Table S1. Sensitivity and specificity of SLα covariance model in detection of HBV.[a]

|		True positive	True negative	False positive	False negative	Sensitivity	Specificity	Accuracy	Truncated hit	MER score[b]	Score threshold
Training (5-fold cross-validation)	1	1034	5009	0	4	99.6	100	99.9	0	4	28.7
	2	1034	5622	0	4	99.6	100	99.9	0	4	28.7
	3	1033	6405	0	6	99.4	100	99.9	0	6	29.0
	4	1031	5724	0	9	99.1	100	99.9	0	9	28.7
	5	1035	5192	0	5	99.5	100	99.9	0	5	28.7
Testing		1290	6969	0	10	99.2	100	99.9	0	10	28.7

Table S2. Sensitivity and specificity of SLβ covariance model in detection of HBV.[a]

		True positive	True negative	False positive	False negative	Sensitivity	Specificity	Accuracy	Truncated hit	MER score[b]	Score threshold
Training (5-fold cross-validation)	1	1033	5009	0	5	99.5	100	99.9	0	5	33.1
	2	1031	5622	0	7	99.3	100	99.9	0	7	33.1
	3	1028	6405	0	11	98.9	100	99.9	0	11	33.2
	4	1030	5724	0	10	99.0	100	99.9	0	10	33.1
	5	1032	5192	0	8	99.2	100	99.9	0	8	33.1
Testing		1292	6969	0	8	99.4	100	99.9	0	8	33.1

[a] Complete genomes of 6495 *Orthohepadnavirus* and 34921 other viruses (retrieved from GenBank viral division at 24 February 2016).

[b] Minimum error rate (MER) score: total number of false positive and false negative.[1,2]
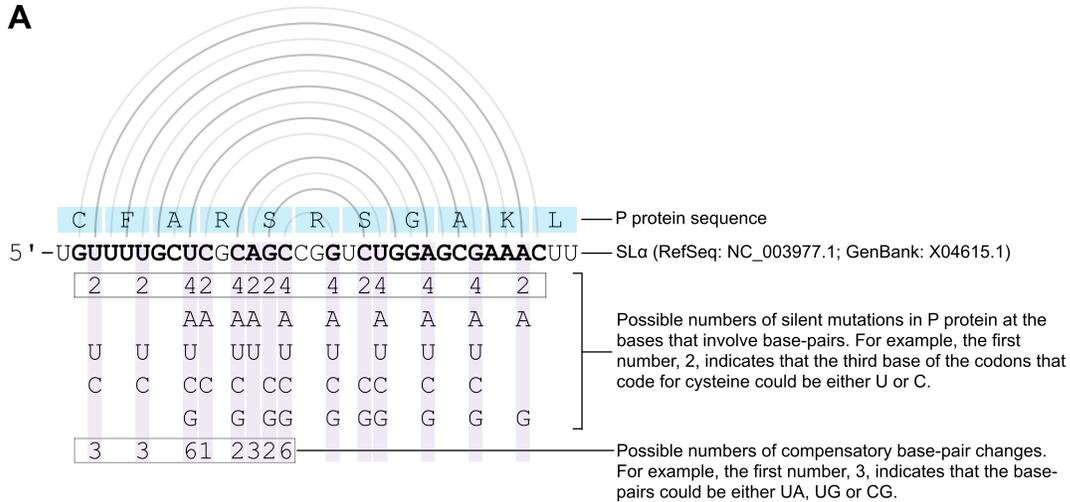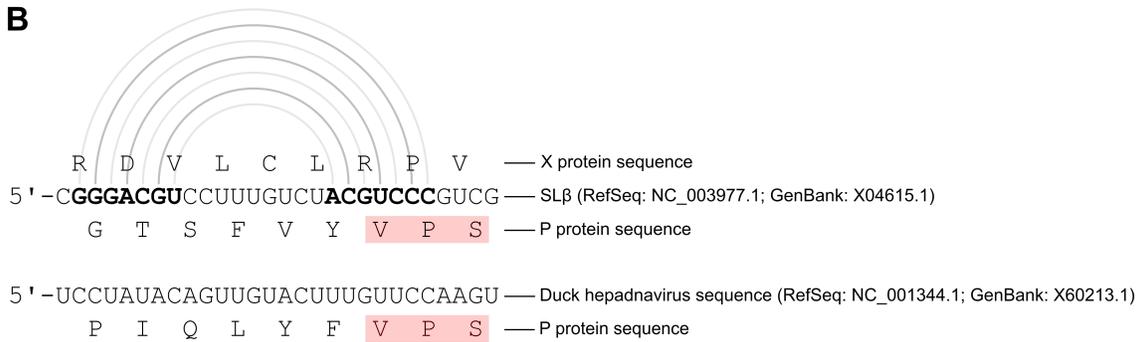
Figure S1. The constraints of P and X reading frames on the base-pairs in SLα and SLβ (RefSeq: NC_003977.1; GenBank: X04615.1). **(A)** SLα is located within the P reading frame (blue shading). There are a total of 3888 possible sequence variants that allow compensatory base-pair changes while preserving the P amino acid sequence ($3 \times 3 \times 6 \times 1 \times 2 \times 3 \times 2 \times 6 = 3888$; purple shading). **(B)** SLβ is located within the overlapping reading frames of P and X proteins. The stem of SLβ would be expected to lack compensatory base-pair changes. The conserved functional motif of RNase H domain of P proteins of mammalian and duck (RefSeq: NC_001344.1; GenBank: X60213.1) hepadnaviruses is highlighted in pink.[3] Bases that involve in base-pairs are indicated in boldface. Base-pairs are indicated as arcs. It should be noted that on 23 October 2015, NC_003977.1 (genotype C; GenBank: X04615.1) was replaced with a different HBV genome, NC_003977.2 (genotype D; GenBank: V01460.1).
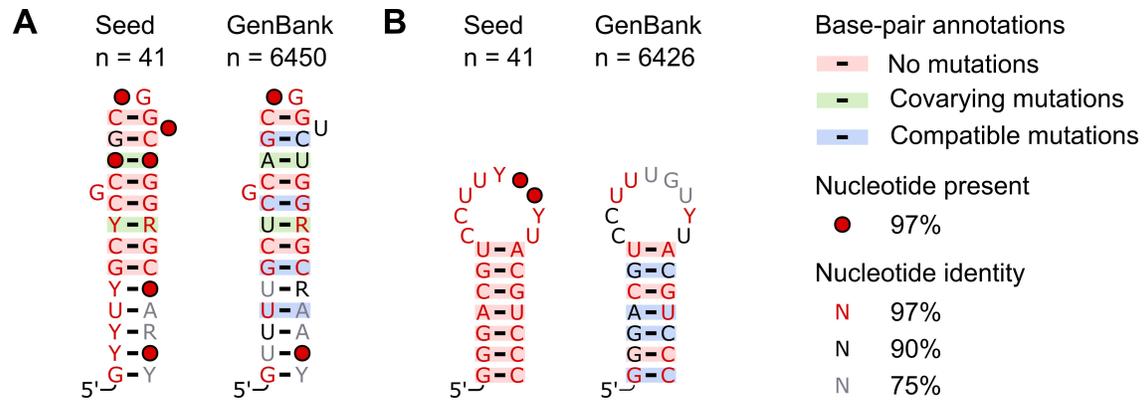
Figure S2. RNA consensus diagrams for **(A)** SLα and **(B)** SLβ based on seed sequences (covariance models) or all mammalian hepadnaviruses in GenBank (retrieved at 24 February 2016). Default settings of R2R 1.0.4 were used.

References

1. Nawrocki EP, Eddy SR. Query-Dependent Banding (QDB) for Faster RNA Similarity Searches. PLOS Comput Biol 2007; 3:e56; PMID:17397253; http://dx.doi.org/10.1371/journal.pcbi.0030056

2. Zhou Y, Lu C, Wu Q-J, Wang Y, Sun Z-T, Deng J-C, Zhang Y. GISSD: Group I Intron Sequence and Structure Database. Nucl Acids Res 2008; 36:D31–7; PMID:17942415; http://dx.doi.org/10.1093/nar/gkm766

3. Chen Y, Marion PL. Amino acids essential for RNase H activity of hepadnaviruses are also required for efficient elongation of minus-strand viral DNA. J Virol 1996; 70:6151–6; PMID:8709240