Figure 1 Temporal graphs displaying the frequency of the glycosylation patterns at HA158 and NA88 with respect to the year. All clade 2.3.2.1 sequences were analyzed globally (A) or in Asia, Africa, Europe, and North America (B-E) till 2022.

Figure 2 Phylogenetic trees of the HA genes of H5N1 AIVs and variation of HA158 and NA1-88 glycosylation sites. Phylogenetic trees were generated from the HA sequences of the clade 2.3.2.1 AIVs that were isolated throughout the world (n=2,565) (A), clade2.3.2.1a in Asia (n=720) (B), clade2.3.2.1c in Asia (n=941) (C) via the maximum likelihood (ML) method. The HA trees were constructed according to strains A/turkey/Ontario/84/1983 (A) and A/Goose/Guangdong/1/96 (B-C). Different HA clades were marked in the phylogenetic tree, while the strains with glycosylation at residue HA158/NA88 and viral isolation regions were marked on the right according to different colors. Temporal graphs displaying the frequency of the glycosylation patterns for HA158 and NA88 in China from 2006-2020 (n=291) (D).

Figure 3 The amino acid molecular polymorphism of HA/NA glycosylations and glycan modeling of HA and NA. Amino acid sequences included HK97, clade 2.3.2.1d YZ, clade 2.3.2.1e DT, clade 2.3.4 SY, A/Hubei/1/2010/clade 2.3.2.1a (old), A/duck/Bangladesh/2020/clade 2.3.2.1a (new), A/barn/swallow/Hong Kong/1161/2010/clade 2.3.2.1b, A/duck/Vietnam/NCVD/0004/2013/clade 2.3.2.1c, and A/bobcat/Wisconsin/2022/clade 2.3.4.4b. The HA number refers to the H5 number (The H5-158/H3 number here is H5-170) (A). NA88 amino acid patterns (NA88

contains NA truncation) (B). The glycan complex was added at the glycosylation sites with a black module (A-B). *In silico* model of N-linked glycan structures mapped onto the HA (HA PDB:6pcx) (C) and NA (NA PDB:3cl2) (D) structures. HA158 is indicated in by dark blue, and NA88 is indicated by purple.

Figure 4 The HI (A-C) and NT (D-F) titers against Re-6, Re-10, and Re-12 antisera for the HA/NA glycosylation H5N1 variants. $^{**}P < 0.01$.

Figure 5 Receptor binding specificity of HA/NA glycosylation H5N1 variants. The direct binding affinity of viruses to sialylglycopolymers containing either 3′SLN-PAA or 6′SLN-PAA was measured. SY(A) and CA09(B) were used as controls. The α-2,3(C) and α-2,6(D) binding affinities of YZ and the variants were determined accordingly.

Figure 6 NA activities of mutant NA proteins and recombinant viruses. Cell lysates were prepared after transfection of pHW-NA plasmid into 293T cells. The sialidase activity of each NA mutant was measured using the soluble NA-XTD substrate (A). Virus samples were standardized to equivalent HA titers. NA activities were compared to those of wild type NA protein or virus samples (B). Two-fold dilutions of each NA mutant were incubated at 37 ℃ with an equal volume of chicken erythrocytes. The samples were then transferred to 37 °C, and their HA titers were recorded every 2 h to evaluate viral elution speed (C). All data were compared to the value of the YZ-HA158-NA88+, $^{**}P< 0.01$ and $^{ns}P> 0.05$.

Figure 7 Thermal stability of the recombinant viruses at 37 ℃ (A), 42 ℃ (B), and 56 ℃ (C) and low-pH stability from pH 4.0-7.0 (D). All data were compared to the value of the YZ-HA158-NA88+, $^{*}P< 0.05$, $^{**}P< 0.01$.

Figure 8 Growth characterization of the recombinant viruses. Plaque morphologies of the recombinant viruses were determined in MDCK cells (A). Growth kinetics of the recombinant viruses were determined in CEF, MDCK, and A549 cells at a MOI of 0.001 (B). Data are presented as the average of virus titrations conducted in triplicates for every 12 h as indicated by $TCID_{50}$ in CEF cells. All data were compared to the value of the YZ-HA158-NA88+, $^{*}P< 0.05$, $^{**}P< 0.01$.

Figure 9 Pathogenicity of the HA/NA glycosylation variants in chickens. Eight five-week-old SPF chickens in each group were inoculated with 200 μL of $10^{4.0}$ EID_{50} mutant virus, and the mortality rates were observed and recorded until 14 days post-infection (A). The heart, liver, spleen, lung, and kidney of three infected chickens were collected at 36 hours post-infection in PBS containing antibiotics. The viral titers of tissue with same weight were determined in CEF cells (B).

Figure 10 Pathogenicity of the HA/NA glycosylation variants in mice. BALB/c mice (n=5) were inoculated intranasally with 10^{1} to 10^{4} EID_{50} of virus. Body weight changes (A-D) and mortality (E-H) of the mice were observed continuously for 14 days.

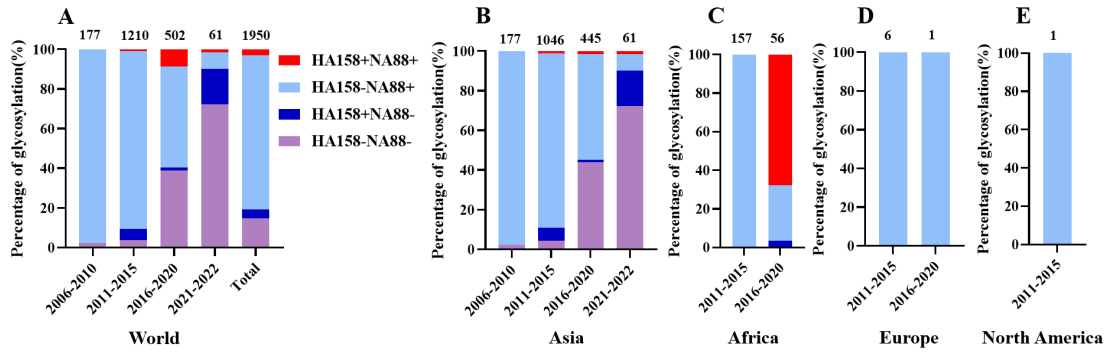


Figure 1

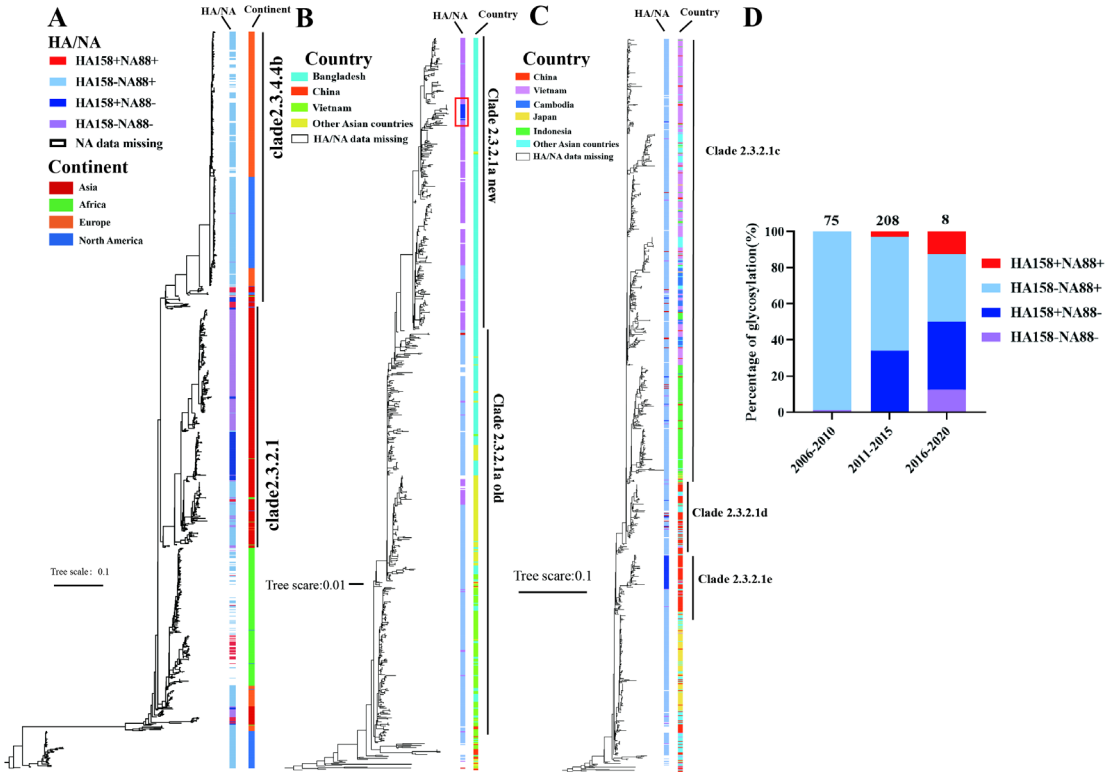
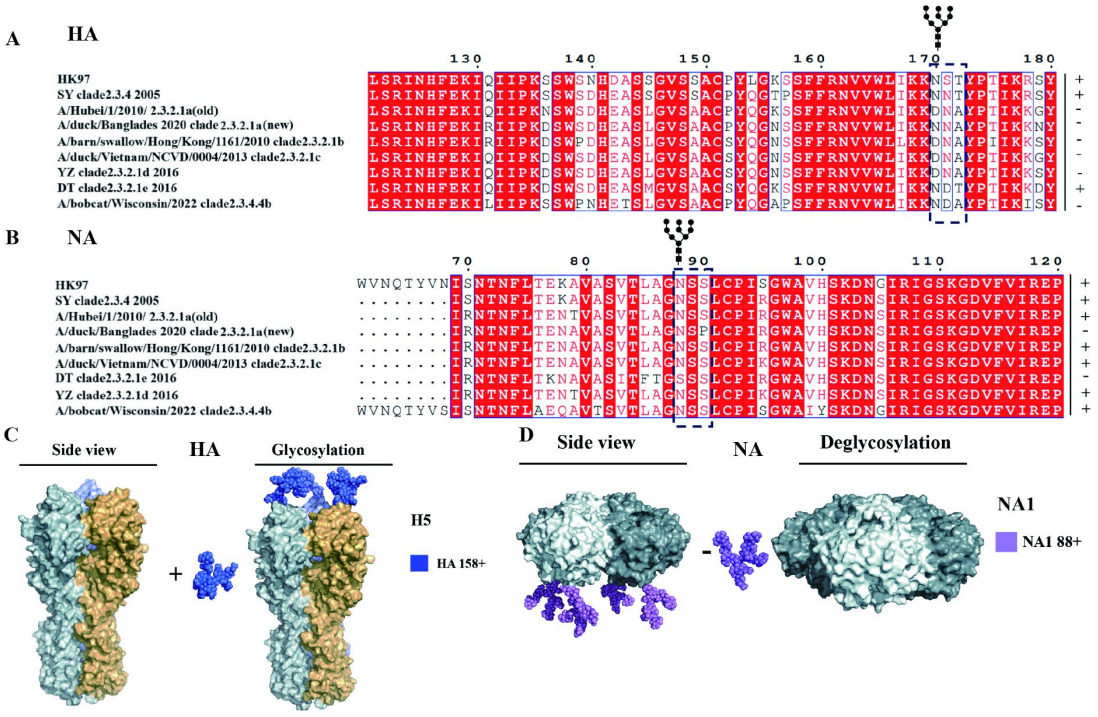
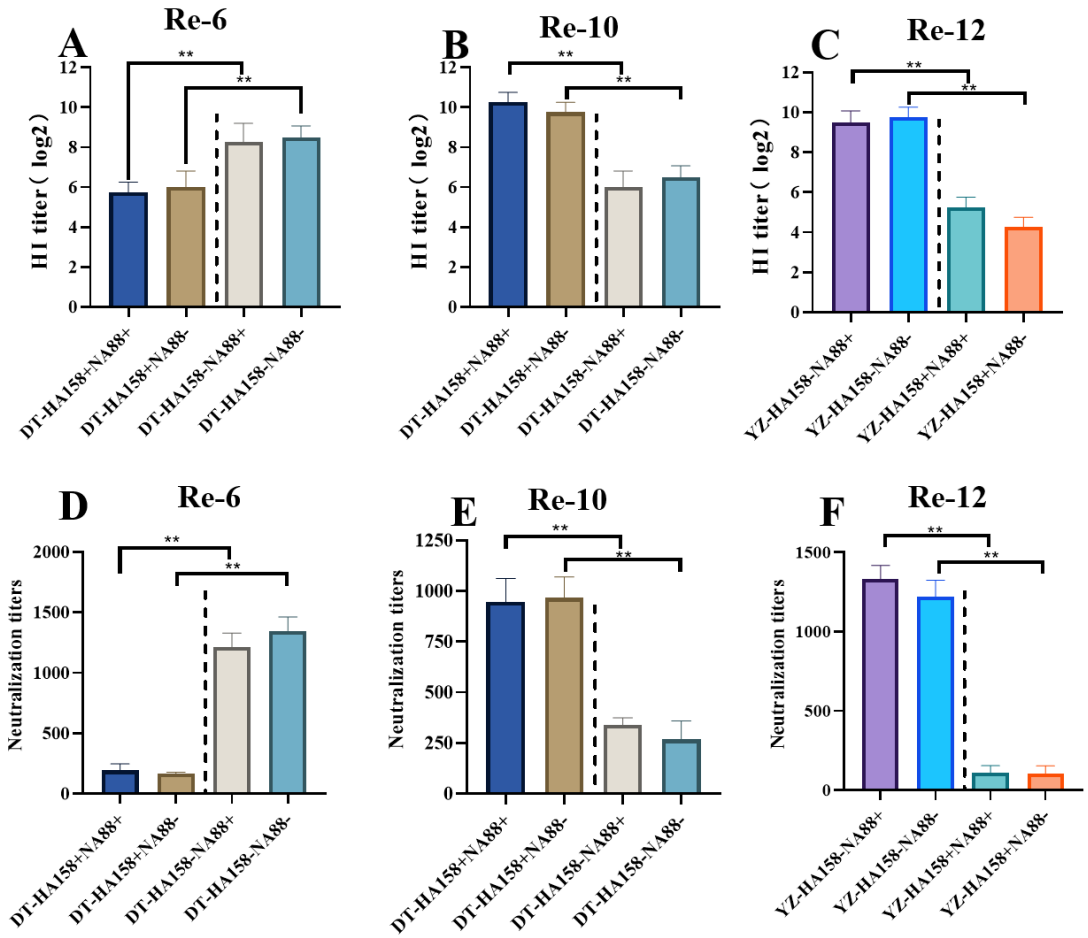


Figure 2



71

72 Figure 3



73

74 Figure 4

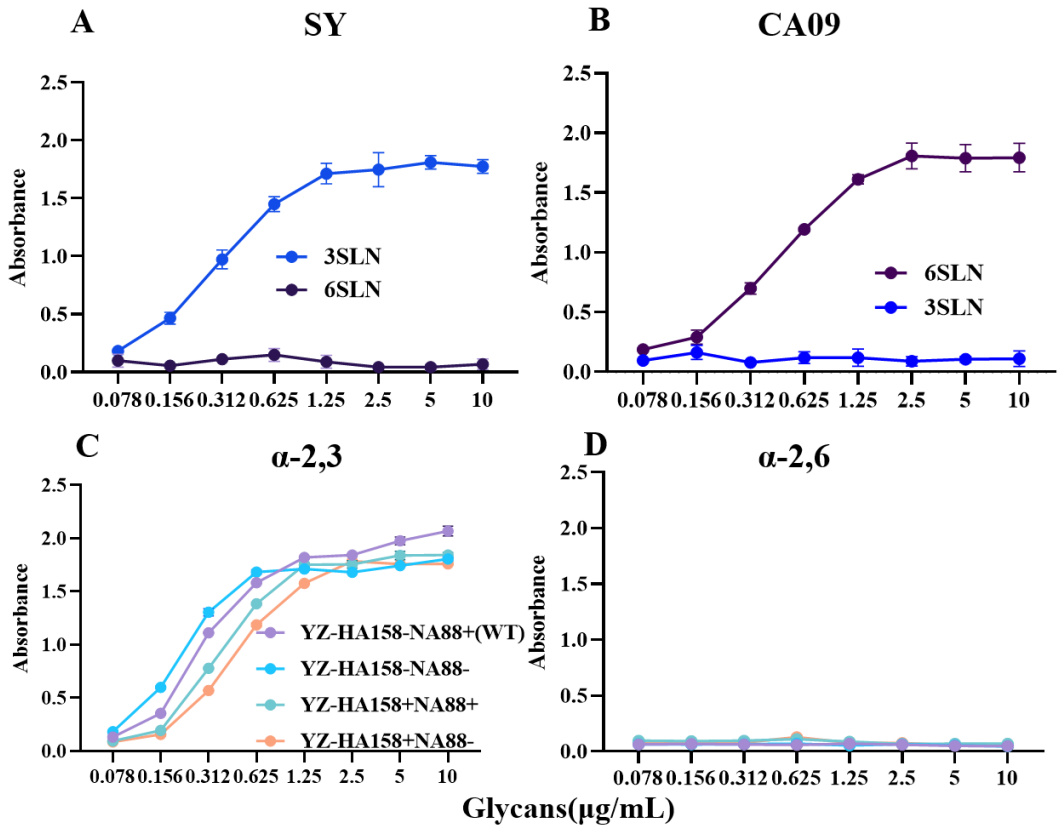


Figure 5

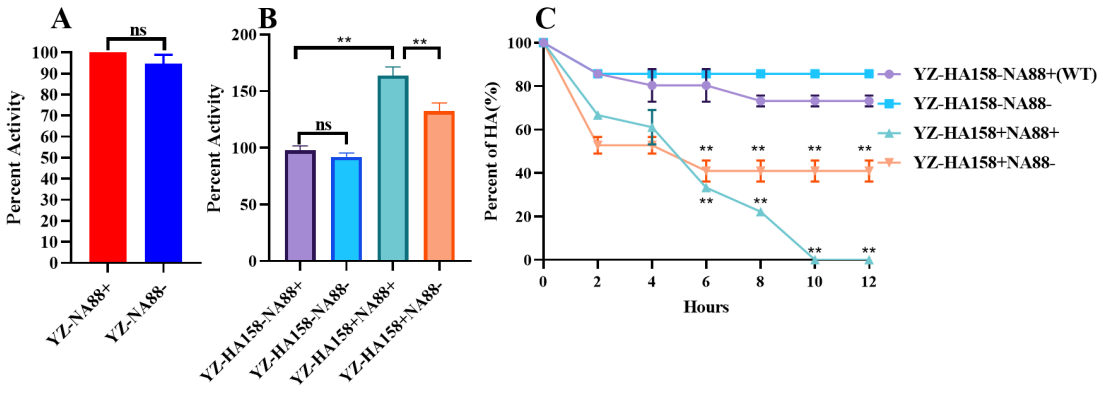


Figure 6

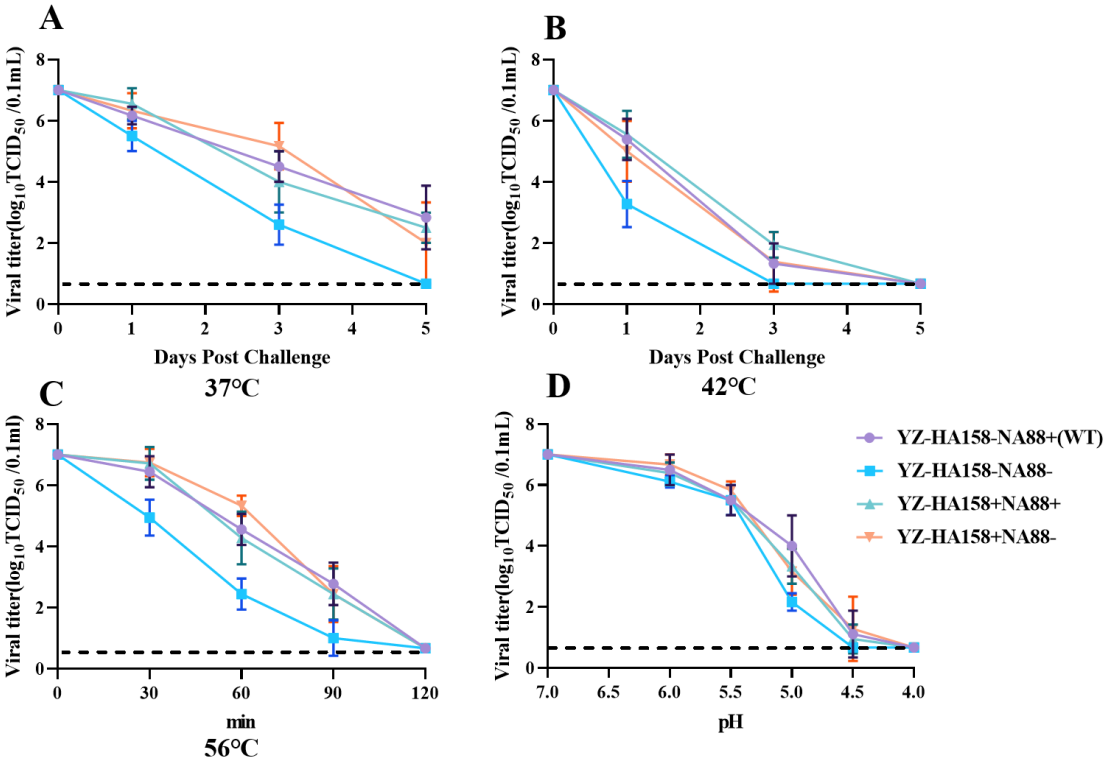


Figure 7

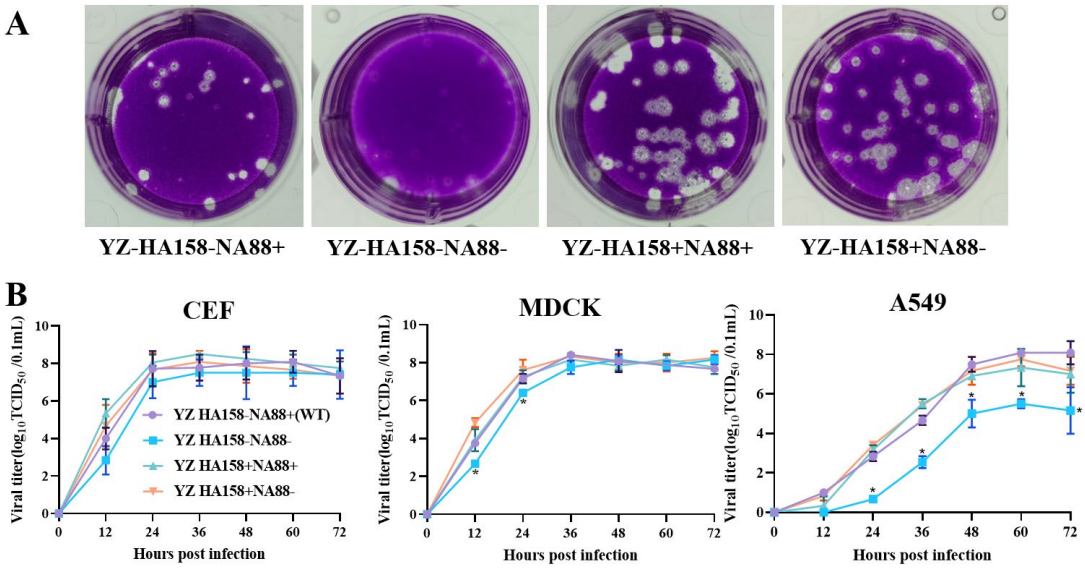
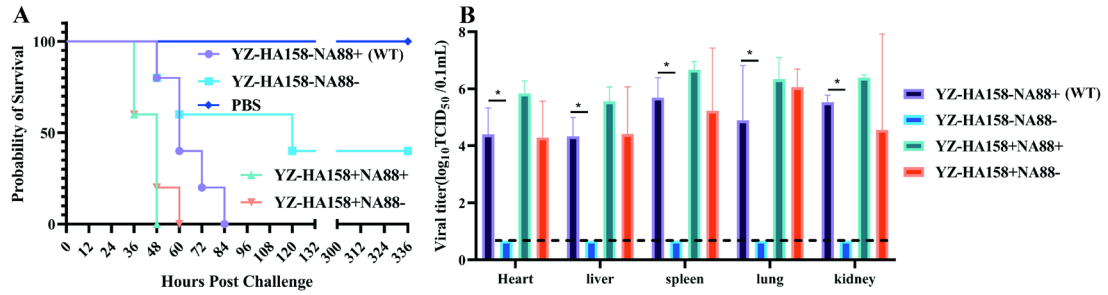
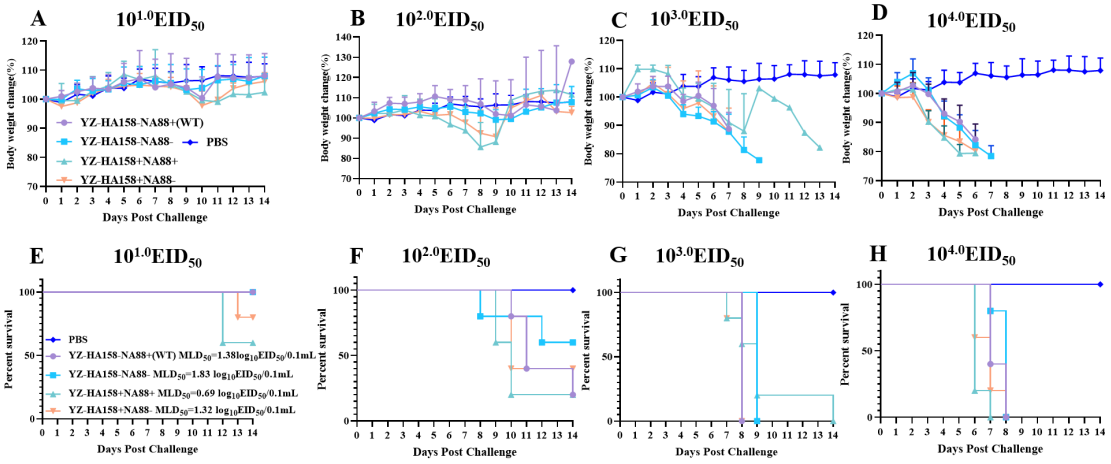


Figure 8



86

87 Figure 9



88

89 Figure 10

90 Figure S1 Verification of glycosylation sites at residues HA158 and NA88 of the mutant

91 viruses

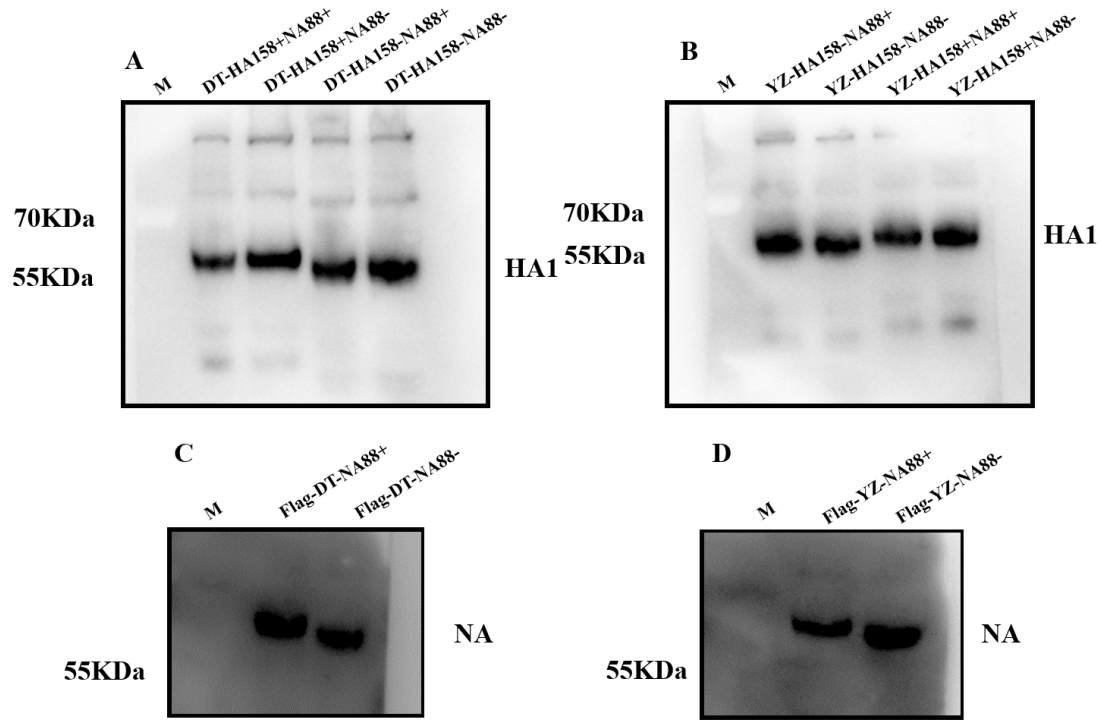


Figure S1

Table S1 Primers for construction of the recombination plasmids

	forward primer	reverse primer
YZ-HA158+	AAAGCTACAATAATACCA ACCGAGAAGATCTCTTG ATACTGTG	TATGTATCGTTCTTTTTTGATA AGCCACACCACATTTCTGAA GA
YZ-NA88-	AGGATGGGCTGTACAAG TAAAGACAACAGTATAA GGATTGGGT	CTGCCAGCTAATGTTACTGAA GCTACAGTGTTCTCAGTAAGG
DT-HA158-	TCAAAAAGGACAATGCA TATCCAACAATAAAGAA AGACTA	GATATGCATTGTCCTTTTTGAT AAGCCACACTACATTTCTGAA
DT-NA88+	ACATTTACGGGCAATTCA TCTCTTTGTCCCATTAGA GGAT	TGGGACAAAGAGATGAATTG CCCGTAAATGTTATTGAA
Flag-N1-88	GGCAAGCTTATGAATCCA	CCGGTACCGGCTACTTGTCAA

	AATCAGAAGATAGTGA	TGGTGAATGGCAAC

96

97 Table S2 Biological characteristic of the recombinant virus

	$TCID_{50}(Log_{10}/mL)$	$EID_{50}(Log_{10}/mL)$
DT-HA158-NA88+	7.5	9.23
DT-HA158-NA88-	7.67	8.83
DT-HA158+NA88+	7.5	8.5
DT-HA158+NA88-(WT)	8.23	9.17
YZ-HA158-NA88+ (WT)	7.83	8.75
YZ-HA158-NA88-	7.35	8.09
YZ-HA158+NA88+	8.28	9.09
YZ-HA158+NA88-	8.22	9.34

98

99