

Electronic Supplementary Material

The Variability of Amino Acid Sequences in Hepatitis B Virus

Jianhao Cao^{1,✉}, Shuhong Luo^{1, ✉}, Yuanyan Xiong^{2,✉}

1. Institute of Antibody Engineering, School of Laboratory Medicine and Biotechnology, Southern Medical University, Guangzhou 510515, China

2. School of Life sciences, Sun Yat-sen university, Guangzhou 510275, China

Supporting information to DOI: 10.1007/s12250-018-0070-x

Fig S1

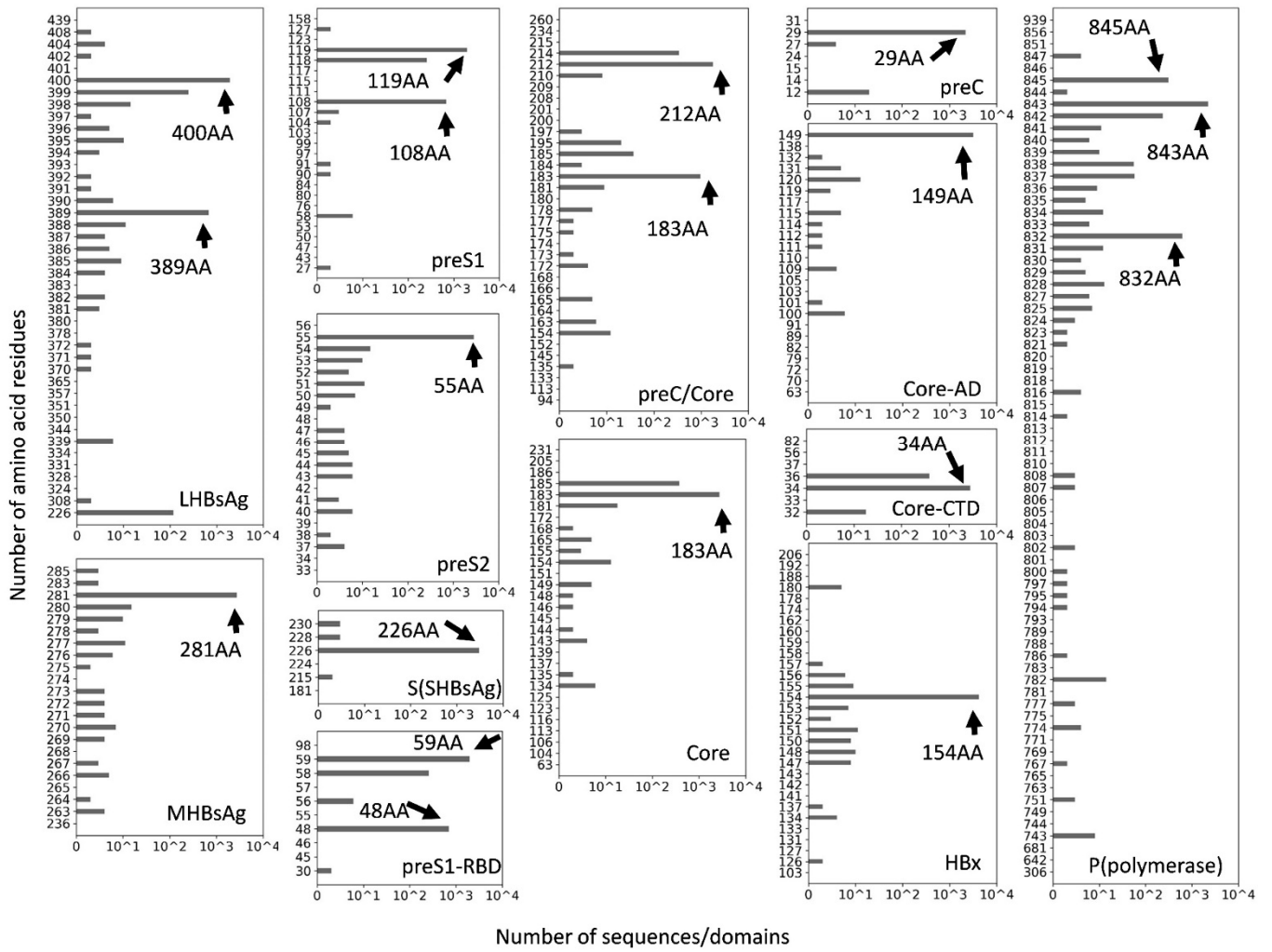


Fig. S1 The detail of adopted length by different sequences/domains. The x-axis is the number of sequences/domains. The y-axis indicates different lengths of sequences/domains. Predominant amino acid lengths are indicated by black arrows.

Fig S2

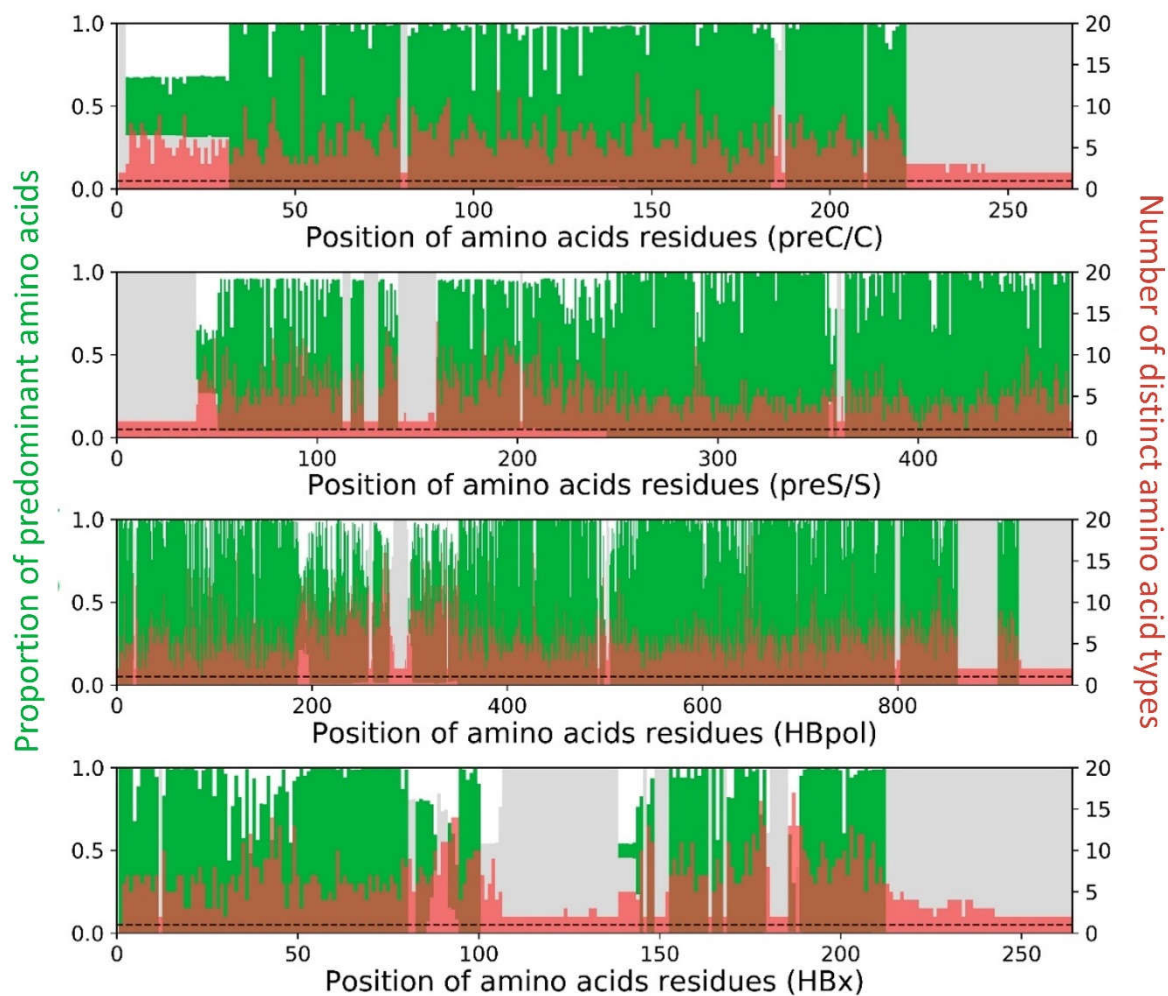
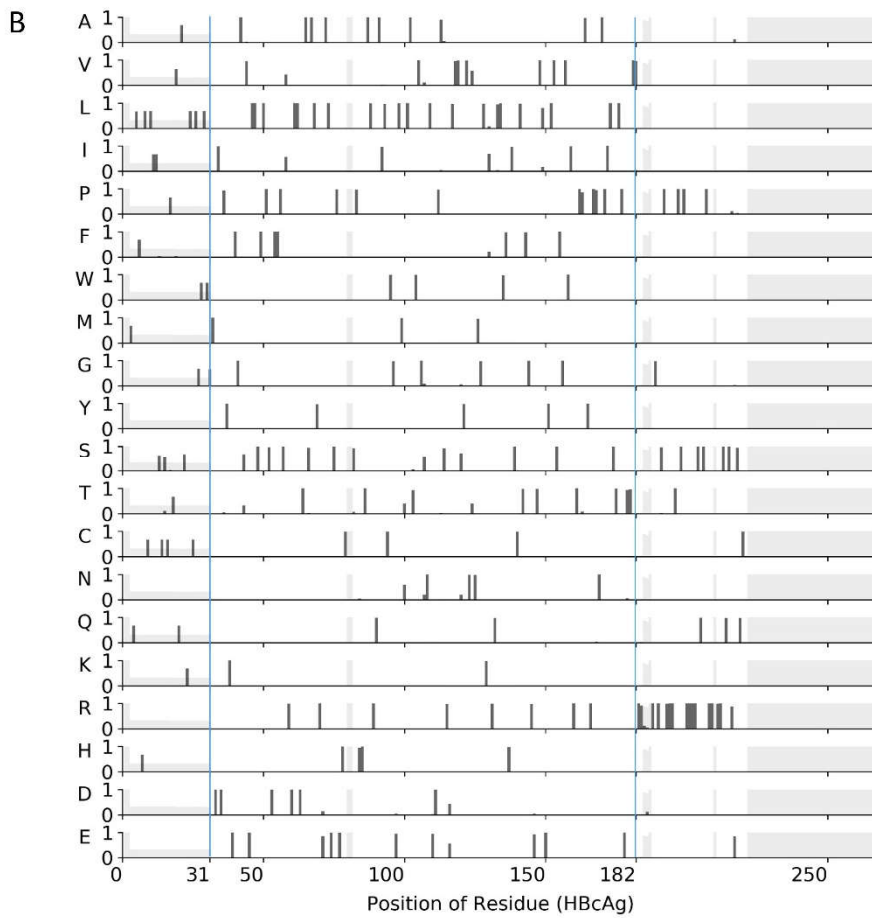
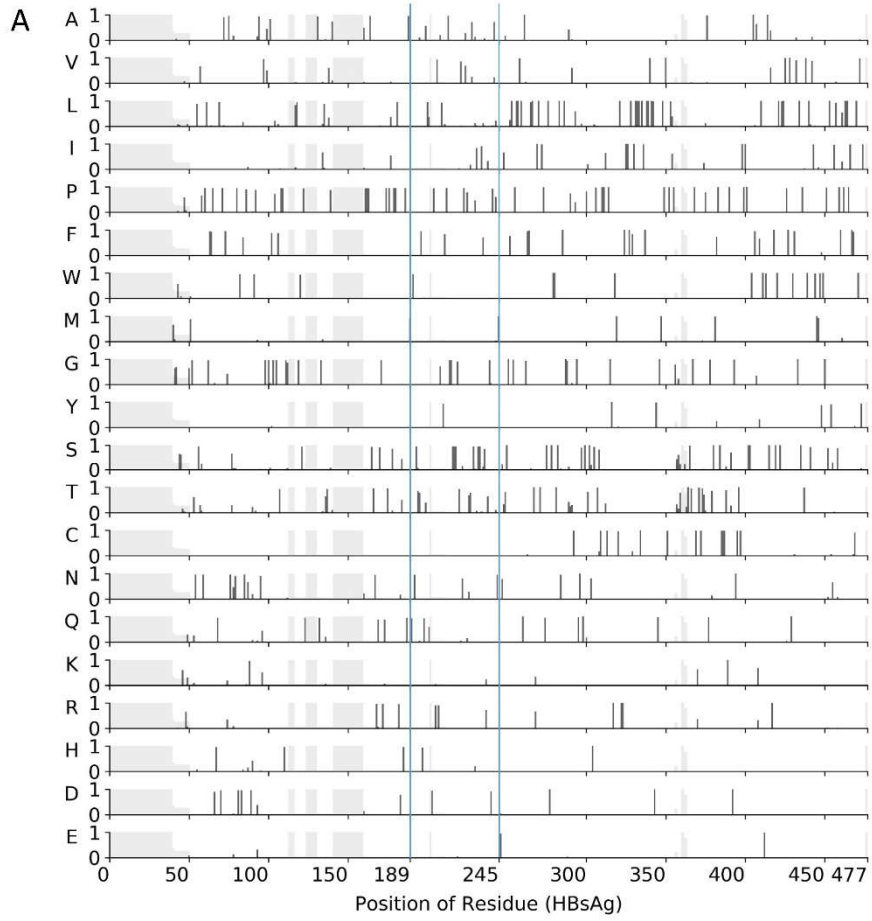


Fig. S2 The profile of predominant amino acids in 4 HBV ORFs. The green bar indicates the proportion of sequences that contain predominant amino acid residue in corresponding sites. The red bar indicates the number of different types of amino acid residues at each position. The grey bar indicates the proportion of sequences containing gaps after alignment.

Fig S3



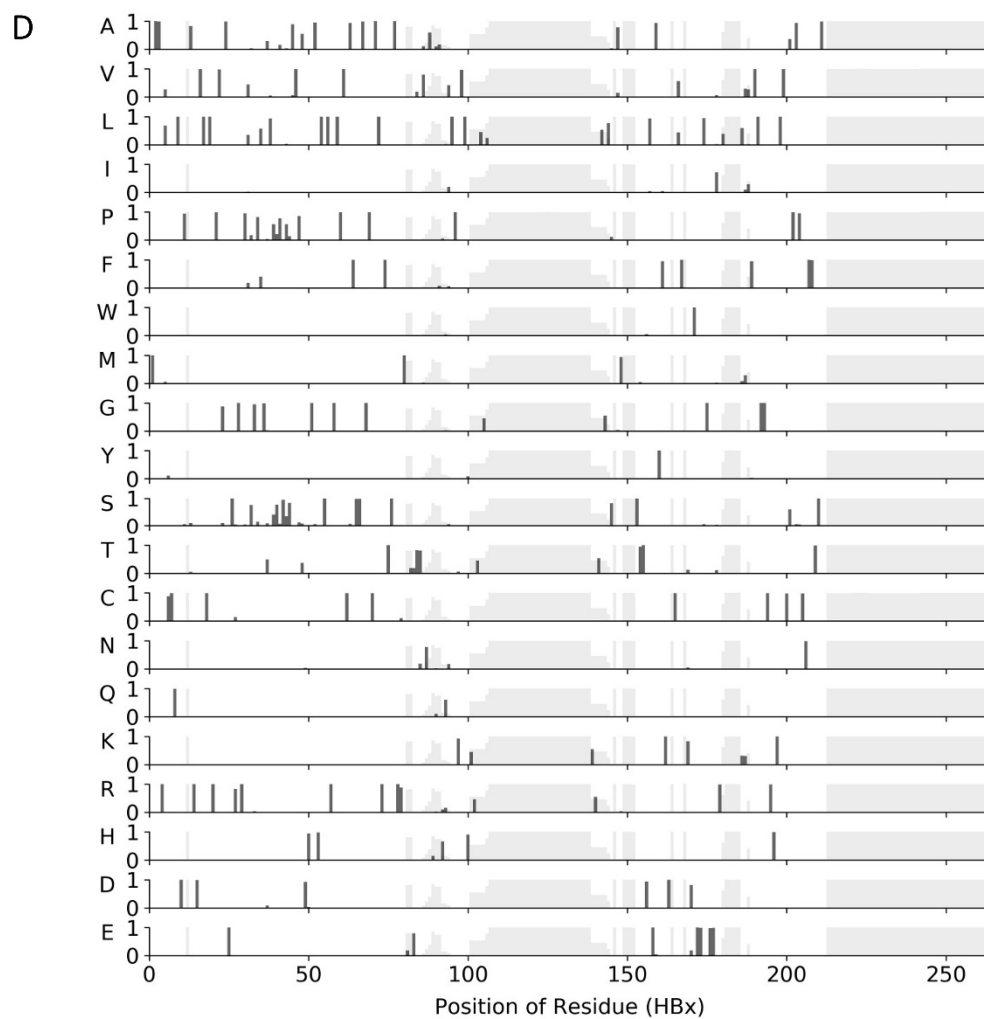
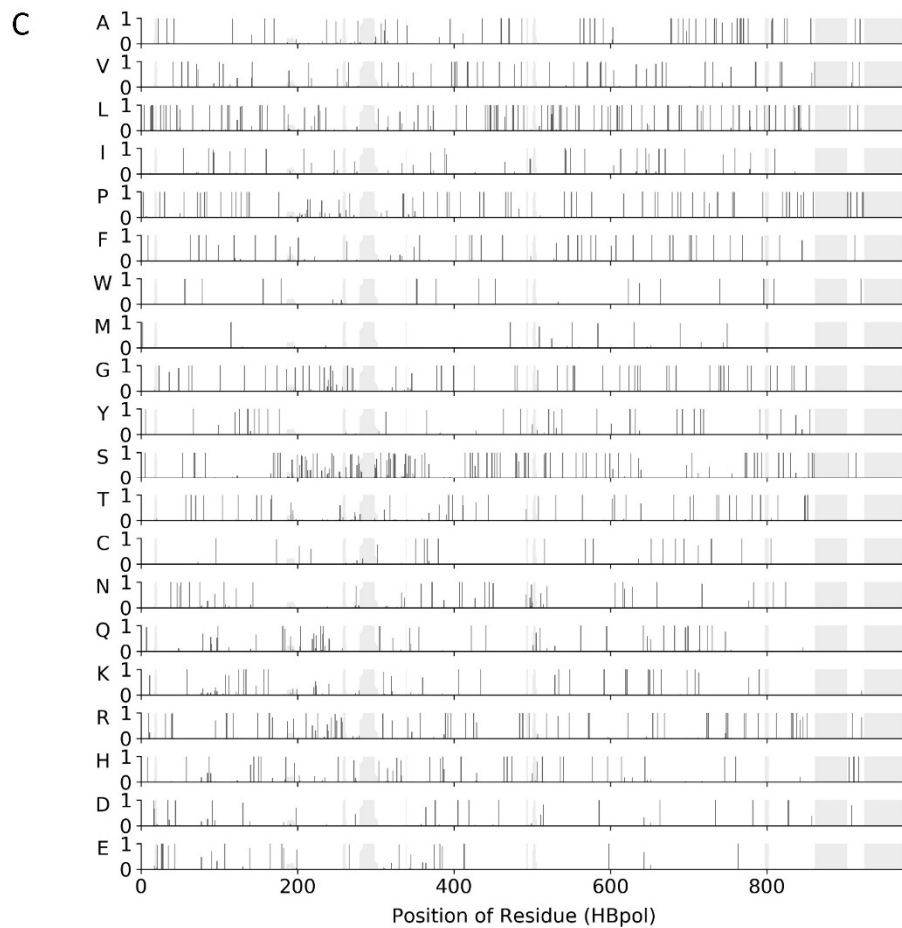


Fig. S3 The frequencies of 20 amino acids in different sites of (A) ORF S (HBsAg), (B) ORF C (preC/HBcAg), (C) ORF P (HBpol) and (D) ORF X (HBx). Positions of different domains are indicated by blue lines: with positions 1–188 indicating the preS1 domain, 189–244 indicating the preS2 domain, and positions 245–477 indicating the S domain in ORF S and with positions 1–30 indicating the preC domain, 31–181 indicating the core AD, and positions 182–268 indicating the core CTD in ORF C. The black bars indicate the proportions of the amino acid used in corresponding sites. The gray bars indicate the proportions of gaps after sequence alignment.