit (LSK109, Oxford Nanopore Technologies, Oxford, UK) and the native barcoding expansion kit (EXP-NBD104, Oxford Nanopore Technologies, Oxford, UK) according to the manufacturer's protocol. Subsequently, the samples were sequenced using a MinION sequencing device and flow cell (MIN-101B and FLO-MIN-106, Oxford Nanopore Technologies, Oxford, UK).

**Data analysis**

Sequences in fastq-format were transformed to fasta-format by seqkit (37). The reads were subjected to blastn (38) analysis against the AAV genomic sequence (Genbank # NC_001401). Lines with no hits and lines with comments were removed (grep -v '#') and tables were read into R 4.0 (The R Foundation for Statistical Computing). Hits with e-values > 0.1 were removed. The sums of all hits were calculated for every read. The reads with sum of hits < 3000 were removed. Dot plots of the remaining reads were drawn. The dot plots were manually categorized.

**Southern blot**

For Southern analysis, extrachromosomal DNA was extracted using the Hirt protocol. *Hin*dIII treatment was performed at 37°C for 1 hour in 1x reaction buffer B with 10U of *Hin*dIII enzyme per reaction (*Hin*dIII, 10656321001, Roche, Switzerland). Mung bean nuclease treatment was performed at 30°C for 30 min in 1x mung bean nuclease reaction buffer with 1U per reaction (Mung Bean Nuclease, M0250S, New England Biolabs, Ipswich, MA, USA). DNA (400 ng per sample) was separated on 0.8% agarose gels and transferred onto nylon membranes (Hybond-N+, RPN119B, Amersham, Little Chalfont, UK). As reference, DIG-labeled marker was used (DNA molecular weight marker III, 11218602910, Roche). Hybridization with a digoxigenin (DIG)-labeled probe specific for AAV2 *rep,* subsequent detection by an anti-DIG antibody conjugated with alkaline phosphatase and activation with the chemiluminescence substrate CDP Star (Roche) were performed according to the manufacturer's protocol. The DIG-labeled probe was synthesized using the PCR DIG probe synthesis kit (11636090910, Roche) and the following primers: 5'-gaa cgc gat atc gca gcc gcc atg ccg gg-3' and 5'- gga tcc gaa ttc act gct tct ccg agg taa tc-3'. Chemiluminescence was imaged using the LI-COR imaging system Odyssey Fc (LI-COR Biosciences, Lincoln, NE, USA).

## RESULTS

### Analysis of replicating AAV genomes using nanopore sequencing

AAV genome replication results in the generation of a range of different replication intermediate structures, which are mostly not detected in bulk analysis approaches such as Southern blotting. Thus, analysis of individual replication intermediates is required. Nanopore sequencing provides unique features, which are useful for the assessment of diverse nucleic acid populations. Sequencing length is (theoretically) unlimited and no fragmentation of sample material is required. Thus, all AAV genome intermediates are sequenced in full-length. For the analysis of different AAV2 genome populations, we extracted extrachromosomal DNA from BJ cells infected with non-replicating (single-infection) or replicating AAV2 (HSV-1 co-infection) using the Hirt protocol and sequenced the isolated DNA with nanopore technology (**Figure 1a**). Notably, extracted DNA was not linearized prior to sequencing and therefore circular genomes are automatically excluded from our analysis.

To facilitate the comparison of the heterogenous DNA replication intermediates, AAV reads were depicted as a function of the included genomic domains. Based on the characteristic patterns obtained, seven categories of genomic species were defined as shown in **Figure 1b**. Category 1 comprises reads showing a single AAV genome. These reads appear as a single straight line crossing over to the other side of the graph from bottom to top (positive polarity) or top to bottom (negative polarity). Category 2 includes duplex genome structures which appear in a "V" or an "A" shape on the dot plot. These reads correspond to two copies of the ss AAV genome, which are inverted relative to each other and covalently linked on one side as shown below the dot plot. These structures would arise from single-to-double-strand conversion. Category 3 contains reads consisting of head-to-tail concatemers. These concatemers appear as parallel lines crossing the plot with multiple genome repetitions with the same

orientation. We expect such structures to be formed upon rolling circle amplification (RCA). Category 4 comprises reads which contain multiple copies of the AAV genome with alternating orientation. These reads appear in the dot plot as zigzag lines. Such concatemers with alternating orientation are the expected product of a rolling hairpin amplification mechanism, where *trs*-nicking did not occur. Category 5 comprises sequences combining both head-to-tail as well as alternating repeats and appear as structures with parallel lines broken up by “V” or “A” shaped structures. We suggest that those concatemers are generated by a combined mechanism with a primary stage of RCA followed by the re-annealing of the 3’-OH ITR initiating the duplex formation of the head-to-tail concatemers, resulting in the inversion of the RCA structure. Alternatively, these structures could also arise from a recombination of two or more head-to-tail concatemers previously formed upon RCA. Category 6 includes reads containing multiple ITR- but no *rep*- or *cap*- sequences. Sequences that could not be assigned to any of the six categories were grouped in category 7. Representative dot plots of reads assigned to category 7 are shown in **Supplementary Figure 1**.

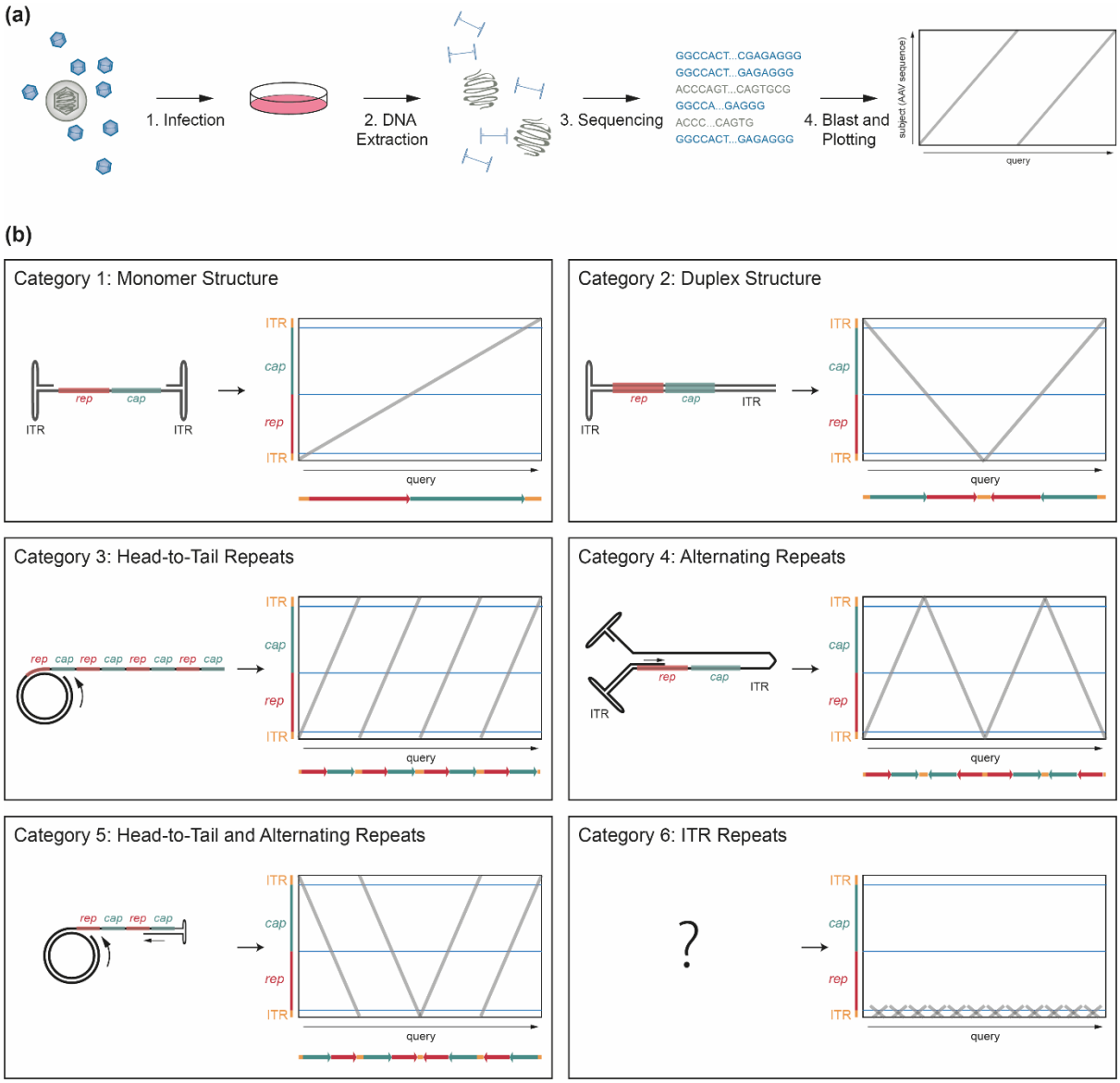


**Figure 1.** Experimental setup and definition of categories. (a) Schematic representation of the experimental setup. (b) Schematic representation of the defined categories with the predicted AAV DNA structure on the left and the corresponding theoretical dot plot on the right. Dot plots were used to visualize the AAV sequence from isolated genomes with the read (query) on the x-axis and the location on the AAV genome on the y-axis. For further clarification, the read sequence is shown below the dot plot. The AAV ITR-ends (orange) and the genome coding regions *rep* (red) and *cap* (green) are indicated.

**Sequence analysis of replicating and non-replicating AAV2**

We sequenced extrachromosomal DNA isolated from AAV2 and HSV-1 co-infected cells and found a wide range of DNA replication intermediates. These reads were categorized as described in the section above (**Figure 1b**). We categorized 200 randomly selected reads where possible or all reads of the sample if we had less than 200 reads. In **Figure 2a-f** we show two representative dot plots of individual reads of each category and the quantification of these reads is shown in **Table 1**. According to Oxford Nanopore Technology, our adapter-ligation approach should exclude any ss DNA for sequencing. Since encapsidated AAV genomes are present as ss DNA of positive and negative orientation in equal amounts (39), we assume that strand-annealing occurred prior to library preparation and thus we were able to detect single AAV genomes.

In samples of extrachromosomal DNA isolated from AAV2 infected BJ cells, the AAV DNA was mainly present as monomeric AAV genomes (category 1) (62% at AAV genome containing particles (gcp)/ cell = 20’000) **(Table 1)**. 8% of AAV reads were present in the duplex form (category 2). Most of the remaining reads were assigned to category 6, which contains reads with ITR-repeats, (19%) or to category 7 (9%), which contains all other structures. These results were in line with our expectation of predominant monomeric AAV genomes and the occasional conversion of ss to the ds DNA duplex structure.

In contrast, AAV2 reads from HSV-1 co-infected cells showed a dramatic change in the observed genome structures **(Figure 2, Table 1)**. The ratio of monomeric AAV genome reads was reduced to 34% for AAV gcp/ cell = 20’000 or 17.4% for AAV gcp/ cell = 500 **(Table 1)**. Formation of duplex structures went up to 17% or 38.5% for AAV gcp/ cell = 20’000 or for AAV gcp/ cell = 500, respectively. The most remarkable result is that HSV-1 co-infection led to the appearance of head-to-tail concatemers at 17% or 21.5% (for AAV gcp/ cell = 20’000 or for AAV gcp/ cell = 500, respectively). We also observed concatemers of alternating orientation (category 4; head-to-head) that would be expected from a rolling hairpin replication mechanism, but these reads were far less abundant than the head-to-tail concatemers (3.5% or 1% for AAV gcp/ cell = 20’000 or for AAV gcp/ cell = 500, respectively). Interestingly, structures presumably formed upon a combination of RCA and rolling hairpin amplification or RCA and recombination (category 5) were more abundant than concatemers with alternating genome orientation (category 4). 6% and 7.5% of reads from samples with AAV gcp/ cell = 20’000 or for AAV gcp/ cell = 500, respectively, were assigned to category 5.

Interestingly, in the presence and absence of HSV-1, we identified a high abundance of reads which contain multiple repetitions of ITR sequences classified as category 6 (**Figure 2f**). Those reads were present in up to 19% of all reads and were generally more abundant in absence of HSV-1. Furthermore, we also identified those ITR-repeat containing reads in the virus stock samples (**Table 1**). The source of those structures remains to be elucidated. AAV DNA, isolated from BJ cells infected with the replication-competent AAV2 variant AAV201, shows a comparable abundance and distribution of structures as seen in BJ cells infected with wtAAV2 **(Supplementary Table 1)**.

Quantification of the AAV2 stock reads showed that the vast majority (83 and 81%) were present as monomeric DNA of approximately 4.7 kb with either positive or negative polarity (category 1). The remaining population of reads were assigned to category 2 (duplex) (4 and 2%), 3 (head-to-tail repeats) (0.5 and 3%), 6 (ITR-repeats) (6.5%) or 7 (other structures) (5.5 and 7.5%) (**Table 1**). The maximum size of most of those rather unexpected reads was 5 kb, corresponding with the maximum packaging size of AAV. This indicates that a small proportion of about 10-15% of viral particles packaged defective genomes. Overall, these sequencing results were in line with our predictions, since we reasoned that most of the viral particles from the AAV2 stock sample contain a monomeric AAV genome.

In conclusion, monomeric AAV genomes were highly abundant in all samples. Complete duplex structures were observed only in DNA samples isolated from infected cells but not from AAV2 stock samples. Furthermore, head-to-tail concatemers were identified at high rates in samples from cells co-infected with AAV2 and HSV-1. Concatemers with a head-to-head structure expected from a rolling hairpin replication mechanism were found only at very low rates. We therefore conclude that HSV-1 co-infection induces the formation of RCA-like head-to-tail concatemers of the AAV genome.

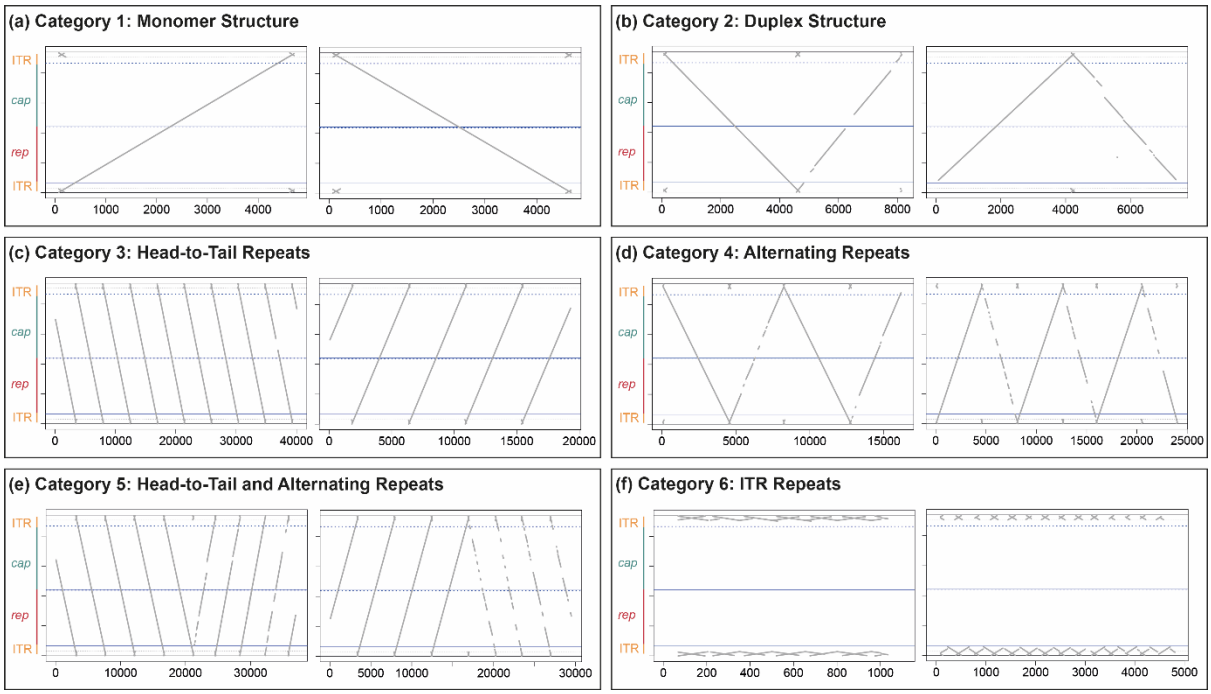


**Figure 2.** AAV2 replication intermediates. (a)-(f) Dot plots of two representative individual reads per category from AAV2/ HSV-1 co-infected BJ cells are shown. Extrachromosomal DNA was isolated at 12 hpi from BJ cells infected with wtAAV2 (gcp/ cell = 20'000) and HSV-1 (pfu/ cell = 1).

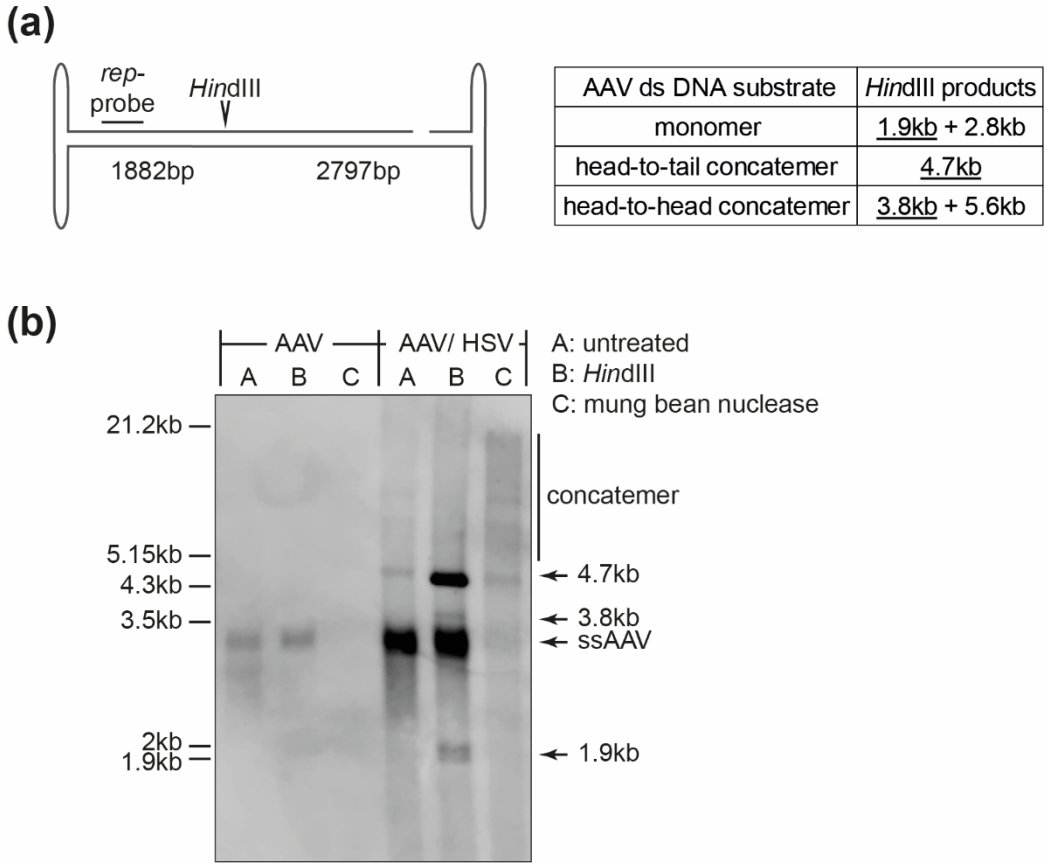
**Table 1.** Read analysis data from genomes isolated from purified AAV2 virus stock or AAV2 single- or HSV-1 co-infected BJ cells at 12hpi.

| Category: | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|---|
| Sample | | Monomer | Duplex Structure | Head-to-Tail Repeats | Alternating Repeats | Head-to-Tail and Alternating Repeats | ITR Repeats | Others | |
| | AAV gcp/ cell | ratio | ratio | ratio | ratio | ratio | ratio | ratio | total reads |
| AAV2 stock | NA | 0.830 | 0.040 | 0.005 | 0.005 | 0.000 | 0.065 | 0.055 | 200 |
| AAV2 stock | NA | 0.810 | 0.020 | 0.030 | 0.000 | 0.000 | 0.065 | 0.075 | 200 |
| AAV2 | 20k | 0.620 | 0.080 | 0.015 | 0.005 | 0.000 | 0.190 | 0.090 | 200 |
| AAV2/ HSV-1 | 20k | 0.340 | 0.170 | 0.170 | 0.035 | 0.060 | 0.100 | 0.125 | 200 |
| AAV2 | 500 | 0.250 | 0.333 | 0.000 | 0.000 | 0.000 | 0.167 | 0.250 | 12 |
| AAV2/ HSV-1 | 500 | 0.175 | 0.385 | 0.215 | 0.010 | 0.075 | 0.010 | 0.130 | 200 |

**Southern analysis of AAV genomes**

We further assessed the structure of AAV genomes isolated from BJ cells infected with AAV2 or co-infected with AAV2 and HSV-1 by Southern analysis. Hirt-DNA extraction of cells infected with AAV2 only resulted in the appearance of a single band with a migration corresponding to about 3 kb ds DNA **(Figure 3b, lane A)**. This band represents the monomeric ss AAV DNA, as it disappeared upon mung bean nuclease treatment, which degrades ss DNA **(Figure 3b, lane C)**. The same band was found in samples isolated from AAV2 and HSV-1 co-infected cells, which also disappeared upon mung bean nuclease treatment **(Figure 3b, lane A and C)**. HSV-1 co-infection led to the appearance of a mung bean nuclease-resistant 4.7 kb band, representing ds AAV2 DNA (duplex) structures. In addition, higher molecular weight bands were observed, which represent concatemers **(Figure 3b, lane A and C)**. Of note, higher molecular weight concatemers are underrepresented on the blot because these molecules transfer much less efficiently from the agarose gel onto the nylon membrane than smaller molecules. To reveal the nature of those concatemers, the samples were treated with *Hin*dIII, which cuts ds AAV2 genomes at nucleotide position 1882, within the *rep* sequence **(Figure 3a)**. *Hin*dIII digestion of monomeric ds AAV2 DNA is expected to yield fragments of 1.9 kb and 2.8 kb **(Figure 3a)**. Since the Southern probe hybridizes to a *rep*-sequence within the 1.9 kb fragment, only this band can be observed. *Hin*dIII digestion of head-to-tail concatemers would result in genome-sized 4.7 kb fragments **(Figure 3a)**. Digestion of head-to-head concatemers would result in a 3.8 kb and a 5.6 kb fragment but the hybridization probe only binds to the 3.8 kb fragment. *Hin*dIII does not digest ss DNA. *Hin*dIII treatment resulted in the disappearance of the concatemers observed in the samples from cells co-infected with AAV2 and HSV-1. Instead, a strong band of approximately 4.7 kb as well as two bands of approximately

3.8 kb and 1.9 kb appeared **(Figure 3b, lane B)**. Since the 4.7 kb band was much stronger than the 3.8 kb band, this demonstrates that AAV2 genome replication intermediates isolated from cells co-infected with HSV-1 are present predominantly as head-to-tail concatemers of ds DNA. These results are in line with our sequencing results described above.




| AAV ds DNA substrate | *Hin*dIII products |
|---|---|
| monomer | 1.9kb + 2.8kb |
| head-to-tail concatemer | 4.7kb |
| head-to-head concatemer | 3.8kb + 5.6kb |

**Figure 3.** Southern analysis of AAV2 DNA. (a) Schematic representation of the AAV2 genome with the *Hin*dIII restriction site and the *rep*-probe binding site indicated. The table on the right shows the expected sizes of the fragments upon *Hin*dIII treatment with the fragment containing the *rep*-probe binding site underlined. (b) Southern blot of Hirt DNA extracted at 12 hpi from BJ cells infected with AAV2 alone (gcp/ cell = 20'000) or co-infected with AAV2 (gcp/ cell = 20'000) and HSV-1 (pfu/ cell = 1) showing untreated (A), *Hin*dIII digested (B) or mung bean nuclease treated (C) samples.

### Validation of nanopore sequencing for the detection of an RCA mechanism

To further validate our sequencing approach for determining viral genome replication mechanisms and support our finding that the predominant product of HSV-1 assisted AAV2 genome replication are head-to-tail concatemers generated by RCA, we determined the replication products from a well-characterized RCA. We analyzed the structures generated upon the isothermal bacteriophage Phi29 polymerase mediated RCA of the plasmid pUC19. As shown in **Table 2** and **Supplementary Figure 2**, we found the same structures of parallel full plasmid repeats as described above. These reads make up 30.5% of all analyzed pUC19 reads. Surprisingly, we also found a substantial number (64%) of reads

showing a structure with a combination of head-to-tail and alternating concatemers as described for category 5. Furthermore, we found structures resembling the aforementioned duplex structures but to a lesser extent (2%). These results clearly show that Phi29-supported RCA results in the formation of structures similar to the ones isolated from AAV2 and HSV-1 co-infected cells. Interestingly, Phi29 also induces the formation of structures showing head-to-tail concatemers, which change orientation to continue further with head-to-tail concatemers. Additionally, we found that 3.5% of reads were present as structures that could not be allocated to any of the other specific categories.

**Table 2.** Read analysis data of Phi29 polymerase mediated RCA of pUC19 DNA.

| **Category:** | **1** | **2** | **3** | **4** | **5** | **7** | |
|---|---|---|---|---|---|---|---|
| **Sample** | **Monomer** | **Duplex Structure** | **Head-to-Tail Repeats** | **Alternating Repeats** | **Head-to-Tail and Alternating Repeats** | **Others** | |
| | **ratio** | **ratio** | **ratio** | **ratio** | **ratio** | **ratio** | **total reads** |
| **pUC19 RCA** | 0 | 0.020 | 0.305 | 0 | 0.640 | 0.035 | 200 |

**Sequence analysis of minute virus of mice replication products**

To validate our approach further, we analyzed sequences of replication intermediates from the autonomous parvovirus minute virus of mice (MVM). MVM genome replication has been characterized in detail and was described to occur as a modified rolling hairpin amplification mechanism similar to the one suggested for AdV-supported AAV genome replication (10, 11, 14, 40). In contrast to AAV, the MVM genome is flanked by two non-identical palindromic terminal (heterotelomeric) sequences and only one of those undergoes nicking during replication (11, 41–43). As a result, packaged MVM genomes are mainly, although not exclusively, present as ss DNA of negative orientation (1). Furthermore, MVM duplex structures have a preference for a covalently linked (left) 5'-OH-end (40). Extrachromosomal DNA isolated at 16 or 20 hpi from MVM-infected A9 cells were sequenced and analyzed as described above. Each read was assigned to one of the previously defined categories (Figure 1b). Two representative dot plots per category are shown in **Supplementary Figure 3**. Monomeric genome structures were less abundant in MVM-samples (14.5-25.5%) when compared to genomes isolated from AAV infected cells (17.5-62%) (**Tables 1 and 3**). However, as mentioned above, quantification of monomeric genome structures should be considered with caution as only ds DNA was described to be sequenced with our approach and thus depends on library preparation rather than the actual amount of ssDNA present in the sample. Nevertheless, the identification of less ss DNA genomes isolated from MVM-infected cells is in line with our previous interpretation. Unlike AAV, MVM-genomes are predominantly packaged as negative sense ssDNA and therefore only a small fraction of the genomes will anneal to a positive ssDNA and subsequently be sequenced. The most abundant structure of isolated MVM DNA was the duplex form (67-79.5%) (**Table 3**). Those duplex structures were mainly present with a covalent link at the 5'-OH-end **("V" structure; Supplementary Table 2)**. No such preference of duplex structure was found in AAV-infected cells **(Supplementary Table 2)**. These

observations are in line with the previously described model of MVM genome replication (14, 40). Importantly, no reads containing head-to-tail concatemers (category 3) were identified as MVM-replication intermediates. In contrast, concatemers of alternating repeats were indeed observed, although with a low abundance (0-2.5%). Since it has previously been described that MVM genome replication requires the formation of a dimer duplex form, it is surprising that only low levels of concatemers with alternating repeats were found. None of the reads showed a combination of head-to-tail and alternating repeats or multiple hairpin end-repeats (categories 5 and 6). No clear difference in genome structure between samples harvested at 16 or 20 hpi was observed. However, MVM genome concatemers of alternating repeats were more abundant at the lower multiplicity of infection (1'000 gcp/ cell versus 10'000 gcp/ cell). Overall, we identified replication intermediates as expected from the previously described replication model of the MVM genome (10, 11, 14, 40) and did not find any indication for the generation of RCA-like replication products. This result further validates the use of nanopore sequencing to study virus replication.

**Table 3.** Read analysis data from extrachromosomal MVM sequences isolated from MVM infected A9 cells.

| **Category:** | | **1** | **2** | **3** | **4** | **5** | **6** | **7** | |
|---|---|---|---|---|---|---|---|---|---|
| **Sample** | | **Monomer** | **Duplex Structure** | **Head-to-Tail Repeats** | **Alternating Repeats** | **Head-to-Tail and Alternating Repeats** | **Hairpin Repeats** | **Others** | |
| **gcp/ cell** | **harvest** | **ratio** | **ratio** | **ratio** | **ratio** | **ratio** | **ratio** | **ratio** | **total reads** |
| **1k** | **16hpi** | 0.225 | 0.730 | 0.000 | 0.025 | 0.000 | 0.000 | 0.020 | 200 |
| **1k** | **20hpi** | 0.145 | 0.795 | 0.000 | 0.025 | 0.000 | 0.000 | 0.035 | 200 |
| **10k** | **16hpi** | 0.184 | 0.767 | 0.000 | 0.010 | 0.000 | 0.000 | 0.039 | 103 |
| **10k** | **20hpi** | 0.255 | 0.670 | 0.000 | 0.000 | 0.000 | 0.000 | 0.075 | 200 |

## DISCUSSION

AAV2 genome replication in presence of AdV was previously described to occur in a rolling hairpin amplification mode (1, 8, 10, 14). We provide evidence for an alternative mechanism of AAV2 genome replication in the presence of HSV-1. Our results clearly show the predominant formation of head-to-tail concatemers in the presence of HSV-1, whereas the head-to-head concatemers were present in very low ratios. This finding indicates that in presence of a helper herpesvirus, the canonical AAV replication based on the rolling hairpin amplification changes to a herpesvirus-like replication mode based on RCA. Nevertheless, we cannot exclude an additional rolling hairpin amplification in presence of HSV-1 as we do observe, albeit at very low levels, concatemers of AAV2 genomes with alternating orientation. Furthermore, there is the possibility that we cannot detect concatemers of alternating repeats for the following reasons: (i) those concatemers are single-stranded and therefore excluded from our library preparation or (ii) efficient *trs*-nicking prevents its formation.

Previous studies identified the helper factors for AAV replication. The essential AdV helper factors include the gene products of *E1a*, *E1b55K*, *E2a*, *E4orf6*, and the VA RNA (44, 45). It was shown that neither the AdV polymerase nor the terminal protein mediate amplification of the AAV2 genome (44). Thus, the AdV helper factors provide an indirect support for AAV2 genome replication. The minimal set of HSV-1 helper factors required to support replication in a transient AAV infection model are the helicase-primase complex (HP; UL5/ UL8/ UL52) and the single-strand DNA binding protein ICP8 (46). Although sufficient to promote replication, addition of the HP complex and ICP8 was only capable of restoring replication to about 1% compared to wildtype HSV-1 supported replication (47). The addition of plasmids encoding the HSV-1 proteins ICP0, ICP4, ICP22 and the HSV-1 polymerase (UL30/UL42) with ICP8 and the HP complex encoding plasmids restored the levels of AAV replication to levels comparable to a wildtype HSV-1 infection (48). The HSV-1 proteins ICP0, ICP4 and ICP22 were found to transactivate AAV *rep* expression, thereby supporting replication (48). Cellular polymerases were found to mediate replication of the AAV genome in the presence of AdV5 (49, 50). In contrast, the HSV-1 polymerase was shown to directly support AAV genome replication (50, 51). The essential factors for HSV-1 genome replication comprise ICP8, UL9, the HP complex and the HSV-1 polymerase (52). It was also shown that these factors mediate a rolling circle amplification of non-viral plasmid DNA in the absence of the origin binding protein UL9 (23, 24). Thus, the minimal factors required for AAV genome replication are the same factors which are required to mediate RCA. It is therefore conceivable that the HSV-1 mediated AAV genome replication can induce an RCA mechanism. We propose that the AAV2 genome is replicated in an RCA mode if HSV-1 is present. This hypothesis is based on our observation that head-to-tail concatemers of the AAV2 genome accumulated in AAV2 and HSV-1 co-infected cells. These head-to-tail concatemers were neither abundant in AAV2 viral stocks nor in cells infected with AAV2 alone and must therefore be induced by the HSV-1 co-infection. There are two possibilities of how the head-to-tail concatemers are formed: intermolecular recombination or RCA. We regard RCA as more likely. If head-to-tail concatemers were generated by intermolecular recombination, the probability of such concatemers containing multiple genome copies with the same direction would be very low. This also means the larger the concatemers, i.e., the more genome repeats generated, the lower the probability of these concatemers to arise from genome end recombination events. Since we observed long head-to-tail concatemers containing 10, and even up to 13, repeats in the same direction, we argue that these concatemers were generated by an RCA mechanism. The reasoning behind these estimations is that every genome end has two possibilities of recombination (either with the *cap*- or the *rep*-proximal end). The likelihood that the large head-to-tail concatemers were formed upon end recombination is very low. For example, the probability for a concatemer containing 13 head-to-tail genome repeats to arise upon multiple recombination events would be 0.0002 (or 0.02%; $0.5^{13-1}$ because of 12 junctions). Furthermore, if concatemers were generated upon genome end recombination, we would expect more variety in the structure of concatemers and a higher abundance of alternating repeats.

In addition to revealing the structure of not previously described AAV2 DNA replication intermediates, we present a high-throughput method to investigate viral replication mechanisms. The unique feature of a, theoretically, endless sequencing length by the Oxford Nanopore Technology can be used to study

replication intermediates on the level of single molecules. In previous approaches, such as Southern analysis, the samples were assessed in bulk and therefore, heterogeneous low-abundant DNA replication intermediates and transient structures were not detected. The high-throughput analysis of replication intermediates supports the detection of a large variety of genomes in a quantitative manner. For example, ss and ds monomeric and concatemeric AAV2 replication intermediates have previously been shown (53, 54). Our approach revealed the formation of head-to-tail concatemers, unexpected concatemers such as long repeats of the ITR-sequence or a combination of head-to-tail and alternating genome concatemers. Further studies are required to address the question of whether, and how, head-to-tail concatemers can give rise to packageable AAV2 genomes.

**ACKNOWLEDGEMENT**

We thank Mathias Ackermann for the scientific discussions. We thank Alban Ramette and Stefan Neuenschwander for their support with nanopore sequencing. We thank Dinis Meier for his support in establishing the design of the figures. We thank Tobias Myland for proofreading the manuscript.

**FUNDING**

This work was supported by the Swiss National Science Foundation [310030_184766/1] to CF.

**CONFLICT OF INTEREST**

The authors declare no conflict of interest.

**REFERENCES**

1. Knipe,D.M. and Howley,P.M. (2013) Fields Virology 6th ed. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, PA 19103 USA.

2. Samulski,R.J. and Muzyczka,N. (2014) AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. *Annu. Rev. Virol.*, **1**, 427–451.

3. Lusby,E.W., Fife,K.H. and Berns,K.I. (1980) Nucleotide Sequence of the Inverted Terminal Repetition in Adeno-Associated Virus DNA. *J. Virol.*, **34**, 402–409.

4. Srivastava,A., Lusby,E.W. and Berns,K.I. (1983) Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J. Virol.*, **45**, 555–564.

5. Atchison,R.W., Casto,B.C. and Hammon,W.McD. (1965) Adenovirus-Associated Defective Virus Particles. *Science*, 10.1126/science.149.3685.754.

6. Ogston,P., Raj,K. and Beard,P. (2000) Productive Replication of Adeno-Associated Virus Can Occur in Human Papillomavirus Type 16 (HPV-16) Episome-Containing Keratinocytes and Is Augmented by the HPV-16 E2 Protein. *J. Virol.*, **74**, 3494–3504.

7. Buller,R.M., Janik,J.E., Sebring,E.D. and Rose,J.A. (1981) Herpes simplex virus types 1 and 2 completely help adenovirus-associated virus replication. *J. Virol.*, **40**, 241–247.

8. Straus,S.E., Sebring,E.D. and Rose,J.A. (1976) Concatemers of alternating plus and minus strands are intermediates in adenovirus-associated virus DNA synthesis. *Proc. Natl. Acad. Sci.*, **73**, 742–746.

9. Cavalier-Smith,T. (1974) Palindromic base sequences and replication of eukaryote chromosome ends. *Nature*, **250**, 467–470.

10. Tattersall,P. and Ward,D.C. (1976) Rolling hairpin model for replication of parvovirus and linear chromosomal DNA. *Nature*, **263**, 106–109.

11. Astell,C.R., Chow,M.B. and Ward,D.C. (1985) Sequence analysis of the termini of virion and replicative forms of minute virus of mice DNA suggests a modified rolling hairpin model for autonomous parvovirus DNA replication. *J. Virol.*, **54**, 171–177.

12. Tattersall,P., Crawford,L.V. and Shatkin,A.J. (1973) Replication of the Parvovirus MVM II. Isolation and Characterization of Intermediates in the Replication of the Viral Deoxyribonucleic Acid. *J. Virol.*, **12**, 1446–1456.

13. Hauswirth,W.W. and Berns,K.I. (1977) Origin and termination of adeno-associated virus DNA replication. *Virology*, **78**, 488–499.

14. Berns,K.I. (1990) Parvovirus Replication. *Microbiol. Rev.*, **54**, 316–329.

15. Im,D.-S. and Muzyczka,N. (1990) The AAV Origin Binding Protein Rep68 is an ATP-Dependent Site-Specific Endonuclease with DNA Helicase Activity. *Cell*, **61**, 447–457.

16. Samulski,R.J., Srivastava,A., Berns,K.I. and Muzyczka,N. (1983) Rescue of adeno-associated virus from recombinant plasmids: Gene correction within the terminal repeats of AAV. *Cell*, **33**, 135–143.

17. Senapathy,P., Tratschin,J.-D. and Carter,B.J. (1984) Replication of Adeno-associated virus DNA. *J. Mol. Biol.*

18. Lusby,E., Bohenzky,R. and Berns,K.I. (1981) Inverted terminal repetition in adeno-associated virus DNA: independence of the orientation at either end of the genome. *J. Virol.*, **37**, 1083–1086.

19. Snyder,R.O., Samulski,R.J. and Muzyczka,N. (1990) In Vitro Resolution of Covalently Joined AAV Chromosome Ends. *Cell*, **60**, 105–113.

20. Brister,J.R. and Muzyczka,N. (2000) Mechanism of Rep-Mediated Adeno-Associated Virus Origin Nicking. *J. Virol.*, **74**, 7762–7771.

21. Boehmer,P.E. and Lehman,I.R. (1997) HERPES SIMPLEX VIRUS DNA REPLICATION. *Annu. Rev. Biochem.*, **66**, 347–384.

22. Boehmer,P. and Nimonkar,A. (2003) Herpes Virus Replication. *IUBMB Life Int. Union Biochem. Mol. Biol. Life*, **55**, 13–22.

23. Skaliter,R., Makhov,A.M., Griffith,J.D. and Lehman,I.R. (1996) Rolling circle DNA replication by extracts of herpes simplex virus type 1-infected human cells. *J. Virol.*, **70**, 1132–1136.

24. Skaliter,R. and Lehman,I.R. (1994) Rolling circle DNA replication in vitro by a complex of herpes simplex virus type 1-encoded enzymes. *Proc. Natl. Acad. Sci.*, **91**, 10665–10669.

25. Duan,D., Sharma,P., Yang,J., Yue,Y., Dudus,L., Zhang,Y., Fisher,K.J. and Engelhardt,J.F. (1998) Circular Intermediates of Recombinant Adeno-Associated Virus Have Defined Structural Characteristics Responsible for Long-Term Episomal Persistence in Muscle Tissue. *J VIROL*, **72**, 11.

26. Sun,X., Lu,Y., Bish,L.T., Calcedo,R., Wilson,J.M. and Gao,G. (2010) Molecular Analysis of Vector Genome Structures After Liver Transduction by Conventional and Self-Complementary Adeno-Associated Viral Serotype Vectors in Murine and Nonhuman Primate Models. *Hum. Gene Ther.*, **21**, 750–761.

27. Radukic,M.T., Brandt,D., Haak,M., Müller,K.M. and Kalinowski,J. (2019) Nanopore sequencing of native adeno-associated virus (AAV) single-stranded DNA using a transposase-based rapid protocol Genomics.

28. Sutter,S.O., Marconi,P. and Meier,A.F. (2020) Herpes Simplex Virus Growth, Preparation, and Assay. In Diefenbach,R.J., Fraefel,C. (eds), *Herpes Simplex Virus : Methods and Protocols*. Springer New York, New York, NY, pp. 57–72.

29. de Oliveira,A.P., Glauser,D.L., Laimbacher,A.S., Strasser,R., Schraner,E.M., Wild,P., Ziegler,U., Breakefield,X.O., Ackermann,M. and Fraefel,C. (2008) Live Visualization of Herpes Simplex Virus Type 1 Compartment Dynamics. *J. Virol.*, **82**, 4974–4990.

30. Vogel,R., Seyffert,M., Strasser,R., de Oliveira,A.P., Dresch,C., Glauser,D.L., Jolinon,N., Salvetti,A., Weitzman,M.D., Ackermann,M., *et al.* (2012) Adeno-Associated Virus Type 2 Modulates the Host DNA Damage Response Induced by Herpes Simplex Virus 1 during Coinfection. *J. Virol.*, **86**, 143–155.

31. Zolotukhin,S., Byrne,B.J., Mason,E., Zolotukhin,I., Potter,M., Chesnut,K., Summerford,C., Samulski,R.J. and Muzyczka,N. (1999) Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther.*, **6**, 973–985.

32. Samulski,R.J., Chang,L.S. and Shenk,T. (1987) A recombinant plasmid from which an infectious adeno-associated virus genome can be excised in vitro and its use to study viral replication. *J. Virol.*, **61**, 3096–3101.

33. Laughlin,C., Tratschin,J.-D., Coon,H. and Carter,B.J. (1983) Cloning of infectious adeno-associated virus genomes in bacterial plasmids. *Gene*, 10.1016/0378-1119(83)90217-2.

34. Leisi,R., Wolfisberg,R., Nowak,T., Caliaro,O., Hemmerle,A., Roth,N.J. and Ros,C. (2020) Impact of the isoelectric point of model parvoviruses on viral retention in anion-exchange chromatography. *Biotechnol Bioeng*, 10.1002/bit.27555.

35. Tattersall,P. (1972) Replication of the Parvovirus MVM I. Dependence of Virus Multiplication and Plaque Formation on Cell Growth. *J. Virol.*, **10**, 586–590.

36. Hirt,B. (1967) Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.*, **26**, 365–369.

37. Shen,W., Le,S., Li,Y. and Hu,F. (2016) SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLOS ONE*, **11**, e0163962.

38. AltschuP,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. Basic Local Alignment Search Tool.

39. Berns,K.I. and Adler,S. (1972) Separation of Two Types of Adeno-Associated Virus Particles Containing Complementary Polynucleotide Chains. *J. Virol.*, **9**, 394–396.

40. Tullis,G., Schoborg,R.V. and Pintel,D.J. (1994) Characterization of the temporal accumulation of minute virus of mice replicative intermediates. *J. Gen. Virol.*, **75**, 1633–1646.

41. Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C. (1983) The complete DNA sequence of minute virus of mice, an autonomous parvovirus. *Nucleic Acids Res.*, **11**, 999–1018.

42. Astell,C.R., Gardiner,E.M. and Tattersall,P. (1986) DNA Sequence of the Lymphotropic Variant of Minute Virus of Mice, MVM(i), and Comparison with the DNA Sequence of the Fibrotropic Prototype Strain. **57**, 14.

43. Willwand,K., Baldauf,A.Q., Deleu,L., Mumtsidu,E., Costello,E., Beard,P. and Rommelaere,J. (1997) The minute virus of mice (MVM) nonstructural protein NS1 induces nicking of MVM DNA at a unique site of the right-end telomere in both hairpin and duplex conformations in vitro. 10.1099/0022-1317-78-10-2647.

44. Weitzman,M.D. and Linden,R.M. (2012) Adeno-Associated Virus Biology. In Snyder,R.O., Moullier,P. (eds), *Adeno-Associated Virus*. Humana Press, Totowa, NJ, Vol. 807, pp. 1–23.

45. Meier,A.F., Fraefel,C. and Seyffert,M. (2020) The Interplay between Adeno-Associated Virus and its Helper Viruses.

46. Weindler,F.W. and Heilbronn,R. (1991) A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication. *J. Virol.*, **65**, 2476–2483.

47. Slanina,H., Weger,S., Stow,N.D., Kuhrs,A. and Heilbronn,R. (2006) Role of the Herpes Simplex Virus Helicase-Primase Complex during Adeno-Associated Virus DNA Replication. *J. Virol.*, **80**, 5241–5250.

48. Alazard-Dany,N., Nicolas,A., Ploquin,A., Strasser,R., Greco,A., Epstein,A.L., Fraefel,C. and Salvetti,A. (2009) Definition of Herpes Simplex Virus Type 1 Helper Activities for Adeno-Associated Virus Early Replication Events. *PLoS Pathog.*, **5**, e1000340.

49. Nash,K., Chen,W. and Muzyczka,N. (2008) Complete In Vitro Reconstitution of Adeno-Associated Virus DNA Replication Requires the Minichromosome Maintenance Complex Proteins. *J. Virol.*, **82**, 1458–1464.

50. Handa,H. and Carter,B.J. (1979) Adeno-associated Virus DNA Replication Complexes in Herpes Simplex Virus or Adenovirus-infected Cells. *J. Biol. Chem.*, **254**, 6603–6610.

51. Ward,P., Falkenberg,M., Elias,P., Weitzman,M. and Linden,R.M. (2001) Rep-Dependent Initiation of Adeno-Associated Virus Type 2 DNA Replication by a Herpes Simplex Virus Type 1 Replication Complex in a Reconstituted System. *J. Virol.*, **75**, 10250–10258.

52. Wu,C.A., Nelson,N.J., McGeoch,D.J. and Challberg,M.D. (1988) Identification of herpes simplex virus type 1 genes required for origin-dependent DNA synthesis. *J. Virol.*, **62**, 435–443.

53. Ward,P. and Linden,R.M. (2000) A Role for Single-Stranded Templates in Cell-Free Adeno-Associated Virus DNA Replication. *J. Virol.*, **74**, 744–754.

54. Ward,P. and Berns,K.I. (1996) In vitro replication of adeno-associated virus DNA: enhancement by extracts from adenovirus-infected HeLa cells. *J. Virol.*, **70**, 4495–4501.