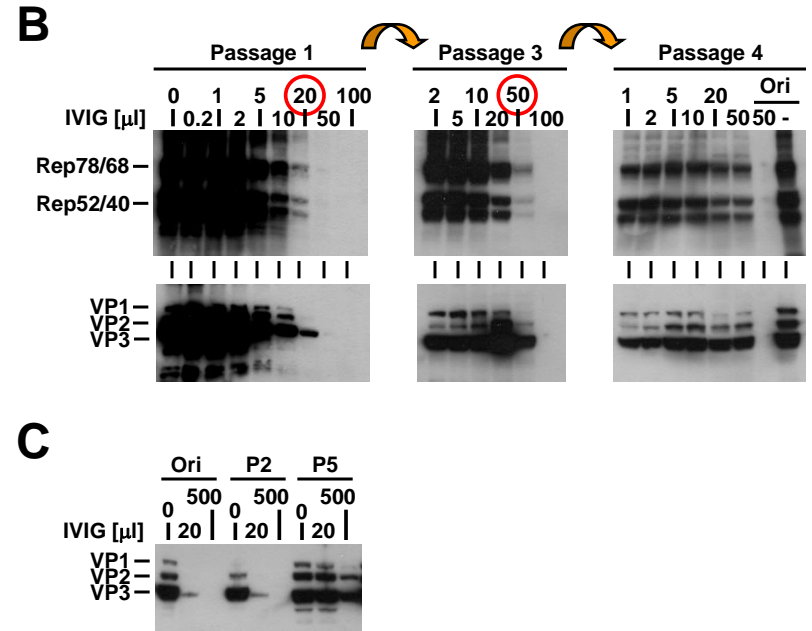
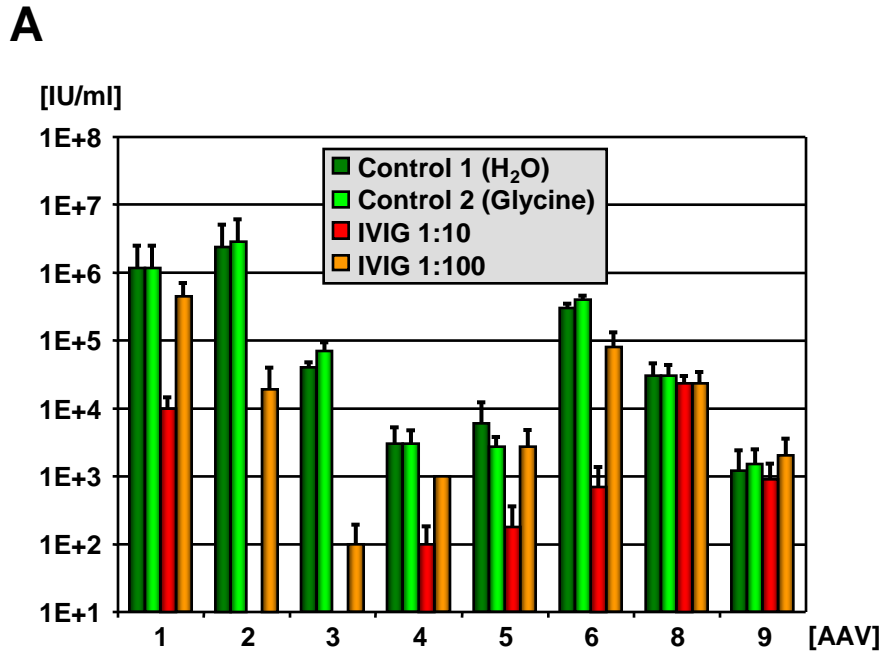
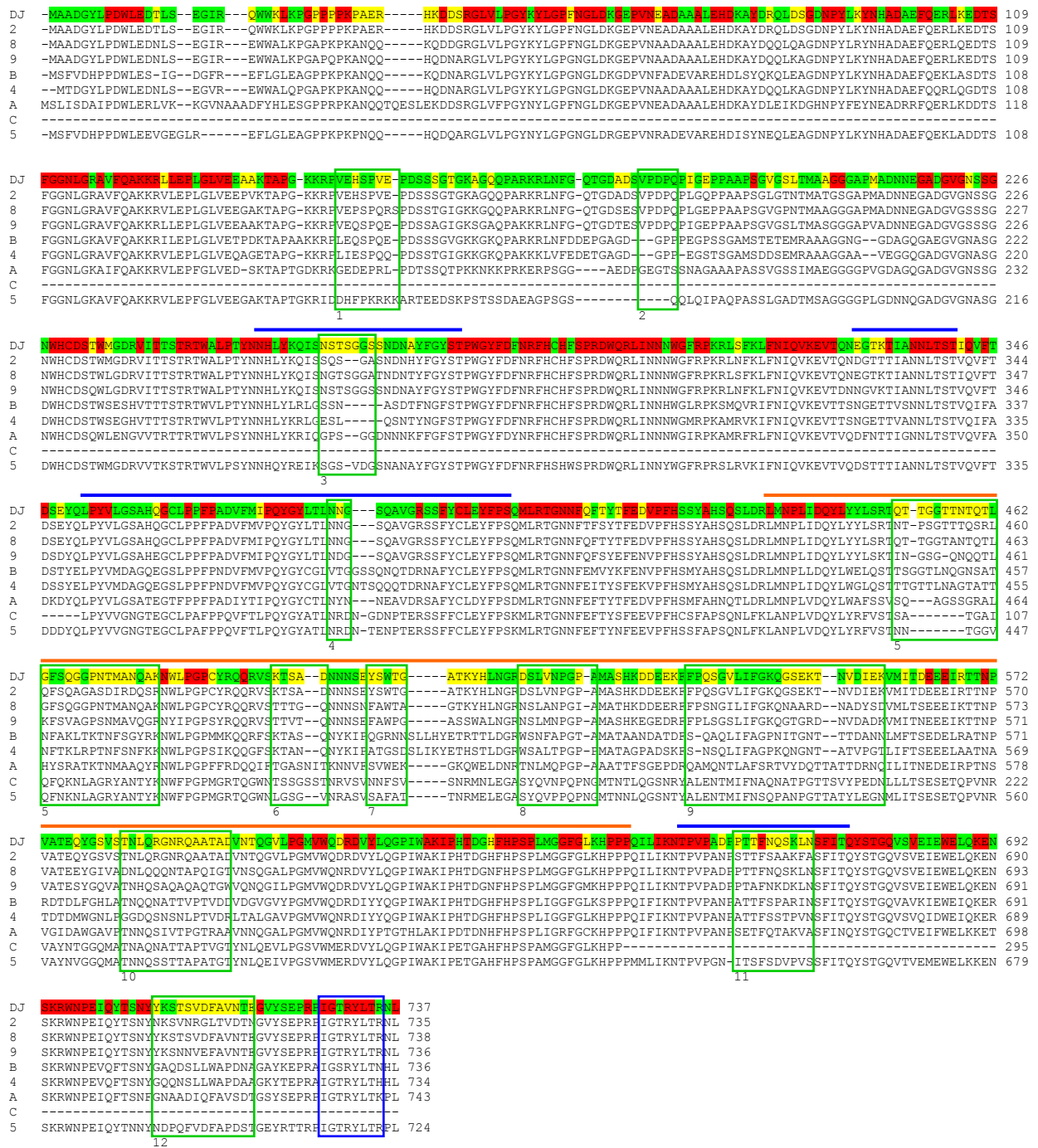


S1 (56%)	<p>TTCAACGCTCAAACGCCACCGCGGAACTACCTCGGTGTACCCAGAGGACAATCTACTGTGACCAGCGAGAGCGAGACTGGCCGTCACACGGGTGG CTTACAAACACGGCGGTGATGGCCACCAACGCCAGCAAGCCACCGCTCCCTTTGTCGGGACCTACCACTCCAGGAAGTCTTCTGGCAGGTT ATGGATGGAGAGGACGTGTACTCCAAAGACCCATCTGGGCCAAGATCCCGAGACAGGAAGCGGCACCTTCACCCCTCACCGCTGATTGGTGGGTTGGG CTGAACACCCCGCTCCCTCATCTCTCAAAAACACACTGTACTCGGGATCTCCCAACGGCTTCAACAAAGGACAGCTGAACCTTTTCATCACC AGTATTCTACTGGCAAGTCAAGCTGGAGTTCAGTGGAGCTGCAGAAAGGAAAACCAACAGCTGGAAACCCGGAGATCCAGTACACTCCCACTATT CAAGTCTAATAATGTTGAATTTGCTGTAACTCAAGGTGTATATAGTGAACCCCGCCCATTTGCCACCAGTACTGACTCGTAATCTGTAA</p>	100 200 300 400 500 594
S2 (66%)	<p>GGCCACAACCCAGAGTGCACCAAGCAGCGGCGAGACCGGCTGGGTTCAAACCAAGGAATACTCCGGGATGGTGGGAAAACAGAGACATTTAC CTACAGGGACCCATTTGGCAAAATTTCCGACACTGACAACTCTCCATCCGTCCCTTTATTGGCTGGTTGGCTGGAAGCATCCCTCCCGCAGA TTTTTCAATAAAACACCCCGTCCCTGGCAACCTTGGAAACGTTCCAGACGGCCAAATGGCTCTTTCATCAACAGTACTCGACCGGACAGTGGAC GGTCAAAATCTTTGGGAACCTCAAGAGGAAACCTGCAAGCTGGAAACCCGAAATCCAGTTCACCTTCCAACTTTGGCAACCGGCGGACATCCAGTT CCGCTCCGACACGGGATCTATTCCGAACCTGCTCCATCGGTACCGTTACCTTACCAATCTCTGTAA</p>	100 200 300 400 472
S3 (87%)	<p>TTTTGGCAAAACGAAGACTGGAAGAGACAACGTTGATCCGGACAAGTCAATGATAACCAACGAGAAGAAATAAAACCTACTAACCCGGTAGCAACGGAG TCCATTTGGACAAGTGGCCACAACCCAGAGTGCACCAAGCAGCGCCAGACCGGCTGGGTTCAAACCAAGGAATACTTCGGGTATGGTTTGGCAGG ACAGAGATGTGTACTTTCAGGGCCCATCTGGGCAAGATTTCCACACACGGAGCAACTTTTCAACCCCTCTCCCTCTATGGGTGATTCGAATTAACA ACCTCTCCACAGATTTCTCATCAAGAACCCCGGTACTCCGCAATCTTCGACACTCTCAGTGGCCAAAGTTTGGTCTCTCATCACAGTACTCC ACGGACAGGTCAAGCTGGAGATCGAGTGGAGCTGCAGAAGGAAAACAGCAACGCTGGAATCCGAAATTCAGTACACTTCCAATACAAACAGTCT TAAATGTAGACTTACTATGAGCTAATGCGGTATTCAGAGCTCGCCCATTTGCCACAGATACTGACTCGTAATCTGTAA</p>	100 200 300 400 500 586
S4 (52%)	<p>CAACGCTCAAACGCCACCGCGGAACTACCTCGGTGTCCAGAGGACAATCTACTGCTGACCAGCGAGAGGAGACTCAGCCCGTCAACCGGTTGG TCAACACCGGGCGGTGATGGCCACCAACGCCAGAACGCCACCGGCTCCACCGTCCGGACCTACAACCTCCAAGAGTCTTCTGGCAGCGTAT GGATGGAGAGGACGTGTACTCCAAAGACCCATCTGGGCCAAGATCCAGAGACGGGGCGACCTTCAACCCCTCACCGTGAATGGTGGGTTGGCT GAAACACCCCGCTCCCAAAATTTTATCAAGAACCCCGGTACTCGGAATCTCGCAACGACCTTCACTCTACTCCGGTAACTCTCATTAATCAG TACAGCTTGGCCAGGTCTGGTGCAGATTTGACTGGGATTCAGAAGGAGCGGTCGAAGATGGAAACCGAGTCCAGTTCAGCTCAGCTACCGATCGAC CACAGACTCGCTTCTGGGCTCCGCAACCGCGGAGCTCAAAAGAGCCAGGGCCATTGGACAGATACTGACTCGTAATCTGTAA</p>	100 200 300 400 500 592
S5 (68%)	<p>CGTTTTTTTCCAGCAACGGGATCTGATTTTTGGCAACAAATGCTGCCAGAGACAATCGGATTACAGCGATGTCTGCTCACCAGCGAGGAGAA TCAAAACCACTAACCTCTGGCTACAGAGAAATACGGTATCGTGGCAGATAAATGGCAGCAGAAAACACGGCTCCCAATTTGAACTGTCAACAGCA GGGGCCTTACCCGATGGTGTGGCAGAACCGGACGTGTACTCGAGGGTCCCATCTGGGCCAAGATTTCTCACAGGACGGCAACTTCACTCTCT CCGCTGATGGGAGGTTTGGAAATGAAACACCCCGCTCTAGATCTCTCAAACACACACTGTACTCTGGGATCTCCAGAGCAATACAGAGCAAT CGGTAACTCTCTACT CCAGTTCAGCTCAACTCAGGACACAGGAGTCTCTCTGGGCTCCCGCAACCGCGGAGCTCAAAAGAGCCAGGGCCATTGGATCCCGATACCT ACCCTGTAATCTGTAA</p>	100 200 300 400 500 600 615
S6 (72%)	<p>CCAGTAAACGGGATCTGATTTTTGGCAACAAATGCTGCCAGAGACAATCGGATTACAGCGATGTCTGCTCACCAGCGAGGAGAAATCAAACCA TAACCTCTGGCTACAGAGAAATACGGTATCGTGGCAGATAAATGGCAGCAGAAAACACGGCTCCCAATTTGAACTGTCAACAGCCAGGGGCGCTTA CCCGTATGGTCTGGCAGAACCGGACGTGTACTCGAGGGTCCCATCTGGGCCAAGATTTCTCACAGGACGGCAACTTCCACCCGCTCTCCGCTGATGG CGGGCTTTGGCTGAAACATCTCCGCTCAGATCTGATCAAGAACCGCTGTACTCGGATCTCCGACCCCTTCAACAGTCAAAAGTGAATCTGACT TTTTATCACCGAGTACCGCGGAGGAGTGTCACTGGGAACTTTGGGAACTCAAGAGCAATACAGAGCAATACAGCAAGCGGTGAAACCCGAGTCA TCCAATCTACTAATAATCAAGTGTGGACTTTGCTGTAAATACAGAAGCGTGTACTTGAACCCCGCCCATTTGGCACCCTTACCTCACCCGTAAT TGTA</p>	100 200 300 400 500 600 605
S7 (46%)	<p>GGATGACAACACGCTCCAAGCGAGCAACCGCTACCGCTGGAAAACCAATGATTTCAACGCTCAAACGCCACCGCGGAACTACTCGGTGTACCC AGGGAACAATCTACTGCTGACCGAGGAGCGAGACTCAGCCGCTCAACCGGTTGGCTTACAACCGCGGTCAGATGGCCCAACAGCAGAGCTCT CCACTGCCCCCGGACCGGCGCTACAACCTCAGAAATCTGTGCCCCGCGGCTGTGGATGGAGGGGACGTGTACTCCAAAGGACCATCTGGGCG AGATCCAGAGAGCGGGCGCACTTCAACCCCTCCCGGCTGGCGGATTTGGATCCAAAGACCCACCGCCATGATGCTCATGAGAAACCGGCTGT CCCGGAAATATACCCGCTCTCTCCGGCCAGAACTCAACGCTTCTATCAACCCAGTACAGCACCGGACAGGTGGCTGTCAAATAGAAATGGAAATCCAG AAGGAGCGGTCCAAGAGATGGAACCCAGAGTCCAGTTTCACTTCAACTACGGAGCACAGGACTCGCTTCTGGGCTCCCGACACCGCGGAGCTACA AAGAGCCAGGGCCATTGGATCCGATACCTCACCACCCCTCTAA</p>	100 200 300 400 500 600 647
S8 (82%)	<p>GGGCAATGCAAGCCACAAGGAGGTAAGAAAATTTTTCCTCAGACCGGGTCTCTCTTTGGGARGCAAGCTCGAGAAAACAAATTTGGACAA TGAARAAGTCAATGATTCAGACGAGAGAGAAATCAAGCAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACCAACCTCCAGAGAGCA AGA AAGCAGCTACCCAGATGTCACACAGCAAGCGTTCTCCAGGATGGTTCGGAGGACAGAGATGTACTCTCAGGGGCCATTTGGCAAAAG TCCACACACAGGATGGAACACTTTCACCCGCTCTCTCTCATTTGGCGGATTTGGACTGAAAGCCCGCTCCAGAAATTTATCAAAAACACTCTCTGAC CGCCAAATCCCGCAAGACCTTCTCTCCGGCCAGAACTCAACAGCTTCTATCAACCCAGTACAGCACCGGACAGGTGGCTGTCAAATAGAAATGGAAATCCAG AAGGAGCGGTCCAAGAGTGGAAACCCAGAGTCCAGTTTACTTCAACTACGGAGCACAGGAACTCTCTGTGTGGGCTCCCGATCGGCTGGGAAATACA CTGAGCTAGGGTATCGGTACCCGCTACCTCACCACCCCTGTAA</p>	100 200 300 400 500 600 647
S9 (73%)	<p>TCTGGACCTGCTATGGCCAGCCACAAGAAGGAGAGGACCGTTTCTTCTTGTCTGGATTTAAATTTTTGGCAACAGGAATCGGAAGAGACAA GTGGTTCGGACAAAGTCAATGAAACCAAGGAGAAATTTAAACTACTAACCCGATGCAACGGAGTCTCTATGGCAAGTGGCCCAACAAACCCAGCA GTGCCAAGCACAGGCGCAGCCGCTGGTTCAAAACCAAGGAATCTTCCCGTATGGTCTGGCAGAAACGGGACGTGTACTCGAGGCTCCATCTG GGCCAGATTTCTCACAGGACGGCAACTTCCACCCGCTCTCCGCTGATGGGAGGTTTGAATGAAACACCCGCTCTCAGATCTCATAAAACACA CTCTGACTCGCGATCTTCAACCGGCTTCAACAGGACAGTGAATCTTCTATCACCCAGTATCTACTGGCCAGTCAAGCTGGAGATCGAGTGGG AGCTGCAGAGGAAAACAGCAAGCGCTTGAACCCGAGATCCAGTACACTTCCAATATTACAAGTCTAATAATGTTGAATTTGCTGTAAATCTGAAG TGTATATAGTGAACCCCGCCCATTTGGCACAGATACCTGACTCGTAATCTGTAA</p>	100 200 300 400 500 600 655
S10 (60%)	<p>AACACCATGATCTTCAACGCTCAAACGCCATTCGCGGAACTACCTCGGTGTACCCAGAGGACAATCTACTGTGACCAGCGAGAGCGAGACTCAGCCG TCAACCCGGTGGCTTACAACACGGCGGTGATGGCCACCAACGCCAGCAAGCCACCGGCTCCACCGTCCGGACCTACAACCTCCAGGAAGTCT TCTGGCAGGCTATGGATGGAGAGGAGGAGCTGTACTTCCAAAGACCCATCTGGGCCAAGATCCAGAGACGGGGGCGACTTTCACCCCTCTCCGCGCATG GGTGGATTCGCAATTAACACCTCTCCACAGATTTCTATCAAGAACCCCGCTGACTTCCGAAATCTCCCAACCGGCTTCAACAAGGACAAGCTGA TTTTCATCCCGAGTATCTACTGGCCAAAGTCAAGCTGGAGTCAAGTGGGAGCTCAGAAAGAAAACAGCAACCGTGGAAACCGAAATTCAGTACA TTCACACTACAGCAAGTCTGTAATGTGGATTTCTGTGGAAGCTAATGGCGGTATTCAGAGCTTGGCCCATTTGCCACAGATACTGACTCGTAAT CTGTAA</p>	100 200 300 400 500 600 606
S11 (54%)	<p>ACGGCAGTCAACCTATTCGCGGAGAACTATCAATTTCAACAGCAGCGGCAACCGGACACCGGACCGTCAAGTCAAGGCAACATCTG ATCACCGGAGGACGAGCAGCGCGGTGAACCCGCTGGCGTACAAGCTCGCGGAGGATGGCCACCAACCAACAGAGCTCCACACTCCCGCCCGG CCGACCTTACACTCCAGAAATCTGGCCCGGACGGTGTGGATGGAGAGGCGGTACTCCCAAGTCAACTCTGGGCAAGGAAAGTCTCCACACAGG GGACATTTTCAACCTTCCGCTCATGGTGGATTCGCAATTAACAACCTCTCCAAAGATTTCTATTAAGAACACCGCTTCCGCAATCTCTG CCACCTTCACTGGGCAAGGTTTGGTCTTCACTACCCAGTACAGCACCGGACAGGTGGCTGTCAAATAGAAATGGAAATCCAGAAAGGAGCGGTCCA AGAGATGGAACCCAGAGTCCAGTTCAAGTCCAACTACGGAGCACAGGACTCGCTTCTTGGGCTCCCGCAACCGCGGAGCTCAAAAGAGCCAGGGC CATTTGATCCCGATACCTCACCACCCCTCTAA</p>	100 200 300 400 500 600 634
S12 (67%)	<p>GGACAGCAAGTTCAGCAACAGCCAGCTCATTTTTGGGGGCTAAACAGAACGGCAACCGGCCACCGTACCCGGGACTCTGATCTTCACTCTGAGGAG GAGCTGGGACAGCTCCGACCCAGTACGGCAGATGTGGGCAACCTCGGCTCGGCTCAGCAGCAACAGCAACCTGCCGCGGTGACAGACTGACAG CCTTGGGACCGGCTCCGAAATGGTCTGGCAAAACAGAGACATTTACTTGCAGAGCCCAATTTGGCCAAAATCTCTCACAGGAGCCACTTTCAC TTCTCCGCTGATGGGAGGTTTGAATGAAACACCGCTCTCAGATCTCTCAAAAACACACTGTACTTGGGATCTCTCCACCGGCTTCAACAG GACAAGCTGAATCTTTCACTACCCAGTATTTCACTGGCCAAAGTCAAGCTGGAGTTCAGTGGGAGCTGCAGAAGGAAAACAGCAACCGTGGAA AAATTCAGTCACTTCCAACTACAACAAGTCTGTAATGTGGATTTACTTGGGCACTAATGGCGGTATTCAGAGCTCGCCCATTTGCCACAGATA CTGACTCGTAATCTGTAA</p>	100 200 300 400 500 600 619

SUPPL. FIG. 2. Nucleotide sequences of 12 randomly chosen clones from the unselected library (3' ends). The clones, labels and colors are identical to Suppl. Fig. 1. The 3' ends of all clones are shown here. Only clone S8 contained the AAV-2-derived HBD (the triplets encoding the two crucial arginines are shown white on black).



SUPPL. FIG. 3. Selection of shuffled AAV capsids with human immunoglobulin (IVIG). (A) To analyse the neutralizing activity of the particular IVIG batch, recombinant *gfp*-expressing AAVs of the shown serotypes were incubated for 1 hour at 37°C with the shown agents, and then titered on 293 cells (all virus stocks were normalized to 2×10^9 particles per ml). IVIG had the strongest neutralizing effect on serotypes 2 and 3, followed by 6, 1, 4, and 5. AAV-8 or -9 were only inhibited ($\sim 10 \times$) with undiluted IVIG (not shown). (B) AAV library amplification on Huh-7 cells under IVIG pressure. For passage 1, 20 μ l of the library were incubated for 1 hour at 37°C with the shown amounts of IVIG, and then left on the cells overnight. The next day, the cells were washed and super-infected with helper Adenovirus. The cells were lysed three days later, and 20 μ l from the supernatant showing minimal AAV protein expression (circled in red) were processed as before. Shown is expression of AAV replication (Rep, top) and capsid (VP) proteins (bottom). Ori, original library. The blot in (C) documents the increasing resistance of the amplified particles to high IVIG doses over the various passages.

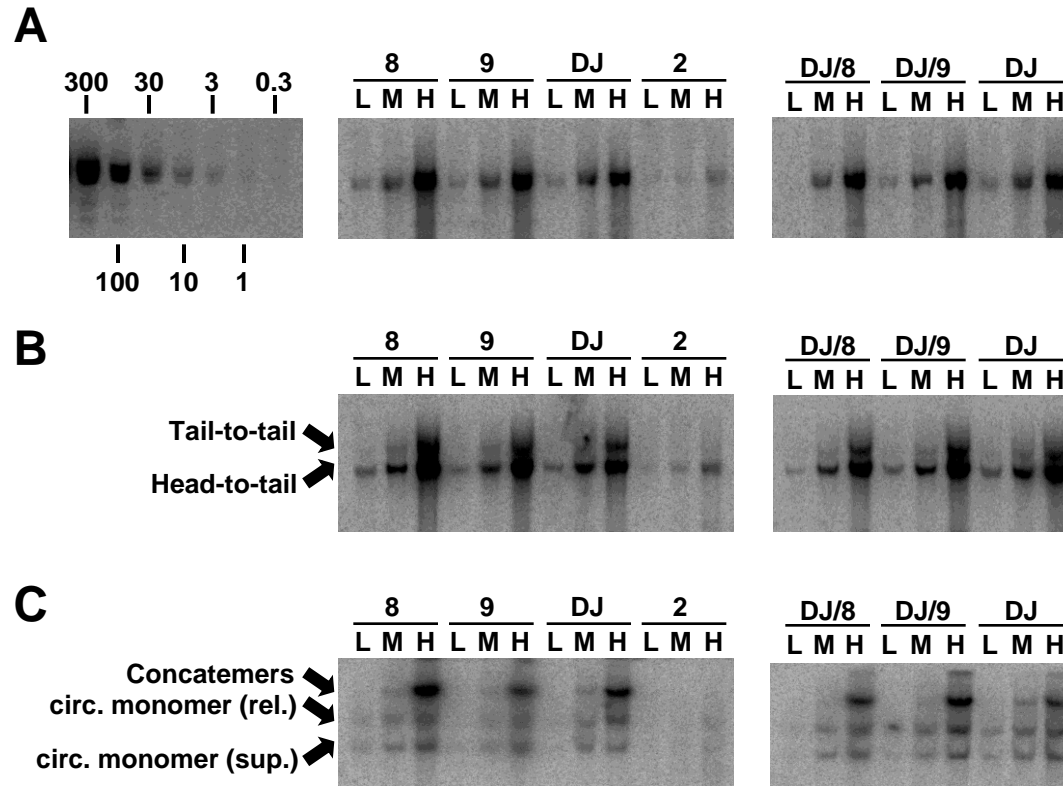


Total DJ-VP1 protein (737 aa):

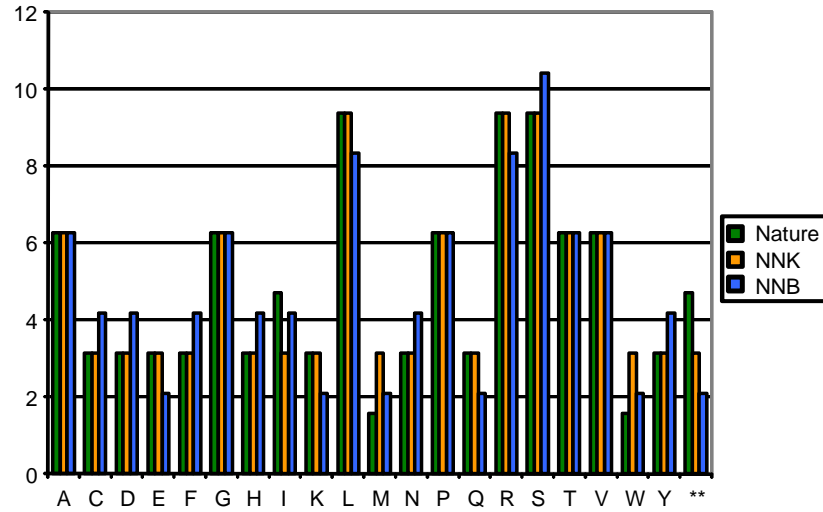
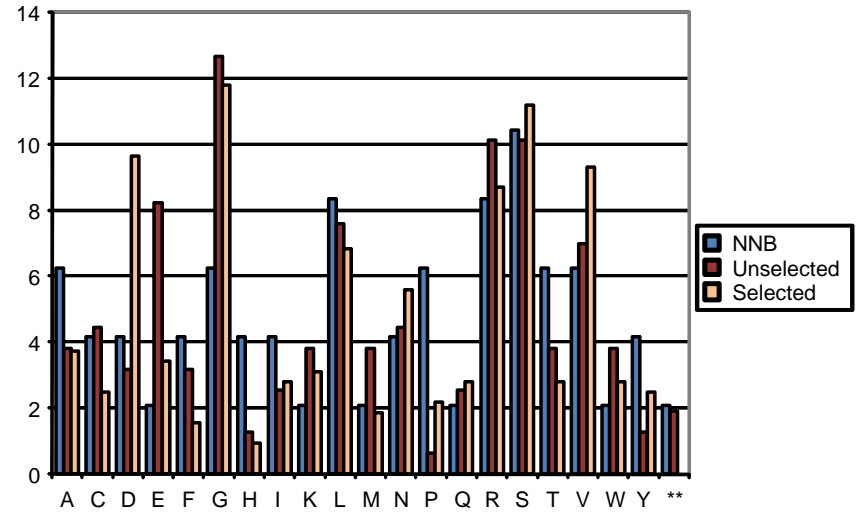
Loop IV (212 aa):

Complete identity :	226 aa / ~30.7%	↔	Complete identity :	39 aa / ~18.4%	↓
Identical in DJ-2-8-9 :	307 aa / ~41.6%	↔	Identical in DJ-2-8-9 :	91 aa / ~42.9%	
Either in DJ / 2 / 8 / 9 :	204 aa / ~27.7%	↔	Either in DJ / 2 / 8 / 9 :	82 aa / ~38.7%	↑

SUPPL. FIG. 4. Protein sequence alignment of AAV-DJ and the 8 parental viruses. The 9 shown protein sequences (full-length VP1 proteins) were aligned using the ClustalW tool. Sequence identities between individual proteins are highlighted, using AAV-DJ as a reference. Individual amino acids are color-coded: red, identical between all 9 AAVs; green, identical between AAV-DJ and serotypes 2, 8 and 9; yellow, not conserved between serotypes 2, 8 and 9. The bars highlight the amino acids constituting the five exposed capsid loops (the largest and most divergent loop IV is highlighted by the orange bar). Green boxes depict 12 hypervariable regions in AAV capsid genes. Note the good correlation with amino acids highlighted in yellow, indicative of a high degree of evolution in these regions. The blue box indicates the B1 antibody epitope (fully conserved in AAV-DJ, -2, -5, -8, -9). Shown at the bottom are comparisons of amino acid identities between the full-length VP1 protein and the major loop IV. Remarkable is the decrease in complete sequence identity in this loop, and the concurrent increase in diversity between AAV-DJ and the other parental serotypes. This suggests the strongest selection pressure was exerted on this loop, resulting in the highest degree of evolution.



SUPPL. FIG. 5. Analyses of liver vector DNA. Mice were injected as described in Fig. 7, with hFIX-expressing AAV-2, -8, -9 or -DJ at low (L, 5×10^{10}), medium (M, 2×10^{11}), or high (H, 1×10^{12}) particle doses. Six weeks later, total liver DNA was extracted and digested with Bam HI and Xho I (A), to determine vector copy numbers, or either Bam HI (single cutter, B) or Nco I (non-cutter, C), to analyze vector DNA forms. There were no differences between vector DNAs delivered by AAV-8, -9, or -DJ (including the HBD mutants) at any dose. All vectors mainly persisted as circular monomers (relaxed or supercoiled), or concatemers at higher doses, in agreement with previous reports for AAV-8 and -9.

A**B**

SUPPL. FIG. 6. Amino acid frequencies in AAV peptide display libraries before and after selection. (A) Shown are natural frequencies of amino acids based on the genetic code ("Nature") versus theoretical frequencies in an NNK or NNB (this paper) library. Note that the frequency of unwanted stop codons ("**") is lowest with the NNB design. (B) Comparison of overall amino acid frequencies in the unselected AAV-DJ-based library versus that in 46 peptides cloned after *in vivo* biopanning in mouse lungs. Note that the frequency of stop codons ("**") is even slightly below the theoretical prediction in the unselected library. Importantly, stop codons were no longer detected after *in vivo* selection, arguing for a 100% coupling of viral genomes and capsids.