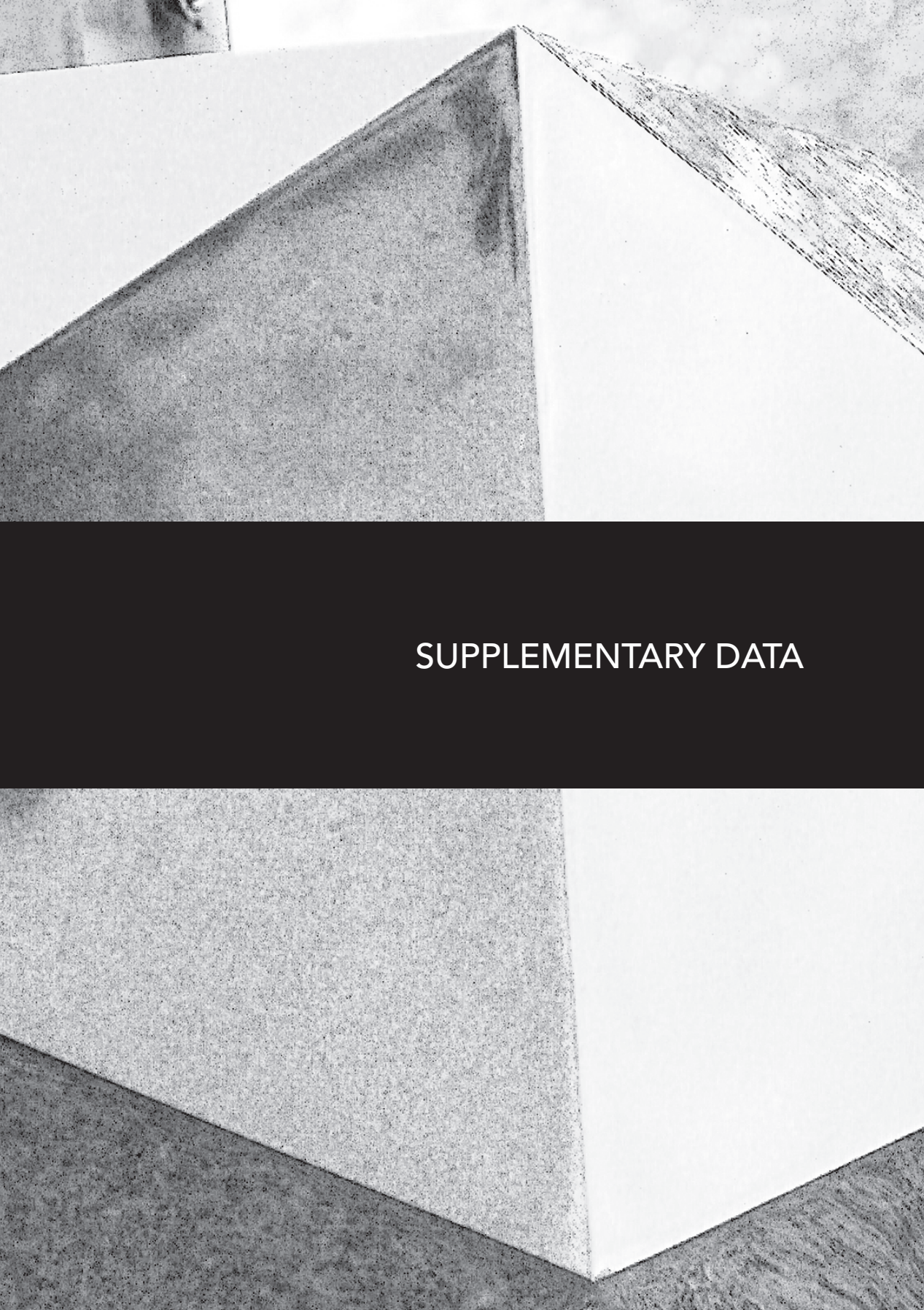
Rational and random approaches to adenoviral vector engineering

Uil, T.G.

5b



SUPPLEMENTARY DATA

CONTENTS

SUPPLEMENTARY FIGURES

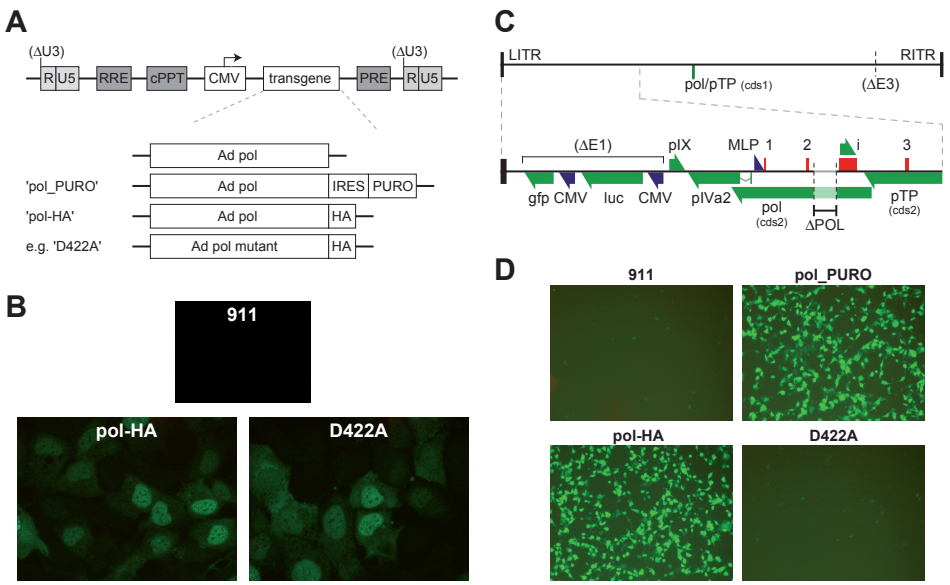
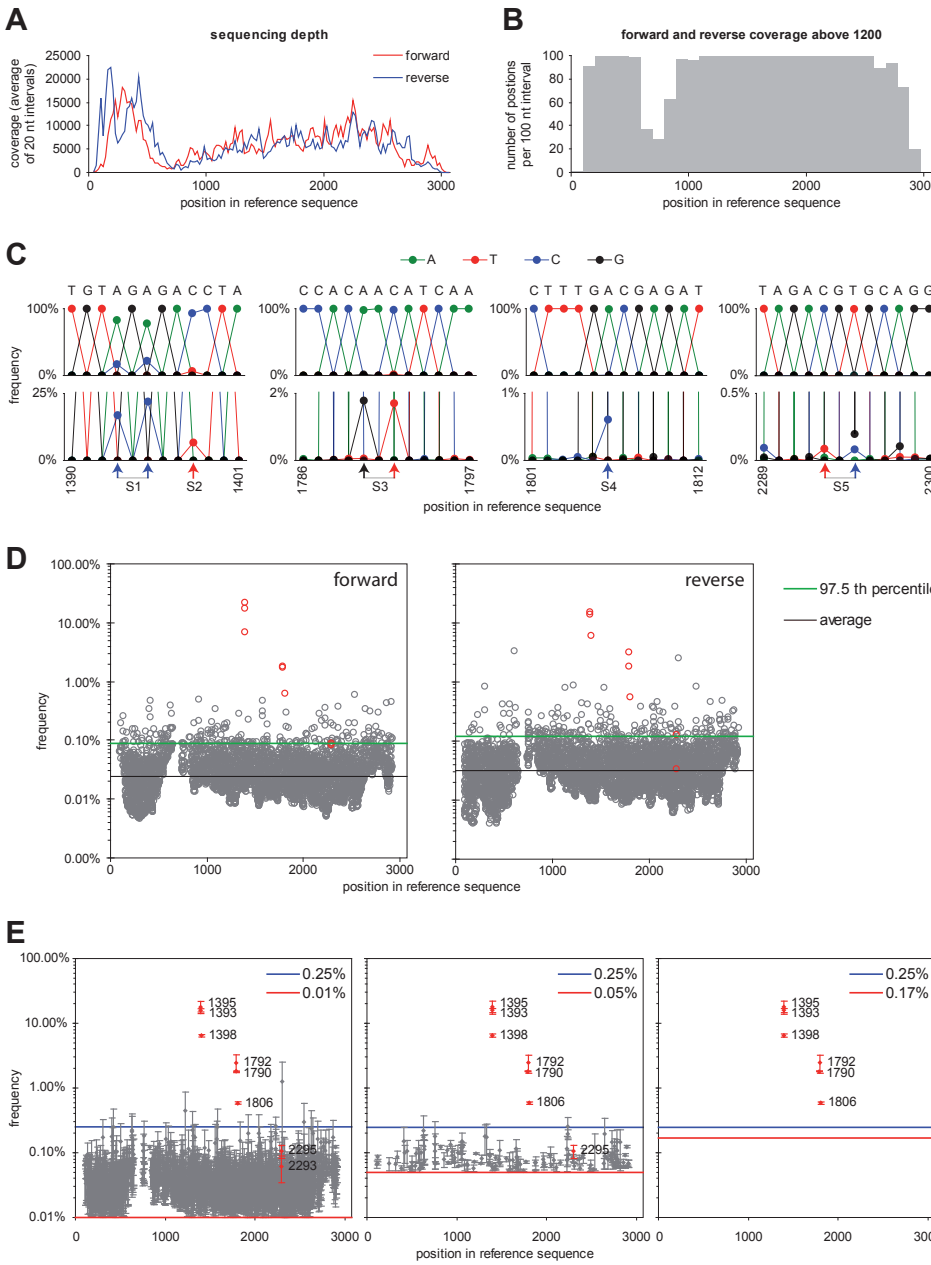


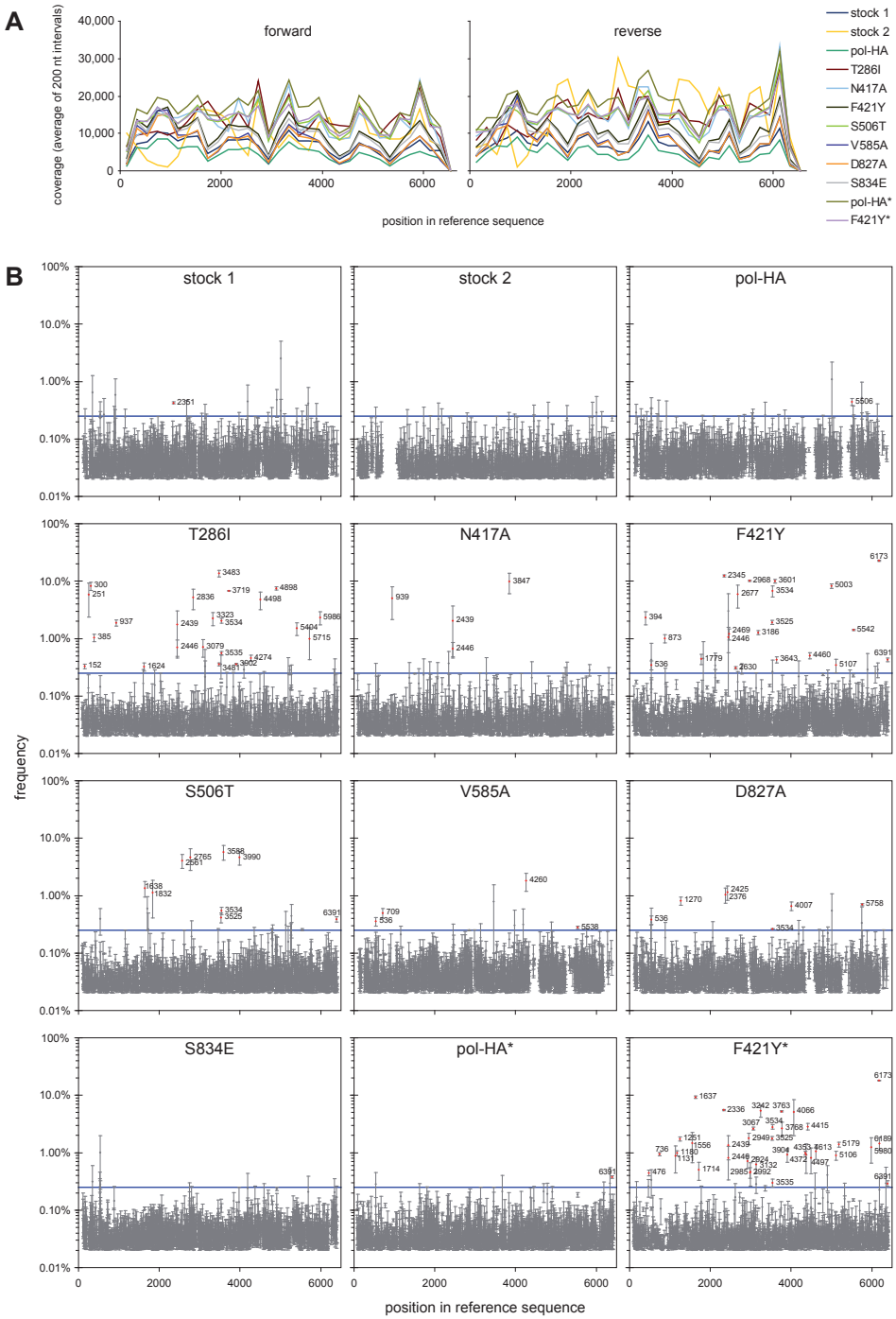
Figure S1. Components of the Ad pol *trans*-complementation system. (**A**) Lentiviral vectors for the stable expression of Ad pol variants in cells. Shown are representations of a lentiviral vector (in the proviral state) and the different transgenic configurations used. IRES, internal ribosomal entry site; PURO, puromycin resistance gene; HA, hemagglutinin (HA) tag; ΔU3, deletion of the U3 region; RRE, HIV-I Rev responsive element; cPPT, HIV-I central polypurine tract; CMV, human cytomegalovirus promoter; PRE, human hepatitis B virus post-transcriptional regulatory element. (**B**) Lentiviral expression – in the E1-complementing cell line 911 – of HA-tagged versions of wild-type Ad pol (pol-HA) and the D422A mutant Ad pol (D422A). Immunofluorescence analysis was performed using an HA-tag-specific primary antibody. (**C**) Representation of the genome of AdGLΔPOL, a polymerase-defective HAdV-5 based reporter vector. E1 and E3 regions are deleted, with the former replaced by GFP and luciferase (luc) gene expression cassettes. The polymerase-disrupting deletion 'ΔPOL' corresponds to positions 7312 to 7882 of the wild-type HAdV-5 genome (accession no. AC_000008). Green and blue arrows respectively represent coding and promoter sequences. Adenovirus genes depicted are polymerase (pol), pTP, pIX, pIVa2, and the i-leader protein (i). Red bars represent leader exons (1, 2, i, and 3) of the major late transcription unit. All components are depicted to scale. MLP, major late promoter. (**D**) Trans-complementation of the polymerase-defective reporter virus by both wild-type Ad pol ('pol_PURO') and pol-HA but not by D422A. 911 cells and 911 cells stably expressing the respective polymerases were infected with AdGLΔPOL. Images showing GFP fluorescence were made 2 days after infection.

5b


A
sequencing depth
forward
reverse
coverage (average of 20 nt intervals)
position in reference sequence
B
forward and reverse coverage above 1200
number of postions per 100 nt interval
position in reference sequence
C
A
T
C
G
frequency
position in reference sequence
S1
S2
S3
S4
S5
D
forward
reverse
97.5 th percentile
average
frequency
position in reference sequence
E
0.25%
0.01%
0.05%
0.17%
frequency
position in reference sequence

◀ **Figure S2. Deep sequencing to detect low-prevalence mutations: analysis of a spiked test sample.** Deep sequencing was performed on a 3.1-kb DNA fragment that was spiked with five minority sequences, each of which contained a single or double nucleotide polymorphism. The sequence reads were mapped against the reference sequence allowing up to two mismatches, after which low-quality bases were masked. (**A**) Forward and reverse sequencing depth distributions over the length of the sequenced fragment. (**B**) Positions in the reference sequence for which both the forward and reverse mappings passed a minimal sequencing depth requirement of 1200. (**C**) Local frequencies of base-calls around the spiked positions. Results represent those of the forwardly mapped reads. S1 to S5 specify the positions of the spiked polymorphisms. The color coding of the arrows correspond to that of the spiked minority bases. (**D**) Minor variant frequencies for the forward and reverse mappings. Depicted per position are the frequencies of all three possible minor variants (i.e. of the non-consensus base occurrences). The spiked minor variants are shown in red. The indicated 97.5th percentiles of the observed minor variant frequencies were used as error rate estimates for the calculation of *P*-values (see Statistics section in the main article). The data shown are for positions for which both the forward or reverse coverage level were above 1200. (**E**) Exemplification of mutation scoring using a minimal coverage requirement of 1200 and minor variant frequency requirements of 0.01%, 0.05%, or 0.17%. Shown – for positions with forward and reverse sequencing depths of minimally 1200 – are the minor variant occurrences for which both the forward and reverse observed frequencies were above a set cutoff value, indicated by the red line. For a given minor variant, the diamond and its associated 'error bars' respectively represent the average and the respective absolute values of the observed forward and reverse frequencies. The spiked minor variants are shown in red; their associated numbers indicate their respective positions in the reference sequence. The blue line (at 0.25%) represents a prevalence cutoff value considered safe for mutation scoring.

5b


A
forward
reverse
coverage (average of 200 nt intervals)
position in reference sequence
stock 1
stock 2
pol-HA
T286I
N417A
F421Y
S506T
V585A
D827A
S834E
pol-HA*
F421Y*
B
frequency
position in reference sequence

◀ **Figure S3. Deep sequencing to detect minor variants in pools of passaged viruses.** Viral DNA fragments obtained from pools of viruses replicated by different Ad pol variants were subjected to deep sequencing. Resultant sequence reads were mapped to the reference sequence allowing up to two edit operations (i.e. mismatches or gaps), after which low-quality base calls were masked. Using the forwardly and reversely mapped base-call distributions, substitutions were scored by imposing a local coverage depth requirement of 1200 and a minor variant frequency cutoff value of 0.25% (both of which requirements were to be met for both the forward and reverse mappings). (**A**) Forward and reverse sequencing depth distributions over the length of the sequenced fragment. (**B**) Minor variant frequency plots showing the scored substitutions. Depicted are all minor variant occurrences for which both the forward and reverse frequencies were above 0.02% (including only those positions for which the minimal sequencing depth requirement was met). Diamonds and their associated 'error bars' respectively represent the average and the respective absolute values of the observed forward and reverse frequencies. Minor variants were scored as bona fide substitutions (indicated in red) when their forward and reverse frequencies were both above the prevalence cutoff value of 0.25% (indicated by the blue line).

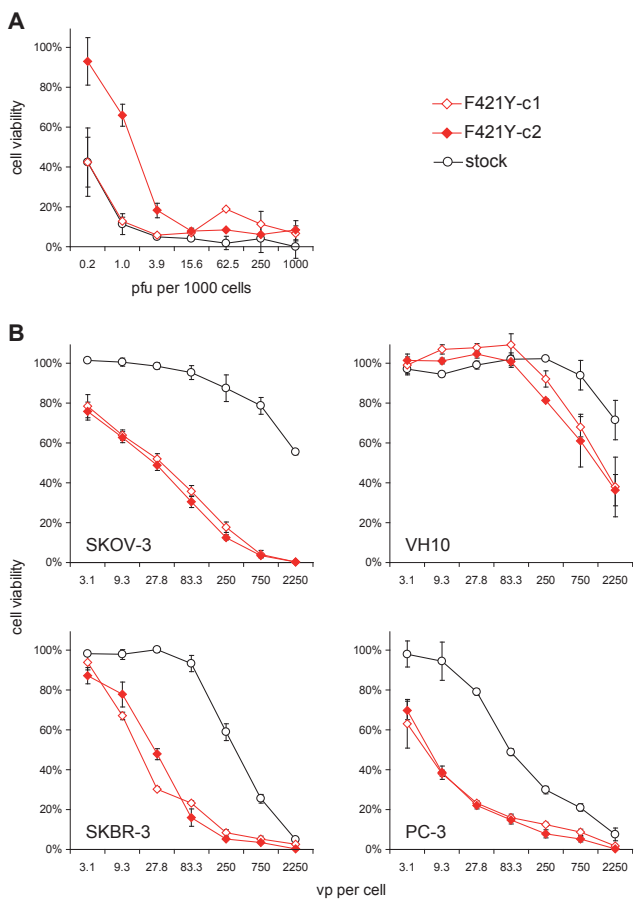


Figure S4. Cell killing abilities of the bioselected viral clones F421Y-c1 and F421Y-c2. Shown are cytotoxicity assays on 911 cells (**A**) and SKOV-3, SKBR-3, PC-3, and VH10 cells (**B**). Cells were infected at the indicated MOI's and cell viabilities were assessed by WST-1 assay.

5b

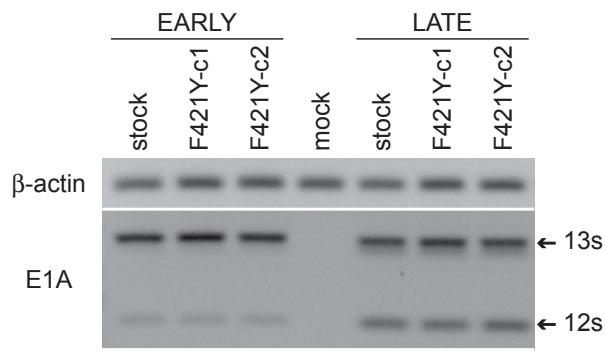


Figure S5. Reverse transcriptase-PCR analysis controls: cellular β-actin and viral E1A. These are controls reactions for the RT-PCRs performed to detect ADP-encoding mRNAs (main article, Figure 6). 12s and 13s are splice variants of E1A.

5b

SUPPLEMENTARY TABLES

Table S1. Ad pol mutants examined in this study

No.[a]	Ad pol mutant	Putative functions of the substituted residue[b]
1	D283A	Exo catalytic activity; metal ion binding; strand displacement (Φ29) (1)
2	E285A	Exo catalytic activity; metal ion binding; strand displacement (Φ29) (1)
3	T286I	Stabilization of the primer terminus at exo active site
4	N417A/D	Stabilization of the primer terminus at exo active site
5	F421A/S/Y	Stabilization of the primer terminus at exo active site
6	D422A	Exo catalytic activity; metal ion binding; strand displacement (Φ29) (1)
7	S506T	Involvement in exo activity; requirement for interaction with the terminal protein (11); primer binding (?)
8	L507T	Involvement in exo activity; primer binding (?)
9	Y580A	Involvement in exo activity; strand displacement (Φ29) (1)
10	D584A	Exo catalytic activity; metal ion binding; strand displacement (Φ29) (1)
11	V585A	Formation of part of the exo active site
12	G666A	dsDNA binding; tuning between pol and exo activities
13	M689K/N	Formation of part of the hydrophobic pocket that binds the base and ribose portions of the incoming nucleotide; role in base discrimination and/or partitioning to exo site
14	Y690F	Interaction with the ribose of the incoming dNTP; role in preventing incorporation of NTPs
15	D827A	Indirect involvement in incoming nucleotide binding (?)
16	S834E	Indirect involvement in incoming nucleotide binding (?)
17	L838A	Formation of part of the nascent base-pair binding pocket; involvement in positioning the templating nucleotide; controlling base-pairing correctness
18	Y844A	Hydrophobic interaction with the base of the incoming – or the templating (*Eco* I) – nucleotide; critical role in checking correctness of base pairing
19	G845A	dsDNA binding (Φ29) (39); role in the insertion of correct nucleotides (*Sce* α)

[a]The residue numbers 1 to 19 correspond to those assigned to the selected Ad pol residues in Figure 1.
[b]Listed are functions ascribed to homologous residues in polymerases of other organisms. Uncertain or debated functions are followed by a parenthesized question mark. (See Supplementary References for literature cited here).
[c]Depicted are those substitutions that were found – in the referenced studies – to either diminish exonuclease activity, affect the pol/exo balance, lower replication fidelity, or give a mutator phenotype *in vivo*. Substitutions of residues not evidently homologous to that of Ad pol are parenthesized. (See Supplementary References for literature cited here).
Abbreviations: exo, exonuclease; pol, polymerase; *Eco*, *Escherichia coli*; *Sce*, *Saccharomyces cerevisiae*; HSV, herpes simplex virus; CMV, cytomegalovirus ; *Taq*, *Thermus aquaticus* ; *Pfu*, *Pyrococcus furiosus*.

Mutator characteristics-yielding substitutions of homologous residues in other polymerases[c]

Φ29: D12A (2,3); T4: D112A/N (4); *Sce* δ: D321G/N (5); HSV: D368A (6)

Φ29: E14A (2,3); T4: E114A (4); *Sce* δ:E323Q (5); HSV: E370A (6)

Φ29: T15I (7); *Sce* δ: C324R (5)

Φ29: N62D (7); PRD-1: N71D (8)

Φ29: F65Y/S (9)

Φ29: D66A (2,3); T4: D219A/N (4); *Sce* δ: D407N (5); HSV: D471A (6); CMV: D413A (10)

Φ29: S122T/N (9)

Φ29: L123T/A (9); T4: L287A (4); *Sce* δ: L479S (5)

Φ29: Y165F/C (3,12); T4: Y320A/Q/F (4); RB69: Y323F (13); HSV: Y577F/H (6,14)

Φ29: D169A (3,12); T4: D324A/N (4); HSV: D581A (6,14)

T4: V325A (15)

Φ29: G228A (16)

T4: L412M (17); *Sce* δ: L612F/M/K/G/N (18); RB69: L415F/G (19); *Eco* I: I709A/F/M/N (20); *Sce* α: L868F/M/W/V (21); *Taq* I: I614K/N/Q (22); *Sce* ε: M644F (23)

Φ29: Y254F/V (24,25); *Eco* I: E710A (26)

(*Pfu*: T471G/A, Q472G/A/P, D473G/A (27))

(*Taq* I: A661E (28))

Φ29: L384R/Q (29); RB69: L561A (30,31); *Taq* I: T664R (28)

Φ29: Y390F/S (24); RB69 Y567A (31-34); *Eco* I: Y766S/A (35-37); *Sce* α: Y951E/P (38)

Sce α: G952Y/A (40)

Table S2. Substitution loads relative to the analyzed amount of viral DNA.

Sample	Pool size[a]	Positions analyzed[b]	Total amount analyzed (bp)[c]	Absolute substitution score[d]	Substitution load per million bp[e]
stock 1	38	5869	2.2×10^5	1	4.5
stock 2	54	5809	3.1×10^5	0	0.0
pol-HA	26	5189	1.3×10^5	0	0.0
T286I	48	6320	3.0×10^5	23	76.3
N417A	42	6296	2.6×10^5	4	15.3
F421Y	47	6162	2.9×10^5	21	72.7
S506T	36	6291	2.3×10^5	8	35.3
V585A	41	5619	2.3×10^5	2	8.7
D827A	33	5670	1.9×10^5	7	37.4
S834E	44	6109	2.7×10^5	0	0.0
pol-HA*	58	6341	3.6×10^5	1	2.7
F421Y*	55	6314	3.5×10^5	35	100.1

[a]Virus pool sizes were estimated based on intra-experiment titrations during pool preparations.
[b]Positions in the reference sequence for which both the forwardly and reversely mapped base distributions had a minimal coverage level of 1200.
[c]The estimated total amount of analyzed DNA sequence is the arithmetic product of 'pool size' and 'positions analyzed'.
[d]The substitution scores as found in Figure 3C. Shown are only the substitution scores for which both the forward and reverse observed minor variant frequencies had associated *P*-values smaller than 1×10^{-4}.
[e]The substitution scores (column 'absolute substitution score') are expressed here relative to the respective DNA amounts analyzed (column 'total amount analyzed').

Table S3. Mutations found in two bioselected clones.

Clone	Position[a]	Base change[b]
F421Y-c1	1,419	G ⇨ A
	7,243	G ⇨ A
	8,251	C ⇨ A
	9,835	C ⇨ T
	16,036	G ⇨ A
	17,576	G ⇨ A
	29,378	C ⇨ T
F421Y-c2	9,114	C ⇨ G
	11,581	G ⇨ A
	20,374	C ⇨ T
	28,638	T ⇨ A
	29,378	C ⇨ T
	31,726	A ⇨ G
	34,301	C ⇨ T
	35,229	A ⇨ Δ

[a]Position in HAdV-5 genome (accession no. AC_000008)
[b]The indicated base change is for the direct strand. 'Δ' stands for a deleted base.

Table S4. Protein coding changes in two bioselected viral clones.

Clone	Position[a]	Base Change[b]	Strand	Protein	Coding Change	Conservation[c]
F421Y-c1	8,251	C ⇨ A	direct	i-leader protein	L92M	C,D
	9,835	C ⇨ T	opposite	pTP	R255H	A,B,C,D,E,F
	17,576	G ⇨ A	direct	pV	V346M	A,B,C,D,E,F
F421Y-c2	9,114	C ⇨ G	opposite	pTP	Q495H	A,B,C,D,E,F
	11,581	G ⇨ A	direct	L1 52/55K	V178M	A,B,C,D,E,F
	28,638	T ⇨ A	direct	E3 6.7K	I31N	C[e]
	31,726	A ⇨ G	direct	fiber	T229A	C
	35,229	A ⇨ Δ	opposite	E4orf1	Frameshift[d]	A,B,C,D,E[f]

[a]Position in HAdV-5 genome (accession no. AC_000008)
[b]The indicated base change is for the direct strand. 'Δ' stands for a deleted base.
[c]HAdV subgroups across which the affected amino acids are conserved.
[d]The base deletion results in a frameshifted, shorter protein: the last 27 amino acids of E4orf1 are replaced by 7 other residues.
[e]E3 6.7K does not exist outside HAdV-5 subgroup C.
[f]Six of the 27 deleted residues are conserved across the indicated subgroups. E4orf1 is not known by us to exist in subgroup F.

SUPPLEMENTARY PROCEDURES AND MATERIAL

Rationale for selecting Ad pol residues to be mutated

A panel of twenty-three single-amino-acid substitution mutants of Ad pol was generated by site-directed mutation of nineteen selected residues thought to be important for governing replication fidelity (Figure 1 and Supplementary Table S1). In our picking of residues to modify, we were led by the consideration that only Ad pol's accuracy of polymerization was to be targeted – any other detrimental effects on functions necessary for faithful adenovirus genome replication were ideally to be avoided. For example, the efficiency of polymerization – or its protein-primed initiation – was preferably not to be compromised. To achieve this goal, in addition to drawing from what is known about Ad pol itself, we tapped into the wealth of data available on DNA-polymerases of other organisms, such as the Φ29, T4, RB69, and PRD1 bacteriophages, herpes simplex virus (HSV), cytomegalovirus (CMV), *Escherichia coli*, *Thermus aquaticus* (*Taq*), *Pyrococcus furiosus* (*Pfu*), and *Saccharomyces cerevisiae*. Eleven of the selected Ad pol residues, all located within the exonuclease domain, have been implicated – mostly through their homologues in other polymerase species – with proofreading function. The eight other targeted residues make part of either the fingers domain or the palm domain and are thought to play important roles in incoming base selection and/or tuning between polymerase and exonuclease activities.

Among the mutations of residues putatively involved in proofreading were those aimed at specifically abolishing the catalytic activity of the exonuclease domain. These substitutions, which were meant to rigorously knock-out exonucleolytic activity, concern the highly conserved residues D283, E285, D422, Y580, and D584 (Figure 1 and Supplementary Table S1, assigned amino acid no. 1, 2, 6, 9 and 10). These are the key residues thought to be involved in metal ion binding and catalysis. Importantly, substitutions of the homologues of these residues in polymerases of Φ29, T4, RB69, and HSV have previously been shown to render these polymerases gravely impaired (or effectively inactive) with respect to the ability to catalyze exonucleolytic hydrolysis (2-4,6,12-14). Furthermore, a D422A mutant of Ad pol has already been shown by Brenkman *et al*. to be exonuclease deficient (41).

Ad pol residues thought to be otherwise (than catalytically) involved in proofreading were also subjected to mutation. Three of these, namely T286, N417, and F421 (no. 3-5), were selected for their suspected direct involvement – based on studies on their homologues in Φ29 – in stabilization of the primer terminus at the exonuclease active site (7,9). Two other residues targeted for mutation were S506 and L507 (no. 7 and 8), both of which make part of to the (S/T)Lx2*h* motif (9) (Figure 1). These residues putatively play important roles in proofreading, possibly through (indirect) interaction with the ssDNA

primer, as has been suggested for their homologous counterparts in Φ29 polymerase (9). Finally, the last residue within the exonuclease domain that was mutated in this study was V585 (no. 11). Substitution of its homologue in T4 polymerase had led to an *in vivo* mutator phenotype, possibly due to functional disturbance of a neighboring catalytic residue, which, in Ad pol, is D584 (no. 10) (15).

Further singled out for mutation were three amino acid residues residing in the conserved Pol II (or Motif A) region of the palm domain (Figure 1). One of these residues, G666 (no. 12), is located in the Y/IxGG/A motif, a region believed to be involved in tuning between polymerase and exonuclease activities (16,42). The other two residues, M689 and Y690 (no. 13 and 14), might make part of the incoming nucleotide binding pocket and, as such, could play important roles in the geometric selection of nucleotides, which has been suggested for their putative homologues in other polymerases (17-26).

Within the fingers domain, with its prominent role in the binding of incoming nucleotides, a total of five residues were selected for mutation. Two non-highly conserved residues were chosen based on the speculation that their respective substitutions could have indirect effects on base selecting properties. One of these residues, D827 (no. 15), is seemingly located in the connecting loop between two helices of the fingers domain (Figure 1). Positioned as such, it might functionally resemble similarly located *Pfu* DNA polymerase residues, mutation of which previously led to strong mutator phenotypes (27). The other lesser-conserved residue that was targeted, S834 (no. 16), which seems to be situated in a second helix of the fingers domain of Ad pol (Figure 1), is positioned just before highly conserved residues thought to be directly involved in catalytic and base-selection functions. Mutation of residues at this approximate location in *Taq* DNA polymerase I has previously given rise to increased *in vitro* and *in vivo* mutation rates (28). Further targeted for mutation in the fingers domain were two highly conserved residues, L838 and Y844 (no. 17 and 18), both of which were selected based on their presumed critical roles in controlling base-pairing correctness (24,28-38). Finally, another target was G845 (no. 19), a highly conserved residue to which different functions have been ascribed for different polymerases. In Φ29 it would be involved in dsDNA binding (39), while in *S. cerevisiae* polymerase alpha, a familiy A polymerase, it is suspected to be critical for the insertion of correct nucleotides (40).

Construction and production of lentiviral vectors

Lentiviruses used in this study were self-inactivating, third-generation HIV-I-derived vectors (43,44). LV plasmids expressing wild-type Ad5 pol and the mutant Ad5 pol D422A were made by insertion of previously described cDNAs

(41) into the basal LV plasmids pRRL-cPPT-CMV-X-PRE-SIN (45) or its derivative pLV.CMV.IRES.PURO (46). C-terminal HA-tagging of these two polymerases involved PCR-amplification using a reverse primer harboring an HA-tag coding sequence (see below for the primer sequence). For the generation of polymerase mutants other than D422A, point mutations were introduced – in a LV plasmid expressing HA-tagged Ad5 pol – by a site-directed mutagenesis technique involving 'inversed' PCR of the parental plasmid and subsequent self-ligation of the PCR product (see below for the primer sequence pairs). Lentivirus production, which was performed as described previously (47), involved the co-transfection of LV plasmids with helper plasmids encoding HIV-1 gag-pol, HIV-1 rev, and the VSV-G envelope. Lentivirus stocks were titrated using a HIV-1 p24 antigen enzyme-linked immunosorbent assay kit (ZeptoMetrix Corp., New York, NY, USA).

Generation of polymerase-defective adenovirus vectors

Generation of the polymerase-defective Ad vectors involved two basal Ad genome-containing plasmids. The first of these, pAdGL, was described previously and contains an E1- and E3-deleted HAdV-5 genome, with, at the place of E1, two tandem expression cassettes encoding GFP and firefly luciferase (46). The second basal plasmid, pAd, carries within its plasmid backbone of pShuttle (48) the complete HAdV-5 genome of pTG3602 (49). These plasmids were modified, in several cloning steps, to contain a 571-bp, polymerase-affecting deletion (ΔPOL; corresponding to nt 7312 to 7882 of the HAdV-5 genome; accession no. AC_000008). First, through splicing by overlap extension-PCR (see below for primer sequences), a chimeric DNA fragment was generated that consisted of two 'deletion-flanking' fragments (of approximately 1 kb each) with, spliced in between them, a chloramphenicol resistance gene (Cam^r), obtained from pGP618 (50). Subsequently, this 'ΔPOL.CAM' fragment served as a selectable donor in homologous recombination [in *E. coli* BJ5183 (51)] with acceptor plasmids pAdGL and pAd, allowing the generation of pAdGLΔPOL.CAM and pAdΔPOL.CAM. Finally, the excision of Cam^r (by virtue of two PCR-introduced *Swa*I sites) yielded pAdGLΔPOL and pAdΔPOL. The viral genomes of these two plasmids were released by *Pac*I digestion and subsequently transfected into 911.AdPol cells for viral rescue of, respectively, AdGLΔPOL and HAdV-5ΔPOL.

Initial setup of the Ad pol *trans*-complementation system

A polymerase complementation system was set up to test Ad pol mutants for their ability to support the functions necessary for productive Ad replication (Supplementary Figure S1). This system employs a polymerase-defective Ad reporter vector that is completely dependent – for its replication – on a functional Ad pol (mutant) being provided in *trans*. Such a polymerase

complementation strategy is analogous to previously taken approaches in studies that investigated variants of HSV (14,52) and RB69 (32) polymerases.

The E1- and E3-deleted polymerase-defective Ad vector generated for this study, AdGLΔPOL, accommodates a polymerase gene-disrupting deletion at a location not known to harbor any other essential elements than the polymerase sequence (Supplementary Figure S1C). Importantly, at the site of this partial polymerase gene-deletion, an extra stop codon was introduced such that only a severely truncated Ad pol (encoded by the first forth of the Ad pol open reading frame) was producible. Of note, Amalfitano *et al*. previously showed that an Ad vector with an essentially similar polymerase deletion was effectively rendered unable to replicate, except when provided with Ad pol *in trans* (53). To facilitate monitoring of viral replication, the AdGLΔPOL genome was additionally equipped with CMV promoter-driven reporter genes for GFP and firefly luciferase. Transfection of the AdGLΔPOL genome into (E1-complementing) cells engineered to express wild-type Ad pol resulted in the rescue of viable AdGLΔPOL virions.

The Ad polymerase complementation system further involved the use of lentivirus (LV) vectors for achieving stable expression of Ad pol variants in cells (Supplementary Figure S1A). In this regard, the above mentioned viral rescuing of AdGLΔPOL was conducted using cells transduced with a LV-vector encoding wild-type Ad pol. Initially, in order to detect the heterologous expression of Ad pol in these cells, immunofluorescence (IF) was performed for which we made use of a lab-generated anti-serum raised against an Ad pol/pTP complex. Although detection by this means proved to be achievable, the procedure suffered from a low signal-to-background ratio (data not shown). Therefore, since more unambiguous detectability was preferred, we opted to equip the Ad pols of this study with HA-tags at their C-termini. Supplementary Figure S1B shows that lentiviraly expressed HA-tagged versions of wild-type Ad pol and a mutant Ad pol, namely D422A, were readily detectable by IF using an HA-tag specific antibody. Both these polymerases were found to localize to both the nucleus and the cytoplasm, with a higher intensity of detection observed in the nucleus.

To ascertain that Ad pol's ability to replicate the Ad genome was not affected by its fusion to an HA-tag, a pilot complementation experiment was carried out (Supplementary Figure S1D). Cells expressing the HA-tagged versions of either wild-type Ad Pol or the D422A mutant were infected with the polymerase-defective vector AdGLΔPOL. Alongside these infections were those of positive control cells expressing (non-tagged) wild-type Ad Pol and negative control cells not expressing any Ad pol. The results of GFP expression analysis at 52 hours post infection show that HA-tagged Ad pol is as efficient as non-tagged Ad pol to complement AdGLΔPOL. Thus, a C-terminal HA-tag proves to be compatible with Ad pol's functions necessary for Ad genome

replication. Interestingly, the results further reveal that the one mutant taken along, D422A, is unable to complement the polymerase-defective vector. Thus, although the D422A mutant had previously been shown to achieve polymerization in an *in vitro* setting (41), it seems not to sustain complete Ad replication.

Pilot deep sequencing run on a spiked test sample

To probe the utility of massively parallel sequencing (MPS) for the detection of minority mutations, we performed a sequencing run on a spiked test sample. A 3.1-kilobase-pair (kb) DNA fragment was spiked with several minority fragments, each of which differing from the original fragment by only a 1- or 2-nucleotide variation. The resulting sample, which contained spikes with theoretical prevalences of 25, 6.25, 1.56, 0.39, and 0.1%, was run on a Solexa/Illumina Genome Analyzer (I) instrument to obtain approximately 20 million 20-mer raw reads. These reads were mapped, allowing up to two mismatches, against the 3.1-kb reference sequence using the ELAND aligner of Illumina's Genome Analyzer data analysis pipeline. Resultant forward and reverse mappings – representing only reads passing the default Illumina chastity filter (threshold = 0.6) – were subjected to single-base masking to exclude bases with phred quality values lower than 29. These operations finally led to forwardly and reversely mapped base-call distributions exhibiting average coverage depths of respectively 7096 and 6134 high-quality bases. Notably, considerable variation in the coverage depth was observed along the length of the reference sequence (Supplementary Figure S2A). Nonetheless, as much as 84% of the reference sequence was found to be covered – in both the forward and the reverse orientation – by at least 1200 high-quality base-calls per position (Supplementary Figure S2B).

Of the spiked mutations, those with prevalences of 25, 6.25, 1.56, and 0.39% were found to be distinctively present in both the forward and the reverse distributions (Supplementary Figures S2C and S2D). Furthermore, each of these spikes was found with approximately its expected prevalence (Table 1). By contrast, the spike with expected prevalence of 0.1% (spike no. 5) did not as prominently stand out from local levels of 'background' base-calls.

Directed by the above-obtained base-call distribution data we defined generalizable filtering rules that allowed for the confident scoring of minority mutations. The first of these rules disallows from the analysis any positions for which either forward or reverse coverage levels are below a certain set minimum. This rule was found necessary to protect against the relatively high background base-call levels seen at positions where the coverage levels dropped. The second filtering rule conditions that a given minor sequence variant may only be scored as a bona fide mutation when present – with

prevalences above a set cutoff value – in both the forward and reverse distributions. Applied to the spiked test sample data, the combination of these rules allowed the scoring of spiked mutations 1 to 4 without picking up any false-positives. Specifically, with the minimal coverage level set arbitrarily at 1200, a mutation prevalence-cutoff value of as low as 0.17% permits scoring of only those four spikes (Supplementary Figure S2E). Of note, for all the thusly 'scored' spikes, both the forward and reverse observed frequencies would have very low estimated probabilities of occurring solely as a consequence of random sequencing errors, i.e. all *P*-values were below 1×10^{-12} (Table 1). By contrast, the 'non-scored' spike could not be significantly distinguished from background sequencing errors (all *P*-values above 0.4).

Oligonucleotides used for cloning, sequencing, and reverse transcriptase-PCRs

Name	Sequence
HA-tag (YPYDVPDYA) addition to the Ad pol C-terminus[a]	
Forward	GACGTATGTTCCCATAGTAACGC
Reverse	GTCACGTGCT**CTA**AGCGTAATCCGGAACATCGTATGGGTACGGCATCTCGATC
Construction of a polymerase-defective adenovirus vector[b]	
Lpol 1	TCGATACGCGTTGGACAGCAACTTGG
Lpol 2	CGGATGAATGGCAGAAATTTAAATGGTCAGGGACACCTTTGC
Cam 1	GGTGTCCCTGACCATTTAAATTTCTGCCATTCATCCGC
Cam 2	CCTTCATGCTGGTCATTTAAATCAGTAAGTTGGCAGCATTC
Rpol 1	GCTGCCAACTTACTGATTTAAATGACCAGCATGAAGGGCA
Rpol 2	CGTATGTATACGCCTTCTCGCAGCTC
Amplification of a 6.5-kb DNA fragment from virus pools for deep sequencing	
Forward	GTAGTTTTGTATCTGTTTTGCAGCAG
Reverse	CGAATTTATCCACCAGACCAC
Reverse transcriptase-PCRs[c]	
β-actin 1	GGCATCCTCACCCTGAAGTA
β-actin 2	GGGGTGTTGAAGGTCTCAAA
E1A 1	GTCCGGTTTCTATGCCAAAC
E1A 2	GATAGCAGGCGCCATTTTAG
E3 ADP 1	CCTGAAACACCTGGTCCACT
E3 ADP 2	GCGTTGGTTGTGTTGGTCAT
ML ADP 1	CGAGAAAGGCGTCTAACCAG
ML ADP 2	GCGTTGGTTGTGTTGGTCAT

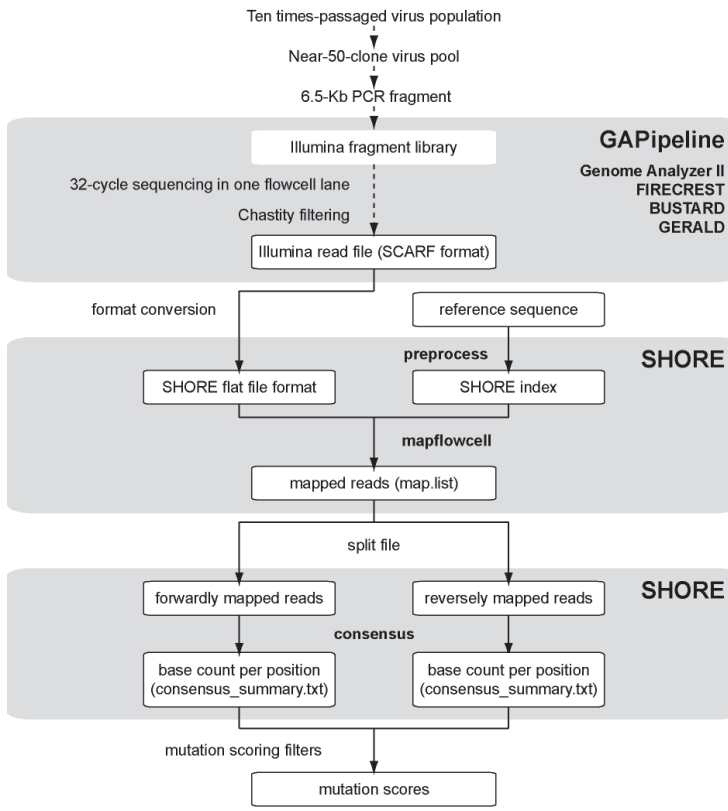
[a]The forward primer is CMV promoter-specific. The reverse primer anneals to the C-terminus of Ad pol. The HA-tag encoding sequence is shown in red font. The new stop codon is in boldface.
[b]The indicated primers were used to construct a recombination donor fragment ('ΔPOL.CAM') carying a chloramphenicol resistance gene (Cam^r) at the site of a partial Ad polymerase gene deletion. First, two 'deletion flanking' fragments were generated using the Lpol and Rpol primer sets, and a Cam^r-containing fragment using the Cam primer set. Then, to generate ΔPOL.CAM, these three fragments were fused by performing two sequential splicing by overlap extension-PCRs using appropriate combinations of primers. The respective sequences providing the overlaps between the three fragments are color-coded blue and green. Underlined sequences indicate introduced restriction enzyme recognition sites.
[c]Primer sets E3 ADP and ML ADP have a common reverse primer.

Oligonucleotides used for the construction of Ad pol mutants

Mutant	Primer 1 (forward)[a]	Primer 2 (reverse)[a]	RE site
D283A	GTACGCTGTAGAGAC-CTATACTTGGATGG	GTAACAAAGAGACGCTCG-GTGC	SnaBI
E285A	CGACCTATACTTGGAT-GGGGGC	CGACATCGTAGGTGACAAA-GAGACG	NruI
T286I	TCTATACTTGGATGGGGGCCT	TCTCTACATCGTAGGT-GACAAAGAG	BglII
N417A	TGCCATCAACGGCTTTGAC-GAGA	TGCCCCACAATGTAAAGT-TCCAAG	SphI
N417D	ATCAACGGCTTTGACGAGATC	ATCGTGGCCCACAATG-TAAAGTTC	EcoRV
F421A	GCCGACGAGATCGTGCTCGCC	GCCGTTGATGTTGTGGCC	KasI
F421S	TCCGACGAGATCGTGCTCGC	TCCGTTGATGTTGTGGCC	BamHI
F421Y	AATGGGTACGACGAGATCGT-GCTCGC	AATGTTGTGGCCCACAAT-GTA	AseI
S506T	GTTGCGGAAGGCCGCGCAG	GTTGTGTGGGTGAGCG-CAAAG	AclI
L507T	ACTCGGAAGGCCGCGCA	ACTGGTGTGGGTGAGCG-CA	ScaI
Y580A	TGCGCCCTAGACGTGCAG	TGCGTCCAGGGTTTCCTT-GATGATG	SphI
D584A	ACAGGTCACCGCCGAGCTG	ACAGCTAGGGCGCAG-TAGTCCAG	BsrGI
V585A	GCAGGTCACCGCCGAGC	GCATCTAGGGCGCAG-TAGTCCAG	FspI
G666A	GGCAGATGCTACCCTACAT-ATCTTGG	GGCGCGGATGCTGGCG	NaeI
M689K	CGTCCGCGCTCACCCACC	CGTATTTGCCGCAAAT-GTCGTAAAC	MluI
M689N	CGTCCGCGCTCACCCACC	CGTAGTTGCCGCAAAT-GTCGTAAAC	MluI
Y690F	CGAGCGCGCTCACCCACCC	CGAACATGCCGCAAAT-GTCG	NruI
D827A	GAACCAAACCCTGCGCTC	TTAGCGCGATCGGCGCG	DdeI
S834E	AGGGAGATCGCCAAGTTGCT-GTCC	CAGGGTTTGGTTTTTGTCGC	Bsu36I
L838A	CTCTCCAACGCCCTCTACGG	GGCCTTGGCGAT-GGAGCGC	EcoO109I
Y844A	AGCAGGGTCGTTTGCCACCA	AGCGCGTTGGACAG-CAACTTG	NheI
G845A	CGTCGTTTGCCACCAAGC	CGTAGAGGGCGTTGGACA	MluI

[a]The indicated primer pairs were used for an 'inversed' PCR on a lentiviral vector plasmid encoding HA-tagged HAdV-5 DNA polymerase. Resultant PCR products were subsequently self-ligated to obtain plasmids encoding the respective mutants. Red fonts indicate base substitutions. Boldfaced nucleotides show the codons mutated to encode the amino acid changes indicated in the first column. Underlined sequences represent the newly generated restriction enzyme recognition sites (depicted in the last column) used for identification of positive clones.

Workflow for the analysis of deep sequencing data obtained from passaged virus pools



Sequencing data obtained from the passaged virus pools were outputted by the Genome Analyzer Pipeline (GAPipeline) in read files of Illumina's SCARF file format. These files included only reads passing Illumina's default chastity filter (i.e. reads that among their first 25 bases have maximally one base with a chastity value less than 0.6). Subprograms 'preprocess' and 'mapflowcell' of the short read analysis pipeline SHORE (http://1001genomes.org/downloads/shore.html) were used for mapping of the reads against the reference sequence. The alignments were performed allowing up to 2 mismatches and/or gaps. The SHORE subprogram 'consensus' was used – separately for the forward and reverse mappings – to generate base counts per position in the reference sequence. Individual base masking was performed (by SHORE consensus) using a quality value cutoff of 30. Resultant forward and reverse 'consensus_summary.txt' files served for mutation scoring using a sequencing depth requirement of 1200 and a prevalence cutoff value of 0.25%. See the table on the next page for further details on the data analysis workflow.

Command line programs and parameters used for the analysis of deep sequencing data obtained from passaged virus pools

Program[a]	Relevant parameters	Notes
Custom Perl script[b]	s_N_sequence.txt	Input read file (Illumina SCARF format)
	reads_0.fl	Output read file (SHORE format)
SHORE preprocess	-f Ref.fa	Reference sequence file (fasta format)
	-i IndexFolder	
SHORE mapflowcell	-o FlowcellFolder	See note below[c]
	-f IndexFolder/Ref.fa.shore	
	-n 2	Maximum no. of mismatches plus gaps
	-g 2	Maximum no. of gaps
Custom Perl script[d]	map.list	Input file (= output of mapflowcell)
	CustomOutputFolder	Output files: for.list and rev.list
SHORE consensus	-n 0001	Arbitrary ID
	-f IndexFolder/Ref.fa.shore	
	-o AnalysisFolder	
	-i for.list *(or rev.list)*	Input: for.list or rev.list (not map.list)
	-q 30	Quality value cutoff for base masking
	-v	Required for 'consensus_summary.txt'
	-r	
Custom Perl script[e]	consensus_summary.txt	Forward input file
	consensus_summary.txt	Reverse input file
	MutationScoreFolder	Output: mutation score files
	1200	Coverage depth requirement
	0.25	Minor variant frequency cutoff (%)

[a]See workflow diagram above for an overview of the data analysis workflow. SHORE is a short read analysis pipeline (http://1001genomes.org/downloads/shore.html) (54) . Its subprogram 'mapflowcell' invokes the aligner GenomeMapper (55). (See Supplemental References for literature cited here)

[b]A Perl script was used to convert between SCARF and SHORE read file formats. Of note, this conversion entailed inferring 'Sanger' base quality scores (needed for SHORE) from the 'Illumina' base quality scores (present in the SCARF format). Furthermore, since SCARF files lack chastity value information (which SHORE files do accommodate), each base in the generated 'reads_0.fl' files was assigned the maximum chastity value.

[c]The read files-containing folder structure normally built by the SHORE subprogram 'illumina2flat' – i.e. ProjectFolder/FlowcellFolder/LaneFolder/ReadFolder/LengthFolder – was created manually and appropriately populated – in its 'LengthFolders'– by the above-generated 'reads_0.fl' files.

[d]This script divides the forwardly and reversely mapped reads of the 'map.list' file (generated by the SHORE subprogram 'mapflowcell') over two new alignment files: 'for.list' and 'rev. list'.

[e]This mutation scoring Perl script uses as its input the forward and reverse mapping-derived 'consensus_summary.txt' files. These files, which were outputted by the SHORE subprogram 'consensus' in the folder 'AnalysisFolder/ConsensusAnalysis/supplementary_data', contain A, G, C, and T base counts – both quality filtered and unfiltered – per position in the reference sequence.

Summary statistics for minor variant frequencies found in deep sequencing data

	Mean[a]		97.5th percentile[a,b]	
	forward	reverse	forward	reverse
Spiked test sample sequencing run				
with spikes	2.4×10^{-4}	3.1×10^{-4}	8.9×10^{-4}	1.2×10^{-3}
without spikes	1.8×10^{-4}	2.7×10^{-4}	8.9×10^{-4}	1.2×10^{-3}
Virus pool sample sequencing runs				
stock 1	1.9×10^{-4}	1.7×10^{-4}	1.0×10^{-3}	9.0×10^{-4}
stock 2	1.6×10^{-4}	1.5×10^{-4}	8.1×10^{-4}	6.8×10^{-4}
pol-HA	1.9×10^{-4}	1.8×10^{-4}	1.1×10^{-3}	1.0×10^{-3}
T286I	1.9×10^{-4}	1.8×10^{-4}	7.5×10^{-4}	6.7×10^{-4}
N417A	1.7×10^{-4}	1.6×10^{-4}	7.9×10^{-4}	6.6×10^{-4}
F421Y	2.2×10^{-4}	1.9×10^{-4}	7.8×10^{-4}	6.8×10^{-4}
S506T	1.9×10^{-4}	1.8×10^{-4}	8.1×10^{-4}	7.2×10^{-4}
V585A	1.8×10^{-4}	1.7×10^{-4}	9.3×10^{-4}	8.3×10^{-4}
D827A	1.7×10^{-4}	1.6×10^{-4}	8.2×10^{-4}	7.5×10^{-4}
S834E	1.8×10^{-4}	1.6×10^{-4}	7.7×10^{-4}	6.8×10^{-4}
pol-HA*	1.8×10^{-4}	1.8×10^{-4}	8.6×10^{-4}	7.6×10^{-4}
F421Y*	2.1×10^{-4}	1.9×10^{-4}	7.4×10^{-4}	6.4×10^{-4}

[a]The statistics are for positions for which both forward and reverse coverage levels were above 1200.
[b]The 97.5th percentiles were taken as error rate estimates used for the calculation of *P*-values for minor variant occurrences (See the statistics section in the main article).

SUPPLEMENTARY REFERENCES

1. Blanco, L. and Salas, M. (1996) Relating structure to function in phi29 DNA polymerase. J.Biol.Chem., 271, 8509-8512.
2. Bernad, A., Blanco, L., Lazaro, J.M., Martin, G. and Salas, M. (1989) A conserved 3'--5' exonuclease active site in prokaryotic and eukaryotic DNA polymerases. Cell, 59, 219-228.
3. Esteban, J.A., Soengas, M.S., Salas, M. and Blanco, L. (1994) 3'-->5' exonuclease active site of phi 29 DNA polymerase. Evidence favoring a metal ion-assisted reaction mechanism. J.Biol. Chem., 269, 31946-31954.
4. Abdus Sattar, A.K., Lin, T.C., Jones, C. and Konigsberg, W.H. (1996) Functional consequences and exonuclease kinetic parameters of point mutations in bacteriophage T4 DNA polymerase. Biochemistry, 35, 16621-16629.
5. Murphy, K., Darmawan, H., Schultz, A., Fidalgo, d.S. and Reha-Krantz, L.J. (2006) A method to select for mutator DNA polymerase deltas in Saccharomyces cerevisiae. Genome, 49, 403-410.
6. Kuhn, F.J. and Knopf, C.W. (1996) Herpes simplex virus type 1 DNA polymerase. Mutational analysis of the 3'-5'-exonuclease domain. J.Biol.Chem., 271, 29245-29254.
7. de Vega, M., Lazaro, J.M., Salas, M. and Blanco, L. (1996) Primer-terminus stabilization at the 3'-5' exonuclease active site of phi29 DNA polymerase. Involvement of two amino acid residues highly conserved in proofreading DNA polymerases. EMBO J., 15, 1182-1192.
8. Zhu, W. and Ito, J. (1994) Family A and family B DNA polymerases are structurally related: evolutionary implications. Nucleic Acids Res., 22, 5177-5183.
9. de Vega, M., Lazaro, J.M., Salas, M. and Blanco, L. (1998) Mutational analysis of phi29 DNA polymerase residues acting as ssDNA ligands for 3'-5' exonucleolysis. J.Mol.Biol., 279, 807-822.
10. Chou, S. and Marousek, G.I. (2008) Accelerated evolution of maribavir resistance in a cytomegalovirus exonuclease domain II mutant. J.Virol., 82, 246-253.
11. de Vega, M., Blanco, L. and Salas, M. (1998) phi29 DNA polymerase residue Ser122, a single-stranded DNA ligand for 3'-5' exonucleolysis, is required to interact with the terminal protein. J.Biol.Chem., 273, 28966-28977.
12. Soengas, M.S., Esteban, J.A., Lazaro, J.M., Bernad, A., Blasco, M.A., Salas, M. and Blanco, L. (1992) Site-directed mutagenesis at the Exo III motif of phi 29 DNA polymerase; overlapping structural domains for the 3'-5' exonuclease and strand-displacement activities. EMBO J., 11, 4227-4237.
13. Wang, C.X., Zakharova, E., Li, J., Joyce, C.M., Wang, J. and Konigsberg, W. (2004) Pre-steady-state kinetics of RB69 DNA polymerase and its exo domain mutants: effect of pH and thiophosphoryl linkages on 3'-5' exonuclease activity. Biochemistry, 43, 3853-3861.
14. Hwang, Y.T., Liu, B.Y., Coen, D.M. and Hwang, C.B. (1997) Effects of mutations in the Exo III motif of the herpes simplex virus DNA polymerase gene on enzyme activities, viral replication, and replication fidelity. J.Virol., 71, 7791-7798.
15. Reha-Krantz, L.J. (1988) Amino acid changes coded by bacteriophage T4 DNA polymerase mutator mutants. Relating structure to function. J.Mol.Biol., 202, 711-724.
16. Truniger, V., Lazaro, J.M., Salas, M. and Blanco, L. (1996) A DNA binding motif coordinating synthesis and degradation in proofreading DNA polymerases. EMBO J., 15, 3430-3441.
17. Reha-Krantz, L.J. and Nonay, R.L. (1994) Motif A of bacteriophage T4 DNA polymerase: role in primer extension and DNA replication fidelity. Isolation of new antimutator and mutator DNA polymerases. J.Biol.Chem., 269, 5635-5643.
18. Venkatesan, R.N., Hsu, J.J., Lawrence, N.A., Preston, B.D. and Loeb, L.A. (2006) Mutator phenotypes caused by substitution at a conserved motif A residue in eukaryotic DNA polymerase delta. J.Biol.Chem., 281, 4486-4494.
19. Zhong, X., Pedersen, L.C. and Kunkel, T.A. (2008) Characterization of a replicative DNA polymerase mutant with reduced fidelity and increased translesion synthesis capacity. Nucleic Acids Res., 36, 3892-3904.

20. Shinkai, A. and Loeb, L.A. (2001) In vivo mutagenesis by Escherichia coli DNA polymerase I. Ile(709) in motif A functions in base selection. J.Biol.Chem., 276, 46759-46764.

21. Niimi, A., Limsirichaikul, S., Yoshida, S., Iwai, S., Masutani, C., Hanaoka, F., Kool, E.T., Nishiyama, Y. and Suzuki, M. (2004) Palm mutants in DNA polymerases alpha and eta alter DNA replication fidelity and translesion activity. Mol.Cell Biol., 24, 2734-2746.

22. Patel, P.H., Kawate, H., Adman, E., Ashbach, M. and Loeb, L.A. (2001) A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. J.Biol.Chem., 276, 5044-5051.

23. Pursell, Z.F., Isoz, I., Lundstrom, E.B., Johansson, E. and Kunkel, T.A. (2007) Regulation of B family DNA polymerase fidelity by a conserved active site residue: characterization of M644W, M644L and M644F mutants of yeast DNA polymerase epsilon. Nucleic Acids Res., 35, 3076-3086.

24. Saturno, J., Blanco, L., Salas, M. and Esteban, J.A. (1995) A novel kinetic analysis to calculate nucleotide affinity of proofreading DNA polymerases. Application to phi 29 DNA polymerase fidelity mutants. J.Biol.Chem., 270, 31235-31243.

25. Bonnin, A., Lazaro, J.M., Blanco, L. and Salas, M. (1999) A single tyrosine prevents insertion of ribonucleotides in the eukaryotic-type phi29 DNA polymerase. J.Mol.Biol., 290, 241-251.

26. Minnick, D.T., Bebenek, K., Osheroff, W.P., Turner, R.M., Jr., Astatke, M., Liu, L., Kunkel, T.A. and Joyce, C.M. (1999) Side chains that influence fidelity at the polymerase active site of Escherichia coli DNA polymerase I (Klenow fragment). J.Biol.Chem., 274, 3067-3075.

27. Biles, B.D. and Connolly, B.A. (2004) Low-fidelity Pyrococcus furiosus DNA polymerase mutants useful in error-prone PCR. Nucleic Acids Res., 32, e176.

28. Suzuki, M., Avicola, A.K., Hood, L. and Loeb, L.A. (1997) Low fidelity mutants in the O-helix of Thermus aquaticus DNA polymerase I. J.Biol.Chem., 272, 11228-11235.

29. Truniger, V., Lazaro, J.M., de Vega, M., Blanco, L. and Salas, M. (2003) phi 29 DNA polymerase residue Leu384, highly conserved in motif B of eukaryotic type DNA replicases, is involved in nucleotide insertion fidelity. J.Biol. Chem., 278, 33482-33491.

30. Zhang, H., Rhee, C., Bebenek, A., Drake, J.W., Wang, J. and Konigsberg, W. (2006) The L561A substitution in the nascent base-pair binding pocket of RB69 DNA polymerase reduces base discrimination. Biochemistry, 45, 2211-2220.

31. Zhang, H., Beckman, J., Wang, J. and Konigsberg, W. (2009) RB69 DNA polymerase mutants with expanded nascent base-pair-binding pockets are highly efficient but have reduced base selectivity. Biochemistry, 48, 6940-6950.

32. Bebenek, A., Dressman, H.K., Carver, G.T., Ng, S., Petrov, V., Yang, G., Konigsberg, W.H., Karam, J.D. and Drake, J.W. (2001) Interacting fidelity defects in the replicative DNA polymerase of bacteriophage RB69. J.Biol.Chem., 276, 10387-10397.

33. Bebenek, A., Carver, G.T., Dressman, H.K., Kadyrov, F.A., Haseman, J.K., Petrov, V., Konigsberg, W.H., Karam, J.D. and Drake, J.W. (2002) Dissecting the fidelity of bacteriophage RB69 DNA polymerase: site-specific modulation of fidelity by polymerase accessory proteins. Genetics, 162, 1003-1018.

34. Yang, G., Wang, J. and Konigsberg, W. (2005) Base selectivity is impaired by mutants that perturb hydrogen bonding networks in the RB69 DNA polymerase active site. Biochemistry, 44, 3338-3346.

35. Bell, J.B., Eckert, K.A., Joyce, C.M. and Kunkel, T.A. (1997) Base miscoding and strand misalignment errors by mutator Klenow polymerases with amino acid substitutions at tyrosine 766 in the O helix of the fingers subdomain. J.Biol. Chem., 272, 7345-7351.

36. Carroll, S.S., Cowart, M. and Benkovic, S.J. (1991) A mutant of DNA polymerase I (Klenow fragment) with reduced fidelity. Biochemistry, 30, 804-813.

37. Polesky, A.H., Steitz, T.A., Grindley, N.D. and Joyce, C.M. (1990) Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from Escherichia coli. J.Biol.Chem., 265, 14579-14591.

38. Ogawa, M., Limsirichaikul, S., Niimi, A., Iwai, S., Yoshida, S. and Suzuki, M. (2003) Distinct function of conserved

amino acids in the fingers of Saccharomyces cerevisiae DNA polymerase alpha. J.Biol.Chem., 278, 19071-19078.

39. Blasco, M.A., Lazaro, J.M., Blanco, L. and Salas, M. (1993) Phi 29 DNA polymerase active site. The conserved amino acid motif "Kx3NSxYG" is involved in template-primer binding and dNTP selection. J.Biol.Chem., 268, 16763-16770.
40. Limsirichaikul, S., Ogawa, M., Niimi, A., Iwai, S., Murate, T., Yoshida, S. and Suzuki, M. (2003) The Gly-952 residue of Saccharomyces cerevisiae DNA polymerase alpha is important in discriminating correct deoxyribonucleotides from incorrect ones. J.Biol. Chem., 278, 19079-19086.
41. Brenkman, A.B., Breure, E.C. and van der Vliet, P.C. (2002) Molecular architecture of adenovirus DNA polymerase and location of the protein primer. J.Virol., 76, 8200-8207.
42. Brenkman, A.B., Heideman, M.R., Truniger, V., Salas, M. and van der Vliet, P.C. (2001) The (I/Y)XGG motif of adenovirus DNA polymerase affects template DNA binding and the transition from initiation to elongation. J.Biol. Chem., 276, 29846-29853.
43. Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D. and Naldini, L. (1998) A third-generation lentivirus vector with a conditional packaging system. J.Virol., 72, 8463-8471.
44. Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L. and Trono, D. (1998) Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J.Virol., 72, 9873-9880.
45. Barry, S.C., Harder, B., Brzezinski, M., Flint, L.Y., Seppen, J. and Osborne, W.R. (2001) Lentivirus vectors encoding both central polypurine tract and posttranscriptional regulatory element provide enhanced transduction and transgene expression. Hum.Gene Ther., 12, 1103-1108.
46. Uil, T.G., de Vrij, J., Vellinga, J., Rabelink, M.J., Cramer, S.J., Chan, O.Y., Pugnali, M., Magnusson, M., Lindholm, L., Boulanger, P. *et al*. (2009) A lentiviral vector-based adenovirus fiber-pseudotyping approach for expedited functional assessment of candidate retargeted fibers. J.Gene Med., 11, 990-1004.
47. Carlotti, F., Bazuine, M., Kekarainen, T., Seppen, J., Pognonec, P., Maassen, J.A. and Hoeben, R.C. (2004) Lentiviral vectors efficiently transduce quiescent mature 3T3-L1 adipocytes. Mol.Ther., 9, 209-217.
48. He, T.C., Zhou, S., da Costa, L.T., Yu, J., Kinzler, K.W. and Vogelstein, B. (1998) A simplified system for generating recombinant adenoviruses. Proc.Natl. Acad.Sci.U.S.A, 95, 2509-2514.
49. Chartier, C., Degryse, E., Gantzer, M., Dieterle, A., Pavirani, A. and Mehtali, M. (1996) Efficient generation of recombinant adenovirus vectors by homologous recombination in Escherichia coli. J.Virol., 70, 4805-4810.
50. Groenen, M.A., Timmers, E. and van de, P.P. (1985) DNA sequences at the ends of the genome of bacteriophage Mu essential for transposition. Proc. Natl.Acad.Sci.U.S.A, 82, 2087-2091.
51. Hanahan, D. (1983) Studies on transformation of Escherichia coli with plasmids. J.Mol.Biol., 166, 557-580.
52. Digard, P., Chow, C.S., Pirrit, L. and Coen, D.M. (1993) Functional analysis of the herpes simplex virus UL42 protein. J.Virol., 67, 1159-1168.
53. Amalfitano, A., Hauser, M.A., Hu, H., Serra, D., Begy, C.R. and Chamberlain, J.S. (1998) Production and characterization of improved adenovirus vectors with the E1, E2b, and E3 genes deleted. J.Virol., 72, 926-933.
54. Ossowski, S., Schneeberger, K., Clark, R.M., Lanz, C., Warthmann, N. and Weigel, D. (2008) Sequencing of natural strains of Arabidopsis thaliana with short reads. Genome Res., 18, 2024-2033.
55. Schneeberger, K., Hagmann, J., Ossowski, S., Warthmann, N., Gesing, S., Kohlbacher, O. and Weigel, D. (2009) Simultaneous alignment of short reads against multiple genomes. Genome Biol., 10, R98.