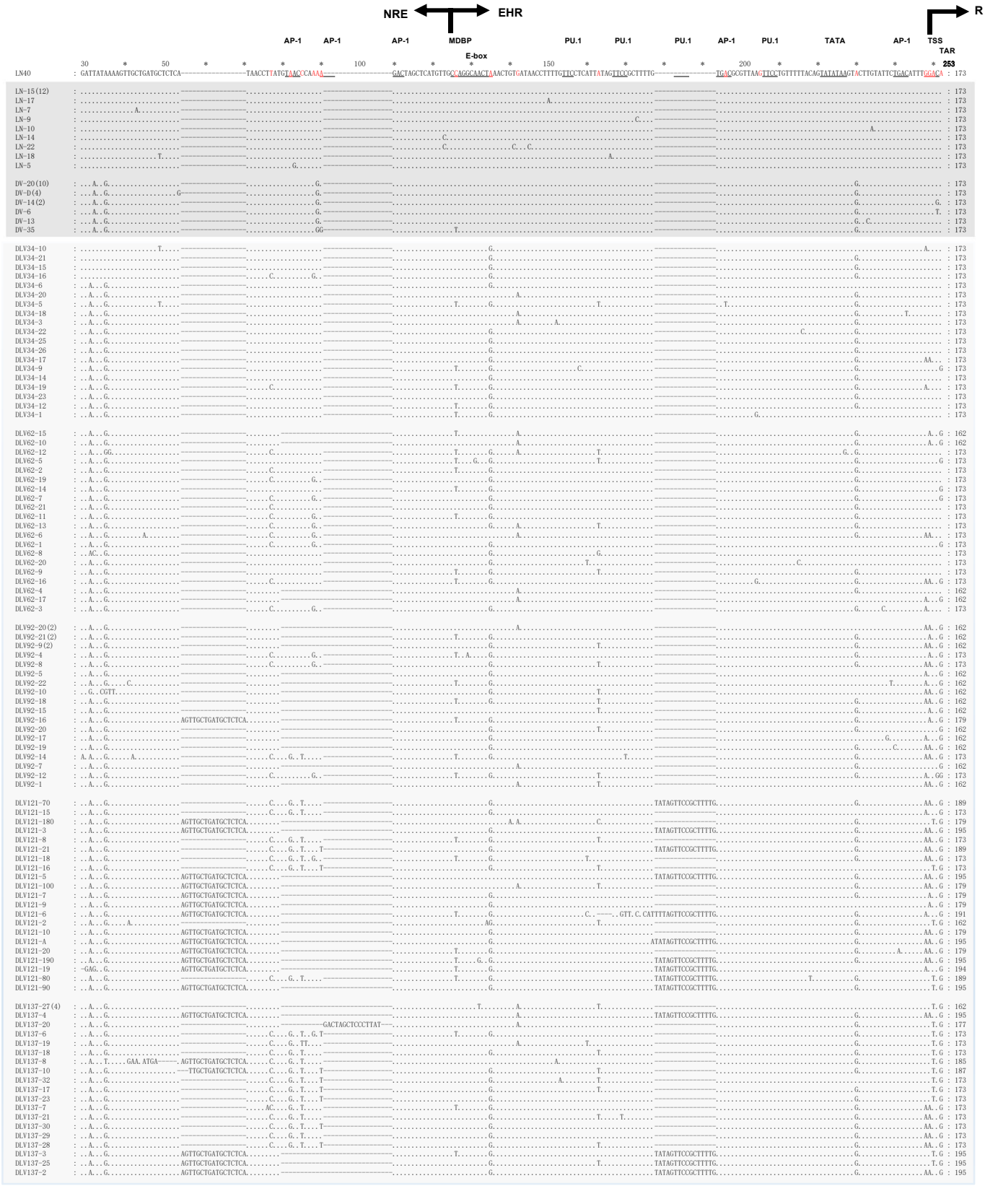

NRE
EHR
R
AP-1
AP-1
AP-1
MDBP
PU.1
PU.1
PU.1
AP-1
PU.1
TATA
AP-1
TSS
TAR
E-box
LN40
DLV34-10
DLV62-15
DLV92-20(2)
DLV121-70
DLV137-27(4)

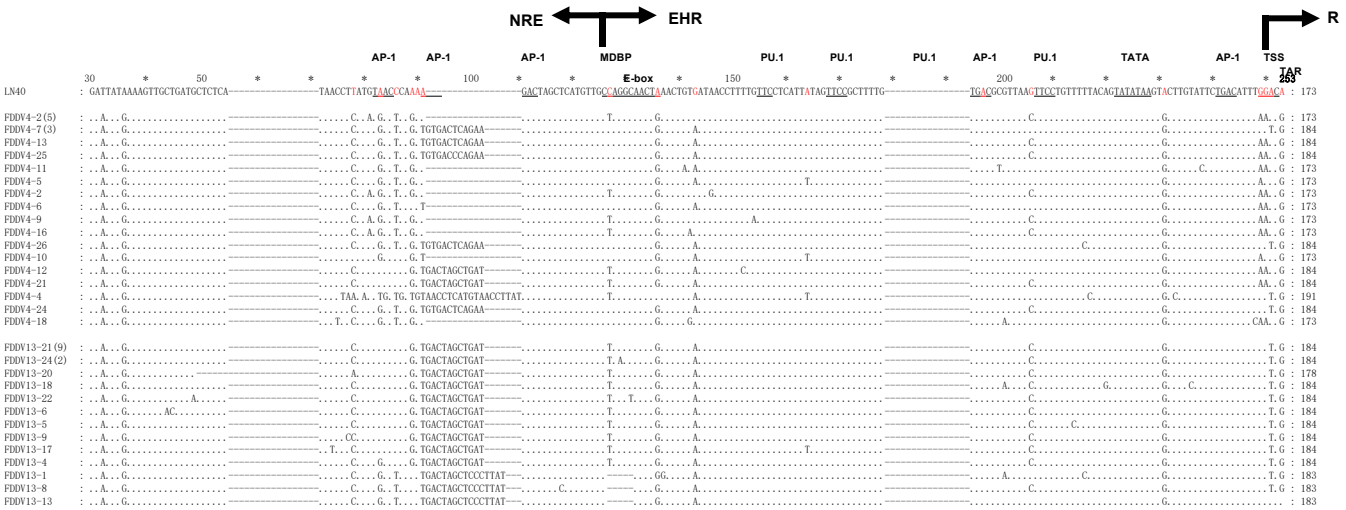
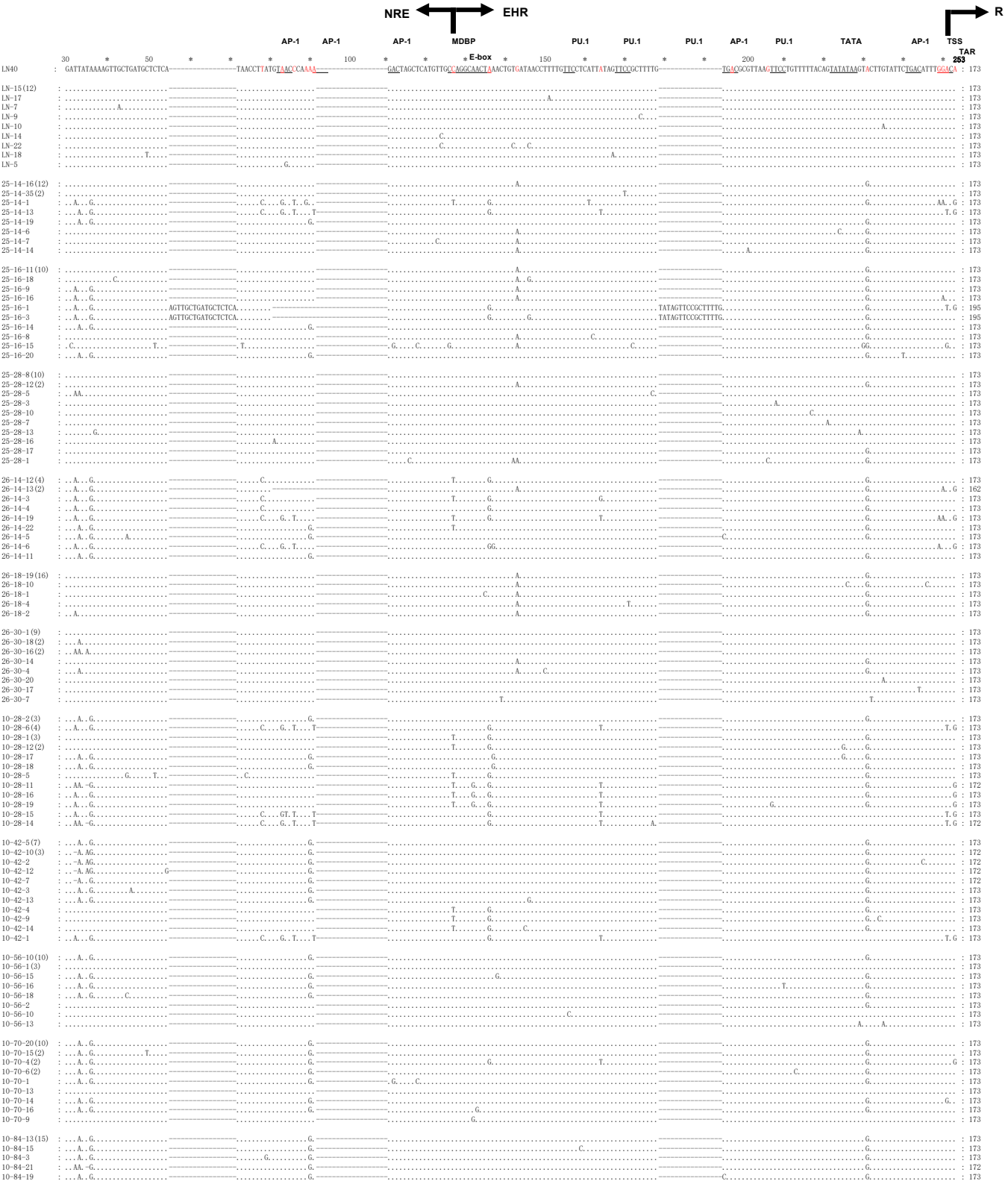


FIG. S1. Comparison of the LTR sequence of various EIAV strains during the attenuation process. All LTR sequences of the samples were aligned with the LN40 consensus sequence, which is the reference sequence. Nucleotides that are the same as those in the reference sequence are represented by dots, and nucleotides differing from those in the reference sequence are represented by the corresponding letter symbols; the dashes indicate deletion mutations. The numbers in brackets on the left side of the figure indicate the number of identical clones in each sample; the red letters in the top LN40 consensus sequence represent the main mutation hotspots of the LTR, and the underlined regions indicate the locations of the various transcription factor binding sites and the TSS.


NRE
EHR
R
AP-1
AP-1
AP-1
MDBP
E-box
PU.1
PU.1
PU.1
AP-1
PU.1
TATA
AP-1
TSS
TAR
LN40

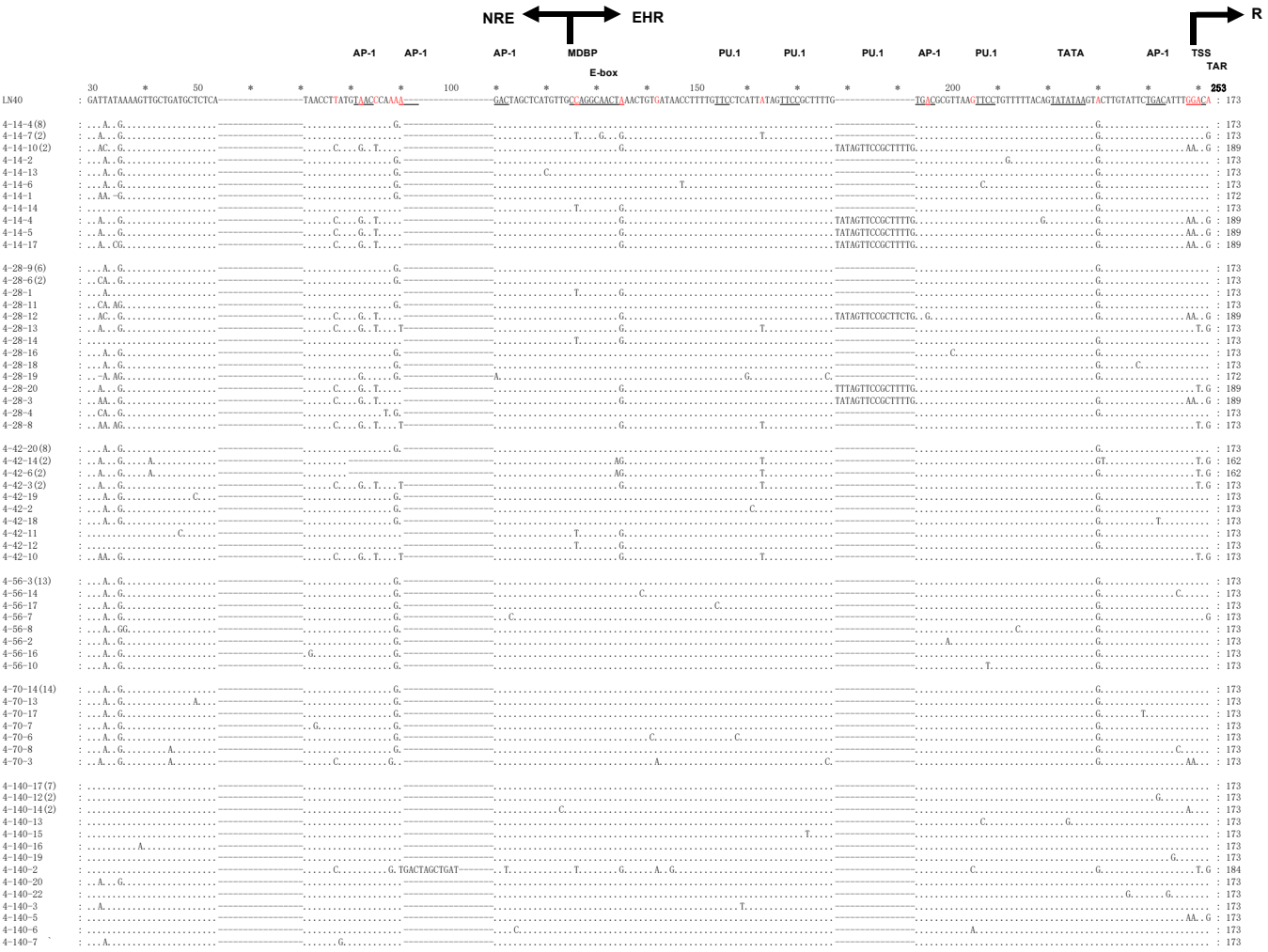
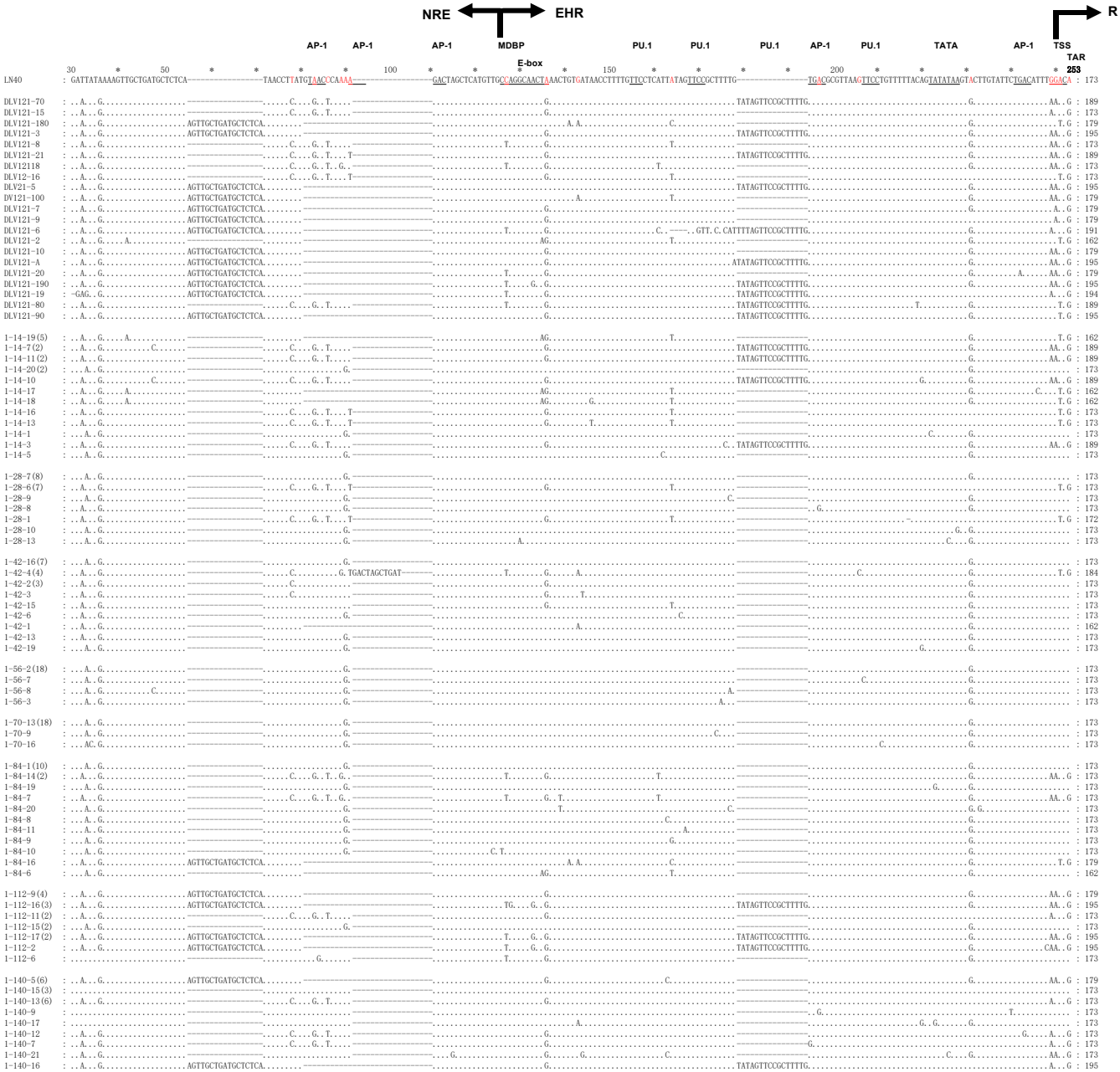


FIG. S2. Sequence comparison of LTRs from LN40-infected horses. All LTR sequences of the samples were aligned with the LN40 consensus sequence, which is the reference sequence. Nucleotides that are the same as those in the reference sequence are represented by dots, and nucleotides differing from those in the reference sequence are represented by the corresponding letter symbols; the dashes indicate deletion mutations. The numbers in brackets on the left side of the figure indicate the number of identical clones in each sample; the red letters in the top LN40 consensus sequence represent the main mutation hotspots of the LTR, and the underlined regions indicate the locations of the various transcription factor binding sites and the TSS.

FIG S3, to be continued-1



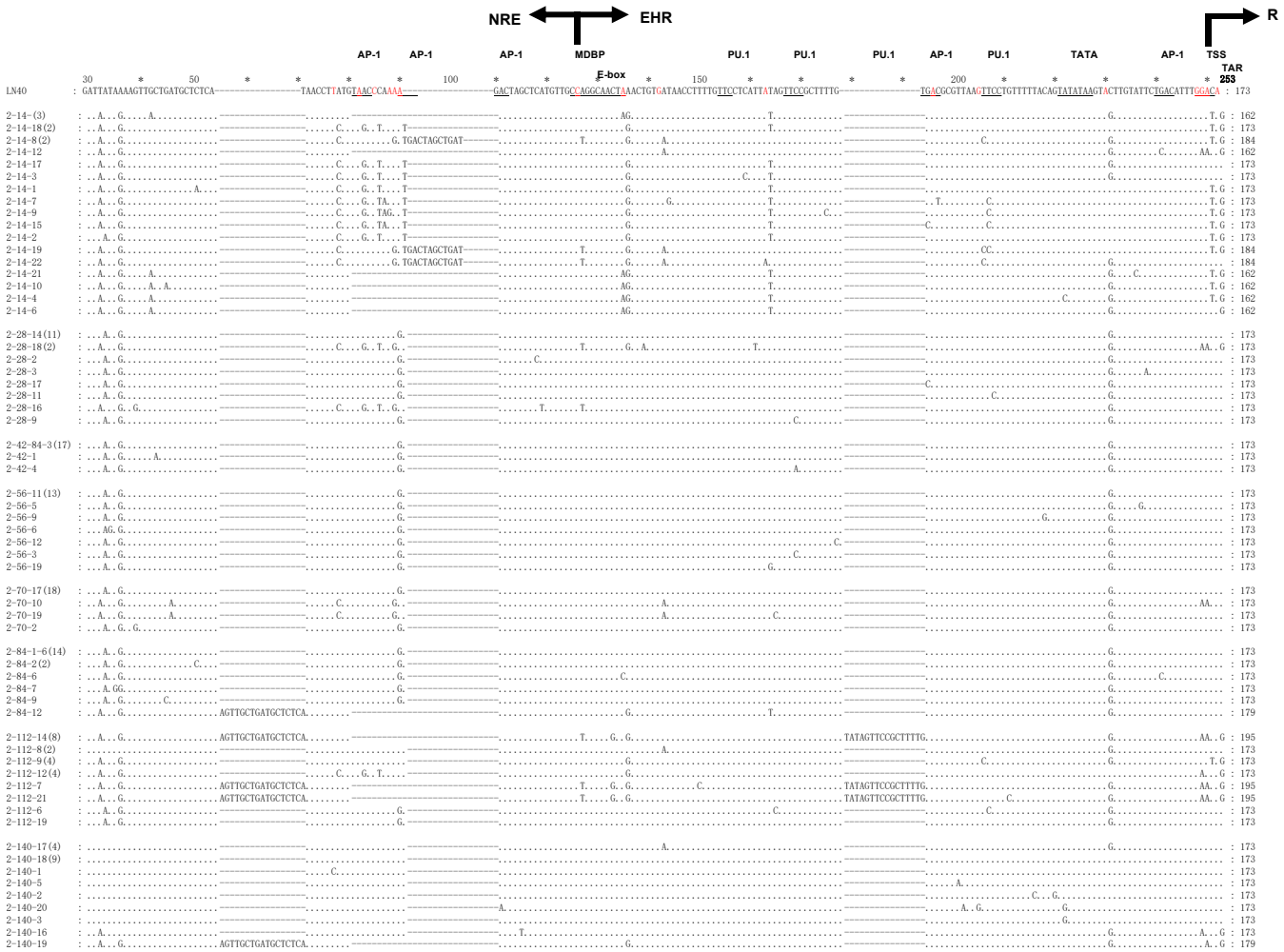


FIG. S3. Sequence comparison of LTRs from DLV121-infected horses. All LTR sequences of the samples were aligned with the LN40 consensus sequence, which is the reference sequence. Nucleotides that are the same as those in the reference sequence are represented by dots, and nucleotides differing from those in the reference sequence are represented by the corresponding letter symbols; the dashes indicate deletion mutations. The numbers in brackets on the left side of the figure indicate the number of identical clones in each sample; the red letters in the top LN40 consensus sequence represent the main mutation hotspots of the LTR, and the underlined regions indicate the locations of the various transcription factor binding sites and the TSS.

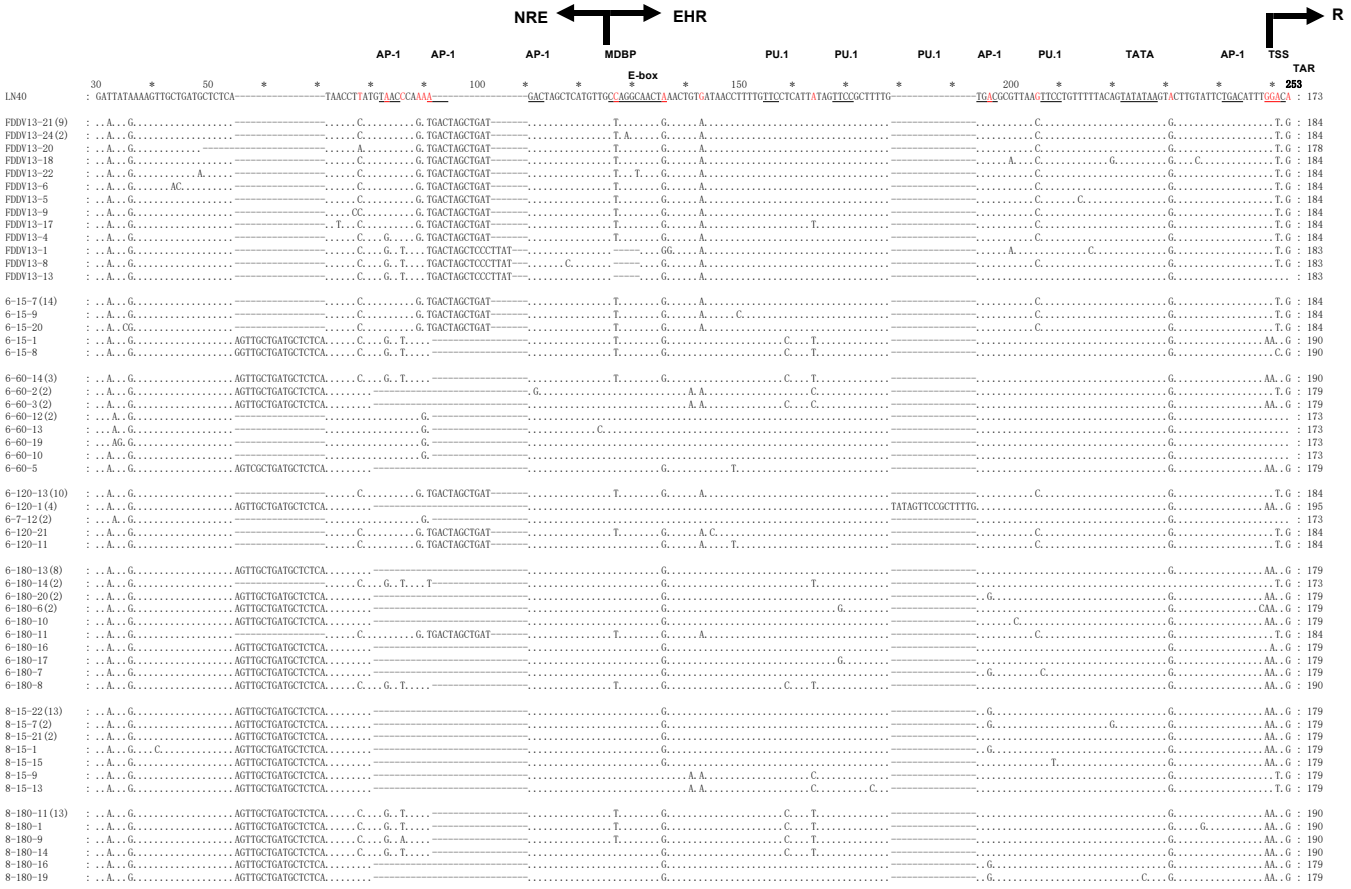


FIG. S4. Sequence comparison of LTRs from FDDV13-infected horses. All LTR sequences of the samples were aligned with the LN40 consensus sequence, which is the reference sequence. Nucleotides that are the same as those in the reference sequence are represented by dots, and nucleotides differing from those in the reference sequence are represented by the corresponding letter symbols; the dashes indicate deletion mutations. The numbers in brackets on the left side of the figure indicate the number of identical clones in each sample; the red letters in the top LN40 consensus sequence represent the main mutation hotspots of the LTR, and the underlined regions indicate the locations of the various transcription factor binding sites and the TSS.