

Guangdi Li<sup>1,2</sup>, Samad Amini-Bavil-Olyaei<sup>3</sup>, Thomas Mina<sup>1</sup>, Philippe Lemey<sup>1</sup>, Piet Maes<sup>1</sup>, Marc Van Ranst<sup>1</sup>, Mahmoudreza Pourkarim<sup>1</sup>

1. Department of Microbiology and Immunology, Laboratory of Clinical and Epidemiological Virology, Rega Institute for Medical Research, KU Leuven, 3000 Leuven, Belgium

2. Metabolic Syndrome Research Center, the Second Xiangya Hospital, Central South University, Changsha, Hunan, China

3. Department of Molecular Microbiology and Immunology, Keck School of Medicine, University of Southern California, Harlyne J. Norris Cancer Research Tower, Los Angeles, CA 90033, USA

## Introduction

1. A holistic picture of HBV whole genome diversity remains unclear.
2. This study aims to reveal the first evaluation of HBV genome heterogeneity from the perspective of large-scale populations around the world.

## Results

1. Sequence numbers: genotype A: 794, genotype B: 1086, genotype C: 2036, genotype D: 1179, genotype E: 330, genotype F: 71, genotype G: 36, genotype H: 34.
2. The lowest genome-wide genetic diversity within HBV genotypes (mean intra-genotypic distance) was 2.8%.
3. Genetic diversity between genotypes (mean inter-genotypic distance) was dramatically jumped to 11.8%.
4. Diversity patterns at protein level across HBV genotypes were consistently similar comprising Polymerase (3.3%), PreS1 (2.7%), PreS2 (5.16%), S (2.1%), PreCore (2.91%), Core (3.05%) and X (4.72%).
5. 3 out of 11 known drug binding sites were overall conserved. 12 out of 26 (46.2%) vaccine-targeted positions were fully conserved.

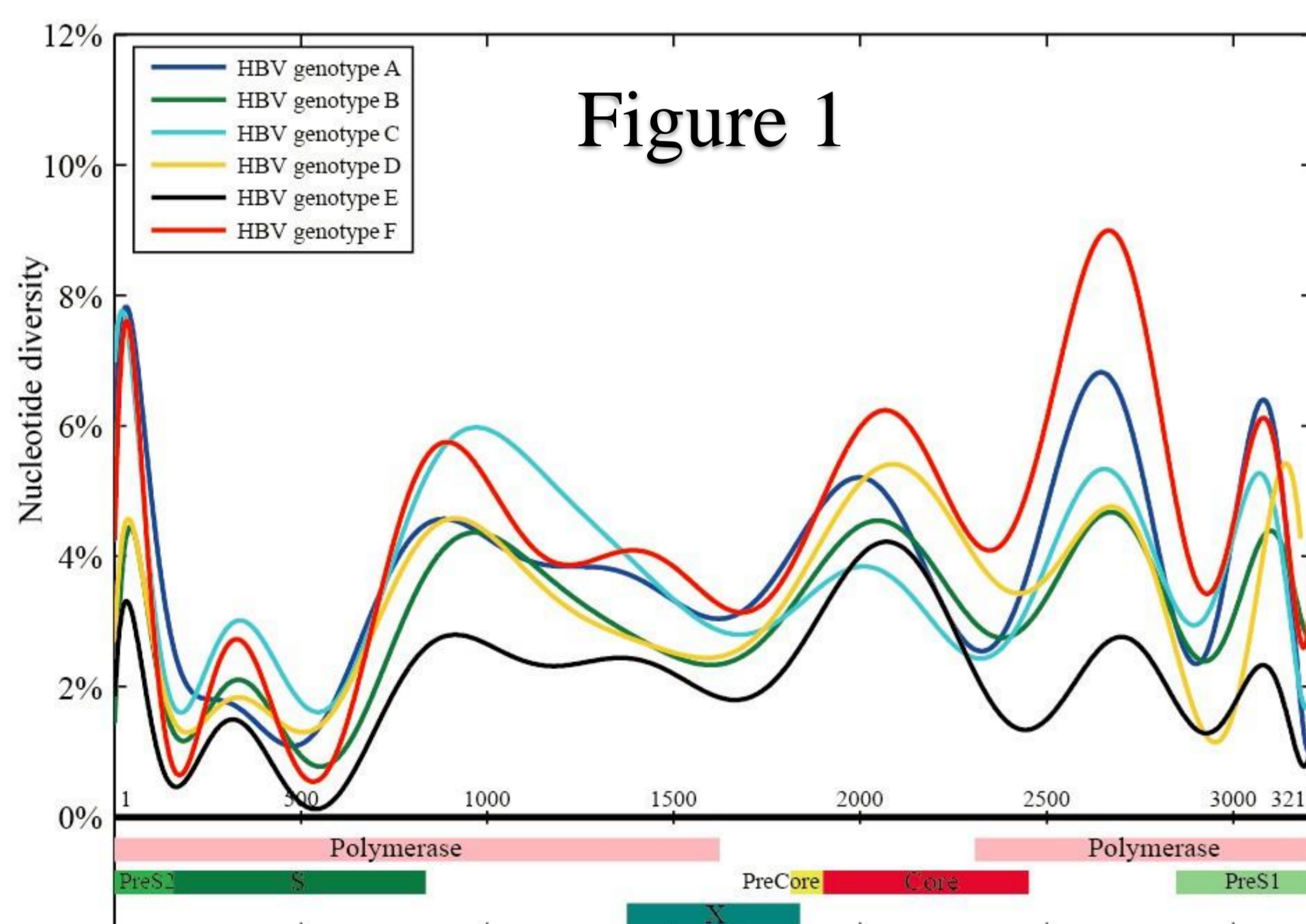


Figure 1  
Nucleotide diversity along the HBV full-length genome. Each colored plot shows the density of nucleotide diversity for one HBV genotype, indicated by the figure legend.

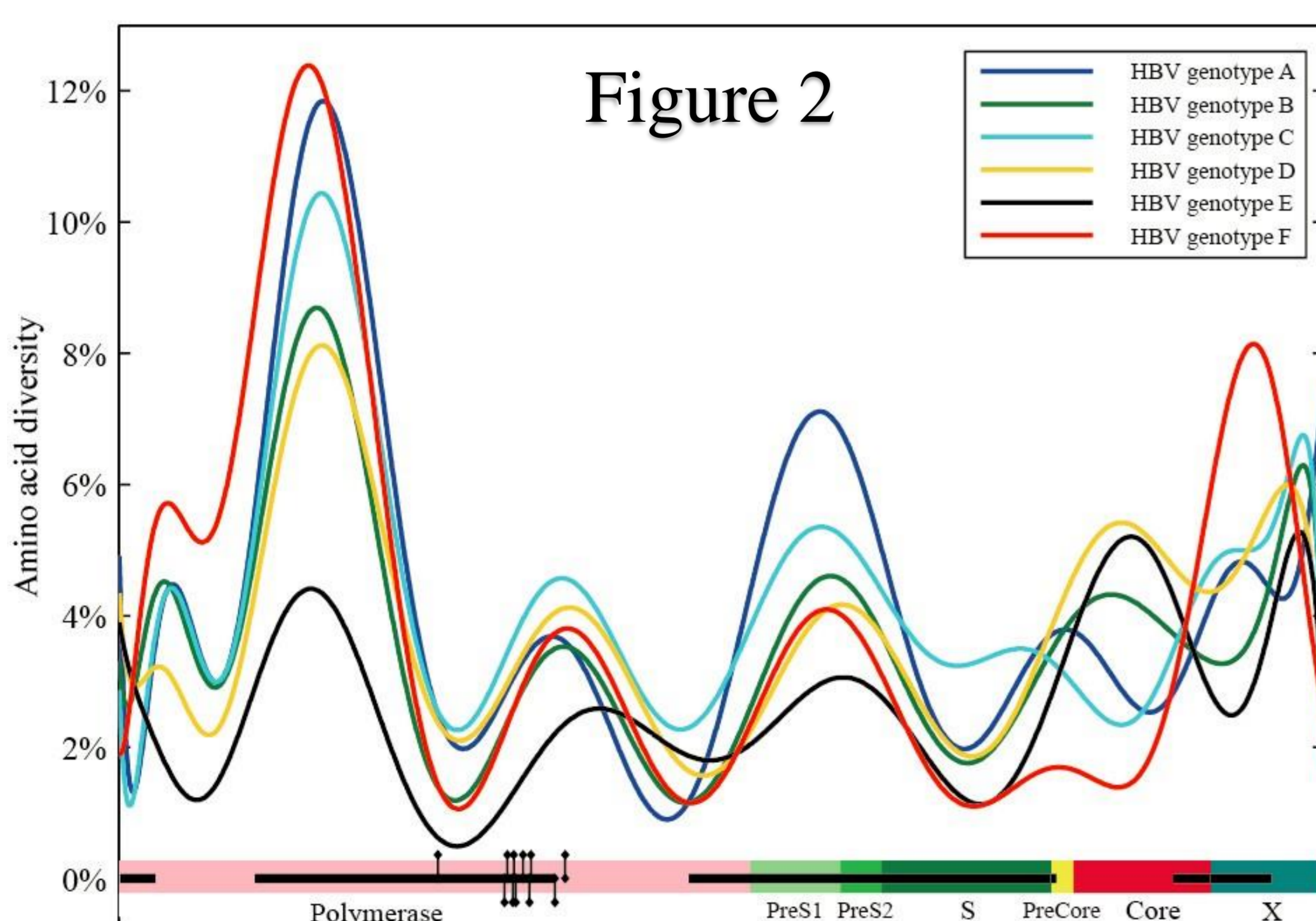


Figure 2  
Amino acid diversity along the HBV full-length genome. Each colored plot shows the density of amino acid diversity for one HBV genotype, indicated by the figure legend.

Figure 3

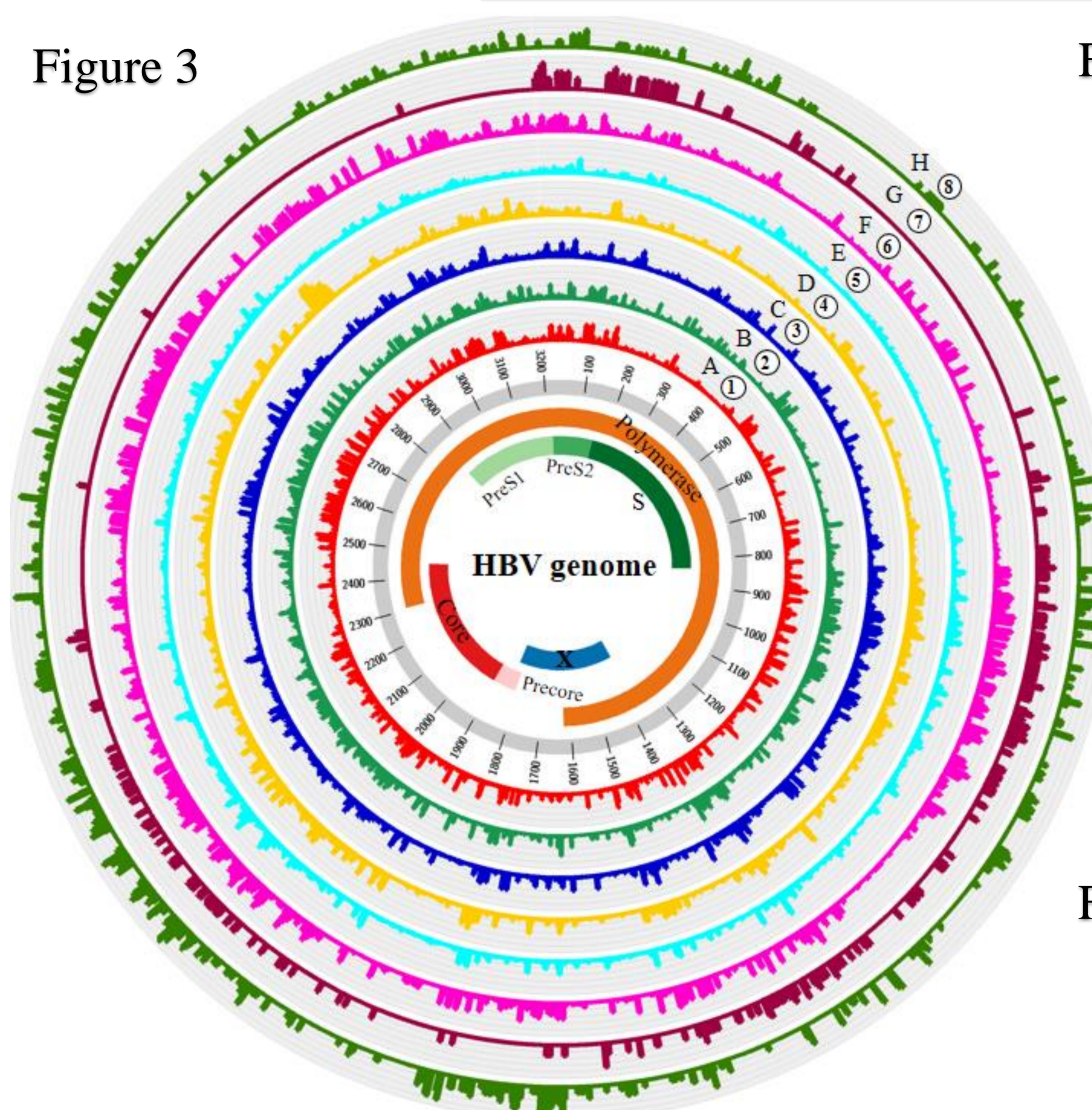


Figure 3  
Nucleotide genomic diversity of 8 HBV genotypes. Schematic view of HBV proteins is mapped in the middle. Reference indices of HBV genotype C genome are indicated. HBV proteins are colored accordingly. Nucleotide diversity of HBV genotype genome is colored in each sub-layer (HBV genotype A-H from inside to outside).

Figure 4

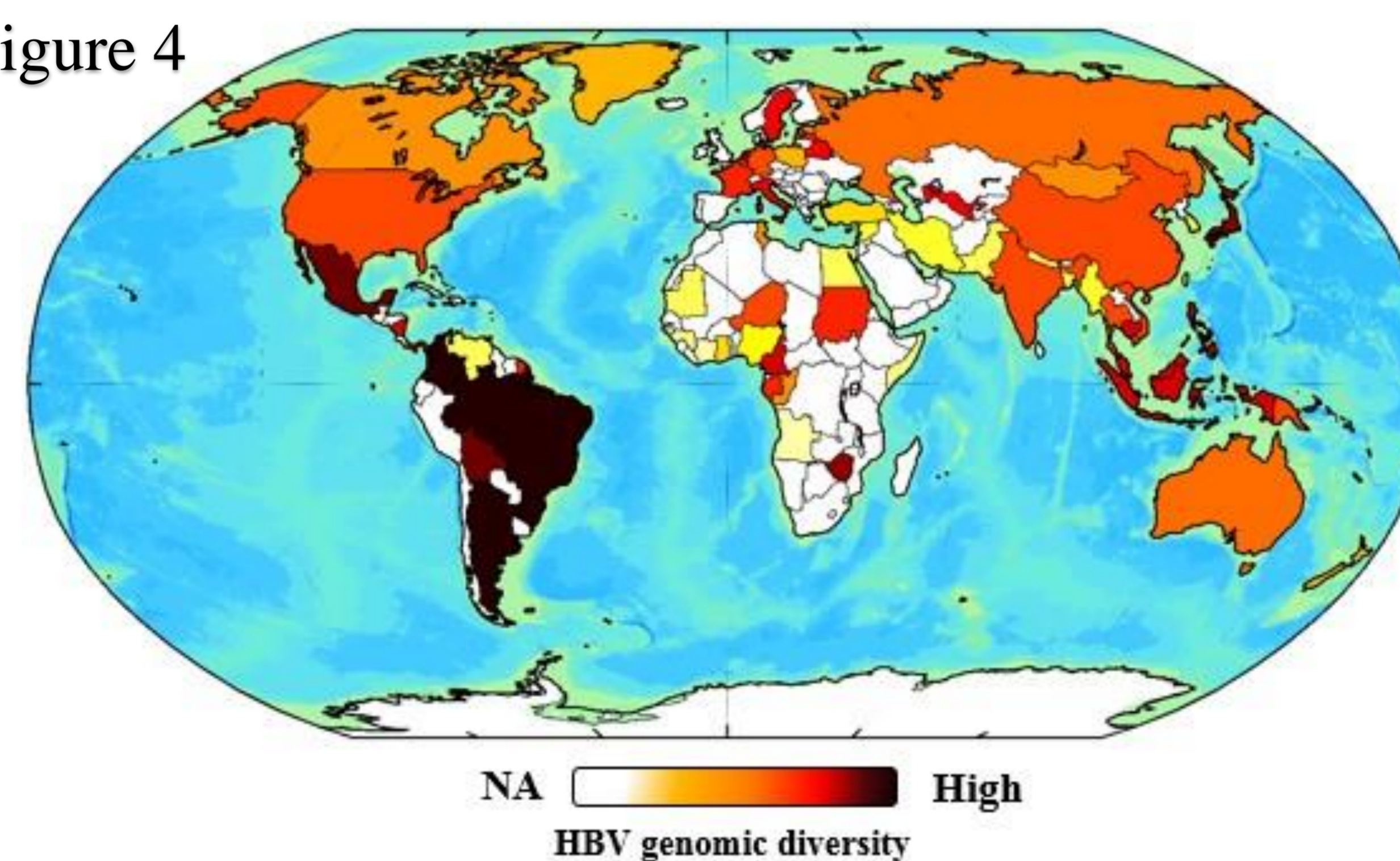
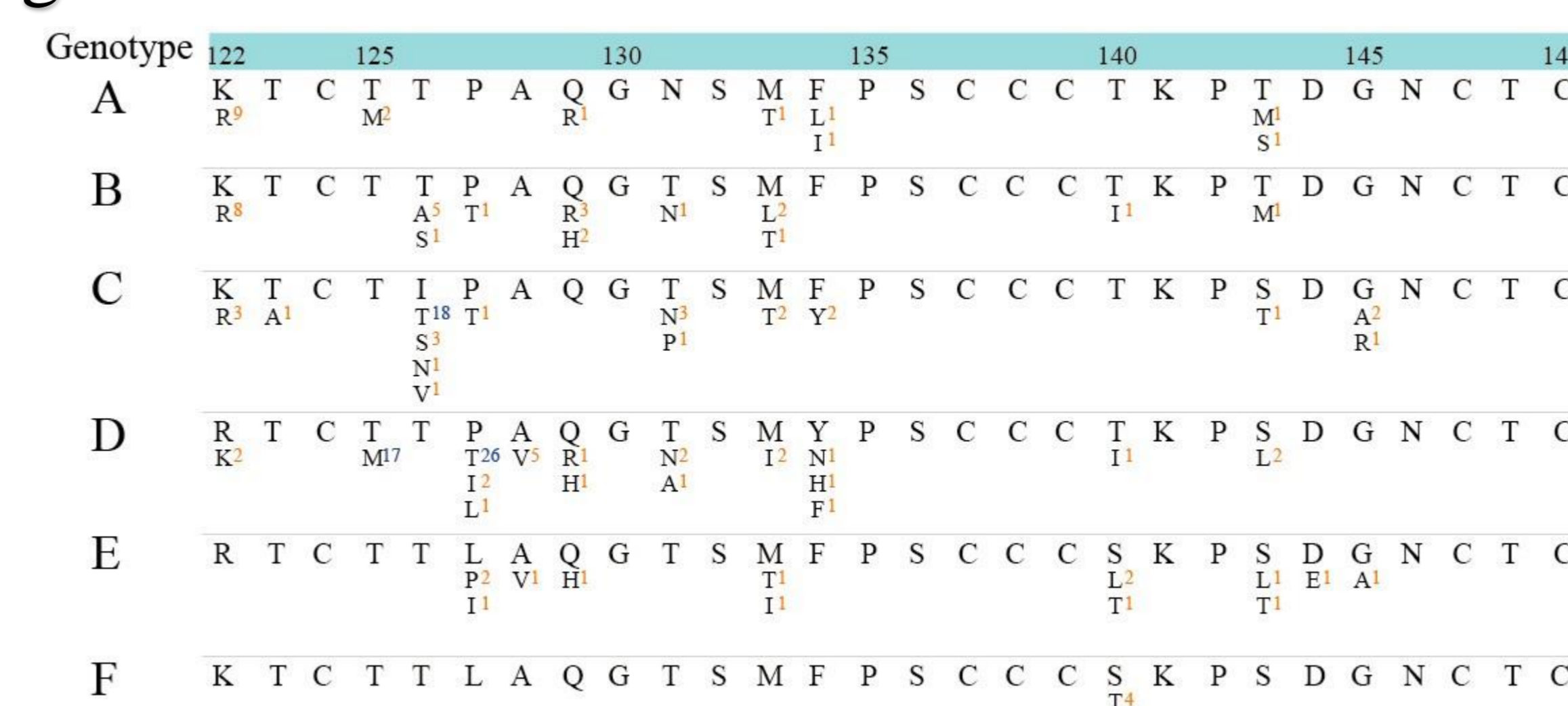


Figure 4  
Geographic distribution of HBV genome diversity. Countries with no more than 3 genomic sequences available (NA) are colored white. Amino acid diversity in individual countries was mapped onto the global cartographic map. Countries with their infections in different genotypes had higher genomic diversity.

Figure 5



## Material & Method

1. 5566 HBV full-length genome sequences from NCBI database.
2. Maximum likelihood phylogenetic tree was constructed by RAxML V7.
3. Vaccine and drug binding positions of FDA-approved HBV inhibitors was retrieved from literature. Protein structures were extracted from the PDB repository.
4. Genetic diversity was calculated by our method described previously (Guangdi Li et al, Retrovirology, 12:18, 2015).

Figure 6

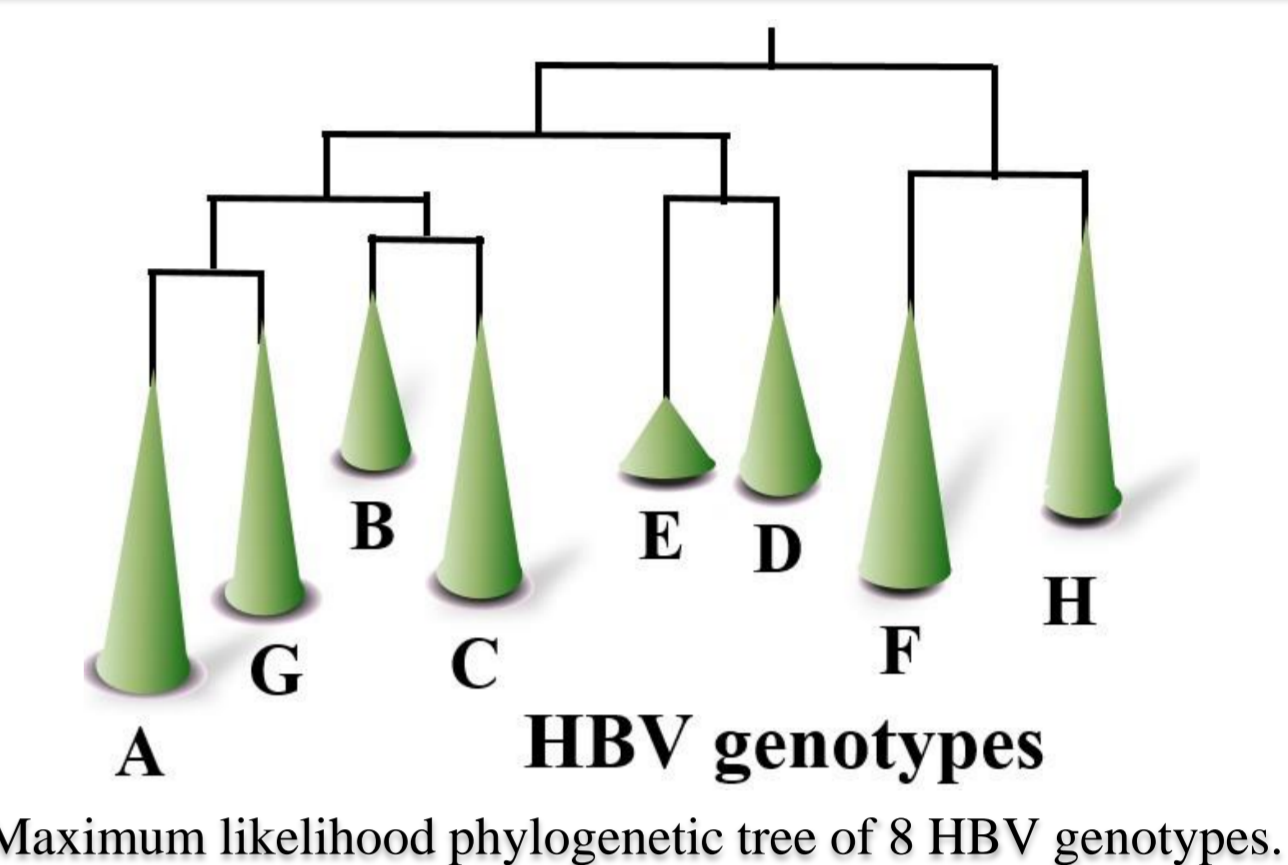


Figure 7

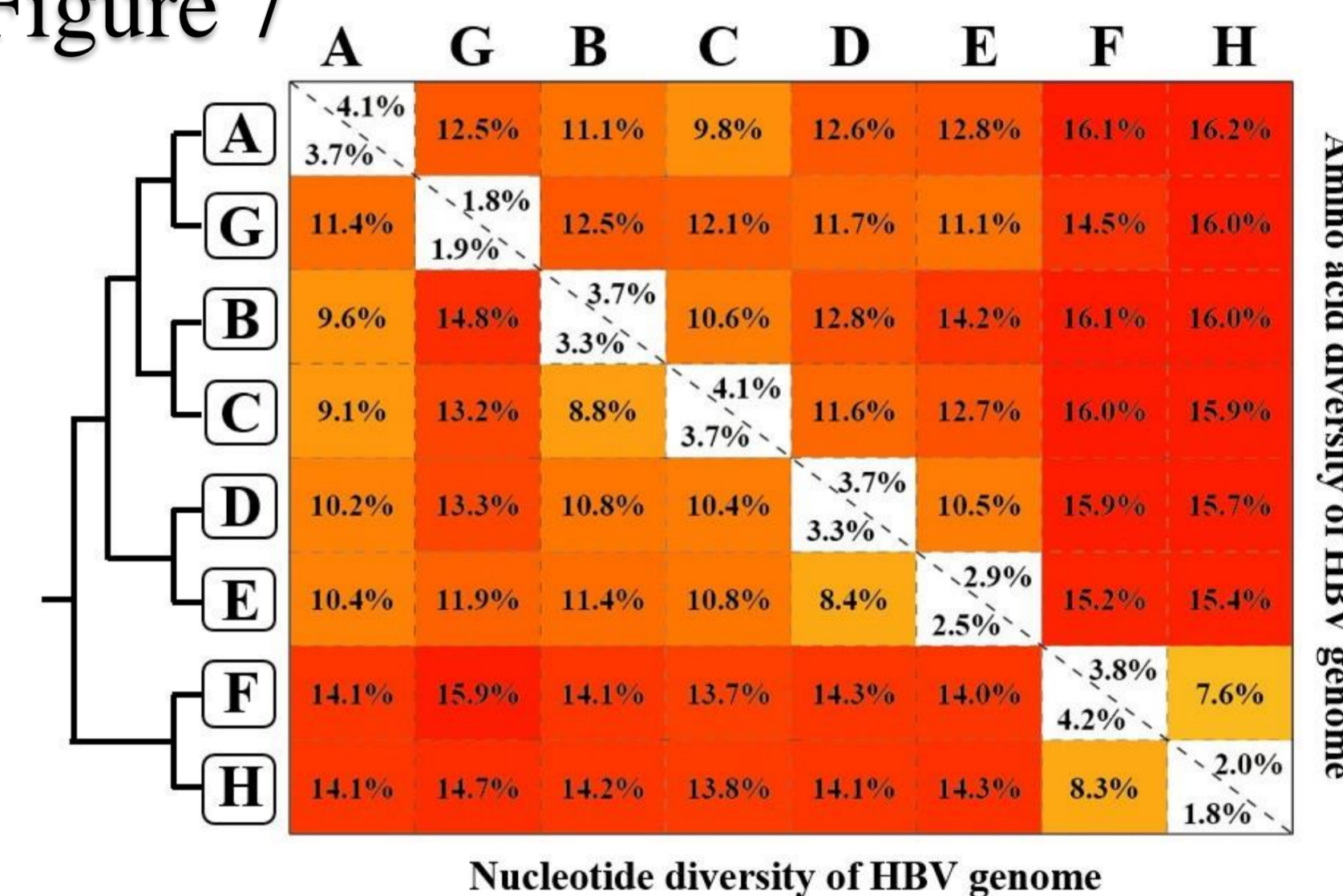


Figure 8

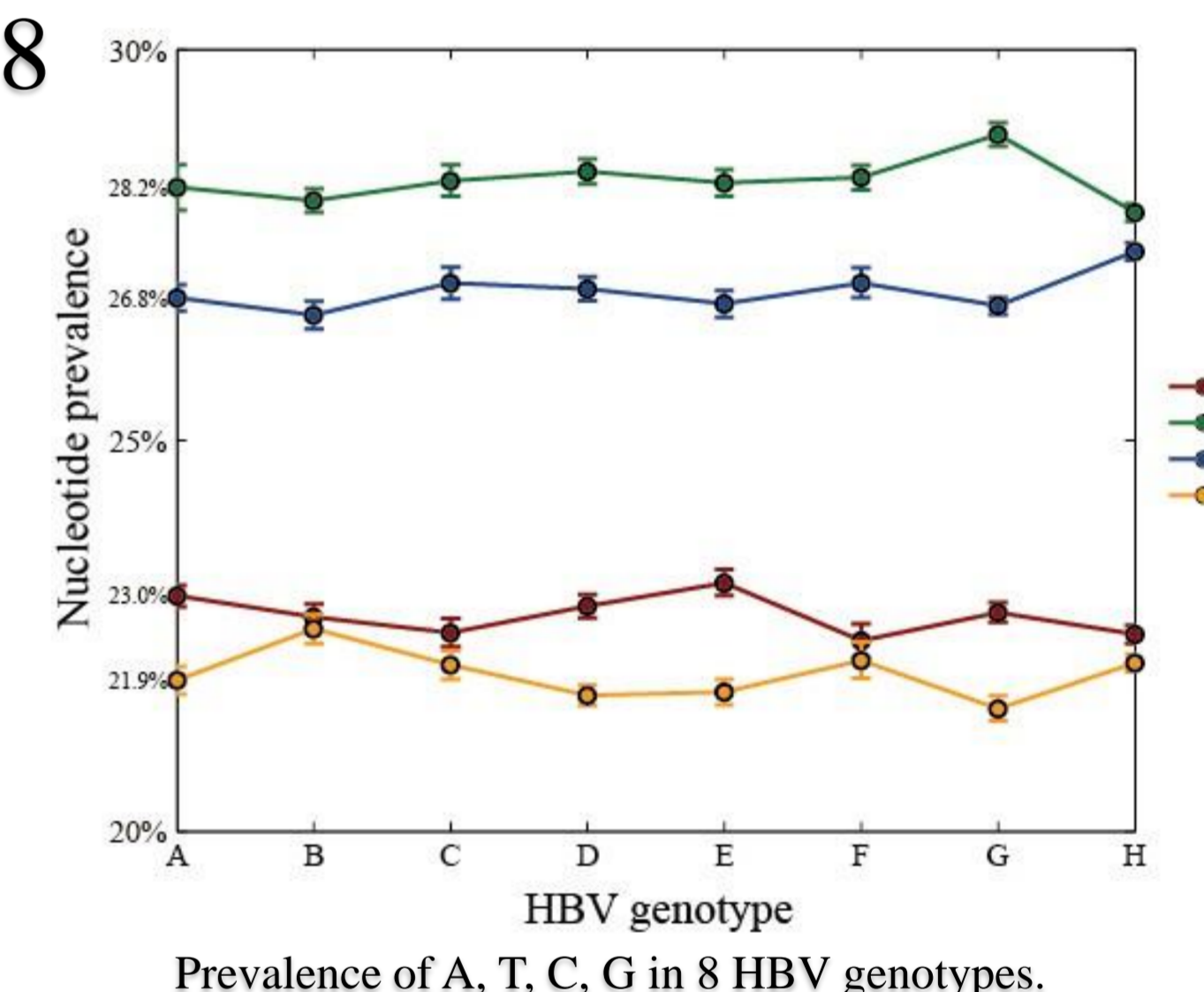


Figure 9

Drug resistance AA positions

Genotype	80	169	173	180	181	184	194	202	204	236	250
A	L	I	V	L	A	T	A	S	M	N	M
B	L	I	V	L	A	T	A	S	M	N	M
C	L	I	V	L	A	T	A	S	M	N	M
D	L	I	V	L	A	T	A	S	M	N	M
E	L	I	V	L	A	T	A	S	M	N	M
F	L	I	V	L	A	T	A	S	M	N	M
G	L	I	V	L	A	T	A	S	M	N	M
H	L	I	V	L	A	T	A	S	M	N	M

## Conclusion

This first large-scale analysis of full-length HBV genome provides a detailed mapping of natural diversity across major HBV genotypes, and highlights the conserved regions in HBV genome.