Molecular Therapy
Methods & Clinical Development
Original Article

Combining CRISPR-Cas-mediated terminal resolution with a novel genetic workflow to achieve high-diversity adenoviral libraries

Julian Fischer,[1] Ariana Fedotova,[1] Lena Jaki,[1] Erwan Sallard,[2] Anja Erhardt,[2] Jonas Fuchs,[1] and Zsolt Ruzsics[1]

[1]Institute of Virology, University Medical Center Freiburg, Medical Faculty, University of Freiburg, 79104 Freiburg, Germany; [2]Virology and Microbiology, Centre for Biomedical Education & Research (ZBAF), Faculty of Health, Witten/Herdecke University, 58453 Witten, Germany

While recombinant adenoviruses (rAds) are widely used in both laboratory and medical gene transfer, library-based applications using this vector platform are not readily available. Recently, we developed a new method, the CRISPR-Cas9 mediated *in vivo* terminal resolution aiding high-efficiency rescue of rAds from recombinant DNA. Here we report on a genetic workflow that allows construction of bacterial artificial chromosome-based rAd libraries reconstituted using highly efficient terminal resolution. We utilized frequent, pre-existing genomic sequences to allow the insertion of a selection marker, complementing two selected target sites into novel endonuclease recognition sites. In the second step, this selection marker is replaced with a transgene or mutation of interest via Gibson assembly. Our approach does not cause unwanted genomic off-target mutations while providing substantial flexibility for the site and nature of the genetic modification. This new genetic workflow, which we termed half site-directed fragment replacement (HFR) allows the introduction of more than 10^6 unique modifications into rAd encoding BACs using laboratory scale methodology. To demonstrate the power of HFR, we rescued barcoded viral vector libraries yielding a diversity of approximately 2.5×10^4 unique rAds per cm^2 of transfected cell culture.

INTRODUCTION

Worldwide, 14 different viral gene transfer solutions are currently approved for clinical use in humans.[1,2] Half of these therapies are based on either human adenoviruses (HAdVs) or chimpanzee adenovirus. There are also numerous ongoing clinical trials and even more research applications utilizing recombinant adenovirus (rAd)-based vectors. Due to their large capacity for transgenes, simple cloning and production principle and the ability to mediate efficient gene transfer both *in vitro* and *in vivo,* rAds are an attractive platform for gene transfer applications.[3–5]

During recent years, recombinant adeno-associated viruses (rAAV) vector technology allowed for randomized recombination of viral serotypes, generating a huge amount of novel viral capsids.[6,7] These feature libraries paved the way to screen for variants with novel tropism, allowing for the targeting of specific cells while minimizing unwanted off-target effects.[8–10] This technological advantage promoted the rAAV platform as the method of choice among the currently ongoing gene therapy clinical trials.[1] For rAds, the method of choice for finding novel vector variants remains rational design.[5,11,12] Despite being successful in some cases, rational design is limited by low throughput and, in the case of vector retargeting, by potentially unknown targeting motifs.

To allow library reconstitution, an effective method is required to achieve high-yield virus preparations carrying diversified genetic material. Recently, we developed a new method to rescue rAd that performs comparable to rAAV or recombinant lentivirus rescue. This technology is based on CRISPR-Cas9-mediated terminal resolution (CTR) and allows an efficient rAd rescue in different cell types.[13] Notably, AAV or lentivirus vector genomes can be easily modified and maintained in small high-copy plasmids in *E. coli,* simplifying library construction. In contrast, modification of large rAd genomes rarely reaches efficiencies that would permit library applications.

Here, we describe a novel method to allow reliable modification of bacmids at any given position. Utilizing CTR, we demonstrate the rescue of infectious and barcoded rAd populations that surpass the variability of currently established rAd library methods.

RESULTS

Establishment of half site-directed fragment replacement for library efficiency manipulation of rAd-bacmids

Among rAd modification technologies, only the red recombineering method allows a free choice of the mutagenesis, site but relies on a counter-selection process, making it unreliable in library applications.[14] Other methods, which can be performed with greater reliability, depend on fixed sites or modification of predetermined fragments.[15–17] This restricts the location of the genetic modification to limited sites either determined by the genomic sequences or inserted

Received 28 November 2023; accepted 14 March 2024;
https://doi.org/10.1016/j.omtm.2024.101241.

Correspondence: Zsolt Ruzsics, Institute of Virology, University Medical Center Freiburg, Medical Faculty, University of Freiburg, 79104 Freiburg, Germany.
E-mail: zsolt.ruzsics@uniklinik-freiburg.de

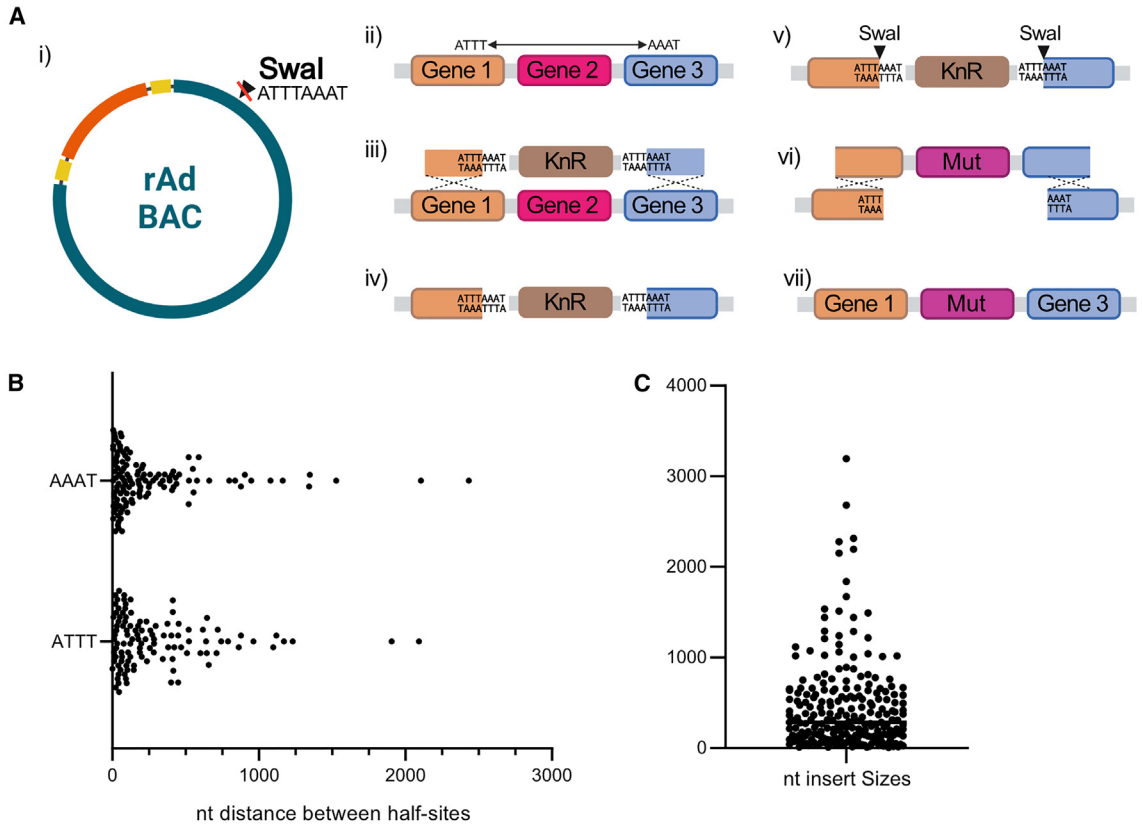

Figure 1. Schematic overview of cloning workflow and applicability on the example of HAdV-C5delE3

(A) (i) Basic cloning principle requires a bacmid containing viral genome, which is not cut by target enzyme (e.g., SwaI). (ii) Insertion site is determined by selection of target region and offset half sites. (iii) RTed recombineering is used to replace target region with Kn resistance, (iv) completing proximal SwaI half sites. (v) Intermediate construct is then digested using target enzyme, linearizing plasmid and (vi) allowing insertion of target DNA fragment, resulting in (vii) scarless modified DNA. (B) Distance in nucleotides between half sites of SwaI in exemplary bacmid containing HAdV-C5delE3 genomic DNA. (C) Minimal required insert size for re-assembly of bacmid (see A, vi).

artificially, since a unique entry point is required. In both cases, it is mandatory to ensure no off-target effects on either adjacent or targeted genomic sequences. We wanted to retain the flexibility of red recombineering for choosing the mutagenesis site, so we decided to use existing genomic features that are found frequently within genomic sequences to provide flexibility in targeting the manipulations. After initial recombination, this workflow would then utilize more reliable methodologies such as Gibson assembly[18] to generate final recombinants. We aimed to tailor the HAdV-C5 genome for Gibson assembly allowing *in vitro* manipulation of linear sequences. For this purpose, restriction enzyme recognition sequences that are natively not found within the genomic constructs have to be used. One of the commercially available restriction endonucleases SwaI cannot cut pBWH-C5-delE3 due to the absence of the ATTTAAAT recognition site (Figure 1A). Instead of artificially inserting the 8-nt site, we analyzed the sequence of the pBWH-C5-delE3 for the frequency of half sites of the SwaI recognition sequence (ATTT and AAAT). Theoretically, one of these sites occurs every 4^4 or 256 nucleotides. Within the analyzed 40,498-bp construct, we found 112 forward-facing half sites (ATTT) and 125 backward-facing half sites (AAAT) within the rAd genomic region. The average distance ranged between 2,093 nt and 6 nt for the forward sides and between 2,423 nt to 5 nt for backward sites, averaging 303 nt and 270 nt, respectively (Figure 1B).

To create the full SwaI sites based on these half sites, one needs to use one forward-facing and one reverse-facing site in combination as the target for red recombineering. Of course, if a SwaI site from two half sites is generated the central sequence between both sites will be lost and needs to be reinserted together with the modified feature(s) to maintain the integrity of the viral genome. In pBWH-C5-delE3, the required minimal re-insertion fragment sizes ranged between 6 nt and 3,191 nt (Figure 1C). Given these size constraints, the re-insertion fragment can be a synthetic dsDNA fragment, directly generated from the bacmid via PCR or derived from sub-cloned transfer vectors if the fragment can be cut by seamless nuclease digest. Using the SwaI half sites analyzed above, the mutagenesis of the rAd genomes would consist of two steps. First, the two SwaI half sites adjacent to the site of the intended mutation are complemented by red recombination and selected via antibiotic resistance testing introduced alongside

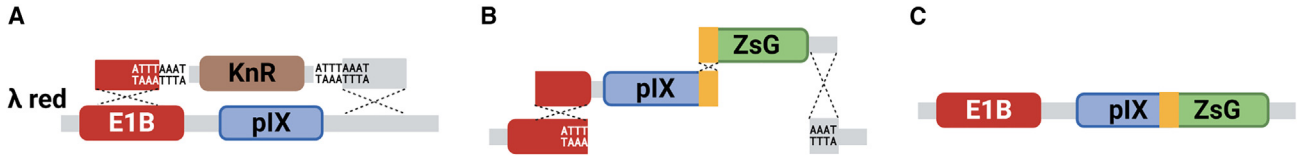

Figure 2. Schematic genetic workflow used for tagging pIX using a T2A peptide linker with ZsGreen

(A) replacement of pIX HFR locus with KnR. (B) Reinsertion of pIX using PCRs fragments for both pIX and ZsGreen, connected via T2A (orange). (C) Finished modified pIX-locus with T2A-linked ZsGreen.

completed half sites, and secondly, the intended sequence alongside the lost wild-type (wt) sequences are inserted (see Figure 1A). As existing genomic features directly determine insertion sites, we decided to coin our methodology half site-directed fragment replacement (HFR).

To illustrate the general applicability of the system for adenoviruses, we chose to further analyze representative types of all HAdV species for the presence of half sites of non-cutting enzymes. All analyzed viral genomes were found to be suitable for our approach, with an average insertion size ranging from 113 to 561 bp and potential maximum insertion sizes being from 1 kb to 4 kb, still permitting easy insertion of synthetic or PCR-derived DNA inserts (Table S1).

Using HFR as a tool to generate single HAdV-C5 constructs expressing ZsGreen by tagging viral structural protein pIX

To demonstrate the utility of our system, we decided to generate recombinant viruses coding a modified capsid protein pIX that is C-terminally fused with a T2A-peptide[19] linked ZsGreen.[20] To achieve this, we first replaced the genomic sequences between a forward SwaI half site found in the E1B region and a backward half site found in the intergenic region between the pIX and the IVa2 genes by a kanamycin-resistance-cassette (KnR) flanked by complemented SwaI sites (Figure 2A). This resulted in an intermediate clone, which can be linearized by SwaI digestion generating a linear genomic vector lacking the KnR fragment and harboring the pre-selected SwaI half sites at each end (Figure 2B). Overlapping insert fragments compatible with Gibson assembly were generated by PCRs using both viral DNA and the synthetic ZsGreen gene as a template (Figure 2B). Notably, such overlapping sequences can be either found in the vector or added by primer design. After generating the intermediate pBWH-C5-pIX-Kan bacmid by red recombineering,[14] it was SwaI digested and Gibson assembled with the two PCR-derived fragments (Figure 2C).

This resulted in a highly efficient assembly of recombinant bacmid, verified by both RFLP and Sanger sequencing. After rescue, four viral clones were selected and expanded. Extraction of viral DNA and subsequent RFLP analysis showed a length shift from 2.9 kb to 3.7 kb, indicating the insertion of the transgene in all four single clones as predicted considering DraIII recognition sites (Figure 3A). A complete overview of all *in silico* predicted fragments, including cut position and size can be found in Table S2. We then verified ZsGreen expression after infection of 293A cells with a representative viral clone by fluorescent microscopy (Figure 3B).

To analyze if our genetic modification affected viral growth or thermostability, growth kinetics and thermostability assays were performed using a representative viral clone in comparison to the unmodified wt virus. The pIX-ZG-derived virus showed no significant growth attenuation as determined by infection of 293A cells and titration of lysates at indicated time points or alteration in thermostability by measuring infectivity of 293A cells after heat-treating viral preparations. This indicates that the genetically modified virus had no apparent fitness disadvantage in cell culture compared to the wt virus (Figures 3C and 3D) as was already reported before for pIX fusion products.[21]

Using HFR to generate libraries of barcoded rAd allows high recovery rates of possible assemblies at high efficiency

As we previously published a method permitting efficient rescue of rAds suitable for libraries coined CTR[13] (a brief overview of CTRs basic principle is provided in Figure S1), we only lacked a method allowing for efficient and clean cloning of bacmid libraries. After our initial successes with HFR outlined above, we decided to use this technology for the generation of viral genomic libraries. Instead of inserting a large library of transgenes or viral mutations, we decided to barcode a single site to assess the library's diversity unbiased by downstream processes and viral fitness.

For the barcoding strategy, we wanted to achieve maximal efficiency and complex randomization. Since the efficiency of Gibson assembly inversely correlates with the number of fragments assembled and the size of the final construct, a single DNA fragment should contain the barcoded region. This ensures that the diversity is not affected by the number of inserts required during cloning. An additional advantage is that a single fragment derived from a small circular vector is more efficiently and reliably cloned and thereby diversified compared with a bacmid-derived DNA.[22–24] Therefore, we designed a shuttle plasmid to release the artificial library in a single fragment. The pO6-fC5-GFP-multiple cloning site (MCS) carried a DNA fragment consisting of the adenoviral packaging signal, a GFP transgene harboring a MCS for the introduction of the barcodes and part of the pIX locus (Figure 4A). The insert was flanked by SapI sites to ensure seamless digestion.

Next, we used overlapping oligonucleotides to assemble randomized barcodes and insert them into pO6-fC5-GFP-MCS (Figure 4B). This was achieved by assembling two, three or four oligonucleotides each carrying a barcode consisting of four randomized nucleotides (8N,

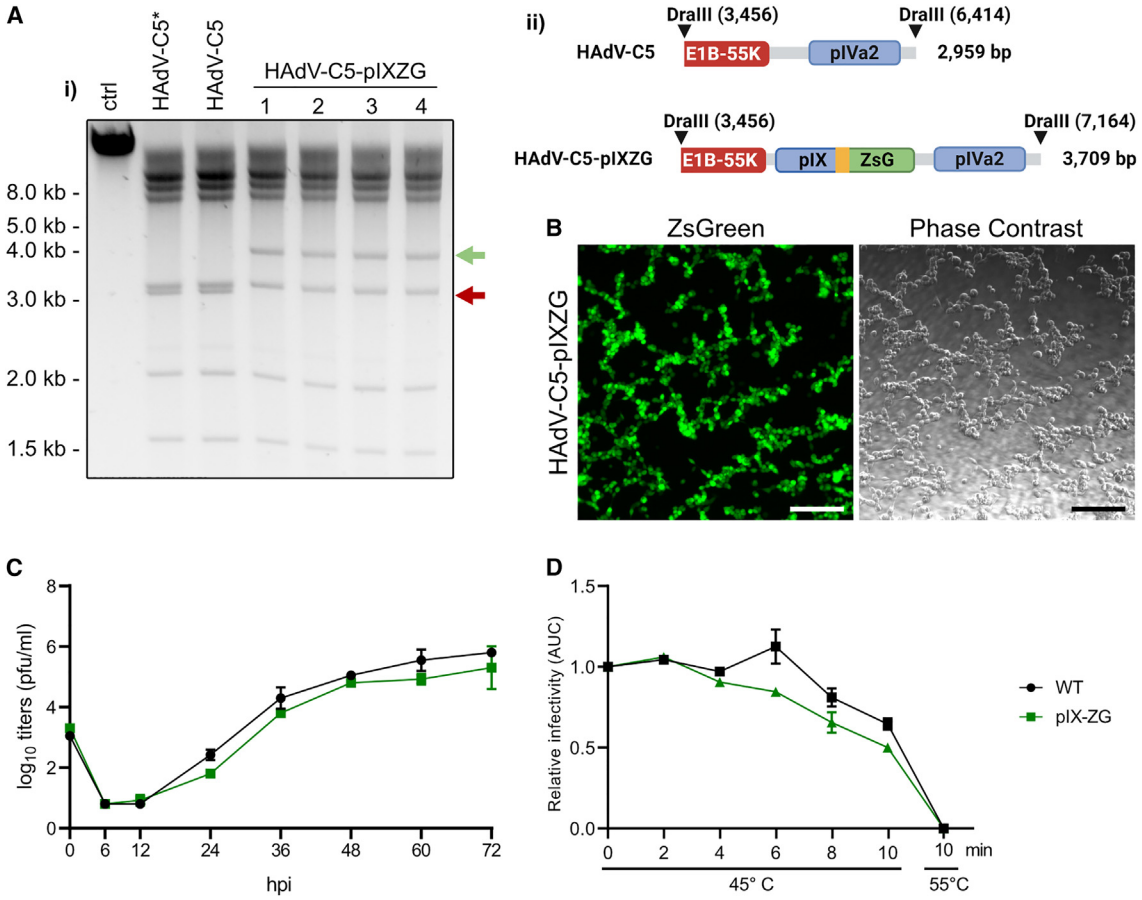

Figure 3. Analysis of fluorescently labeled wt HAdV-C5

(A) (i) RFLP analysis of extracted viral DNA, digested using DraIII. ctrl, undigested HAdV-C5 DNA derived directly from inoculum; HAdV-C5*, wt HAdV-C5 DNA derived directly from inoculum; HAdV-C5, wt-like HAdV-C5 DNA generated using CTR; HAdV-C5-pIXZG, recombinant HAdV-C5 DNA carrying ZsGreen-tagged pIX protein; red arrow indicates height of expected wt fragment, green arrow indicates height of expected fusion fragment, size difference visualized in (ii) schematic. (B) Microscopy of HAdV-C5-pIXZG infected 293A cells using fluorescent channel and phase contrast. Scale bar, 200 μm. (C) Growth curve of recombinant HAdVs, comparing HAdV-C5 and HAdV-C5-pIXZG. (D) Thermostability experiment of recombinant HAdVs, comparing HAdV-C5 and HAdV-C5-pIXZG.

12N, and 16N). Using the 8N barcode gives a theoretical library diversity of 65,536. According to our previously published data,[13] this diversity should not be limited by the virus rescue within smaller rescue experiments. To compare the theoretical diversity with that of the HFR-generated libraries, we first generated an 8N-barcoded shuttle plasmid and transferred the insert into our bacmid. We sequenced each step in the library propagation process by next-generation sequencing (NGS) and analyzed the diversity with capture-recapture population modeling for heterogeneous populations as described by Chao,[25] giving us an estimate for minimal diversity predicted within the analyzed sample (Figure 4C). Sequencing was carried out on 200-bp PCR products of the barcoded region. Both pO6-subcloning and assembled bacmid-libraries showed similar levels of minimal expected barcode diversity at 5.61×10^4 and 5.90×10^4 unique barcodes, respectively (see Figure 4D). The diversity was close to the theoretically achievable number of barcodes indicating that the diversity was likely not limited by the assembly workflow. Next, we rescued recombinant viruses carrying the barcoded transgenes in either 20 cm^2 or 60 cm^2 cell culture dishes. Since our viral DNA samples were contaminated with a low level of bacmid-derived DNA even after benzonase treatment, we further purified our samples by DpnI digestion (Figure 4E). As the DpnI recognition sequence is present in the PCR product used for NGS, it should only affect non-encapsidated, bacmid-derived sequences due to the dam-methylated-GATC-specificity of the enzyme. Indeed, this treatment completely abolished PCR detection of bacmid- but not virus-derived DNA (Figure 4E). The library diversity in the virus population was again assessed by NGS. The estimated diversity was with 4.69×10^4 and 5.97×10^4 unique barcodes for the 20 cm^2 and 60 cm^2 rescues highly similar to that of the pO6-subcloning and bacmid libraries (Figure 4F). Nucleotide frequencies of the viral-encoded barcodes were randomly distributed indicating that libraries were not biased to specific barcodes (Figure 4G). In each step of the library generation workflow, barcodes were consistently detected and reached a near-to-theoretical diversity (Figure 4H).

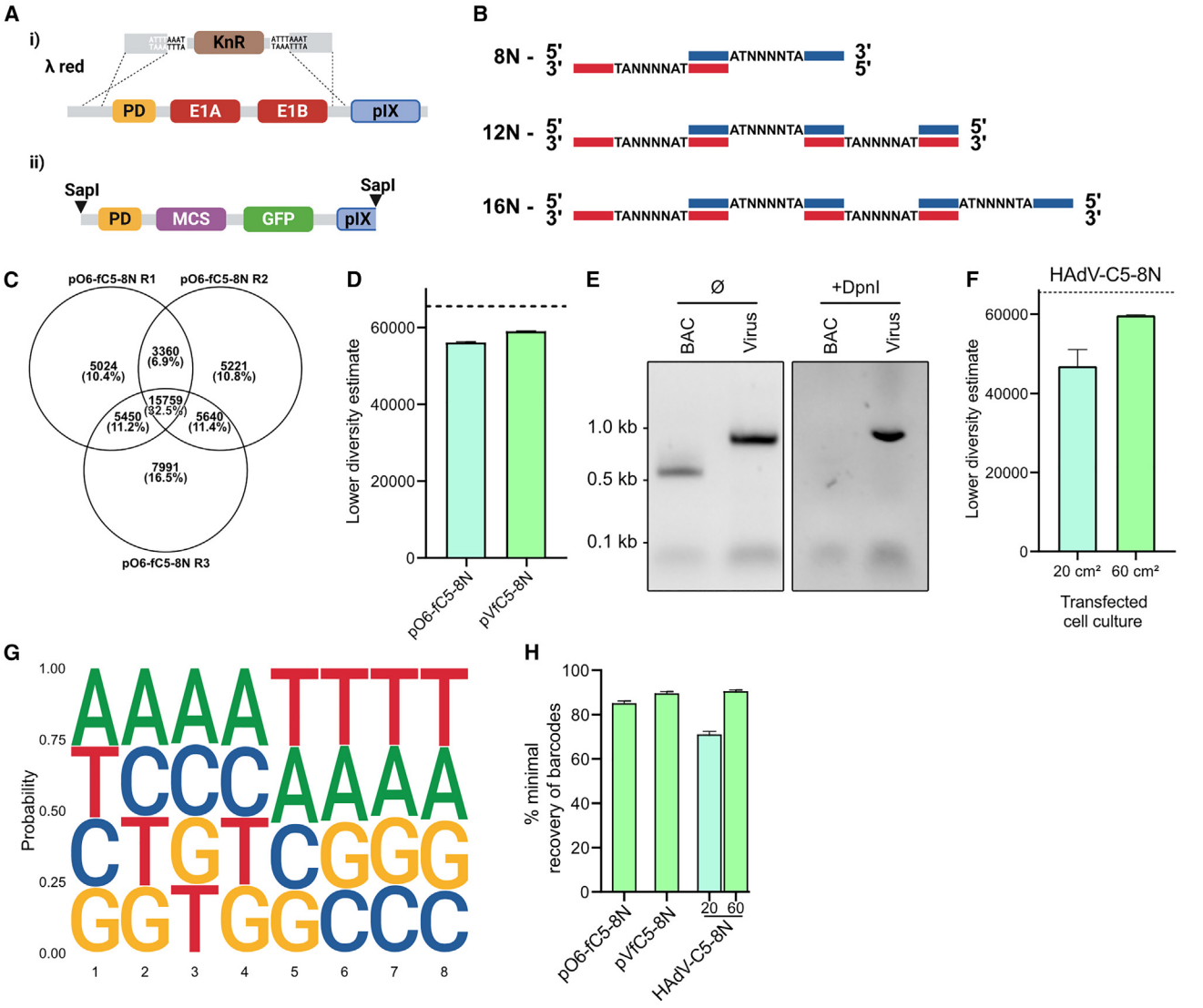

Figure 4. Overview of library cloning and analysis principle and 8N-diversified plasmid, bacmid, and virus libraries

(A) Schematic of (i) targeted region for transgenic insertion in bacmid, showing target removed region resulting in pBWH-fC5-E1Kn and (ii) fragment released from pO6-fC5-GFP-MCS via SapI digestion, designated for bacmid insertion. (B) Overview of oligonucleotide-assemblies used for generation of diversified libraries, containing barcodes of 8, 12, or 16 nucleotides in length. (C) Schematic view of sequencing overlaps of barcodes recovered via NGS workflow, shown on three sequencing runs of pO6-fC5-8N. (D) Lower diversity estimate of plasmid (pO6-fC5-8N) and bacmid (pVfC5-8N) containing 8N-nucleotide diversified barcodes, the dashed line indicates the maximum possible diversity. (E) Agarose gel showing PCR of extracted encapsidated DNA, using primers targeting either specifically bacmid or viral genomic DNA, pre- and post-DpnI digestion. (F) Lower diversity estimate of viral libraries containing 8N-nucleotide diversified barcodes, rescued in either 20 cm^2 or 60 cm^2 of transfected cell culture, the dashed line indicates the maximum possible diversity. (G) Frequency plot showing nucleotide frequency per position in barcode based on barcodes recovered from viral libraries. Sequences were aligned to strands shown in Figure 4B. (H) Overview of minimal estimated recovery rates for 8N-nucleotide-based barcodes for different steps in the process, presented as percent of maximal possible diversity.

Therefore, we analyzed the limits of HFR by increasing the barcodes from 8N to 12N or 16N, (Figure 4B), theoretically leading to diversities of 1.7×10^7 and 4.3×10^9 unique barcodes, respectively. Intriguingly, for both barcoding strategies, we estimated between 10^6 and 10^7 unique barcodes for the pO6-subcloning and the bacmid libraries (Figure 5A). As the library workflow reaches comparable levels of diversity in both bacmids, it is reasonable to assume that the bacmid libraries act as a limiting factor. This was substantiated by the recombinant virus rescue. The 12N and 16N virus libraries produced from 20 cm^2 or 60 cm^2 dishes resulted in diversity estimates of $5.17 \times 10^5/4.33 \times 10^5$ and $1.44 \times 10^6/1.36 \times 10^6$, respectively (Figure 5B). Unlike 8N-coding libraries, these higher diversity libraries showed a clearer relation between the surface area of transfected cell culture and barcode diversity. Notably, libraries from these

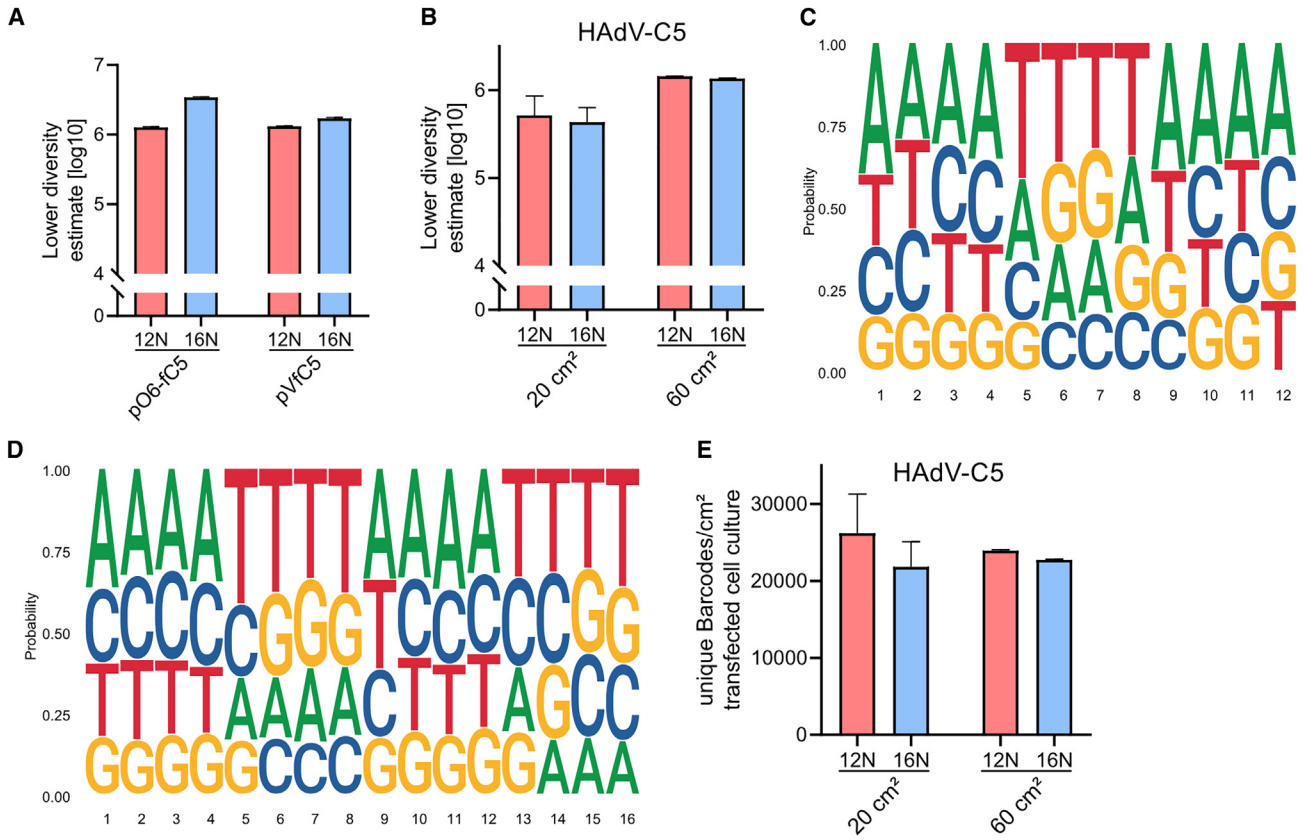

Figure 5. Analysis of 12N- and 16N-diversified plasmid, bacmid, and viral libraries

(A) Lower diversity estimates of pO6-based and bacmid cloning steps for libraries containing 12N- and 16N-diversified barcoded libraries. (B) Lower diversity estimate of viral libraries containing 12N- and 16N-nucleotide diversified barcodes, rescued in either 20 cm^2 or 60 cm^2 of transfected cell culture. (C) Frequency plot showing nucleotide frequency per position in barcodes based on barcodes recovered from 12N-diversified viral libraries. (D) Frequency plot showing nucleotide frequency per position in barcode based on barcodes recovered from 16N-diversified viral libraries. (E) Efficiency of barcode recovery for 12N- and 16N-diversified viral libraries normalized to cm^2 of transfected cell culture.

transfections showed a position-independent bias toward thymine or the reverse complement adenine which likely represents a slight bias of the PCR or the synthesized nucleotides rather than a bias by the workflow itself (Figures 4B, 5C, and 5D). As both library preparations reach similar diversity estimates, despite having vastly different potential diversities, we can normalize these estimates by the area of transfected cell culture used for generation libraries. This reveals a clear upper boundary for the here presented barcoding strategy of approximately 2.5×10^4 unique barcodes per cm^2 of transfected cell culture (Figure 5E). In summary, HFR combined with CTR produced highly diverse transgene libraries that were consistent over all steps in the assembly and rescue workflow.

HFR modification and CTR rescue provide highly reliable viral rescues while preserving DNA integrity

One major concern during the generation of libraries is the presence of wt sequences that might be preferably selected compared to tailored particles. Using HFR as a tool for generating libraries has the benefit of generating intermediate constructs, removing the possibility of contaminations with wt viruses. Attempting CTR with the intermediate pBWH-fC5-E1Kn construct yielded no viral particles due to the removal of the essential packaging domain during the construction of the assembly vector. Using a clonal intermediate plasmid allows for efficient screening for the successful selection of correct clones, removing potential background from recombineering. Utilizing transfer plasmids allows greater freedom in design, allowing the addition of additional features such as restriction sites to ensure the purity of the plasmid library, removing the background of miss-assembled libraries as well. Indeed, sequencing of viral DNA obtained from our library preparations against the original plasmid revealed a complete lack of wt sequences within mapped reads (Figure S2). Furthermore, we also tested the overall genome integrity of the viral libraries to analyze for the presence of off-target mutations. We were only able to detect mutations already present in the initial wt stock compared with the HAdV-C5-delE3 reference genome. Furthermore, frequencies of the detected mutations were highly similar in all samples indicating that the barcoded and wt rescued viruses were near-to-identical (Figure 6). A detailed overview of sequenced polymorphisms

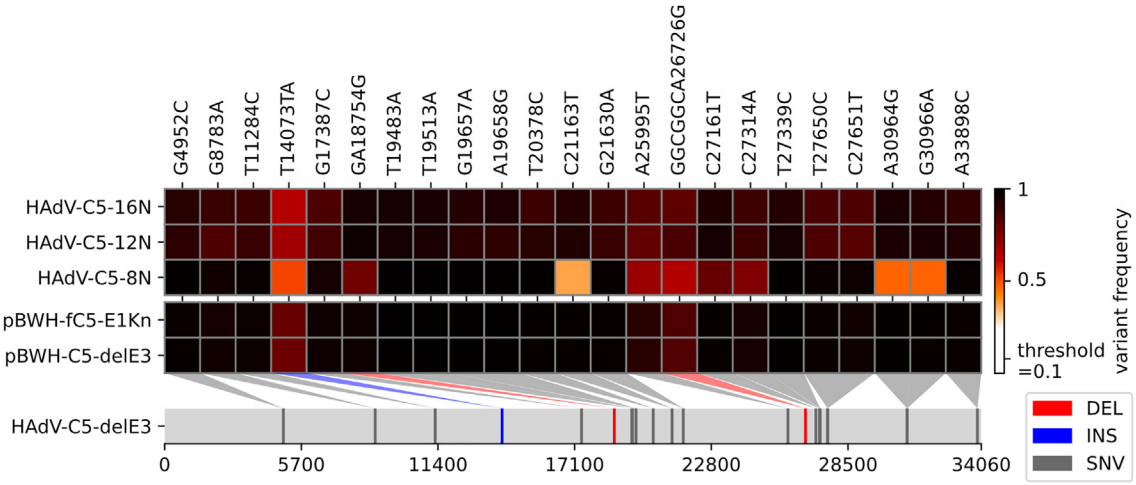

Figure 6. HFR does not cause any mutational hotspots off-target

Variant frequency plot of analyzed viral libraries showing variant frequency for detected SNPs and indels relative to HAdV-C5-delE3 viral reference genome.

and their effect on CDS can be found in Table S3. Notably, variants not occurring at 100% allele frequency were mostly called in highly repetitive genomic regions, which are notoriously difficult to amplify and sequence using established methods.[26–28] In combination with the data obtained from single mutants analyzed for pIX-targeted mutagenesis, HFR did not cause off-target effects during library generation.

DISCUSSION

Handling and modifying large DNA genomes is not a trivial task. Bacterial artificial chromosome (BAC) technology substantially aided this approach for large DNA viruses.[11,29–32] During recent years, several advances were made in this field by further improving the usability and flexibility of the bacmid system.[33–36] These innovations also spread to the field of rAds, often utilizing the high efficiency of Gibson assembly for streamlined adenoviral modification.[16,17,37,38] The major advantages of these new approaches are increased speed and reliability. In this work, we now further push the rAd technology by combining recent advancements with the flexibility of red recombination into a method coined HFR. We then apply this method in combination with our previously established CTR[13] to construct highly diverse rAd transgene libraries. CTR takes advantage of the higher transfection efficiency of circular DNA molecules,[39] a principle already used previously to facilitate more efficient rAd reconstitution,[40] and combines it with the idea of intracellular genome linearization, also utilized previously to further increase rAd reconstitution efficiency.[41] Unlike previous work, CTR does not rely on endonuclease digest but rather facilitates linearization via gRNA-targeted double-strand breaks by Cas9. We found that this increased primary rescue efficiency by approximately 50-fold.[13] To achieve high-complexity and high-purity libraries we were, however, lacking a method efficiently modifying bacmids, which HFR is now aiming to rectify.

While previous work already investigated the applicability of libraries in the adenoviral context, the workflow presented here improves both the complexity of libraries by 10- to 100-fold and the purity of generated rAd libraries as the whole viral population was correctly recombined while previous work focused mostly on either high efficiency or high purity.[42–44] An overview of previously published rAd libraries can be found in Table S4. A key feature of this novel technology is the flexibility in the modification site while minimizing off-target effects, such as scars within the viral genome. Moreover, all steps of the library generation can be performed using commercially available standard reagents, also permitting a wide range of applications, as inexpensive oligonucleotides can be used for both diversification and linkage of individual fragments via PCR overlaps.

Using different lengths of randomized barcodes, the theoretical diversity varies by several orders of magnitude. Interestingly, we reached very similar levels of diversity in both higher variance approaches. Therefore, we hypothesize that we have reached the maximal possible diversity of approximately 7.5×10^4 unique clones per ng of transformed subcloning vector DNA and approximately 1.9×10^4 unique clones per ng of bacmid DNA in our experimental setting (Figure 5). Since we used multiple fragments for the assembly of our shuttle plasmid, which is a less efficient approach compared to a two-fragment assembly,[45] the efficiency might be improved further in future applications. During the cloning of single recombinant bacmids, we found the primary determinant of cloning success is the nature of the linear DNA fragment used for bacmid re-assembly (Figure 1A, vi). Using enzymatically released fragments provided the most success here, so we suggest using the shown method of generating specific sub-cloned libraries first. The bacmid assemblies on the other hand are already optimized, but linear upscaling could still increase diversity. However, great care must be taken when applying this method, as larger or smaller inserted cassettes might also impact efficiency.

Based on our prior work, the observed reduced diversity after CTR compared to the prior assembly was expected as rescue of rAd virus remains the major limiting factor.[13] However, the here reached diversities are comparable to concurrent viral vector based-library platforms utilizing lentivirus or AAV-based vectors[46,47] and are linearly scalable. In contrast with lentiviruses or AAVs, first-generation adenoviruses are replication competent in the presence of E1 *trans*-complementation. Therefore, linear scalability allows for an increase in library diversity similar to the expansion-based approaches used to overcome the limitations of bacteriophage vectors.[48] With this in mind achieving 10^8 complexity libraries for rAd seems feasible even at the laboratory scale considering the here demonstrated complexity of approximately 2.5×10^4 per cm^2 of cell culture. Higher diversities might yet be achieved by scaling even further, however suspension cultures could become interesting to apply at such scales as they allow high-yield and scalable solutions for adenovirus production.[49–51] While classical rAd reverse genetics approaches might still be used for lower diversity library applications CTR seems to be the only viable option to achieve this at a reasonable scale based on rescue efficiencies observed in our previous study.[13]

Using the lower variability 8N approach for assembly, we were also able to recover a vast majority of available barcodes. We included this setting as a proof-of-concept that HFR is usable for full-scale recovery of genomic libraries as long as the total diversity of input is suitable for assembly and to show that the methodology to estimate the diversity was sound. The here used statistical capture/recapture model is an estimate of total population size and typically under rather than over-estimates population sizes.[25] As such, the actual recovery rate might be even higher than anticipated. However, manual inspection of the primary data obtained from the sequencing results seems to collaborate with the obtained estimates.

Here, we used a transgene barcoding strategy which has no direct applicability as a proof-of-principle. Nevertheless, this technology can potentially be used for various library generation approaches such as CRISPR-forward screens or functional immunoglobulin screening. Even at the lowest estimate, we can achieve a comparable diversity to previously used AAV technology already applied for phenotypic screenings.[10,52] Using the same principles of feature-oriented mutagenesis utilized for AAV, we predict that rAds can be used similarly to further improve vector performance. HFR might be directly applicable for *in vitro* evolution-based approaches of adenovirus capsids and might be a feasible alternative to currently available indirect applications based on phage displays aiming to find fiber loop insertions or mutations providing additional or adjusted tropism.[53,54] Direct screening of these modifications within rAd would eliminate false-positive hits of binding fiber proteins unable to functionally integrate into the viral capsid, decreasing the overall workload required downstream for the selection of functional variants. Previous studies used rational design on de-targeted vectors[55] to find novel rAd vector variants with target tropism. Our library workflow could here instead be used to find binding partners where interacting peptide sequences are not fully known or understood. Due to a common bacmid intermediate being the basis for this screening, subsequent tests of different structures or insertion sites could also be performed without requiring time-intensive recloning. In AAV, such library applications were already successfully performed to identify novel tropism variants out of approximately 10^8 unique variants.[52] Other possible applications would be the identification of cell-type-specific viral promotor variants as was performed previously to allow selective HSV1 replication in cancer cells[56] or the identification of immune escape variants by modifying immune epitopes displayed on the virion surface, permitting repeat treatment of patients with the same virus. Selection and screening for gaining fitness by serial passages in restrictive conditions seem to be feasible, since using the presented methodology rAd libraries can be built on replication-competent rAds. For library applications outlined above aiming to find fiber variants, this would consist of passaging on cells not usually permissive to rAd infection transfected to display target receptors. As only viruses displaying the desired phenotype will be able to enter cells and propagate, a few repeat rounds of passaging should yield desired variants. For the selection of immune escape variants, passaging can be performed on regular target cells (e.g., 293A) under the presence of neutralizing antibodies or patient sera. Such screening approaches were already applied successfully for rAd to find variants with improved oncolytic potential[57,58] or variants resistant to antibody-mediated neutralization.[59] In contrast with these works, the here presented workflow allows more precise and directed modification of target genes rather than a randomized genome-wide screening. Judging from previous works performed on AAV we assume that the here showcased levels of diversification should be sufficient to generate several hits to be investigated further for these applications via randomized sequence insertion or mutation.[52] However, reaching the levels of diversified functional virus shown here for *in vitro* evolution is unlikely as potentially vast amounts of variants might prove non-functional or display notably reduced fitness. Rational design still needs to be applied to library generation to ensure maximizing the possibility of selecting desired and viable virus variants.

Other high-diversity applications, such as phage display, could also be tailored into rAd-libraries, allowing eukaryotic expression of target protein.[60–62] This would improve both the speed and applicability of the screening platform to allow faster advances in both research and medical purposes. Lentiviral libraries are already used in similar applications[63] however, rAd provides a broad tropism, allowing for wide phenotypic screenings and monogenomic transductions.[64,65] Screening for these variants would be performed similarly to viral libraries outlined above and were already performed previously,[66] but require more consideration concerning the selection process as transgenes do not necessarily impact viral fit. As such, viral fitness is not a valid readout as a baseline for transgenic libraries. Depending on the target application, the large capacity of rAd can also provide further benefits. As vectors for library applications used in this study were designed following the design of first-generation adenoviral vectors lacking both E1 and E3 genes,[67] approximately 5 kb of original viral DNA were removed. Considering previous works showing

Table 1. Overview of bacterial strains used in this study

Strain	Features	Product ID
E. coli NEB10beta	Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14- Φ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL (StrR) rph spoT1 Δ(mrr-hsdRMS-mcrBC)	NEB C3020K
E. coli Pir1	F- Δlac169 rpoS(Am) robA1 creC510 hsdR514 endA recA1 uidA(ΔMluI)::pir-116	Invitrogen™ C101010

adenoviruses being able to stably package and maintain genomes up to 105% of their original size[68] for a total capacity of 38 kb gives us a transgene capacity of approximately 7 kb while retaining the simple production principle and scalability in cell culture helping rAds ease of use. Applications such as screening for oncolytic viruses might require replication-competent rAd vectors. In these cases, E1 should be retained within the vector still allowing for insertions of transgene cassettes approximately 4 kb in size due to removal of E3. While possible, larger insertions are generally not recommended due to the instability of resulting constructs.[68]

MATERIAL AND METHODS

Tables for diversified oligonucleotides, primers and plasmids used in this study can be found in Tables S5–S7.

Cell and viruses

We cultured 293A (Thermo Fisher Scientific, Waltham, MA, USA, Invitrogen #R70507) and A549 (ATCC, Manassas, VA, USA, CCL-185) cells using DMEM (Anprotec, Bruckberg, Germany, #AC-LM-0014) supplemented with 10% fetal calf serum (PAN-Biotech, Aidenbach, Germany, #P30-3306) and penicillin-streptomycin (100 U/mL; Thermo Fisher Scientific, Gibco #15140-122). The wt HAdV-C5 was cultured from inoculum provided by Albert Heim (German Adenovirus Reference Laboratory, Hannover, Germany). HH-Ad5-VI-wt[6] was cultured from an inoculum provided by Harald Wodrich (Microbiologie Fondamentale et Pathogénicité, Bordeaux, France).

Bacterial strains

Table 1 lists all the bacterial strains used in this study for the maintenance and modification of plasmids.

Construction of recombinant bacmids and transfer plasmids used for library assembly

First, CTR-compatible bacmids for replication-competent and first-generation recombinant HAdV C5 genomes were generated as previously described.[13] To construct the bacmid carrying the wt genome (pBWH-C5), we fused PCR amplified vector fragments based on pKSB2[69] backbone using primer pairs BWHC05for/GHBrev and GHBfor/BWHC05rev with viral genomic DNA isolated from infected cells. We used the same approach to clone an E3-deleted genome using viral DNA isolated from cells infected with HH-Ad5-VI-wt, which is a replication-competent HAdV-C5 strain, which carries an E3 deletion.[70] Finished constructs were verified using restriction fragment-length polymorphism (RFLP) analysis and sequencing and checked for viability using CTR.

For completing the half sites to new recognition sequences of restriction endonuclease SwaI, which we utilized in our new genetic workflow, we used recombineering as described by Datsenko and Wanner.[14] As linear fragments for recombination, gene cassettes encoding for Kn resistance flanked by the completed SwaI sites were amplified via PCR on a pGPS1.1 template. Primers were designed to contain 30-nt 5′-homologous regions to genomic regions targeted for insertion, as well as 4-nt sequences completing SwaI restriction sites in addition to their priming sites for the Kn cassette. For construction of the pIX-modification designated intermediate, primers C5-pIXKan_for and C5-pIXKan_rev, for barcoding library designated intermediate primers fC5-Kan-L_for and fC5-Kan-L_rev were used. Recombination-competent bacteria carrying the target genomes were transformed by the linear fragments. Resulting recombinants were selected for both Kn and chloramphenicol resistance and analyzed using RFLP. Correct insertion was furthermore confirmed using Sanger sequencing. Verified molecular clones for the pIX tagging were coined pBWH-C5-pIX-Kn. In generating pBWH-fC5-E1Kn for the library construction, we also deleted the E1 region, which resulted in a new bacmid to construct first-generation rAd vectors in the next step.

To provide high-quality linear inserts for efficient assembly into genomic bacmids, we generated an optional transfer vector for use with barcoding libraries, which could release its transgenic insert by restriction endonuclease digest. For this, synthetic dsDNA inserts containing SapI site flanked genomic sequences between the targeted SwaI half sites, including adjacent homologies, which need to be retained in the first generation vectors, a GFP expression cassette as well as a MCS for the insertion of the randomized sequences have been inserted into pO6-A5-GFP[71] resulting in pO5-fC5-GFP. SapI sites allowed the release of the exact fragment that could be used directly for the genomic assembly. Correct construction was confirmed using RFLP and Sanger sequencing.

Construction of pO6-fC5-GFP-MCS-based diversified barcoded shuttle plasmid libraries

To generate a controlled source of barcoded linear fragments we constructed a barcoded pO6-fC5-GFP-MCS shuttle plasmid library by Gibson assembly.[18] Linearized vector fragments were amplified by PCR using primers PCR-LibBB_for and PCR-LibBB_rev. The inserted barcoding sequences were annealed using randomized overlapping single-stranded oligonucleotides by adding 3 μL of each oligonucleotide (100 μM) to 500 μL 1× 2.1 NEB-buffer (New England Biolabs, Ipswich, MA, USA, #B6002S) followed by boiling at 100°C for 3 min to finally let them associate during the cooling to room temperature. We used 2–4 diversified oligonucleotides (Table S3) for each library. The general design of oligonucleotides consisted of 30-nt homologous regions on both ends facing either adjacent

oligonucleotides or vector backbone, as well as a 6-nt central diversity cassette (ATNNNNTA). Finally, 100 ng purified linear vector fragments were assembled with 5 μL annealed oligonucleotide pool using NEBuilder HiFi DNA Assembly (New England Biolabs, #E2621L) according to the manufacturer instructions. The assembly mixes were then dialyzed against deionized water, and competent *E. coli* Pir1 cells were electrotransformed by 2 μL of the dialyzed assembly mixes. The transformed bacteria were struck out on Kn-containing LB-agar plates and incubated overnight. The entire pool of these bacteria was harvested by adding liquid LB media onto the plate, tapping it lightly and transferring the media to inoculate 100-mL cultures for large-scale plasmid DNA preparation of the recombinant transfer plasmid libraries.

The DNA extracted and purified from these pools was then digested using HincII (New England Biolabs, #R0103S) and SapI (New England Biolabs, #R0569S) overnight to release transgenic insert and remove vector background. After heat inactivation of the restriction endonucleases (20 min at 65°C) these DNA preparations were directly used in the bacmid assemblies as barcoding inserts.

Construction of genomic bacmid assemblies

The targeted intermediate bacmids pBWH-C5-pIX-Kan or pBWH-fC5-E1Kn were digested by SwaI (New England Biolabs, #R0604S) overnight and heat-inactivated at 65°C, after which these preparations were used as vectors in the genomic assembly reactions. The genomic assemblies were set up by combining approximately 400 ng bacmid vector preparation and either 15 ng of each PCR-derived insert fragment (for pIX modifications) or 100 ng barcoding insert (for the library assemblies) at a final volume of 20 μL in the presence of 1× NEBuilder HiFi DNA Assembly Master Mix on ice. The reaction was incubated for 60 min at 50°C and dialyzed against deionized water. Electrocompetent *E. coli* NEB10beta (New England Biolabs, #C3020K) were transformed by 4 μL dialyzed bacmid assemblies according to the manufacturer instructions and plated on chloramphenicol (25 μg/mL)-containing LB-agar plates and incubated overnight. For the pIX-modification experiments, single clones were isolated, analyzed by RLFP and finally verified by Sanger sequencing. For the barcoding libraries, the entire pool of plated bacteria was harvested by adding liquid LB media onto the plates, tapping it lightly and transferring the media to inoculate 200-mL cultures for large-scale low-copy plasmid DNA preparation by column purification using NucleoBond Xtra Midi kit (Macherey Nagel, Düren, Germany, #740410.50), analyzed by NGS and used directly for rAd rescue by CTR.

Rescue of viral libraries and extraction of viral genomes

For rescuing rAds from the assembled bacmids, Cas9-mediated terminal resolution (CTR) was performed as previously described.[13] Briefly, 10 cm^2 cell culture surface of subconfluent 293A cells were co-transfected by 1 μg purified genomic bacmid DNA (gained either from single clones or bacmid library preparations) and 0.5 μg pARgRNA-Int5 Cas9-expressing helper plasmid via Lipofectamine 2000 (Thermo Fisher Scientific, #11668019). After 24 h, the transfected cells were split either 1:16 into 96-well plates for isolation of single rescues or completely transferred to a 10-cm dish (for library rescues). Cells were monitored using fluorescent microscopy. For rAd library applications, the cultured cells were harvested 4 days after transfection. The cells were collected by centrifugation and re-suspended in PBS and frozen using liquid nitrogen. After two freeze-thaw cycles, samples for titration of infectious viruses were saved. The bulk of the lysates was used for viral DNA preparation. Benzonase (Sigma-Aldrich, St. Louis, MO, USA, #E1014-25KU) was added to the lysates to remove both genomic DNA and the remaining transfected bacmid. Benzonase was inactivated using 50 mM EDTA and DNA was then extracted using Monarch Genomic DNA purification kit (New England Biolabs, #T3010S).

The single rescues were harvested from single wells containing a single focus of reconstituted virus after full lysis of the identified wells occurred (the first foci appeared 2–3 days after transfections and lysis was observed regularly at 7–8 days after transfection). These primary lysates were expanded to high-titer preparations.

Virological methods

The viral titers were determined by TCID-50 assay using an AAV-based replicon reporter.[72] Shortly, serial dilutions of viruses were used to infect cells treated with an AAV vector expressing Gaussia luciferase. At 4 days after infection, supernatants were collected and expression was tested via luminescence. Luminescence of more than 100-fold than uninfected control cells was used as a threshold to determine infection, allowing for the calculation of infectious titers as described by Lock (2019).[73]

The thermostability of the purified viruses was determined as follows: Viral lysates were incubated at 45°C for up to 10 min as well as an aliquot at 55°C for 10 min as described by Vellinga et al.[74] Afterward, the infectivity of the treated lysates was compared to untreated controls using the TCID-50 assay.

Multistep viral growth analysis was performed using 293A cells which were infected at MOI 0.1 of the respective rAd and incubated for 90 min. Zero-point samples were collected using both culture supernatant and infected cells, remaining samples for later time points were washed with PBS and fresh media was added. Thereafter, supernatants and cells were both collected at time points of up to 72 h after infection, the samples were lysed by three cycles of freezing and thawing and total titers were determined by titration.

NGS

For NGS, NEBNext ULTRA II FS DNA Library Prep (New England Biolabs, #E6177) was used. For diversity analysis on viral bacmid or plasmid DNA, a PCR was performed using primers M13A_for/M13A_rev targeting only N-diversified regions to maximize reads. Indexed pair-end libraries were prepared for Illumina sequencing and normalized and pooled sequencing libraries were denatured with 0.2 M NaOH. Sequencing was performed on an Illumina MiSeq

instrument using the 300-cycle MiSeq Reagent Kit v2 (Illumina, San Diego, CA, USA, MS102-2002).

Data analysis and statistics

NGS data was initially pre-processed and analyzed using a previously published[75] Galaxy[76] workflow. Sequencing adapters were trimmed and low-quality basecalls (phred Q < 30) were discarded. Reads were then mapped to a reference of the diversified genomic regions and resulting sequences were extracted. Using an in-house R script, single 4-nt diverse regions were extracted and matched to each other, resulting in barcodes 8–16 nt in length, depending on library assembly. Reads not providing a full barcode were discarded. The remaining barcodes were then analyzed further. This was performed a total of three times for each pool. Using lists of unique barcodes from each run for a single sample, the total diversity of the library pool could be estimated using a heterogeneous capture-recapture model established by Chao (1987).[25] Variant call frequencies of full genome sequencing were cut off at a minimum coverage of 25 reads and showing variants with a frequency of greater than 0.1.

Software

In silico sequence analysis and planning were performed using Geneious Prime.[77] NGS data were processed using custom Galaxy[76] and R[78] workflows. Capture/Recapture models were performed using the R package Rcapture.[79] Sequence logos were generated using the R package ggseqlogo.[80] Variant frequency plots were generated using VirHEAT.[81] Mapping data was generated using QualiMap 2.[82] Graphical Abstract was generated using BioRender.com.

DATA AND CODE AVAILABILITY

The data underlying this article are available in the article and its online supplemental information. R code used for analysis is available online at https://doi.org/10.5281/zenodo.10040806. Raw data underlying these analyses are available online at https://doi.org/10.5281/zenodo.10090475.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2024.101241.

ACKNOWLEDGMENTS

We thank Harald Wodrich (Microbiologie Fondamentale et Pathogénicité, Bordeaux) for providing us with HH-Ad5-VI-wt and Albert Heim (Medizinische Hochschule Hannover, Institute of Biology, Hannover) for the wt HAdV-C5 and Hans-Gerhard Burgert for critical reading of the manuscript and his great mentorship along the way. We also want to acknowledge Simone Gruber and Dominique Gütle for providing excellent technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – TRR 359 – Project number 491676693 and the German Ministry of Education and Research (BMBF) – Project identifier ADENOTHECA 16LW0230. The Galaxy server that was used for some calculations is in part funded by Collaborative Research Center 992 Medical Epigenetics (DFG grant SFB 992/1 2012) and German Federal Ministry of Education and Research (BMBF grants 031 A538A/A538C RBC, 031L0101B/031L0101C de.NBI-epi, 031L0106 de.STAIR (de.NBI)). Funding for open access charge: Deutsche Forschungsgemeinschaft TRR 359.

AUTHOR CONTRIBUTIONS

J.Fischer and Z.R. conceptualized the project and all experiments and wrote the paper. J. Fischer, A.F., and L.J. performed the experiments and analyzed the data. L.J. and J.Fuchs contributed analysis tools. J.Fischer, E.S., A.E., and Z.R. validated the presented information. All authors assisted by critical review and revisions of the manuscript.

DECLARATION OF INTERESTS

J.Fischer and Z.R. are co-inventors in patent application EP23199021.9 by the Albert-Ludwigs-University of Freiburg, which describes a two-step workflow to seamlessly modify circular BAC-/plasmids at high efficiency. J.Fischer and Z.R. are co-inventors in patent application PCT/EP2021/076757 by the Albert-Ludwigs-University of Freiburg, which describes a novel way of generating rAds by utilizing CRISPR-Cas9 linearization.

REFERENCES

1. Zhao, Z., Anselmo, A.C., and Mitragotri, S. (2022). Viral vector-based gene therapies in the clinic. Bioeng. Transl. Med. *7*, e10258. https://doi.org/10.1002/BTM2.10258.

2. U.S. Food & Drug Administration (2022). Approval Letter - ADSTILADRIN. https://www.fda.gov/media/164030/download?attachment.

3. Lee, C.S., Bishop, E.S., Zhang, R., Yu, X., Farina, E.M., Yan, S., Zhao, C., Zheng, Z., Shu, Y., Wu, X., et al. (2017). Adenovirus-mediated gene delivery: Potential applications for gene and cell-based therapies in the new era of personalized medicine. Genes Dis. *4*, 43–63. https://doi.org/10.1016/j.gendis.2017.04.001.

4. Wold, W.S.M., and Toth, K. (2013). Adenovirus Vectors for Gene Therapy, Vaccination and Cancer Gene Therapy. Curr. Gene Ther. *13*, 421–433. https://doi.org/10.2174/15665232136661311250950 46.

5. Jt, S., M, H., Wam, B., Ac, B., and Sa, N. (2022). Adenoviral vectors for cardiovascular gene therapy applications: a clinical and industry perspective. J. Mol. Med. *100*, 875–901. https://doi.org/10.1007/s00109-022-02208-0.

6. Weinmann, J., Weis, S., Sippel, J., Tulalamba, W., Remes, A., El Andari, J., Herrmann, A.-K., Pham, Q.H., Borowski, C., Hille, S., et al. (2020). Identification of a myotropic AAV by massively parallel *in vivo* evaluation of barcoded capsid variants. Nat. Commun. *11*, 5432. https://doi.org/10.1038/s41467-020-19230-w.

7. Wang, D., Tai, P.W.L., and Gao, G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. Nat. Rev. Drug Discov. *18*, 358–378. https://doi.org/10.1038/s41573-019-0012-9.

8. Maiakovska, O., Baumgartl, C., and Grimm, D. (2022). Determination of AAV properties by single amino acids: Go(o)d is in the details. Mol. Ther. Methods Clin. Dev. *27*, 93–95. https://doi.org/10.1016/J.OMTM.2022.09.006.

9. El Andari, J., Renaud-Gabardos, E., Tulalamba, W., Weinmann, J., Mangin, L., Pham, Q.H., Hille, S., Bennett, A., Attebi, E., Bourges, E., et al. (2022). Semirational bioengineering of AAV vectors with increased potency and specificity for systemic gene therapy of muscle disorders. Sci. Adv. *8*, eabn4704. https://doi.org/10.1126/SCIADV.ABN4704.

10. Becker, J., Fakhiri, J., and Grimm, D. (2022). Fantastic AAV Gene Therapy Vectors and How to Find Them—Random Diversification, Rational Design and Machine Learning. Pathogens *11*, 756. https://doi.org/10.3390/pathogens11070756.

11. Syyam, A., Nawaz, A., Ijaz, A., Sajjad, U., Fazil, A., Irfan, S., Muzaffar, A., Shahid, M., Idrees, M., Malik, K., and Afzal, S. (2022). Adenovirus vector system: construction, history and therapeutic applications. Biotechniques *73*, 297–305. https://doi.org/10.2144/btn-2022-0051.

12. Sallard, E., Zhang, W., Aydin, M., Schröer, K., and Ehrhardt, A. (2023). The Adenovirus Vector Platform: Novel Insights into Rational Vector Design and Lessons Learned from the COVID-19 Vaccine. Viruses *15*, 204. https://doi.org/10.3390/v15010204.

13. Riedl, A., Fischer, J., Burgert, H.-G., and Ruzsics, Z. (2022). Rescue of Recombinant Adenoviruses by CRISPR/Cas-Mediated *in vivo* Terminal Resolution. Front. Microbiol. *13*, 854690. https://doi.org/10.3389/fmicb.2022.854690.

14. Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. USA *97*, 6640–6645. https://doi.org/10.1073/pnas.120163297.

15. Luo, J., Deng, Z.L., Luo, X., Tang, N., Song, W.X., Chen, J., Sharff, K.A., Luu, H.H., Haydon, R.C., Kinzler, K.W., et al. (2007). A protocol for rapid generation of recombinant adenoviruses using the AdEasy system. Nat. Protoc. *2*, 1236–1247. https://doi.org/10.1038/NPROT.2007.135.

16. Miciak, J.J., Hirshberg, J., and Bunz, F. (2018). Seamless assembly of recombinant adenoviral genomes from high-copy plasmids. PLoS One *13*, e0199563. https://doi.org/10.1371/JOURNAL.PONE.0199563.

17. Ni, N., Deng, F., He, F., Wang, H., Shi, D., Liao, J., Zou, Y., Wang, H., Zhao, P., Hu, X., et al. (2021). A one-step construction of adenovirus (OSCA) system using the Gibson DNA Assembly technology. Mol. Ther. Oncolytics *23*, 602–611. https://doi.org/10.1016/J.OMTO.2021.11.011.

18. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345. https://doi.org/10.1038/NMETH.1318.

19. Liu, Z., Chen, O., Wall, J.B.J., Zheng, M., Zhou, Y., Wang, L., Vaseghi, H.R., Qian, L., and Liu, J. (2017). Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Sci. Rep. *7*, 2193–2199. https://doi.org/10.1038/s41598-017-02460-2.

20. Matz, M.V., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zaraisky, A.G., Markelov, M.L., and Lukyanov, S.A. (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. Nat. Biotechnol. *17*, 969–973. https://doi.org/10.1038/13657.

21. Pfitzner, S., Hofmann-Sieber, H., Bosse, J.B., Franken, L.E., Grünewald, K., and Dobner, T. (2020). Fluorescent protein tagging of adenoviral proteins pV and pIX reveals 'late virion accumulation compartment. PLoS Pathog. *16*, e1008588. https://doi.org/10.1371/journal.ppat.1008588.

22. Galka, P., Jamez, E., Joachim, G., and Soumillion, P. (2017). QuickLib, a method for building fully synthetic plasmid libraries by seamless cloning of degenerate oligonucleotides. PLoS One *12*, e0175146. https://doi.org/10.1371/journal.pone.0175146.

23. Festa, F., Steel, J., Bian, X., and Labaer, J. (2013). High-throughput cloning and expression library creation for functional proteomics. Proteomics *13*, 1381–1399. https://doi.org/10.1002/pmic.201200456.

24. Davidsson, M., Diaz-Fernandez, P., Schwich, O.D., Torroba, M., Wang, G., and Björklund, T. (2016). A novel process of viral vector barcoding and library preparation enables high-diversity library generation and recombination-free paired-end sequencing. Sci. Rep. *6*, 37563. https://doi.org/10.1038/srep37563.

25. Chao, A. (1987). Estimating the Population Size for Capture-Recapture Data with Unequal Catchability. Biometrics *43*, 783–791. https://doi.org/10.2307/2531532.

26. Hommelsheim, C.M., Frantzeskakis, L., Huang, M., and Ülker, B. (2014). PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. Sci. Rep. *4*, 5052. https://doi.org/10.1038/srep05052.

27. Treangen, T.J., and Salzberg, S.L. (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat. Rev. Genet. *13*, 36–46. https://doi.org/10.1038/nrg3117.

28. Tørresen, O.K., Star, B., Mier, P., Andrade-Navarro, M.A., Bateman, A., Jarnot, P., Gruca, A., Grynberg, M., Kajava, A.V., Promponas, V.J., et al. (2019). Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. Nucleic Acids Res. *47*, 10994–11006. https://doi.org/10.1093/nar/gkz841.

29. Ruzsics, Z., Lemnitzer, F., and Thirion, C. (2014). Engineering adenovirus genome by bacterial artificial chromosome (BAC) technology. Methods Mol. Biol. *1089*, 143–158. https://doi.org/10.1007/978-1-62703-679-5_11.

30. Paredes, A.M., and Yu, D. (2012). Human Cytomegalovirus: Bacterial Artificial Chromosome (BAC) Cloning and Genetic Manipulation. Curr. Protoc. Microbiol. *14*, Unit14E.4. https://doi.org/10.1002/9780471729259.mc14e04s24.

31. Krueger, B., Plaisance, K., Sangani, R., Lanier, C., Jain, V., Hu, J., and Renne, R. (2012). A core laboratory for the generation of quality-controlled g-herpesvirus bacmids: generation of KSHV microRNA mutants. Infect. Agent. Cancer *7*, P27. https://doi.org/10.1186/1750-9378-7-S1-P27.

32. Hahn, G., Jarosch, M., Wang, J.B., Berbes, C., and McVoy, M.A. (2003). Tn7-mediated introduction of DNA sequences into bacmid-cloned cytomegalovirus genomes for rapid recombinant virus construction. J. Virol. Methods *107*, 185–194. https://doi.org/10.1016/S0166-0934(02)00232-X.

33. Hobom, U., Brune, W., Messerle, M., Hahn, G., and Koszinowski, U.H. (2000). Fast Screening Procedures for Random Transposon Libraries of Cloned Herpesvirus Genomes: Mutational Analysis of Human Cytomegalovirus Envelope Glycoprotein Genes. J. Virol. *74*, 7720–7729. https://doi.org/10.1128/JVI.74.17.7720-7729.2000.

34. Chiu, Y.F., Tung, C.P., Lee, Y.H., Wang, W.H., Li, C., Hung, J.Y., Wang, C.Y., Kawaguchi, Y., and Liu, S.T. (2007). A comprehensive library of mutations of Epstein-Barr virus. J. Gen. Virol. *88*, 2463–2472. https://doi.org/10.1099/VIR.0.82881-0.

35. Osoegawa, K., and de Jong, P.J. (2004). BAC library construction. Methods Mol. Biol. *255*, 1–46. https://doi.org/10.1385/1-59259-752-1:001.

36. Kanda, T., Yajima, M., Ahsan, N., Tanaka, M., and Takada, K. (2004). Production of High-Titer Epstein-Barr Virus Recombinants Derived from Akata Cells by Using a Bacterial Artificial Chromosome System. J. Virol. *78*, 7004–7015. https://doi.org/10.1128/JVI.78.13.7004-7015.2004.

37. Guo, X., Sun, Y., Chen, J., Zou, X., Hou, W., Tan, W., Hung, T., and Lu, Z. (2022). Restriction-Assembly: A Solution to Construct Novel Adenovirus Vector. Viruses *14*, 546. https://doi.org/10.3390/v14030546.

38. Hamdan, F., Martins, B., Feodoroff, M., Giannoula, Y., Feola, S., Fusciello, M., Chiaro, J., Antignani, G., Grönholm, M., Ylösmäki, E., and Cerullo, V. (2021). GAMER-Ad: a novel and rapid method for generating recombinant adenoviruses. Mol. Ther. Methods Clin. Dev. *20*, 625–634. https://doi.org/10.1016/j.omtm.2021.01.014.

39. Chancham, P., and Hughes, J.A. (2001). Relationship between plasmid DNA topological forms and *in vitro* transfection. J. Liposome Res. *11*, 139–152. https://doi.org/10.1081/LPR-100108458.

40. Graham, F.L. (1984). Covalently closed circles of human adenovirus DNA are infectious. EMBO J. *3*, 2917–2922. https://doi.org/10.1002/j.1460-2075.1984.tb02232.x.

41. Gao, G., Zhou, X., Alvira, M.R., Tran, P., Marsh, J., Lynd, K., Xiao, W., and Wilson, J.M. (2003). High throughput creation of recombinant adenovirus vectors by direct cloning, green-white selection and I-Sce I-mediated rescue of circular adenovirus plasmids in 293 cells. Gene Ther. *10*, 1926–1930. https://doi.org/10.1038/sj.gt.3302088.

42. Elahi, S.M., Oualikene, W., Naghdi, L., O'Connor-McCourt, M., and Massie, B. (2002). Adenovirus-based libraries: efficient generation of recombinant adenoviruses by positive selection with the adenovirus protease. Gene Ther. *9*, 1238–1246. https://doi.org/10.1038/sj.gt.3301793.

43. Hillgenberg, M., Hofmann, C., Stadler, H., and Löser, P. (2006). High-Efficiency System for the Construction of Adenovirus Vectors and Its Application to the Generation of Representative Adenovirus-Based cDNA Expression Libraries. J. Virol. *80*, 5435–5450. https://doi.org/10.1128/JVI.00218-06.

44. Lupold, S.E., Kudrolli, T.A., Chowdhury, W.H., Wu, P., and Rodriguez, R. (2007). A novel method for generating and screening peptides and libraries displayed on adenovirus fiber. Nucleic Acids Res. *35*, e138. https://doi.org/10.1093/nar/gkm914.

45. New England BioLabs (2022). NEBuilder® HiFi DNA Assembly – Benefits over GeneArt Gibson Assembly® and In-Fusion® Snap Assembly. https://www.neb.com/en/applications/cloning-and-synthetic-biology/dna-assembly-and-cloning/nebuilder-hifi-dna-assembly/nebuilder-hifi-dna-assembly—benefits-over-neb-gibson-assembly.

46. Adachi, K., Enoki, T., Kawano, Y., Veraz, M., and Nakai, H. (2014). Drawing a high-resolution functional map of adeno-associated virus capsid by massively parallel sequencing. Nat. Commun. *5*, 3075. https://doi.org/10.1038/ncomms4075.

47. Dong, W., and Kantor, B. (2021). Lentiviral Vectors for Delivery of Gene-Editing Systems Based on CRISPR/Cas: Current State and Perspectives. Viruses *13*, 1288. https://doi.org/10.3390/v13071288.

48. Zhang, Y. (2023). Evolution of phage display libraries for therapeutic antibody discovery. mAbs *15*, 2213793. https://doi.org/10.1080/19420862.2023.2213793.

49. Liu, H., Liu, X.-M., Li, S.-C., Wu, B.-C., Ye, L.-L., Wang, Q.-W., and Chen, Z.-L. (2009). A high-yield and scaleable adenovirus vector production process based on high density perfusion culture of HEK 293 cells as suspended aggregates. J. Biosci. Bioeng. *107*, 524–529. https://doi.org/10.1016/j.jbiosc.2009.01.004.

50. Moreira, A.S., Silva, A.C., Sousa, M.F.Q., Hagner-McWhirterc, Å., Ahlénc, G., Lundgren, M., Coroadinha, A.S., Alves, P.M., Peixoto, C., Carrondo, M.J.T., et al. (2020). Establishing Suspension Cell Cultures for Improved Manufacturing of Oncolytic Adenovirus. Biotechnol. J. *15*, e1900411. https://doi.org/10.1002/biot.201900411.

51. Lee, D.-K., Park, J., and Seo, D.-W. (2020). Suspension culture of Vero cells for the production of adenovirus type 5. Clin. Exp. Vaccine Res. *9*, 48–55. https://doi.org/10.7774/cevr.2020.9.1.48.

52. Naumer, M., Ying, Y., Michelfelder, S., Reuter, A., Trepel, M., Müller, O.J., and Kleinschmidt, J.A. (2012). Development and Validation of Novel AAV2 Random Libraries Displaying Peptides of Diverse Lengths and at Diverse Capsid Positions. Hum. Gene Ther. *23*, 492–507. https://doi.org/10.1089/hum.2011.139.

53. Hong, S.S., and Boulanger, P. (1995). Protein ligands of the human adenovirus type 2 outer capsid identified by biopanning of a phage-displayed peptide library on separate domains of wild-type and mutant penton capsomers. EMBO J. *14*, 4714–4727.

54. Pereboev, A., Pereboeva, L., and Curiel, D.T. (2001). Phage Display of Adenovirus Type 5 Fiber Knob as a Tool for Specific Ligand Selection and Validation. J. Virol. *75*, 7107–7113. https://doi.org/10.1128/JVI.75.15.7107-7113.2001.

55. Uusi-Kerttula, H., Davies, J.A., Thompson, J.M., Wongthida, P., Evgin, L., Shim, K.G., Bradshaw, A., Baker, A.T., Rizkallah, P.J., Jones, R., et al. (2018). Ad5NULL-A20: A Tropism-Modified, αvβ6 Integrin-Selective Oncolytic Adenovirus for Epithelial Ovarian Cancer Therapies. Clin. Cancer Res. *24*, 4215–4224. https://doi.org/10.1158/1078-0432.CCR-18-1089.

56. Shah, A.C., Parker, J.N., Shimamura, M., and Cassady, K.A. (2009). Spontaneous and Engineered Compensatory HSV Mutants that Counteract the Host Antiviral PKR Response. Viruses *1*, 510–522. https://doi.org/10.3390/v1030510.

57. Kuhn, I., Harden, P., Bauzon, M., Chartier, C., Nye, J., Thorne, S., Reid, T., Ni, S., Lieber, A., Fisher, K., et al. (2008). Directed Evolution Generates a Novel Oncolytic Virus for the Treatment of Colon Cancer. PLoS One *3*, e2409. https://doi.org/10.1371/journal.pone.0002409.

58. Uil, T.G., Vellinga, J., de Vrij, J., van den Hengel, S.K., Rabelink, M.J.W.E., Cramer, S.J., Eekels, J.J.M., Ariyurek, Y., van Galen, M., and Hoeben, R.C. (2011). Directed adenovirus evolution using engineered mutator viral polymerases. Nucleic Acids Res. *39*, e30. https://doi.org/10.1093/nar/gkq1258.

59. Myers, N.D., Skorohodova, K.V., Gounder, A.P., and Smith, J.G. (2013). Directed Evolution of Mutator Adenoviruses Resistant to Antibody Neutralization. J. Virol. *87*, 6047–6050. https://doi.org/10.1128/JVI.00473-13.

60. Alfaleh, M.A., Alsaab, H.O., Mahmoud, A.B., Alkayyal, A.A., Jones, M.L., Mahler, S.M., and Hashem, A.M. (2020). Phage Display Derived Monoclonal Antibodies: From Bench to Bedside. Front. Immunol. *11*, 1986. https://doi.org/10.3389/FIMMU.2020.01986.

61. Hammers, C.M., and Stanley, J.R. (2014). Antibody Phage Display: Technique and Applications. J. Invest. Dermatol. *134*, 1–5. https://doi.org/10.1038/JID.2013.521.

62. Bashir, S., and Paeshuyse, J. (2020). Construction of Antibody Phage Libraries and Their Application in Veterinary Immunovirology. Antibodies *9*, 21. https://doi.org/10.3390/antib9020021.

63. Schmidt, F.I., Hanke, L., Morin, B., Brewer, R., Brusic, V., Whelan, S.P.J., and Ploegh, H.L. (2016). Phenotypic Lentivirus Screens to Identify Functional Single Domain Antibodies. Nat Microbiol. *1*, 16080. https://doi.org/10.1038/nmicrobiol.2016.80.

64. Sharon, D., and Kamen, A. (2018). Advancements in the design and scalable production of viral gene transfer vectors. Biotechnol. Bioeng. *115*, 25–40. https://doi.org/10.1002/bit.26461.

65. Bulcha, J.T., Wang, Y., Ma, H., Tai, P.W.L., and Gao, G. (2021). Viral vector platforms within the gene therapy landscape. Signal Transduct. Target. Ther. *6*, 53. https://doi.org/10.1038/s41392-021-00487-6.

66. Berman, C.M., Papa, L.J., Hendel, S.J., Moore, C.L., Suen, P.H., Weickhardt, A.F., Doan, N.-D., Kumar, C.M., Uil, T.G., Butty, V.L., et al. (2018). An Adaptable Platform for Directed Evolution in Human Cells. J. Am. Chem. Soc. *140*, 18093–18103. https://doi.org/10.1021/jacs.8b10937.

67. Danthinne, X., and Imperiale, M.J. (2000). Production of first generation adenovirus vectors: a review. Gene Ther. *7*, 1707–1714. https://doi.org/10.1038/sj.gt.3301301.

68. Bett, A.J., Prevec, L., and Graham, F.L. (1993). Packaging capacity and stability of human adenovirus type 5 vectors. J. Virol. *67*, 5911–5921. https://doi.org/10.1128/jvi.67.10.5911-5921.1993.

69. Ruzsics, Z., Wagner, M., Osterlehner, A., Cook, J., Koszinowski, U., and Burgert, H.G. (2006). Transposon-Assisted Cloning and Traceless Mutagenesis of Adenoviruses: Development of a Novel Vector Based on Species D. J. Virol. *80*, 8100–8113. https://doi.org/10.1128/JVI.00687-06.

70. Schreiner, S., Martinez, R., Groitl, P., Rayne, F., Vaillant, R., Wimmer, P., Bossis, G., Sternsdorf, T., Marcinowski, L., Ruzsics, Z., et al. (2012). Transcriptional Activation of the Adenoviral Genome Is Mediated by Capsid Protein VI. PLoS Pathog. *8*, e1002549. https://doi.org/10.1371/journal.ppat.1002549.

71. Wodrich, H., Henaff, D., Jammart, B., Segura-Morales, C., Seelmeir, S., Coux, O., Ruzsics, Z., Wiethoff, C.M., and Kremer, E.J. (2010). A Capsid-Encoded PPxY-Motif Facilitates Adenovirus Entry. PLoS Pathog. *6*, e1000808. https://doi.org/10.1371/journal.ppat.1000808.

72. Ruzsics, Z., Hoffmann, K., Riedl, A., Krawczyk, A., Widera, M., Sertznig, H., Schipper, L., Kapper-Falcone, V., Debreczeny, M., Ernst, W., et al. (2020). A Novel, Broad-Acting Peptide Inhibitor of Double-Stranded DNA Virus Gene Expression and Replication. Front. Microbiol. *11*, 601555. https://doi.org/10.3389/fmicb.2020.601555.

73. Lock, M., Korn, M., Wilson, J., Sena-Esteves, M., and Gao, G. (2019). Measuring the Infectious Titer of Recombinant Adenovirus Using Tissue Culture Infection Dose 50% (TCID 50) End-Point Dilution and Quantitative Polymerase Chain Reaction (qPCR). Cold Spring Harb. Protoc. *2019*, pdb.prot095562. https://doi.org/10.1101/pdb.prot095562.

74. Vellinga, J., van den Wollenberg, D.J.M., van der Heijdt, S., Rabelink, M.J.W.E., and Hoeben, R.C. (2005). The Coiled-Coil Domain of the Adenovirus Type 5 Protein IX Is Dispensable for Capsid Incorporation and Thermostability. J. Virol. *79*, 3206–3210. https://doi.org/10.1128/JVI.79.5.3206-3210.2005.

75. Weigang, S., Fuchs, J., Zimmer, G., Schnepf, D., Kern, L., Beer, J., Luxenburger, H., Ankerhold, J., Falcone, V., Kemming, J., et al. (2021). Within-host evolution of SARS-CoV-2 in an immunosuppressed COVID-19 patient as a source of immune escape variants. Nat. Commun. *12*, 6405. https://doi.org/10.1038/s41467-021-26602-3.

76. Afgan, E., Nekrutenko, A., Grüning, B.A., Blankenberg, D., Goecks, J., Schatz, M.C., Ostrovsky, A.E., Mahmoud, A., Lonie, A.J., Syme, A., et al. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. *50*, W345–W351. https://doi.org/10.1093/NAR/GKAC247.

77. Biomatters Developement Team (2019). Geneious Prime at Biomatters Ltd. http://www.geneious.com/.

78. R Core Team (2022). R: A Language and Environment for Statistical Computing atR Foundation for Statistical Computing. https://www.r-project.org/.

79. Rivest, L.-P., and Baillargeon, S. (2022). Rcapture: Loglinear Models for Capture-Recapture. https://cran.r-project.org/web/packages/Rcapture/index.html.

80. Waigh, O. (2017). ggseqlogo: A "ggplot2" Extension for Drawing Publication-Ready Sequence Logos. https://ggplot2.tidyverse.org/.

81. Fuchs, J. (2021). VirHeat. https://doi.org/10.5281/zenodo.8271929.

82. Okonechnikov, K., Conesa, A., and García-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics *32*, 292–294. https://doi.org/10.1093/bioinformatics/btv566.