Supplementary Figure 4. Sequence logos of the DHHRN, HNH, and His-Cys box endonuclease families

Sequences near the active site of an endonuclease family are used to generate sequence logos. At each position of the sequence logo, the total height of a stack of letters shows the information content in bits that is calculated from a profile hidden Markov model, and the height of a letter relative to the total height of letters at a position represents the letter’s frequency. The red lines indicate gaps in the multiple sequence alignment. Above each sequence logo, the corresponding residue numbers of Endo VII, I-HmuI, and I-PpoI are shown, which are representatives of the DHHRN, HNH, and His-Cys box families, respectively. Black boxes show the corresponding catalytic residues used by Endo VII, I-HmuI, and I-PpoI.