SUMMARY

[0005a] In one aspect, provided here is an AAV capsid comprising a polypeptide sequence of:

(a) L-A-L-G-X3-X1-X1-X4 (SEQ ID NO: 94);

(b) L-K-L-G-X3-X1-X1-X4 (SEQ ID NO: 95); or

(c) V-K-L-G-X3-X1-X1-X4 (SEQ ID NO: 96),

wherein:

X1 is Alanine (A), Asparagine (N), Glutamine (Q), Serine (S), Threonine (T), Glutamic Acid (E), Aspartic Acid (D), Lysine (K), Arginine (R), or Histidine (H);

X3 is E, S, or Q; and

X4 is K, R, E, or A, and

wherein the polypeptide sequence is in a VP domain of the AAV capsid.

[0005b] In another aspect, provided herein is a vector comprising a polynucleotide sequence encoding the AAV capsid as described herein.

[0006] Provided herein is a modified adeno-associated virus (AAV) capsid that comprises an exogenous polypeptide sequence in a VP domain of the AAV capsid as compared to an otherwise comparable unmodified AAV capsid. In an aspect, the exogenous polypeptide sequence comprises a sequence of formula 1: X0-X1-X2-X1-X3-X1-X1-X4-X5 (SEQ ID NO: 93), wherein X0 is Valine (V), Isoleucine (I), Leucine (L), Phenylalanine (F), Tryptophan (W), Tyrosine (Y) or Methionine (M), wherein X1 is Alanine (A), Asparagine (N), Glutamine (Q), Serine (S), Threonine (T), Glutamic Acid (E), Aspartic Acid (D), Lysine (K), Arginine (R), or Histidine (H), wherein X2 is V, I, L, or M, wherein X3 is E, S, or Q, and wherein X4 is K, R, E, or A, optionally wherein formula 1 further comprises X5 being Proline (P) or R. In an embodiment, formula 1 comprises the X_5, and wherein the X_5 is the Proline (P) or R. In an embodiment, the capsid is of serotype AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV13, and any combination thereof. In an embodiment, the capsid comprises AAV2. In an embodiment, the capsid comprises at least two serotypes. In an embodiment, the at least two serotypes are selected from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV8, AAV9, AAV11, AAV12, and AAV13. In an embodiment, formula 1 comprises: L-A-L-G- X3-X1-X1-X4 (SEQ ID NO: 94), L-K-L-G- X3-X1-X1-X4 (SEQ ID NO: 95), or V-K-L-G-X3-X1-X1-X4 (SEQ ID NO: 96). In an embodiment, formula 1 comprises the V-K-L-G-X3-X1-X1-X4 (SEQ ID NO: 96), and wherein the V-K-L-G-X3-X1-X1-X4 (SEQ ID NO: 96) comprises: V-K-L-G-X3-X1-T-X4 (SEQ ID NO: 97) or V-K-L-G-X3-X1-X1-K (SEQ ID NO: 98). In an embodiment, formula 1 comprises the L-A-L-G- X3-X1-X1-X4 (SEQ ID NO: 94), and wherein the L-A-L-G- X3-X1-X1-X4 (SEQ ID NO: 94) comprises: L-A-L-G- X3-X1-T-X4 (SEQ ID NO: 99) or L-A-L-G-X3-X1-S-X4 (SEQ ID NO: 100). In an

embodiment the L-A-L-G- X3-X1-T-X4 (SEQ ID NO: 99) comprises: a) L-A-L-G- X3-X1-T-R (SEQ ID NO: 101); b) L-A-L-G- X3-X1-T-K (SEQ ID NO: 102); c) L-A-L-G- X3-X1-T-E (SEQ ID NO: 103); or d) L-A-L-G- X3-X1-T-A (SEQ ID NO: 104). In an embodiment, the L-A-L-G- X3-X1-S-X4 (SEQ ID NO: 100) comprises L-A-L-G- X3-X1-S-K (SEQ ID NO: 105). In an embodiment, formula 1 comprises the L-K-L-G- X3-X1-X1-X4 (SEQ ID NO: 95), and wherein the L-K-L-G- X3-X1-X1-X4 (SEQ ID NO: 95) comprises L-K-L-G- X3-X1-T-X4 (SEQ ID NO: 106). In an embodiment, the L-K-L-G- X3-X1-T-X4 (SEQ ID NO: 106) comprises L-K-L-G- X3-X1-T-K (SEQ ID NO: 107). In an embodiment, formula 1 comprises a polypeptide sequence having at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% identity with a sequence of Table 2. In an embodiment, formula 1 comprises a polypeptide sequence of Table 2. In an embodiment, the VP domain of the AAV capsid is VP1. In an embodiment, the VP domain of the AAV capsid is VP2. In an embodiment, the VP domain of the AAV capsid is VP3. In an embodiment, the AAV