

Genome-Wide Characterization of QYYZ-Like PRRSV During 2018–2021

Hu Xu^{1†}, Lirun Xiang^{2†}, Yan-Dong Tang^{2†}, Chao Li², Jing Zhao², Bangjun Gong², Qi Sun², Chaoliang Leng³, Jinmei Peng², Qian Wang², Guohui Zhou², Tongqing An², Xuehui Cai², Zhi-Jun Tian², Hongliang Zhang^{2*} and Mingxin Song^{1*}

¹ College of Veterinary Medicine, Northeast Agricultural University, Harbin, China, ² State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin, China, ³ Henan Key Laboratory of Insect Biology in Funiu Mountain, Henan Provincial Engineering Laboratory of Insect Bioreactors, China-UK-NYNU-RRes Joint Laboratory of Insect Biology, Nanyang Normal University, Nanyang, China

OPEN ACCESS

Edited by:

Zhou Mo, Heilongjiang University, China

Reviewed by:

Jianguo Dong, Xinyang Agriculture and Forestry University, China Yifeng Jiang, Shanghai Veterinary Research Institute (CAAS), China

*Correspondence:

Mingxin Song songmx@neau.edu.cn Hongliang Zhang zhanghongliang01@caas.cn

[†]These authors have contributed equally to this work and share first authorship

Specialty section:

This article was submitted to Veterinary Infectious Diseases, a section of the journal Frontiers in Veterinary Science

> Received: 17 May 2022 Accepted: 06 June 2022 Published: 30 June 2022

Citation:

Xu H, Xiang L, Tang Y-D, Li C, Zhao J, Gong B, Sun Q, Leng C, Peng J, Wang Q, Zhou G, An T, Cai X, Tian Z-J, Zhang H and Song M (2022) Genome-Wide Characterization of QYYZ-Like PRRSV During 2018–2021. Front. Vet. Sci. 9:945381. doi: 10.3389/fvets.2022.945381 In the last decade, the emergence of QYYZ-like porcine reproductive and respiratory syndrome virus (PRRSV) has attracted increasing attention due to the high incidence of PRRSV mutation and recombination. However, the endemic status and genomic characteristics of the QYYZ-like strains are unclear. From 2018 to October 2021, 24 QYYZ-like PRRSV isolates were obtained from 787 PRRSV-positive clinical samples. Only one QYYZ-like positive sample was from a northern province, and the rest were from central and southern provinces. We selected 9 samples for whole-genome sequencing, revealing genome lengths of 15,008-15,316 nt. We retrieved all the available wholegenome sequences of QYYZ-like PRRSVs isolated in China from 2010 to 2021 (n = 28) from GenBank and analyzed them together with the new whole-genome sequences (n = 1)9). Phylogenetic tree analysis based on the ORF5 gene showed that all QYYZ-like PRRSV strains belonged to sublineage 3.5 but were clustered into three lineages (sublineage 1.8, sublineage 3.5, and sublineage 8.7) based on whole-genome sequences. Genomic sequence alignment showed that QYYZ-like strains, have characteristic amino acids insertions or deletions in the Nsp2 region (same as NADC30, JXA1 and QYYZ) and that thirteen strains also had additional amino acid deletions, mostly between 468 and 518 aa. Moreover, QYYZ-like strains (sublineage 3.5) have seven identical characteristic amino acid mutations in ORF5. Recombination analysis revealed that almost all QYYZlike complete genome sequences (36/37) were products of recombination and mainly provided structural protein fragments (GP2-N) for the recombinant strains. Overall, QYYZ-like strains were mainly prevalent in central and southern China from 2018 to 2021, and these strains provided recombinant fragments in the PRRSV epidemic in China.

Keywords: QYYZ-like PRRSV, whole-genome analysis, complex patterns of recombination, recombination hotspots, epidemiological characteristics

INTRODUCTION

Porcine respiratory and reproductive syndrome (PRRS) is a major disease in the pig industry that causes huge economic losses to the swine industry worldwide (1). The causative agent, porcine reproductive and respiratory syndrome virus (PRRSV), is an enveloped, positive-sense, single-stranded RNA virus belonging to the genus *Betaarterivirus* and family *Arteriviridae* of order

1

Nidovirales (2). Due to its high degree of genetic diversity, PRRSV has been further divided into two species, PRRSV-1 (formerly known as European genotype 1) and PRRSV-2 (formerly known as North American genotype 2) (3). Since PRRSV was first discovered in China in 1996, the PRRSV-2 strain has been the main circulating strain in China (4).

Phylogenetic analyses of ORF5 of PRRSV-2 strains showed that PRRSV-2 could be divided into nine distinct lineages, with each lineage containing several sublineages (5, 6). Four different PRRSV-2 lineages have become widespread in China, including lineage 8 (JXA1-like/CH-1a-like), lineage 5 (VR-2332like), lineage 1 (NADC30-like/NADC34-like), and lineage 3 (QYYZ-like) (7). The first PRRSV-2 strain CH-1a of China was isolated in 1996 in lineage 8 (8). HP-PRRSV (JXA1like) was recognized in 2006 and originated from CH-1a-like strains (9). In 2013, some NADC30-like strains were isolated in China, and they have gradually become the dominant strains in recent years (10). NADC34-like strains were first reported in China in 2017 and became one of the major epidemic strains in 2020 (7, 11). Lineage 3 strains were initially reported in Taiwan and have emerged in Hong Kong (12). The FJ-1 strain was the first lineage 3 PRRSV detected in mainland China in 2005. The representative isolate of lineage 3 was QYYZ, which was identified in 2010 in mainland China and gradually became prevalent in southern China (12, 13). More importantly, lineage 3 viruses with greater virulence have been reported in southern China, and these viruses, recombining with lineage 1 and 8 PRRSV, pose a great threat to the Chinese pig industry (14-17).

Several studies have summarized the origin, classification, epidemic history and population dynamics of QYYZ-like strains based on the ORF5 gene (18, 19). However, the prevalence and genomic characteristics of these strains in recent years remain unclear. In this study, we carried out molecular epidemiological investigations for QYYZ-like PRRSV surveillance from 2018 to 2021. Meanwhile, the genome-wide characteristics of QYYZ-like strains and the role of these strains in the PRRSV epidemic were explored by comparing the latest strains and all reported genome-wide sequences.

MATERIALS AND METHODS

Sample Collection and Genome Sequencing

From 2018 to 2021, we collected 1,803 clinical samples (including lung, lymph node and serum samples) of suspected PRRSV infection from different pig farms in 16 provinces in China (Heilongjiang, Jilin, Liaoning, Shandong, Henan, Guangdong, Guangxi, Zhejiang, Hebei, Hubei, Xinjiang, Inner Mongolia, Tianjin, Sichuan, Jiangxi and Jiangsu). Tissue sample processing, RNA extraction, cDNA preparation, RT–PCR and genome sequencing were performed as described in previous reports (8, 11). The primers used to detect PRRSV and amplify entire gene sequences were reported previously (20).

Sequence Analysis and Phylogenetic Analysis

Sequence analysis was performed with DNASTAR (version 7.1) software. Phylogenetic trees and molecular evolutionary analyses were conducted by MEGA 7 software using the neighbor-joining method with 1,000 bootstrap replications (21). The generated phylogenetic tree was annotated using the online software ITOL (https://itol.embl.de/) (22).

Recombination Analysis

To determine whether recombination screening occurred in the generation of QYYZ-like PRRSV strains, recombination events were considered only when supported by at least four of seven recombination detection algorithms (RDP, GENECONV, BootScan, MaxChi, Chimera, SiScan and 3Seq) in the Recombination Detection Program version 4.8 (RDP v.4.8). Finally, the pictures of recombination events were drawn by SimPlot v3.5.1 within a 500-bp window sliding along the genome alignment (20-bp step size).

RESULTS AND DISCUSSION

From 2018 to 2021, 1,803 clinical samples were collected from 16 provinces of China; 787 (43.64%) tested positive for PRRSV according to RT-PCR. Of the 787 positive samples, 191 were from central or southern provinces (Henan, Guangdong, Guangxi, Zhejiang, Hubei, Sichuan, Jiangsu, and Jiangxi), and the remaining 596 were from northern provinces (Heilongjiang, Jilin, Liaoning, Shandong, Hebei, Xinjiang, Inner Mongolia, and Tianjin) (Figure 1). Through ORF5 phylogenetic analysis, 24 samples were confirmed to have QYYZ-like PRRSV. The results showed that QYYZ-like PRRSV did not cause a pandemic in China but persisted during 2018-2021. Interestingly, almost all samples with QYYZ-like PRRSV were from central and southern provinces, including Henan (4), Guangdong (9), Zhejiang (1) and Guangxi (9), and only one QYYZ-like strain came from Heilongjiang Province (Table 1). The QYYZ-like strains accounted for approximately 12% (23/191) of the cases in the central and southern provinces (Figure 1). Therefore, QYYZlike PRRSV was mainly prevalent in central and southern China. To further study the complete genome characteristics of QYYZlike PRRSV in China, we selected 9 strains (HNLCL15-1903, GXXNF10-1803, GXXNF53-1805, GDXNF60-1805, GXXNF74-1806, GXXNF78-1806, GDXNF229-1811, HNTZJ1714-2011, GXTZJ2325-2112) from 24 newly identified QYYZ-like isolates based on large homology differences and different branches of an ORF5 phylogenetic tree for whole-genome sequencing. The genomes of these isolates were 15,008–15,316 nt in length, excluding 3' poly (A) tails. The significant difference in gene length between the newly QYYZ-like PRRSV strains may be due to the different deletion or insertion patterns in the Nsp2 region.

To evaluate the genomic characteristics of the newly identified QYYZ-like PRRSV strains, the genomes of the novel PRRSV isolates were compared with those of different lineage viruses, including JXA1 (lineage 8), NADC30 (lineage 1) and QYYZ (lineage 3) (**Table 2**). Genome alignment revealed that the



TABLE 1 Information on 24 strain	s of the newly identified QYYZ-like PRRSV.
------------------------------------	--

No.	Isolates	Accession no.	Time	Isolation source	Province	Gene region
1	GDXNF8-1802	ON462023	2018.02	Lung/lymph nodes	Guangdong	ORF5
2	GXXNF10-1803	ON462046	2018.03	Serum	Guangxi	Whole genome
3	GDXNF17-1804	ON462024	2018.04	Lung/lymph nodes	Guangdong	ORF5
4	GXXNF53-1805	ON462047	2018.05	Lung	Guangxi	Whole genome
5	GDXNF59-1805	ON462025	2018.05	Lung	Guangdong	ORF5
6	GDXNF60-1805	ON462048	2018.05	Lung	Guangdong	Whole genome
7	GXXNF74-1806	ON462049	2018.06	Lung	Guangxi	Whole genome
8	GXXNF75-1806	ON462026	2018.06	Lung	Guangxi	ORF5
9	GXXNF76-1806	ON462027	2018.06	Lung	Guangxi	ORF5
10	GXXNF77-1806	ON462028	2018.06	Lung/lymph nodes	Guangxi	ORF5
11	GXXNF78-1806	ON462050	2018.06	Lung/lymph nodes	Guangxi	Whole genome
12	GXXNF79-1806	ON462029	2018.06	Lung	Guangxi	ORF5
13	GDXNF134-1809	ON462030	2018.09	Lung	Guangdong	ORF5
14	ZJWK211-1809	ON462022	2018.09	Serum	Zhejiang	ORF5
15	GDXNF229-1811	ON462051	2018.11	Lung/lymph nodes	Guangdong	Whole genome
16	HNLCL15-1903	ON462043	2019.03	Serum/lung	Henan	Whole genome
17	HNLCL19-1904	ON462031	2019.04	Serum/lung	Henan	ORF5
18	HNLCL25-1905	ON462032	2019.05	Serum/lung	Henan	ORF5
19	HNLCL43-1906	ON462033	2019.06	Serum/lung	Henan	ORF5
20	GDHSW97-2001	ON462034	2020.01	Lung	Guangdong	ORF5
21	GDHSW100-2001	ON462035	2020.01	Lung	Guangdong	ORF5
22	HLJTZJ1224-2011	ON462036	2020.11	Lung	Heilongjiang	ORF5
23	HNTZJ1714-2011	ON462044	2020.11	Serum	Henan	Whole genome
24	GXTZJ2325-2112	ON462045	2021.12	Lung	Guangxi	Whole genome

ORF5 gene of 24 QYYZ-like PRRSV strains shared 91.2–96.0% nucleotide homology with that of QYYZ, which was higher than the homology shared with that of JXA1 (81.9–85.7%) and NADC30 (82.2–85.2%). The homology among the 24 QYYZ-like strains was 86.4–100% (**Table 2**), and most of the QYYZ-like PRRSV ORF5 genes had low homology. The complete

genome results also showed that the 9 newly identified QYYZlike PRRSV strains shared 84.6–96.1% identity with JXA1, 82.3– 90.8% identity with NADC30, and 83.6–93.8% identity with QYYZ (**Table 2**). Among them, the isolates GXXNF53-1805, GDXNF60-1805, GXXNF78-1806 and HNTZJ1714-2011 showed the highest identity (93.2, 95.4, 96.1, and 89.0%) with JXA1; TABLE 2 | Nucleotide and amino acid sequence similarity between the 24 new QYYZ-like PRRSVs and the reference strain.

Amino acids/nucleotides	JXA1	NADC30	QYYZ	QYYZ-like (newly)
5'UTR	-/96.8-98.9	-/91.0-93.1	-/94.1-95.2	-/94.1-99.5
Nsp1a	94.4-99.4/93.0-98.1	93.9-96.7/85.6-89.4	95.6-97.2/89.6-92.2	92.8-98.9/88.7-96.9
Nsp1β	87.6-95.5/89.3-97.0	73.7-75.8/75.7-80.2	79.7-82.7/81.5-84.3	84.7-95.0/88.3-96.2
Nsp2	70.9–96.3/76.0–97.3	68.6-88.1/73.6-90.4	68.0-87.3/74.7-91.4	67.3-92.9/73.4-94.3
Nsp3	86.1–99.1/81.3–98.1	87.0-96.5/80.6-93.6	87.8-98.3/80.9-94.3	86.1-99.1/80.1-96.1
Nsp4	90.7-100/81.9-98.5	89.2-93.1/81.4-93.8	91.7-97.5/83.5-95.4	90.2-99.0/81.0-97.5
Nsp5	87.6-95.9/80.2-97.1	86.5-96.5/80.2-94.3	87.6-97.1/80.6-95.3	86.5-95.3/78.8-94.1
Nsp6	93.8-100/85.4-100	93.8-100/83.3-97.9	93.8-100/83.3-97.9	93.8-100/81.2-100
Nsp7α	87.2-99.3/81.7-98.2	91.3-94.0/81.0-93.7	86.6-98.0/80.3-94.0	85.9-98.0/79.9-86.2
Nsp7β	71.8-100/77.9-99.1	70.9–90.9/77.6–92.7	75.5–94.5/78.5–94.5	69.1-100/75.8-98.2
Nsp8	93.3-100/88.1-98.5	91.1-95.6/83.0-94.8	93.3-100/88.1-94.1	88.9-100/83.7-97.8
Nsp9	96.6-99.5/85.4-98.9	96.0-98.3/85.3-95.0	95.0-98.1/84.4-92.5	94.9-99.1/84.3-97.5
Nsp10	93.7-99.5/84.5-98.6	94.6-97.7/84.8-95.1	93.0-99.8/84.6-98.1	92.7-99.3/84.1-97.2
Nsp11	92.8-98.2/85.9-95.7	91.9-96.3/84.0-91.2	93.7-97.8/86.1-96.7	91.5-96.9/83.3-96.9
Nsp12	96.1-100/86.4-98.9	92.9-96.1/85.3-90.1	93.5-99.4/85.7-97.4	92.9-99.4/84.0-98.1
ORF2a	86.0-90.3/87.8-92.0	87.2-88.3/84.2-87.4	92.6-96.5/91.7-97.0	88.7-94.6/89.0-94.7
ORF2b	86.5-91.9/91.0-94.6	83.8-91.9/88.7-91.4	91.9-98.6/94.1-96.4	87.8-94.6/90.5-94.6
ORF3	83.5-87.5/84.4-89.5	78.0-95.9/82.6-86.5	85.1–95.3/86.7–95.3	83.1-93.3/86.4-93.3
ORF4	85.5-96.1/84.7-94.4	88.3–91.6/86.0–88.5	87.2-96.1/86.2-95.2	87.2-94.4/84.5-92.7
ORF5	82.1-85.1/83.4-85.4	84.1-87.6/82.8-86.7	92.0-96.5/91.2-96.0	88.6-100/86.4-100
ORF5a	72.5-78.8/81.0-85.9	73.1-80.8/82.4-86.5	84.6-98.1/89.7-98.1	80.4-98.1/85.0-96.2
ORF6	94.3-96.6/89.0-91.2	92.0-95.4/88.2-90.9	96.0-98.9/92.8-97.5	95.4-98.9/90.3-96.4
ORF7	87.9-90.3/85.2-89.2	88.7-99.2/84.7-95.2	90.3-99.2/90.9-96.5	87.1-97.6/84.1-95.2
3'UTR	81.2-88.9	83.6–93.0	90.1–96.0	83.8–94.5
Whole genome	84.6–96.1	82.3–90.8	83.6–93.8	87.3–94.7

The ORF5 homology results were for 24 newly identified PRRSV strains, and the other regional homology results were for 9 whole-genome sequences.

the isolates GXXNF10-1803, GXXNF74-1806, GDXNF229-1811 and GXTZJ2325-2112 displayed the highest identity (87.8, 93.8, 91.0, and 86.3%) with QYYZ; and HNLCL15-1903 shared the highest identity with NADC30, but the homology was only 90.8%, suggesting that the newly identified QYYZ-like PRRSVs may have undergone large variation or recombination.

To establish genetic relationships between Chinese QYYZlike PRRSV strains and other PRRSV isolates, we constructed phylogenetic trees based on both the ORF5 gene and complete genomic sequence. Phylogenetic analysis based on ORF5 gene sequences demonstrated that all 24 new QYYZ-like PRRSV strains could be classified into sublineage 3.5 (Figure 2A). To analyze the genetic diversity of the novel QYYZ-like PRRSV in as much detail as possible, we expanded a data set (n =470) including all ORF5 sequences of lineage 3 collected from GenBank and submitted before December 2021. As shown in Figure 2B, all 24 newly identified QYYZ-like strains in China (deep red labeled) clustered deeply within sublineage 3.5 from mainland China and did not form a separate clade. They are distantly related to sublineages 3.1-3.3 circulating in Taiwan and sublineage 3.4 circulating in Hong Kong (Figure 2B). To understand the genome-wide characteristics of QYYZ-like strains in China, we collected the whole-genome sequences of all QYYZ-like strains (ORF5 classified into sublineage 3.5) from GenBank (n = 28) (Table 3) and submitted before December 2021 and constructed phylogenetic trees together with the new strains (n = 9). Since the homology of QY2010 (Accession: JQ743666.1) and QYYZ (Accession: JQ308798.1) was up to 100% and the collection time of the two strains was very close, we regarded them as one strain and expressed them as QYYZ (Table 3). As shown in Figure 2C, 37 isolates were clustered into several branches with viruses belonging to different lineages. A total of 21 strains (GXXNF53-1805, GDXNF60-1805, GXXNF78-1806, HNTZJ1714-2011, etc.) were closely related to JXA1 and Hun4 (HP-PRRSV sublineage 8.7), while ten strains (GXXNF10-1803, GXXNF74-1806, GDXNF229-1811, etc.) were clustered into a separate branch close to QYYZ (sublineage 3.5). Additionally, HNLCL15-1903, GXTZJ2325-2112, SCya18 and FJLIUY-2017 were closer to NADC30 (sublineage 1.8). There was a large difference between the ORF5 gene-based and whole genome-based phylogenetic analyses. The epidemiology of PRRSV has been investigated largely by sequencing the ORF5 gene and classifying virus lineages based on ORF5 phylogenetic analysis (5, 6, 33). With the increasing number of recombinant strains, it is necessary to sequence the whole genome of key strains in order to classify them.

Nsp2 is the most significant variable region, with remarkable mutations, insertions and deletions, and is recognized as an



FIGURE 2 | constructed based on the Nsp1 gene of 37 QYYZ-like PRRSV isolates and reference PRRSV strains from each lineage. The QYYZ-like prototype strain QYYZ is labeled with a red star (\bigstar). Newly obtained sequences in this study are labeled with red squares (\blacksquare). The QYYZ-like strains with complete genome sequences obtained from GenBank are labeled with red triangles (\blacktriangle).

TABLE 3 | Information on the whole genomes of all reported QYYZ-like PRRSV strains in China.

Таха	Isolation date	Recombination with	QYYZ-like regions	Accession no.	References
GXXNF10-1803	2018.03	JXA1	Nsp2-Nsp8, ORF2-7	ON462046	This study
GXXNF53-1805	2018.05	JXA1	ORF2-ORF7	ON462047	This study
GDXNF60-1805	2018.05	JXA1	ORF2-ORF7	ON462048	This study
GXXNF74-1806	2018.06	JXA1	Nsp2-Nsp7, NSP9-ORF7	ON462049	This study
GXXNF78-1806	2018.06	JXA1	ORF2-ORF3, ORF5-ORF7	ON462050	This study
GDXNF229-1811	2018.11	JXA1	Nsp2-Nsp5, Nsp9-ORF7	ON462051	This study
HNLCL15-1903	2019.03	SH/CH/2016+NADC30	Nsp10-ORF6	ON462043	This study
HNTZJ1714-2011	2020.11	JXA1+NADC30	ORF2-ORF7	ON462044	This study
GXTZJ2325-2112	2021.12	JXA1+NADC30	Nsp4-ORF7	ON462045	This study
QYYZ	2010	-	-	JQ308798	(12)
QY2010	2010	-	-	JQ743666	(23)
GM2	2011	VR2332	Nsp1-Nsp8, Nsp11-ORF7	JN662424	(12)
SH1211	2012	JXA1	Nsp12-ORF2, ORF5-ORF6	KF678434	(24)
FJFS	2012	JXA1	Nsp2-ORF7	KP998476	(19)
GD1404	2014	JXA1	ORF4-ORF7	KT961415	(25)
HiNZWQ	2014	JXA1	ORF3-ORF6	KY373215	Not found
XJzx1-2015	2015	CH-1a	ORF5-ORF7	KR080330	(26)
GDsg	2015	JXA1	Nsp2-Nsp9, NSP10-ORF7	KX621003	(27)
SDqd1501	2015	JXA1	Nsp2-Nsp8, Nsp11-ORF7	MN642099	(28)
GD-KP	2015	JXA1+VR2332	Nsp2-Nsp3, Nsp5-NSP7, NSP11-ORF7	KU978619	(19)
SCcd16	2016	NADC30+JXA1	ORF5-ORF7	MF196905	(14)
GDZS2016	2016	JXA1	ORF2-ORF6	MH046843	(13)
ZJnb16-2	2016	JXA1	ORF2-ORF7	MH174986	(15)
SH/CH/2016	2016	JXA1	NSP12-ORF7	KY495781	(29)
JX/CH/2016	2016	JXA1	NSP12-ORF7	KY495780	(30)
SD110-1608	2016	JXA1	Nsp2-Nsp9, Nsp11-ORF7	MK780825	(19)
GDYDZZZ	2016	JXA1	NSP12-ORF7	KY745901	Not found
GZgy17	2017	JXA1	ORF2-ORF6	MK144542	(29)
FJNP2017	2017	JXA1	ORF2-ORF7	MH046842	(13)
SCya17	2017	JXA1+NADC30	ORF3-ORF5	MH324400	(14)
FJLIUY-2017	2017	NADC30+BJ4+JXA1	ORF5-ORF7	MG011718	(16)
SDWH27-1710	2017	JXA1	Nsp2-ORF7	MK780824	(19)
FJDJQ-2017	2017	NADC30	ORF2-ORF7	MG011719	Not found
LN-DB87	2017	JXA1	ORF2-ORF7	MN046242	(31)
PRRSV2/CN/F1004/2017	2017	JXA1	ORF4-ORF7	MT416544	(32)
SCya18	2018	SH/CH/2016 +NADC30	NSP11-ORF6	MK144543	(29)
PRRSV2/CN/N9185/2018	2018	NADC30+JXA1	ORF4-ORF7	MT416542	(32)
GXNN202010	2020	JXA1	ORF2-ORF7	MW561593	Not found

Since the homology of QY2010 and QYYZ was up to 100% and the collection time of the two strains was very close, we regarded them as one strain and expressed them as QYYZ in the paper.

important target gene for analyzing the genetic variation and molecular epidemiology of PRRSV (34). Partial Nsp2 sequence alignment showed that 37 Nsp2 sequences (from complete genome sequences of sublineage 3.5 PRRSVs) were divided into 5 large patterns (**Figure 3A**). Compared to Nsp2 of VR2332, the Nsp2 proteins of pattern A strains had a deletion pattern that was identical to that of JXA1 (1 aa +29 aa). Pattern B strains not only have the same 30-amino acid (aa) deletion in Nsp2 as JXA1 (1 aa +29 aa) but also have a 36-aa insertion at position 817-852, which was the same as QYYZ (**Figure 3A**). Pattern C strains also have the same 36-aa insertion in Nsp2 as QYYZ (**Figure 3A**). Additionally, Pattern

							+ 491			ene .com		.561					
LDNNSVPL TAFS LANYYYRAQGDE VRHRER LTAV	/LSKLEKWREEYGLMPTE	PGPRPTLPRGLDELK	DQMEEDLLK LANAQ	TTSDMMAWAVEQVDLKTWVKNYP	RWTPPPPPPKVQP	SPVSLGGDVPNSWI	DLAVSSPFDLPTP	PEPATPSSELVIVS	PQCIFRPATPL	SEPAPIPAPRGTVSRPV	PLSEPIPVPAPRRKFQQV	RLSSAAAIP	GGAG			SFTDL	VR2332
LDNNSVPL TAFS LANYYYRAQGDE VRHRER LTAV	LSKLEKWREEYGLMPTE	PGPRPTLPRGLDELK	DOMEEDLLK LANAG	T TSDMMAWA VE QVDL KTWVKNYP A TSEMMAWA AE OVDL KAWYSYP	RWTPPPPPPKVQP	SPVSLGGDVPNSWI	DLAVSSPFDLPTP	PEPATPSSELVIVS	PQCIFRPATP	SEPAPIPAPRGTVSRPV	PLSEP IP VPAP RRKF QQ V	RUSSAAAIP	GGAG			SFTDL	RespPRRS MLV
SREDSVPL TAFS LSNCYYPAQGDE IHHRDR LNSV	/LRKLEEWLEEYGLMPAG	FGPRPVLPSGLDELK	DQMEDDLLKLANT(ATSEMMARTVEQVDLKAW/KKYP	RNTPPPPSPRVQP	SPDFL NGS	DLAIGGPLKFPTS	SEPVTSMSEPALTPA	SQYVPKLMTPL	SE LAPVPAPRR TVPPPM	PLSEPVFVSALRHKHQQVC	EVDLAATTL	GGTG			SFTDS	XJzx12015
LGKD SVPL TAFS LSNC YYPAQGDE VHHRER LNSV	/LSKLEEWLEEYGLMSTG	LGPRPVLPSGLDELK	DQMEEDLLKLANTQ DOMEEDLLKLANTQ	ATSEMMAWAAE QVDL KAWVKSYP ATSEMMAWAAE OVDL KAWVKSYP	RWTPPPPPPRVQP	SPVLMGDN/PNGS	E TVGGPLNFPTP	SEPMTPHSEPVLVP/	SRRVPKLMTPL	SGSAPVPAPRRTVT		TTL TTL	GGVG			SFTDL	JXA1 HIN4
GKDSVPL TAFS LSNCYYPAQGDE VHHR ER LNSV	/LSKLEEWLEEYGLMSTG	LGPRPVLPSGLDELK	DIMEEDLIKLADTO	ATS EMMAWA AE QVDL KAWVKS YP	RWTPPPPPPRVQP	SMVSVGDNVPNSSI	EE TVGGPLNFPTP	VEPITPHNEPVLVP	SRCVPKLMTP	SGSAPVPAPRR TVT		πι	CGVC			SSADS	GXXNF53-1805
LGKDSVPL TAFS LSNCYYPAQGDE VHHR ER LNSV LGKD LVPL TAFS LSNCYYPAQGDE VHHR ER LNSV	/LSKLEEVILEEYGLMSTG /LSKLEEWLEEYGLMSTG	LGPRPVLP SGLDELK	DOMEEDILLK LANTO DOMEEDILLK LVNAG	ATS EMMARAVE QVDL KAME KNYP ATS EMMARAAE OVDL KVWVKS YP	RWTPPPPPPRVQP RWTPPPPPPRVQP	SPVLMSDNVPNGST SPVSMGDDVPNGST	E TVGGPLNFPTP	SEPMTPHSEPVLMP/	SRRAPKLMIPL	SGSAPVPAPRRTVT		πι πι	SGVG			SFTDL SFTDL	GDXNF68-1885 GXXNF78-1886
GKDSVPL TAFS LSNCYYPAQGDE VHHRER LNSV	/LSKLEEWLEEYGL#STG	LGPRPVLPSGLDELK	DQMEEDLLKLTNTQ	ATS EMMAWAAE QVDL KAWVKS YP	RWTPPPPPPRVQP	S PV LMGD NV PNDS I	EE TVGGPLDFPTP	SEPMTPHSEPVLVP/	SRRVPKLMTPL	SG SA PVPAPRR TVT		<mark>πι</mark>	GGVG			SFTDL	GD 14 04
LGRDSVPL TAFS LSNGYYPAQGDE VLHRER LNAV LGRDSVPL TAFS LSNCYYPAQGDE VRHRER LNSV	/LSKLEEWLEEYGLMPTG /LSKLEEWLEEYGLMSTG	LGPRPVLP SGLDEL K	DKMEEDLLK LADAQ DOMEEDLLK LANTO	A TS DHMAWAAE RVNL KAWVKS YP A TS EMMAWAAE RVDL KAWVKN YP	RWTPPPPPP <mark>RVQP</mark> RWTPPPPPPRVOP	S LV SPVLMGDN/PNGSI	E TVGGPLNFPTP	SEPVTPHSEPVLMPA	SRRVSTLMIP SRRVPKLMTPI	SGSAPVPAPRR TVT		πι πι	GGVG			SFTDL	HINZHO GDZS2016
LGKDSVPL TAFS LSNCYYPAQGDE VHHR ER LNSV	/LSKLEEWLEEYGFMSTG	LGPRPVLPSGLDELK	DOMEEDLLKLTNTO	ATSEMMAWAAEQVDLKAWVKSYP	RWTPPPPPPRVQP	S PV LMGN NV PNGS I	E TVGGPLNFPTP	SEPMTPMSEPVLVPA	SRRVPKLMTPL	SGSAPVPAPRR TVT		<mark>π</mark> ι	GGVG			SFTDL	FJNP2017
LGKDSVPL TAFSLSNCYYPAQGDEVHHRERLNSV LGKDSVPL TAFSLSSCYYPAQGDEVRHRERLNSV	/LSKLEEW/LEEYGLMSTG /LSKLEEAV/LEEYGLMSAG	LGPRPVLPSGLDELK	DQMEEDLLKLTNTQ DQMEEDLLKLANTQ	A TSEMMAWA AE QVDL KAWVKSYP A TSEMMAWA AE QVDL KAWVKRYP	RWTPPPPPPRVQP RWTPPPPPPRVQP	S PV LMGD NV PNGS I S SV SMGD NV PNDS I	EE TVGGPLNLPTP EN TVGGLPNFPTL	PELMAPISEPVLVP/	SRRVPKLMTPL	SGSAPVPAPRRTVT SGSAPVPAPRRTVT		πι πι	GGVG			SFTDL SFTDF	E SH1211
LGKDSVPL TAFS LSNCYYPAQGDE VHHR ER LNSV	/LSKLEEWLEEYGLVPIG	LGPRPMLPSGLNELK	DQMEEDLLKLANTQ	ATS EMMAWAAE QVDL KAWVKS YP	RWTPPPSPPRVLP	slv		SEPVLVPV	SRCDPKLMTPS	GE SAPVPAPRK TVT		<mark>πι</mark>	DGAG			SSTDL	5H/CH/ 2016
GKDSVPL TAFS LSNC YYPAQGDEVHHRER LNSV .GKDSVPL TAFS LSNC YYPAQGDEVHHRDR LNSV	/LSKLEEWLEEYGLMSTG /LSKLEEWLEEYGLMSTG	LGPRPVLP SGLDDLK	DQMEKDLLE LANTO DQMEEDLLK LANTO	ASSEMMAWAAE QVDL KAWVKSYP ATSEMMARAAE QVDL KAWVKSYP	RWTPPPPPPRVQP	SPVLMQDNAPDGL1 SPVSMQDNVPNSS1	E TVGGPFNFSTS	SEPMITPHS EPVEVPV	SRRVPKPMTPL PRRAPKLMMPL	SVSAPIPAPRRTVT		ATL	6656			SFTDL	Z3nb16-2
GKGPVPL TAFS LSNC YYPAQGDE VHHR ER LNSV	LSKLEEWLEEYGLISTV	LGPRPVLPSGLSELK	DQMEEDLLK LANSO	T TS EMMARVAE QVDL KAWVKS YP	RWTPPPPLPRVQP	SPVS		EPVLVPA	GRRDPMLMTPS	SGSAPVPAPRK TVT		TML	GGVG			SFADL	<u>3X/CH/2016</u>
LGKDSVPL TAFS LSNCYYPAQGDE VHHR ER LNSV	LSKLEEWLEEYGF TP TG	LGPRPELPGGLDELK	DQMEEDLLKLTNTQ	A TS EMMAWA AE QVDL KAWVKS YP	RWTPPPPPPRAQP	SPILMDDNVSSGS	DE IVSGPLNFPTP	SOPMITLVSEPALVPA	SRRVPKLMTPL	SGSAPVPAPRKTVT		πι	GRAD			LLADL	SCya17
LGKNSIPL TAFSLSNCYYPAQGDE VHHRER LNTV LGKNSVPL TAFSLSNCYYPAQGDE VHHRER LNSV	/LSKLEEWLEEYGLVSTG /LSKLEEWLEEYGLISTG	LGSRPVLPSGLNELK	DQMEEDLLKLADTQ DOMEEDLLKLANTQ	A TS EMMAWA AE QVDL KAWVKNYP A TS EMMAWA AE QVDL KAWVKNYP	RWTPPPPPPRVQP RWTPPPPPTRVR	NPVLTGDNAPTGSE SPILVGDDVPNDRC	E TVGGPLDSPTP	SEPMTPL TRPVPVPA	PRIMTPI	SGSAPVPAPRRTVT		TL ΠL	GGVG			LSTDL	GZ gy 17 LND887
LGKDSVPL TAFS LSNCYYPAQGDE VRHR ER LNSV	/LSKLEEWLEEYGLMPTE	LGPRPALPSGLRELK	DOMEEDLLKLANTO	TTS ERMAWAAE QVDL KTWIKS YP	RWTPPPPPPRVQP	SPVLMGDNVPNGSI	E TVGGPLNFPTP	SEPMTPHSEPVLVP/	SRRVPKLMTPL	SGSAPVPAPRRTVT		<mark>πι</mark>	GGVG			PFTDL	PRRSV2/CN/N9:
LGKESVPLTAFSLSNCYYPAQGDEVHHRERLNSV +322	/LSKLEEWLEEYGLMSTG	LGPRPVLPSGLDEIK	DQMEEDLLK LANTQ	A TS EMMAWAAE QVDL KAWVKS YP	RWTPPPPPPRAQP +432	SPVLMGDNVPDGSI	<pre>CC TVGGPFNYPTP +483</pre>	SEPMTPHSEPVLVPV +SB4	SRRVPKLITPU	SGSAPVPALRRAAT		TTP	GGVS			SFTDL +852	<u>GXNN202010</u>
LGRDSVPL TAFS LSNCYYPAQGDE VHHR ER LNSV	/LSKLEEWLEEYGLMSTG	LGSRPVLPSGLNELK	DQMEEDLLK LANTQ	ATS EKMAWAAE QVDL KAWVKS YP	RWIPPPPPPRVQP	SPVLMGDNVPNGS	TVGGPLNFPTP	SEPMTPMSEPVLMP	SRRVPKLMTPU	SG SA PVPAPRR TVT		<mark>π</mark> ι	GGT GL VSDP GV	KRWLHEIGL	AK ENEQPL VPDDGT CA	SAPTSTEL	E SDgd1501
LGKDSVPL TAFS LSNCYYPARGNE VRHRER LNSV LGRDSVPL TAFS LSNCYYPAQGDE VHHRER LNSV	/LSKLEEVFLKEYGLTSIG /LSKLEEVVLEEYGLNSTG	PGLRSALPSGLDELK	DRMEEDLLKLVNAQ DQMEEDLLKLANTQ	A TSEQMA LAAE QVDLEAWVRGYP A TSEKMAWAAE QVDL KAWVKSYP	RWVPPLPPPRAQP RWIPPPPPPRVQP	S LVLMGDN/PNGS	G TVGGL LNFP TS	SEPMTPHSEPVPAP	SQAVPRLMTPL	SE LAPVPAPRRAVP		RLM TTL	GGTGL VSDPGV GGTGL VSDPGV	KRVVSHEIGL KRVVLHEIGL	AK GDEQPF VPNDGT CA: AK ENEQPL VPDDGT CA	SASTSTEL SAPTSTEL	GD-KP SD110-1608
LDRDSVPL TAFS LSNCYYPAQGDE VRHRER LNSV	/LSKLEEWLEEYGLTSVG	PGLRSALPSGLDELK	DQMEEDLLKLVNAQ	V TS EEMA SAAE RVDL KAWVQS YP	RWTPPLLPPIVQP	SPI LMGDNV LNGR	DSTVSGSFGLPAP	PEPMAPPGGPTLMPA	LQHEPKSTTPL	SGPAPVPAPRRTVPRLM	PLSEPVFVSAPRHEFRQVE	GVNL TATAL	GGT GL VSDP GV	KRVVSHEIGL	AK GDEQPF VPNDGT CA	SAPTSTEL	OY YZ
WKDSVPL TAFS LSNCYYPAQGDE VRHRER LNFV	LSKLEEVVLEEYGLISVG	PGERSALP SGEDEEK	DQMEEDLLK LVNAQ DQMEEDLLK LVNTQ	V TS EEMA SAAE KVDL KAWVQS YP V TS EKMA SAAE QVDL KVWVKS YP	RWTPPLLPPT VQ RWTPPPPPPRVQP	SLSLMGDNVLNGWE	EDST ISGP FS	PEPMAPPGGPTLMPA	LUHEPKSTIPU	SGPAP VPAP RR TV PR LM RGPV PVPAP RR TL SR LM	PLSEPVFVSAPRHEFRQV PPSEPILVSAPRHELQQV	GVNLTATAP	GGT GL VSDPGV	KRGVSHEIGL KRGVSHEIGL	WKRDEQPF VPNDGTCA	SAPTSTEL	C FJES
LDKDLVPL TAFS LSNCYYPAQGDE IHHRER LNSV	/LSKLEEWLEEYGLNSAE	FGPRPVLPSGLAELK	DQMEEDLLK LANTO	ATS EMMARAAE QVDL KTIWKNYP	RWAPSPPPPRVQP		SLT IGGPSGFPTS	PEPTTPMSESVPTR	PORVPKLMTP	SGSAPVPATRRTVSRPM	PLSEPVFSSTTRYKFQRV6	GTNPATTTL GVNLATTAP		KR VVSC GAGL	AK GDEQPC AP SGGV CA	SAPANNEQ	GXXNF18-1883
LGKDSVPL TAFS LSNCYYPAQGEE VHHRER LNSV	/LSKLEEVFLEEYGLTPVG	PGLRPALP SGLDEL K	DQVEEDLLKLVNAG	VTS EEMA SAAE RVDL KAWVQS YP	RWTPPPLPPRVQP	GPTQKGDDVPNGPG	KSVI	TANK FOOP IL MPY	PKLATL	GGPAPVPAPRR TVPR LM	PSSEPWUPAPQHEFQELE	GANNVTATL	LAPDPGV	EKVVSHEVGH	IAE GD KQLF VP NGDT CA	STPTSTES	GDXNF229-181
LGKDSVPL TAFS LSNCYYPAQGDEVHHRER LNSV PDRDSVPL TAFS LSNCYYPAQGDEVHHRER LNSV	/LSKLEEWLEEYGLTSTE	LDPRPVLPSGLGELK	DQMEEDLLK LANTO	ATS EMMANAAE QVDL KANWKGYP	RWTPPPPPTRFRP	SPI		EPMAPPGEL TOWN		SGPAPVPAPRRTVPRLT	PSSEPIFVFAPRHEFQQW PSSGPILLSADBWFEPOW	GANPMATTL			AT GDEQPE VP SC GV CA	PASPNTEL	GD 5g SDWH27-1710
					SPEFQS	SLVSSGGNFPDSR	DOSA GAPFHSPVL	.PK SVAR SNE		PVPVPAPRR TVSR PR	SPTVTTPAPAPRCGLQRV	GVNSAVGT L	EGLA			EPAGS	HNLCL15-1983
					SPEFQ SSPGF0	SLAALDGNSPDSWE SLVLLGDNFLGSW	DLA SGPFHSPVL	PESVARSDE		PVP IPAPRRTVFR LM	SPTASTPVPAPRRWLQQAG	KMDLAVLTP	GGPP			ELADL	E HNT231714-20
					SPVFLS	SLVSLGGNPPDSR	ENLA GGPFHFPGL	PEAVARSNE		PAPVPAPRRTVSQLK	SPTTSAPVPAPRHGLQQVE	GMNLAVGTL	GRSP			ELADL	FJLIUY2017
					SPEFQS	SLVSLGGNPPDSR	DSV-GAPFCSPVL	.PEAVARSNE		PAPVPAPRR TVSQLK	SPT TSAPVPAPRHOLQQV8 SPTMTAPAPAPRCGLQRV8	GVNSAVRAP	EGLP			EPAGS	SCya18
					SPEFQA	S LV SL GGNS PD SWI	EDLA <mark>-</mark> GGPFHSPVL	.PEPVARSNE		PVPVPAPRRTVSRLK	SPINSTPVPAPRCGLQQV8	GMDLAVGTL	GGSP			ELADL	F3203
Epite	ope A Epitor		main Epitope C	SEV TEDVI THI VSVCA	ITTSHELDTVAL	VTVSTACEVH	RVVI SS IVAN	T cell epito	pe 1		Iomain T cell epitope	2		B c	ell epitope	VP2	222
Epity EKCLTAGCYSQLLSLWCIVPFCFAVI	ope A Epitor	Ectodo	main Epitope C DWLANKFDWA1	RSFV IFPVLTHI VSYGA		VTVSTAGFVHG	RYVLSSIYAV	T cell epito CALAALTCFVII	pe 1 RFAKNCMSV	Endo	Iomain T cell epitope KGRLYRWRSPVIIE	2 Krgkveve	GHL IDLKRVV	B c LDGSVAT	ell epitope PITRVSAEQWGRP	VR2	332
Epite EKCLTAGCYSQLLSLWCIPPFCFAVL 10 20 EKCLTAGCYSQLLSLWCIPPFCFAVL	ope A Epitop LVNASNDSSSHLQLJ 30 40 LVNASNDSSSHLQLJ	Ectodo	main Epitope C DWLANKFDWA1 60 DWLANKFDWA1	RESEVIEPVLTHIVSYGA 70 80 RESEVIEPVLTHIVSYGA	LTTSHFLDTVAL 90 LTTSHFLDTVAL	VTVSTAGFVHG 100 VTVSTAGFVHG	RYVLSSIYAV 110 RYVLSSIYAV	T cell epito CALAALTCFVII +	pe 1 RFAKNCMSV 130 RFAKNCMSV	Endoo RYACTRYTNFLLD 140 RYACTRYTNFLLD	Iomain T cell epitope KGRLYRWRSPVIIE 150 160 KGRLYRWRSPVIIE	2 KRGKVEVE 1 KRGKVEVE	GHL IDLKRVV	B c LDGSVAT 80 LDGSVAT	ell epitope PITRVSAEQWGRP +	VR2 00 VR2	332 332
ERCLTAGCYSQLLSLWCIVPFCFAV 10 20 ERCLTAGCYSQLLSLWCIVPFCFAVI 6. Y. T. C. R	ope A Epitop LVNASNDSSSHLQLJ 30 40 LVNASNDSSSHLQLJ NSNQF NN T	Ectodo pe B IYNLTLCELNGT 50 IYNLTLCELNGT	main Epitope C DWLANKFDWA1 60 DWLANKFDWA1	ESFVIFPVLTHIVSYGA 70 80 TESFVIFPVLTHIVSYGA .T	UTTSHFLDTVAL 90 UTTSHFLDTVAL G. G	VTVSTAGFVH 100 VTVSTAGFVH VTVSTAGFVH VTVSTAGFVH VTVSTAGFVH	RYVLSSIYAW 110 RYVLSSIYAW	T cell epito CALAALTCFVII 120 CALAALTCFVII 120 CALAALTCFVII I	pe I RFAKNCMSV 130 RFAKNCMSV L	Endoe RYACTRYTNFLLD 140 RYACTRYTNFLLD SQ. S	Iomain T cell epitope KGRLYKWRSPVIIE 150 160 KGRLYRWRSPVIIE 	2 KRGK VEVE 1 KRGK VEVE . G	GHL IDLKRVV 	B c LDGSVATI 80 LDGSVATI	ell epitope PITRVSAEQWGRP +	VR2 00 VR2 CH1 IXA	332 -R 1
ExclTAGCYSQLLSLWCIVPFCFAV +	ope A Epitor UNASNDSSSHUQLJ 30 40 UVNASNDSSSHUQLJ 	Ectodo pe B IYNLTLCELNGT 50 IYNLTLCELNGT	main Epitope C DWLANKFDWAY 60 DWLANKFDWAY	ESFV IFPVLTHI VSYGA 70 80 ESFV IFPVLTHI VSYGA T. T. T.	TTSHFLDTVAL 90 LTTSHFLDTVAL 	VTVSTAGFVHG 100 VTVSTAGFVHG Y. AYY. AYY.	RYVLSS IYAW 110 RYVLSS IYAW	T cell epito CALAALTCFV II 120 CALAALTCFV II 120 CALAALTCFV II I. I.	PE 1 RFAKNCMSV 130 RFAKNCMSV L L	Endou RYACTRYTNFILD 140 RYACTRYTNFILD 	tomain T cell epitope KGRLYKWRSPVIIE 150 160 KGRLYKWRSPVIIE 	2 KRGKVEVE 1 KRGKVEVE . G . G	GHLIDLKRVV 	B c LDGSVATI 80 LDGSVATI	tell epitope PITRVSAEQWGRP 	VR2 VR2 CH1 JXA HuN	332 -R 1 4
ERCLTAGCYSQLLSLWCIPPGFAV 10 20 EXCLTAGCYSQLLSLWCIPPGFAVI G. Y. T. C. R. N	ope A Epitor LVNASNDSSSHLQLJ 30 40 LVNASNDSSSHLQLJ NSNQF NNI NNI S.NGNYS oven ve	Ectodo pe B IYNLTLCELNGT 50 IYNLTLCELNGT	main Epitope C DWLANKFDWA1 60 DWLANKFDWA1	ESFV IFPVLTHI VSYGA 70 80 TSFV IFPVLTHI VSYGA T. T. T. C. e v v	90 91 1755HFLDTVAL 	VTVSTAGFVH 100 VTVSTAGFVH Y. AYY. AYY. 	RYVLSSITAN 110 RYVLSSITAN	<u>T cell epito</u> CALAALTCFVII 120 CALAALTCFVII I I 	De 1 RFAKNCMSV 130 RFAKNCMSV L	Endou rtyactrytnfilld 140 rtyactrytnfilld 	domain T cell epitope KKGRLY RWRSPV I IE 150 160 KGRLY RWRSPV I IE	2 KRGKVEVE 1 KRGKVEVE G G G	GGHLIDLKRVV 70 1 3GHLIDLKRVV 3	B c LDGSVAT1 50 LDGSVAT1 A. A.	tell epitope PITRVSAEQWGRP 190 21 PITRVSAEQWGRP L L L	VR2 VR2 CH1 JXA HuN QYY	332 -R 1 4 2
ERGLATAGCYSQLLSLWCUPFGFAV EKCLTAGCYSQLLSLWCUPFGFAV 10 20 EKCLTAGCYSQLLSLWCUPFCFAV G., C, C, R, F,, YL, G., C, C, R, F,, YL, G., S., C, F, F,, STA, G. S., C, F, F, STA, G. S., C, F, F, STA, G. J, HS, F, STA,	ope A Epitor LVNASNDSSSHLQLJ 30 40 LVNASNDSSSHLQLJ NSNQF NNI S.NGNYS S.NGNYS V. NONYS	Ectodo pe B IYNLTLCELNGT 50 IYNLTLCELNGT	main Epitope C DWLANKFDWA1 60 DWLANKFDWA1 Q TN TN	ESFV IFFVLTHI VSYGA 70 80 ESFV IFFVLTHI VSYGA T. T. T. C. V. C. V. V. V. V. V. V. V. V. V. V. V. V. V.	TTSHFLDTVAL 90 LTTSHFLDTVAL G. G. G. S. S. S. S.	VTVSTAGFVH 100 VTVSTAGFVH YY. AYY. AH. AH.	RYVLSS I VAW 110 RYVLSS I VAW	T cell epito CALAALTCFV II 120 CALAALTCFV II 	De 130 RFAKINCMSV L L LV LV LV	Endou RYACTRYTNFLLD 140 RYACTRYTNFLLD 	Iomain T cell epitope KKGRLYRWRSPVIIE 150 160 KKGRLYRWRSPVIIE	2 KRGKVEVE 1 KRGKVEVE 6 7.	3GHL IDLKRVV ⊢	B c LDGSVAT1 80 LDGSVAT1 A. A. A. A.	tell epitope PITRVSAEQWGRP 190 24 PITRVSAEQWGRP LL.L.L.L.L.L.L.L.L.L.L.L.L.L.L.L.L	VR2 VR2 CH1 JXA HuN QYY GM2 SH1	332 -R 1 4 2 211
EKCLTAGCYSQLLSLWCIVPFCFAY 10 20 EKCLTAGCYSQLLSLWCIVPFCFAY 40 EKCLTAGCYSQLLSLWCIVPFCFAY 6C.C.R.FYLL. 6C.C.R.FYLL. 6S.C.F.F.STA 6LHS.F.STA 6I.C.RS.F.STA	ope A Epitor UVNASNDSSSHQL1 30 40 UVNASNDSSSHQL1 	Ectodo pe B IYNLTLCELNGT 50 IYNLTLCELNGT	main Epitope C DWLANKFDWA1 60 DWLANKFDWA1 Q TN TN TR	ESFV IFPVLTHI VSYGA 70 80 70 80 71 70 71 70 71 70 71 70 70 70 70 70 70 70 70 70 70 70 70 70 7	90 LTTSHFLDTVAL 90 LTTSHFLDTVAL 6. 	VTVSTAGFVH 100 VTVSTAGFVH 4Y. AYY. AYY. A.H. YY. A.H.	RYVLSS IYAW 11D RYVLSS IYAW	T cell epito CALAALTCFV II 120 CALAALTCFV II 	De 1 130 RFAKNCMSV L L L L L L L L L L	Endor RYACTRYTNFLLD 140 RYACTRYTNFLLD S	Iomain T cell epitope ISO 150 ISO KIRRLYRWRSPVIE V. V. V. V. V. V. V. KI. KI. V.	2 KRGK VE VE G G G G G G C	3GHL IDLKRVV → → → → → → → → → → → → → → → → → → →	B c LDGSVAT1 80 LDGSVAT1 A. A. A. A.	tell epitope PTTRVSAEQWGRP 190 21 PITRVSAEQWGRP LL.L.L.L.L.L.L.L.V.KIR.H. V.KIR.H. V.KIR.H. V.KIR.H.	VR2 VR2 CH1 JXA HuN QYY GM2 SH1 FJF	332 -R 1 44 211 S
ERIL EKCLTAGCYSQLLSLWCIVPFCFAV 10 EKCLTAGCYSQLLSLWCIVPFCFAV 40, Y. T. C. R. 6, C. C. R. F. YLL. 6, C. C. R. F. YLL. 6, C. C. R. F. SIA. 6, HS, F. SIA. 6, HS, F. SIA. 6, KS, F. SIA. 6, KS, F. SIA.	Ope A Epitor UVNASNDSSSHLQLJ 30 40 UVNASNDSSSHLQLJ 	Ectodo pe B IYNLTLCELNGT 50 IYNLTLCELNGT	main Epitope C DWLANKFDWAY 60 DWLANKFDWAY Q TN TN TR TN SK.	SFV1FPVLTHLVSYGA 70 80 PSFV1FPVLTHLVSYGA T. T. C. C. C. C. C. C. C. C. C. C. C. C. C.	90 90 175HFLDTVAL 	VTVSTAGFVH 100 VTVSTAGFVH AYY. AYY. AYY. AH. YY. A. H. YY. A. H. YY.	RYVLSSJYAW 11D RYVLSSJYAW	<u>T cell epito</u> CALAALTCFVII + 120 CALAALTCFVII I.I. F.I.I.I. F.I.I.A. F.I.I.F. FV.	De 1 130 RFAKNCMSV L L LV	Endor RYACTRYTNFLLD 140 RYACTRYTNFLLD S	domain T cell epitope KGRLYRØRSPI 116 KGRLYRØRSPI 116 KGRLYRØRSPI 116 KI KI KI KI KI	2 KRGKVEVE G	BGHLIDLKRVV → → → → → → → → → → → → → → → → → → →	B c LDGSVAT1 80 LDGSVAT1 A. A. A. A. A. A.	tell epitope + 178VSAEQWGRP + 190 22 PITRVSAEQWGRP LL LL V.KH V.KH V.KH V.KH V.KH V.KH	VR2 VR2 CH1 JXA HuN QYY GM2 SH1 FJF GD1 HiN	332 -R 1 44 2 2 11 5 404 2004 2700
Epit EKCLTAGCYSQLLSLWCIVPPCFAV 10 20 EKCLTAGCYSQLLSLWCIVPCFAV G. Y. T. C. R	ope A Epitor two.syndossitu.el. a0 40 two.syndossitu.el. . NSN. off. . NN. I. . NN. J. . NN. J. . NN. YS. . S. NON. YS. . S. NY. YS.	Ectodo pe B IYNLTLCELNGT	main Epitop C DWLANKFDWA 60 DWLANKFDWA Q Q TN TN TR TR TR SK SK SK	ESFV IFPVLTHI VSYGA 70 80 TSFV IFPVLTHI VSYGA T	ITSHFLDTVAL 90 TTSHFLDTVAL 6. 6. 6. 6. 7. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8.	VTVSTAGFVH 100 VTVSTAGFVH 	RYVLSSJYAW 11D RYVLSSJYAW	T cell epito CALAALTCFV I 120 CALAALTCFV I F. I. I. F. I. I. F. I. I. F. I. I. F. T. V. F. V. I.	De 1 RFAKNCMSV 130 RFAKNCMSV L L LV LV LV LV LV LV LV LV LV LV LV LV	Endoo RYACTRYTNFLLD 140 RYACTRYTNFLLD 	domain T cell epitope (KGR1YKRSPY11E 150 160 (KGR1YRRSPY11E V. V. V. KI. KI. KI. KI. KI. KI. KI. KI	2 KRGK VEVE 1 KRGK VEVE 6	EGHLIDLKRVV → → → → → → → → → → → → → → → → → → →	B c LDGSVAT1 50 LDGSVAT1 A. A. A. A. A. A. A. A. A.	tell epitope PITRVSAEQWGRP 190 20 PITRVSAEQWGRP LL.L LL.L V.KI.R.H. V.KI.R.H. V.KI.R.H. V.KI.R.H. V.KI.R.H. V.KI.R.H. V.KI.R.H. V.KI.R.H.	VR2 VR2 CH1 JXA HuN QYY GM2 SH1 FJF GD1 HiN XIz	332 -R 1 4 2 2 111 <u>S</u> 404 2 WQ x1-2015
EXCLTAGCYSQLLSLWCUPFGCAY EXCLTAGCYSQLLSLWCUPFGCAY 6. Y. T. C. R	ope A Epitor two.syndossitu.gl. a0 40 two.syndossitu.gl. . NSN. qF. . NN. , I. . NN. , I. . NN. , YS. . NON. , YS. . S. NYN. , YS.		main Epitope C DWLANKFDWA 60 DWLANKFDWA Q TN TR TR SK G TN.N. TN.N.	25FV IFPVLTHI VSYGA 70 80 DSFV IFPVLTHI VSYGA 7	UTTSHFLDTVAL 90 UTTSHFDJTVAL G. G. G. S. S. S. S. S. S. S. S. S. S. S. S. S.	VTVSTAGFVH 100 VTVSTAGFVH 	RYVLSS YAW 110 RYVLSS YAW	T cell epito ccaLALTCFVI 120 CCALALTCFVI I. I. I. I. F. I. I. F. I. I. F. T. V. F. T. V. F. T. V. F. I. I. F. I. I.	De I 130 RFAKNCMSV LL. LL. LVLV. LVLV. LVLV. LV.	Endou RYACTRYINFLLD 140 RYACTRYINFLLD S	domain T cell epiope T cell epiope 150 160 KI	2 KRGK VEVE 6 . G 6 . G 6 	3GHLIDLKRVV 70 1 3	B c LDGSVAT1 50 LDGSVAT1 A.	tell epitope PITRVSAE0wGRP 190 24 PITRVSAE0wGRP LL. LL. V.KIR.H. V.KIR.H. V.KIR.H. V.KIR.H. V.KIR.H. V.KIR.H. V.KIR.H. V.KIR.H. V.KI.R.R.H. V.KI.R.R.H. V.KI.R.R.H. V.KI.R.R.H. V.KI.R.R.H. V.KI.R.R.H. V.KI.R.R.H. V.KI.R.R.H.	VR2 OH UN UN UN UN UN UN UN UN UN UN UN UN UN	332 -R 1 4 2 2 11 5 404 2 8 404 2 8 409 2 8 409 2 8 409 2 8 41501
ERCLTAGCYSQLLSLWCUPPCFAY 10 20 EXCLTAGCYSQLLSLWCUPPCFAY 4 EXCLTAGCYSQLLSLWCUPPCFAY 6, C, C, R, F, YL, 6, C, C, R, F, YL, 6, C, C, R, F, STA, 6, C, C, R, F, STA, 6, C, RS, F, STA, 6, L, RS, F, STA, 6, L, RS, F, STA, 6, LQF, F, STA, 6, LQF, F, STA, 6, LQF, F, STA, 6, C, C, F, STA, 6, LQF, F, STA, 6, C, C, F, STA, 6, C, C, F, STA, 6, STA, STA, STA, STA, STA, STA, STA, STA,	ope A Epitor 100		main Epitope C DULANKEDWA 60 DULANKEDWA Q Q TN TN TN SK G TN. N. TN. N. TN. N.	25FV IFPVLTHI VSYGA 70 80 SEVV IFPVLTHI VSYGA T	UTTSHFLDTVAL 90 UTTSHFDDTVAL G. G. G. G. G. G. G. G. G. G. G. G. G.	VTVSTAGFVH + 100 VTVSTAGFVH - - - - - - - - - - - - -	RYVLSSTYAW 110 RYVLSSTYAW	T cell epito cCaLALTCFVI 120 CCALALTCFVI I. I. F. I. I. F. I. I. F. I. I. F. T. V. F. T. V. F. V. I. F. I. I. F. I. I. F. V. I. F. I. I. F. I. I. F. V. I. F. I. I. F. I. I. F. I. I. F. I. I. F. I. I. F. V. I. F. I. I. F. I. I. F. I. I. F. V. I. F. I. I. F. I. I. F. V. V. F. I. I. F. I. I. F. I. I. F. V. V. F. I. I. F. I. I. F. V. V. F. V. V. F. I. I. T. T. T. T. T. T. T. T. T. T	De 1 RFAKNCMSV 130 RFAKNCMSV L L LV	Endoo RYACTRYTNFLLD + + + - - - - - - - - - - - - -	Jomain T cell epitope 150 160 160 160 KI V KI KI	2 KRGKVEVE G	3GHLIDLKRVV 70 1 3	B c LDGSVAT1 50 LDGSVAT1 A. A. A. A. A. A. A. A. A. A. A. A. A. A.	tell epitope PTTRVSAEQWGRP 190 24 190 24 1	VR2 CH1 JXA HuNY GM2 SH1 FJF GD1 HiN XJz GDs SDq GD-	332 -R 1 4 4 211
ERLTAGCYSQLLSLWCUVPFCFAY, 10 20 ERCLTAGCYSQLLSLWCUVPFCFAY, 40, T. T. C. R. 6, C. C. R. F. YLL., 6, C. C. R. F. YLL., 6, C. C. R. F. STA, 6, C. F. STA, 6, C. R. F. STA, 6, MS, F. STA, 6, MS, F. STA, 6, RS, F. STA, 6, CL, F. F. STA, 6, CL, F. F. STA, 6, CL, F. STA, 6, CL, F. STA, 6, CL, F. STA, 6, CL, F. STA, 6, S,	Ope A Epitop 80 40 80 40 80 40 80 40 80 50 80 50 80 50 80 40 80 50 80 50 80 50 80 50 80 50 80 50 80 50 90 50 90 50 90 50 90 50 90 50 90 50 90 50 90 50 90 50 90 50 90 50 90 50	Ectodo pe B I YNLTLCELNGT 50 I YNLTLCELNGT	main Epitope C DWLANKFD#A 60 DWLANKFD#A Q Q TN TR SK G TN TN TR SK G TN TN A DR	ESFV IFPVLTHI VSYGA + + 70 80 ESFV IFPVLTHI VSYGA T. C. C. V C. C. C. C. C. C. C. C. C. C.	TTSHFLDTVAL 90 TTSHFJDTVAL 6. 6. 6. 6. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7.	VTVSTAGFVE 100 VTVSTAGFVE AYY. AYY. AH. YY. AH. TA.H. IC. TY. I.A.H. AH. AH. Y. AH.	RYVLSSJYAW 110 RYVLSSJYAW I I.	T cell epititic CALAALTCFVII + 120 CALAALTCFVII -	De 1 RFAKNCMSV 130 RFAKNCMSV L L LV L	Endov RYACTRY INFILLD 140 RYACTRY INFILLD S	Iomain T cell epitopie T cell epitopie 160 T cell epitopie 170 T cell epitopie 180 T cell epitopie 18	2 KRGK VE VE 1 KRGK VE VE 6	BGHLIDLKRVV 70 1. 3	B c LDGSVAT1 50 LDGSVAT1 A.	ell epitope PTTRVSAEQWGRP 190 22 190 22 190 24 190 24 1	VR2 CH1 JXA HuN GM2 SH1 FJF GD1 HiN XJz GDs SDq GD- GD- GD- GD2	332 332 1 4 <u>7</u> <u>8</u> <u>404</u> <u>7</u> <u>8</u> <u>404</u> <u>7</u> <u>8</u> <u>404</u> <u>7</u> <u>8</u> <u>404</u> <u>7</u> <u>8</u> <u>404</u> <u>7</u> <u>8</u> <u>404</u> <u>7</u> <u>8</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>404</u> <u>7</u> <u>8</u> <u>1</u> <u>1</u> <u>1</u> <u>1</u> <u>1</u> <u>1</u> <u>1</u> <u>1</u> <u>1</u> <u>1</u>
ERIT EKCLTAGCYSQLLSLWC1VPFCFAV 10 20 EKCLTAGCYSQLLSLWC1VPFCFAV 0, Y. T. C. R G C. C. R. F G C. C. R. F G C. C. R. F G	Ope A Epitop 10 HASDDSSSHL01. 10 HASDDSSSHL01. 10 HASDDSSSHL01. 10 HASDDSSSHL01. 10 HASDDSSHL01. 10 HASDDSSHL01. 10 HASDDSSHL01. 10 HASDDSSHL01. 11 HASDDSSHL01. 12 HASDDSSHL01. 13 HASDDSSHL01. 14 HASDDSSHL01. 15 HASDDSSHL01. 16 HANDLSSHL01. 17 HASDDSSHL01. 18 HANDLSSHL01. 19 HASDDSSHL01. 10 HASDDSSHL01. 11 HASDDSSHL01. 12 HASDDSSHL01. 13 HANDLSSHL01. 14 HASDDSSHL01. 15 HANDLSSHL01. 16 HANDLSSHL01. 17 HASDDSSHL01. 18 HANDLSSHL01. 19 HANDLSSHL01. 19 HANDLSSHL01. 19 H	Ectodo pe B IYNLTLCELING 50 IYNLTLCELING	main Epitope C DBLANKFDWA 60 DBLANKFDWA Q TN TN TN SK G TN.N. SK G TN.N. A A A	ESFV IFPVLTHI VSYGA T 4 4 70 80 ESFV IFPVLTHI VSYGA T	LTTSHFLDTVAL 900 LTTSHFLDTVA 5, 6, 6, 7, 7, 8, 7, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8,	VTVSTAGFVH 100 VTVSTAGFVH Y. AYY. AYY. AYY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. AH. YY. YY. AH. YY. AH. YY. AH. YY. AH. YY. YY. AH. Y.	RYVLSS YAW 110 RYVLSS YAW	T cell epititic CALAALTCFVII + 120 CALAATCFVII -	DC 1 RFAKNCMSV 130 RFAKNCMSV L	Endow RYACTRYTNFLLD 140 RYACTRYTNFLLD S	Iomain T cell epispe More and the second	2 KRGK VE VE G	BGHLIDLKRVV → → → → → → → → → → → → → → → → → → →	B c LDGSVAT1 30 LDGSVAT1 A.	ell cpitope PTTRVSAEQWGRP 1900 24 PTTRVSAEQWGRP LL LL LL V.KI., R.H. V.KI., R.H.	VR2 VR2 CH1 JXA HuN <u>VYY</u> <u>GM2</u> SH1 FJF GD1 HIN XJ2 GD2 GD2 CC CD2 ZJn	332 332 R 1 4 4 Z 2 111 S 2111 S 2111 S 3 404 404 404 404 404 404 405 8 1 1 5 2015 8 8 1 1 5 2016 1 5 2016 1 8 1 8 1 1 8 1 1 8 1 1 1 8 1 1 1 1 8 1
ERIT EKCLTAGCYSQLLSLWCIVPFCFAV 10 20 EKCLTAGCYSQLLSLWCIVPFCFAV, G. Y. T. C. R G C. C. R. F. YLL. G C. C. R. F. YLL. G C. F. F. STA. G HS. F	OPE A E-Filip 50 40 40 40 40 40 40 40 40 40 40 40 40 40 40	Ectodo pe B iyn, tlceling 50 iyn, tlceling	main Epitope C DWLANKEDWA 60 DWLANKEDWA 9 7.N 7.N 7.N 8 7.N 7.N 7.N 9 7.N	ESFV IFPVLTH VSYGA 70 80 TSFV IFPVLTH VSYGA T C C V.C. V.C. C C. C. C. C. C. C. C. C. C. C. C	TTSHFLDTVAL 90 TTSHFLDTVAL 6, 6, 6, 5, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8,	VTVSTAGFVH 100 VTVSTAGFVH 	RYVLSS IYAW 110 RYVLSS IYAW 110 110 110 110 110 110 110 11	T cell spitou CALAALTCFV II 120 CALAALTCFV II . I. . I. . F. I. I. . F. I. I. . F. I. I. . F T. V. . F T. V. . F. I. V. . F. I. V. . F. I. . T. . T.	DE 1 RFAKNCMSV 130 RFAKNCMSV L	Endou RYACTRYINFLLD 140 S	Iomain T cell epitype 150 160 150 160 160 160 17 V. V. V. XI. XI.	2 KRGKVEVF 1 1 KRGKVEVF G	SGHL IDLKRVV 70 1. 3GHL IDLKRVV 3	B c LDGSVAT1 30 LDGSVAT1 A. 	CII epitope PITRVSAEQWGRP PITRVSAEQWGRP LL. LL. LL. V.KI. R. H. V.K R. H. V.K R. H. V.K R. H. V.K.I. R. H. V.KI. R. R. H. V.KI. R. H.	VR2 VR2 CH1 JXA HuN <u>VYY</u> <u>GM2</u> SH1 FJF GD1 HIN XJ2 GD2 GD2 ZJn SU/ SV/ VY/	332 -R 1 4 Z - 211 221 230 240 412015 § 404 40501 52006 41501 52006 52006 516-2 CH/2016 CH/2016
ERCLTAGCYSQLLSLWCUTVPECAV 10 EKCLTAGCYSQLLSLWCUTVPECAV 4 EKCLTAGCYSQLLSLWCUTVPECAV G. Y. T. C. R G C. C. R. F. YLL G C. C. R. F. YLL G C. C. R. F. STA G. S. C. F. F. STA G. Y. RS. F. STA G. H. S. F. STA G C. RS. F. STA G G. F. STA G S. STA S	Ope A Epitop 100 Sciences Sciences 101 Sciences Sciences 102 Sciences Sciences 103 NSN. NS. 104 NSN. NS. 105 NSN. NS. 106 NSN. NS. 107 NSN. NS. 108 NNN. NS. 109 NSN. NS. 100 NSN. NS. 100 NSN. NS. 101 NSN. NS. 102 NSN. NS. 103 NSN. NS. 104 NSN. NS. 105 NSN. NS. 106 NSN. NS. 107 NSN. NS. 108		main Epitope C DMLANKFDWA 60 DMLANKFDWA Q TN TN TR SK G TN.N A A A A T A T A T A T A	25FV IFPVLTHI VSYGA 70 80 15FV IFPVLTHI VSYGA T CV CV CV CV CC.	TTSHFLDTVAL 90 TTSHFLDTVAL 6. 6. 6. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8.	VTVSTAGFVH 10 VTVSTAGFVH 	RYVLSS / YAW 110 RYVLSS / YAW 	T cell epito CALAALTCFV II 120 CALAALTCFV II . I. . I. . I. . I. . I. . I. . I.	DE 1 RFAKNCMSV 130 RFAKNCMSV L L LV	Endov RYACTRYINFLLD 140 140 SQ. SQ. SS.	Iomain T cell epilope H 150 H 150 KI	2 1 1 1 1 1 1 1 1 1 1 1 1 1	3GHL IDLKRVV 70 1 3	B c LDGSVAT1 30 LDGSVAT1 A. 	Vell epitope P1TRVSAEQWGEP P1TRVSAEQWGEP P1TRVSAEQWGEP LL. LL. V.K.L. R. H. V.K.L. R. H.	VR2 CH1 JXA HuN QYY GM2 GM2 GM2 GM2 GM2 GM2 GM2 GM2 GM2 GM2	332 332 -R 1 4 Z 211 200 200 200 200 200 200 200
EKCLTAGCYSQLLSLWCIVPFCFAY + 20 EKCLTAGCYSQLLSLWCIVPFCFAY + 20 EKCLTAGCYSQLLSLWCIVPFCFAY G C. C. R. F YL G C. C. R. F YL G C. C. R. F YL G F	Ope A Epitor 50 40 50 40 50 40 50 50 50 40 50 50 50 40 50 50 50 50 50 40 50 50	Ectodo	main Epitope C DBLANKED#A 0.0	ESFV IFPVLTHI VSYGA + + 70 80 ESFV IFPVLTHI VSYGA T T GV GV GV GV GV GV G	TTSHFLDTVAL 90 TTSHFLDTVAL 6. 6. 6. 7. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8.	VTVSTAGFVH 100 VTVSTAGFVH 	RYVLSS YAW 110 RYVLSS YAW	T cell epido CALAALTCFYI 120 CALAALTCFYI I. CALAALTCFYI I. F. I. F. F. F. T. T. F. T. T. F. T. T. F. T. T. F. T. T. F. T. T. T. F. T. T. T. F. T. T. T. F. T. T. T. T. T. T. T. T. T. T. T. T. T.	DE 1 130 130 RFAKOCUSY 1 L. 1 LV 1	Endov RYACTRYINFLLD 140 RYACTRYINFLD S	Jomain T cell epilope H cell epilope H cell service H cell service Kil	2 1 1 1 1 1 1 1 1 1 1 1 1 1	3GHL IDLX RVV 70 1 3 K. 3 K. 3 K. 3 K. 3 K. 3 K. 3 K. 3 K. 3 K. 3	B c LDGSVAT1 30 LDGSVAT1 A. AA. AA. A. A. A. A. A. A. A. A. 	tell epitope P1TRVSAEQWGRP 190 21 P1TRVSAEQWGRP L1 L1 L1 V.KI	VR2 VR2 (H1 JXA HuM VY YV YV YV YV YV YV YV YV YV YV YV XV SD 0 0 0 2 CC C 0 0 Z C C 0 0 Z C C 1 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 1 XA 1 XA 1 XA 1 1 XA 2 XA 1 XA 1	332
ERIT EKCLTAGCYSQLLSLWCIVPFCFAY 10 EKCLTAGCYSQLLSLWCIVPFCFAY 0, Y. T. C. R. G., C. C. R. F. YLL., G., C. C. R. F. YLL., G., C. C. R. F. STA, G. S. C. F. F. STA, G. J HS, F. STA, G HS, F. STA, G RS, F. STA, G ST, STA, G ST, STA, G	Ope A Epitor 50 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 40 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50 100 50	Ectodo	main Epitope C DBLANKEDJAA 0 Q Q TN TN TR SK G TN.N. N.N. N.N. A. A. A. A. A. A. A. A. A. A. A. A. A.	ESFV IFPVLTH USYGA + + + 70 80 ESFV IFPVLTH USYGA T CV C	DTSHFLDTVAL 90 UTSHFLDTVAL 6	VTVSTAGFVH 100 VTVSTAGFVH 4YY. 4YY. 4YY. 4YY. 4YY. 4YY. 4YY. 4YY. 4YY. 4Y.	RYVLSS I YAW 110 RYVLSS I YAW . I.	T cell epito CALAALTCF1 120 CALAALTCF1 I. . I. . I. . I. . I. F. I. F. I. F. I. I.	PE I ISO ISO ISO ISO IL ISO IL ISO IV ISO	Endov RYACTRYTNFLLJ 140 RYACTRYTNFLJ S	Iomain T cell episepe KGRL 18985971 II: 150 160 NSRL 19985971 II: 150 170 NSRL 1998 NSRL 1997 NSRL	2 RRGKVEVE 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	SGHL IDLKRVV 	B c C LDGSVAT 80 LDGSVAT A. 	cell epitope PITRVSAEQUCEP image: point of the second	VR22 VR2 VR2 VR2 VR2 VR2 VR2 VR2 VR2 VR2	332 -R 1 2 2 211 5 404 404 404 2 2 2 2 2 2 1 5 404 404 404 5 5 405 1 5 404 404 404 1 5 404 404 1 5 404 404 1 5 404 405 1 1 1 1 1 1 1 1 1 1 1 1 1
ERIT EKCLTAGCYSQLLSLWCIVPFCFAY 10 20 EKCLTAGCYSQLLSLWCIVPFCFAY 0, Y. T. C. R G C. C. R. F YL G C. C. R. F YL G C. C. R. F G S C. F. F STA. G R. F G G G G G STA. G	Ope A Epitop 10 Hold Status 11 Hold Status 12 Hold Status 13 Hold Status 14 Hold Status 15 Hold Status 16 Hold Status 17 Hold Status 18 Hold Status <t< td=""><td>Ectodo</td><td>main Epitope C DRLANKED#A 60 DRLANKED#A Q Q Q Q TN TN TN TN TN TN TN TN A DR A A A A A A A A B</td><td>ESFV IFPVLTHI VSYGA T</td><td>TTSHFLDTVAL 90 TTSHFJLDTVAL 6,, G, 5,, G, 5,, S, 8,, S, 8,, S, 9,, S, 9,</td><td>VTVSTAGFVE 100 VTVSTAGFVE AYY. AYY. AH. IC. YY. AH. AY. AY. AY. AY. AY. AY. AY. AY. AY. AA. AH. AH. AH. AH. AH. AH. AH. AH. AY.</td><td>RYVLSS YAW 110 RYVLSS YAW . I. . I.</td><td>T cell epito CALAALTCFY I 120 CALAALTCFY I 120 CALAALTCFY I </td><td>DE 1 ISO ISO ISO IL IL IL IV IV</td><td>Endov RYACTRYINFLLD 140 RYACTRYINFLLD S</td><td>Iomain T cell epitope KIGU.109859V1 IE KIGU.109859V1 IE KIGU.109859V1 IE KIGU.109859V1 IE KIGU.1000 KIG</td><td>2 RRGKVEVF 1 1 1 RRGKVEVF 6 6 6 6 6 6 6 6.</td><td>SGHL IDLKRVV </td><td>B c C LDGSVAT 30 LDGSVAT A. A. </td><td><pre>cll epitope PITRVSAE0wGRP ITRVSAE0wGRP L</pre></td><td>VR2 VR2 VR2 VR2 VR2 VR2 VR2 VR2 VR2 VR2</td><td>332 332 -R 1 4 7 2 2 2 11 5 2 4 4 4 4 4 4 4 4 4 4 4 4 4</td></t<>	Ectodo	main Epitope C DRLANKED#A 60 DRLANKED#A Q Q Q Q TN TN TN TN TN TN TN TN A DR A A A A A A A A B	ESFV IFPVLTHI VSYGA T	TTSHFLDTVAL 90 TTSHFJLDTVAL 6,, G, 5,, G, 5,, S, 8,, S, 8,, S, 9,, S, 9,	VTVSTAGFVE 100 VTVSTAGFVE AYY. AYY. AH. IC. YY. AH. AY. AY. AY. AY. AY. AY. AY. AY. AY. AA. AH. AH. AH. AH. AH. AH. AH. AH. AY.	RYVLSS YAW 110 RYVLSS YAW . I. . I.	T cell epito CALAALTCFY I 120 CALAALTCFY I 120 CALAALTCFY I 	DE 1 ISO ISO ISO IL IL IL IV	Endov RYACTRYINFLLD 140 RYACTRYINFLLD S	Iomain T cell epitope KIGU.109859V1 IE KIGU.109859V1 IE KIGU.109859V1 IE KIGU.109859V1 IE KIGU.1000 KIG	2 RRGKVEVF 1 1 1 RRGKVEVF 6 6 6 6 6 6 6 6.	SGHL IDLKRVV 	B c C LDGSVAT 30 LDGSVAT A. 	<pre>cll epitope PITRVSAE0wGRP ITRVSAE0wGRP L</pre>	VR2 VR2 VR2 VR2 VR2 VR2 VR2 VR2 VR2 VR2	332 332 -R 1 4 7 2 2 2 11 5 2 4 4 4 4 4 4 4 4 4 4 4 4 4
ExcLTAGCYSQLLSLWCUPFCFAV EXCLTAGCYSQLLSLWCUPFCFAV G. Y. T. C. R G C. C. R. F G C. C. R. F G C. C. R. F G C. C. R. F G C. C. R. F G G. S. C. F. F STA G G G G G. F. R. S. C. S. F STA G G G. F. R. S. STA G G. F. R. S. STA G G G. F. STA G G. F. STA G G. F. STA G G. F. STA G S. F STA G S. S. C. F. STA G G. F. STA G S. S. C. F. STA G G. F. STA G S. S. C. F. STA G S. S. C. S. F STA G S. S. C. S. STA G S. S. C. S. STA G S. S. C. S. S. S. STA G S. S. C. S. S. S. S. STA G. S. L. S. F S. S. S. S. S. S. S. STA G. V. O. S. F. STA G G. V. O. S. F. STA G G. V. O. S. F. STA G S. S	Open A Epitopic UMASNDSSSHULL	Ectodo	Sk. Sk. G0 TN. TN. TN. TN. TN. TN. TN. SK. G. G. TN. TN. TN. A. A. A. TN. TN. TN. E. TR. E. TR. DR. TN.	ESFV IFPVLTHI VSYGA T 70 80 TSFV IFPVLTHI VSYGA T	TTSHFLDTVAL 90 7TSHFJDTVAL 6	VTVSTAGFVE 100 VTVSTAGFVE AYY. AYY. AYY. AH. IC. IYY. AH. AH. AH. AH. AH. AH. AYY.	RYVLSSI YAW 110 RYVLSSI YAW II 0 I 1 I 1 I 1	T cell epito CALAALTCFY I 120 CALAALTCFY I I. I. I. I. F. I. I. F. T. F. I. F. F. I. F. I. F. I. F. I. F. I. F. I. F. I. F. I.	DE 1 130 1 14 1 15 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1	Endov RYACTRYINFLLD 140 RYACTRYINFLLD S	Iomain T cell epitype 150 160 KKN KINSEVI II KKN KINSEVI II KKN KINSEVI II KKN KI KKI KKI KKI KKI KKI KKI KKI	2 KRGK VEVE KRGK VEVE	EGHL IDL KRVV 	B c C LDGSVATI 30 LDGSVATI 	<pre>vell epitope P1TRVSAEQWGRP P1TRVSAEQWGRP vester veste</pre>	VR22 (H1) JXA HNN VY2 SH1 HNN VY7 SH1 HNN VY7 SH1 HNN VY7 SH1 HNN VY7 SH1 HNN VY7 SH1 SH2 SC SC SC ST VY7 STW	332 332 -R 1 4 2 - 2 1 2 1 2 2 2 2 2 2 2 2 4 0 4 1 2 2 2 2 2 2 2 2 2 2 2 2 2
$\begin{array}{c} {\rm Epit} \\ \\ {\rm EXCLTACCYSQLLSLWCUVPFCFAV} \\ \hline 1 \\ {\rm EXCLTACCYSQLLSLWCUVPFCFAV} \\ \hline 0 \\ {\rm EXCLTACCYSQLLSLWCUVPFCFAV} \\ {\rm G. V. C. C. R. F. VL \\ {\rm G C. C. R. F. VL \\ {\rm G C. C. R. F. STA. \\ {\rm G C. F. F. STA. \\ {\rm G F. F. STA. \\ {\rm G \\ {\rm G \\ {\rm G$	Ope A Epitor 50 40 50 40 50 50 50 50 50 40 50 50 50 50 50 40 50 50 50		Sk. Sk. 0. 0. 0.	ESFV IFPVLTHI VSYGA + + 70 80 ESFV IFPVLTHI VSYGA T. T. G. V G. V G. V G. V G. C. C. C. C. C. C. C. C. C. C	TTSHFLDTVAL 90 TTSHFLDTVAL 6. 6. 6. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8. 8.	VTVSTAGFVE 100 VTVSTAGFVE AYY. AYY. AYY. AYY. AYY. AYY. A.	RYVLSSI YAW 110 RYVLSSI YAW III 0 III 0 IIIII 0 III 0 IIIII 0 III 0 III 0 III 0 III 0 III 0 IIII 0 IIII 0 III 0 IIII 0 III 0 IIII 0 III 0 III 0 III 0 III 0 III 0 III 0	T cell epito CALAALTCFY1 120 CALAALTCFY1 I. I. I	DE 130 IFARACIAS WILL 130 REPARICIAS WILL 130 IL IL IV IL IV IL IL IL IV IV	Endov RYACRYINFLLD 140 RYACRYINFLD SQ. S. S. S. S. S. S. S. S. S. S	Jomain T cell epilopie 150 160 KKI. Menissey II. II KKI. Service V. V. V. KKI. V. KKI. KI. KKI. KKI.	2 KRGK VEVE 1 RRGKVEVE 6. 6. 6. 6. 6. 6. 6. 6	BGHL IDL KRVV 	B c c LDGSVAT1 30 LDGSVAT A. A. A. A. A. A. A. A. A. A. A. A. A.	<pre>vell epitope P1TRVSAEQWGEP P1TRVSAEQWGEP vester veste</pre>	00 VR22 CHI JXA HNY VI VI VI VI VI VI VI VI VI VI	332 -R -R -1 44 -2
ERIT EKCLTAGCYSQULSLWCUVPFCFAY + 10 EKCLTAGCYSQULSLWCUVPFCFAY - 20 EKCLTAGCYSQULSLWCUVPFCFAY 	PP A Epitory 50 40 50 40 50 40 50 40 50 40 50 40 50 40 50 50 50	Ectodo	Sk Sk G TN TN TN SK G G TN.N A TN A TN A TN.N A TN BR A C T SDR TR SDR TR	ESFVIFPVLTHUSYGA + + 70 80 ESFVIFPVLTHUSYGA T. C. C. C. C. C. C. C. C. C. C	DTSHFLDTVAL 00 UTSHFLDTVAL G. G. G. S. S. S. S. S. S. S. S. S. S	VTVSTAGFVE 100 VTVSTAGFVE 4YY. 4Y	RYVLSS YAW 110 RYVLSS YAW 111 110 110 100 100 100 100 10	T cell epito CALAALTCF1 120 CALAALTCF1 I F. I.I. F. I. F. F. I. F. V. V. F. I. F. F. V. V. F. V. I. F. F. V. I. F. V. I. F. V. I. F. V. I. F. V. I. I. I. F. V. I.	DE 1 130 RFAROCUSY L L L L LV DV LV LV LV SV SV SV LV LV LV LV LV LV LV LV SV SV LV LV SV LV LV SV LV LV SV SV LV LV LV SV LV LV SV LV LV LV LV SV SV LV LV LV LV LV LV SV SV LV	Endov RYACTRYTNFLLJ 140 RYACTRYTNFLJ 5	Iomain T cell episepe KGRL VRWSPY IIE 150 160 KGRL VRWSPY IIE 150 160 KGRL VRWSPY IIE KGRL VRWSPY IIE	2 KRGKVEVE CG.	SGHL IDL KRVV 70 1 3 GHL IDL KRVV 3 K. 3	B c c LDGSVAT1 30 10GSVAT1 30 10GSVAT1 30 1A. A. A. A. A. A. A. A. A. A	CII epitope PTTRVSAEQUCEP i90 21 PTTVSAEQUCEP i00 21 PTTVSAEQUCEP i	00 VR22 CH1 JXA HNN 002 00 00 00 00 00 00 00 00 0	332 -R -R 1 2 211 2 211 2 211 3 2 211 3 4 4 4 4 4 4 4 4 4 4 4 4 4
ERIT EKCLTAGCYSQLLSLWC1VPFCFAY 10 20 EKCLTAGCYSQLLSLWC1VPFCFAY 0. C. C. R. F. YLL. G C. C. R. F. YLL. G C. C. R. F. YLL. G C. C. R. F. STA G HS. F. STA G HS. F. STA G	Ope A Epitor 50 40 50 40 50 40 50 40 50 50 50 40 50 40 50 40 50 40 50 50 50	Ectodo	Sk. Sk. G. TN. TN. Sk. G. TN. TN. Sk. G. TN. TN. TN. Sk. G. G. TN. Sk. G. G. TN. A. DR. A. A. G. TN. N. A. DR. DR. DR.	ESFVIFPULTHIVSYGA T	TTSHFLDTVAL 90 1TSHFLDTVAL 6 	VTVSTAGFVE + 100 VTVSTAGFVE AYY. AYY. AH. YY. A.R. YY. A.H. YY. A.H. YY. A.R. YY. A.H. YY. Y.	RYVLSS YAW 110 RYVLSS YAW 110 10 10 10 10 10 10 10 10 1	T cell epito CALAALTCFY I 120 CALAALTCFY I I	DE 1 TRFARCISS* 130	Endov RYACTRYINFLLD 140 RYACTRYINFLD S	Iomain T cell epitope Kord. IransSPT II E 150 160 Kord. IransSPT II E Kord. IransSPT II E Kord. Irans V. V. V. V. V. KI. KI. KI. KI. KI. KI. KI. KI	2	BGHL IDL K RVV 70 1 3 K. 3	B c c IDGSVAT 10DSV	CII epitope PITRVSAE00GR 1100 20 211rVSAE00GR 211rVSAE00GR 211rVSAE00GR 211rVSAE00GR 211rVSAE00GR 211rVSAE00GR 211rVSAE00GR 211rVSAE0GR 2111rVSAE0GR 2111rVSAE0GR 2111rVSAE0GR 2111rVSAE0GR 2111rVSAE0	00 VR22 (11) JXA HuND (11) JXA HUND (11) JXA (11) JXA (11	332 332 -R 1 4 2 2 2 2 2 2 2 4 4 4 4 4 4 4 4 4 4 4 4 4
Epit EXCLTAGCYSQLLSLWCUPPGCAY + 0 2 EXCLTAGCYSQLLSLWCUPPGCAY G. Y. T. C. R	Open A Epirote Construction E	Ectodo	Sk G TN. TN. TN. TN. TN. TN. TN. TN. TN. TN. SK. G G. TN. TN. TN. TN. TN. TN. TN. G. TN. A. A. A. TN. E. TR. DR. DR. SDR DR.	ESFV IFPVLTHI VSYGA T	TISHFLDTVAL 90 91TISHFLDTVAL 6	VTVSTAGEVE 100 VTVSTAGEVE 4	RYVLSS YAW 110 RYVLSS YAW 110 10 10 10 10 10 10 10 10 1	T cell epito CALAALTCFY I 120 CALAALTCFY I I . CALAALTCFY I I . CALAALTCFY I I . I . I . I . I . I . I . I . I . I	DE 1 RFAROCISS 130 RFAROCISS 131 130 RFAROCISS 132 134 135 1 5 1 5	Endou RYACTRYINFLLD 140 RYACTRYINFLLD S	Iomain T cell epitype 150 160 KKN KINSEVI II E KKN KINSEVI II E KKN KINSEVI II E KKI KI KKI KI	2 (RRGKVEVE (RGKVEVE (RGKVEVE (G.)) (G.) (G.) (G.) (G.) (G.) (G.) (G.	EGHL IDL KRVV 	B c c 1050X11 DISSVAT1 DISSVAT1 DISSVAT1 DISSVAT1 A. A. A. A. A. A. A. A. A. A. A. A. A.	<pre>vell epitope PITRVSAEQWGRP PITRVSAEQWGRP LL. LL. LL. L</pre>	00 VR22 GHI JXA HaNN VI VI VI VI VI VI VI VI VI VI	332 332 -R 1 4 2 2 1 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2
ERITAGUSQUISUFUTPECFAY + 2 EKCLTAGCYSQUISUFUTPECFAY + 2 EKCLTAGCYSQUISUFUTPECFAY - 2 EKCLTAGCYSQUISUFUTPECFAY - 3 C. C. C. R. F. YLL., G C. C. R. F. YLL., G C. C. R. F. YLL., G C. C. R. F. STA, G. S. C. F. F. STA, G M. S. F. STA, G	Ope A Epitopic 50 40 50 40 50 50 50 50 50 40 50 50 50 50 50 50 50 50 50 40 50 50 50	Ectodo	Sk. Sk. G. TN TR TR SK. G. TN TR SK. G. TN TN TR TN SK. G. TN TN SK. G. TN.N. TN A. TN A. TN A. TN A. TN SBR. SDR DK. E. K. R. A. TN	ESFV IFPVLTHI VSYGA + + 70 80 ESFV IFPVLTHI VSYGA T I C	TTSHFLDTVAL 90 TTSHFLDTVAL 6,, G, 8,, S, 8,, S, 8,, S, 8,, S, 9,, S, 5,, S, 5,, S, 5,	VTVSTAGFVE 100 VTVSTAGFVE AYY. A	RYVLSS YAW 110 RYVLSS YAW 110 I. I. I.	T cell epito CALAALTCFYI 120 CALAALTCFYI I. I. I. F. V. V. F. I. I. I. F. V. V. F. I. I. I. F. V. V. F. I. I. I. F. V. V. F. I. I. I. F. V. V. F. I. I. F. V. V. F. V. V. V. F. V. V. V. F. V. V. F. V. V. F. V. V. V. F. V. V. V. F. V. V. V. F. V. V. V. F. V. V. V. F. V. V. V. T. V.	DE 1 130 1474KNCMS 131 132 134 134 135 134 134 134 134 134 134 134 134	Endov RYACRYINFLLD 140 RYACRYINFLD S	Jomain T cell epitype it 50 160 KKI, Itomissey II. it 50 160 KKI, Itomissey II. 160	2 (RRGKVEVF (RGKVEVF (G	BGHL IDL KRVV 	B c c LDSSVAT 30 LDSSVAT 1 	<pre>vell epitope P1TRVSAEQWCEP P1TRVSAEQWCEP vest vest vest vest vest vest vest vest</pre>	00 VR22 (11) JAA HuN HI JI JAA HU HI HI JI JI JI HI HI JI JI JI HI HI HI JI JI HI HI HI HI HI HI HI HI HI H	332 332 -R -R 1 4 4 211 -2 -2 -2 -2 -2 -2 -2 -2 -2 -2
Enti EXCLTAGCYSQLLSLWC1VPFCFAY + 20 EXCLTAGCYSQLLSLWC1VPFCFAY G. C. C. R. F. YLL., G C. C. R. F. YLL., G C. C. R. F. YLL., G C. C. R. F. STA, G. S. C. F. F. STA, G. Y. C. R. S. STA, G. Y. C. S. F. STA, G L. M. S. F. STA, G S. F. STA, G S. F. STA, G S. S. STA, G	Open A Epitor 50 40 50 40 50 40 50 40 50 40 50 40 50 40 50 50 50 40 50 50 50	Ectodo	Sk. Sk. G. TN TN TN TN TN TN TN TN TN SK G G TN TN TN SK G TN A DR A DR TN.N A TN.N A TN.N A TN.N A TN.N A TN.N A TN B TR DR DR DR DR DR DR DK DK E. R A T	ESFV IFPVLTH USYGA + + 70 80 SSPV IFPVLTH USYGA T. C. C. C. C. C. C. C. C. C. C	DTSHFLDTVAL 00 UTSHFLDTVAL G. G. G. S. S. S. S. S. S. S. S. S. S	VTVSTAGFVE 100 VTVSTAGFVE 4Y	RYVLSS YAW 11D RYVLSS YAW 11D 11D 12D 12D 12D 12D 12D 12D	$\begin{array}{c} T \ cell \ epito \\ CALAAL TCFY I \\ 120 \\ CALAAL TCFY I \\ 120 \\ T \\ R \\ R$	INPE I ISPAROSCI ISO ISPAROSCI ISO ISPAROSCI ISO IL ISO IV ISO	Endov RYACTRYTNFLLJ 140 RYACTRYTNFLJ 5	Iomain T cell episepe KGRL VRWRSPY II 150 160 NSRL VRWRSPY II 150 160 NSRL VRWRSPY II 150 160 NSRL VRWRSPY II 150 160 NSRL VRWRSPY II 160 NSRL	2 (RRGKVEVE 1 1 1 (RGKVEVE 6 .	SGHL IDL KRVV 70 1 3 GHL IDL KRVV 3 K. 3	B c c IDCSVAT 30 1DCSVAT A A A A A A A A A A A A A A A	cell epitope PITRVSAEQUCEF i90 20 2171XVSAEQUCEF i	00 VR22 01 VR22 01 01 01 01 01 01 01 01 02 02 02 02 02 02 02 02 02 02	332 -R -R 1 1 2 2 211 2 2 2 2 2 2 2 2 4 4 4 4 4 4 4 4 4 4 4 4 4
ERITAGUSQUISLWCIVPFCFAY 10 20 EKCLTAGCUSQUISLWCIVPFCFAY 0. C. C. R. F. YLL., G C. C. R. F. YLL., G C. C. R. F. YLL., G C. C. R. F. STA, G. S. C. F. F. STA, G. S. C. F. F. STA, G RS, F. STA, G ST, STA, G	Open A Epitopic 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10 50 10	Ectodo	Sk. Sk. G. TN. TN. TN. Sk. G. TN. TN. TN. TN. Sk. G. TN. TN. Sk. G. TN. TN. A. DR. DR. DR. DR. DR. TN.N. A. TN. A. TN. A. C. TR. DR. DR. DR. DR. DR. DR. Str. TR. Str. TR. Str. TR. Str. DR. Str. DR. Str. TR.	ESFVIFPULTHUSYGA T	TTSHFLDTVAL 90 TTSHFLDTVAL 6 	VTVSTAGFVE 100 TVSTAGFVE AYY. AYY. AH. YY. A. H. YY. A. H. YY. I. A. H. YY. I. A. H. YY. A. H. YY. YY. A. H. YY. A. H. YY. A. H. YY. A. H. YY. YY. YY. A. H. YY. A. H. YY. A. H. YY. A. H. YY. A. H. YY. A. H. YY. Y. YY. Y. YY. Y. YY. Y. YY. Y. YY. Y. Y. YY. Y.	RYVLSS YAW 110 RYVLSS YAW 110 110 RVVLSS YAW 110 100 100 100 100 100 100 10	T cell epito CALAALTCFY I 120 CALAALTCFY I 120 CALAALTCFY I . I. . I.	DE 1 130 GFAROCUS 1 13. 13. 13. 13. 13. 13. 13. 13	Endou RYACTRYINFLLD 140 RYACTRYINFLD S	Iomain T cell epitope Kord. IransSPT II E Kord. IransSPT II E Kord. IransSPT II E Kord. IransSPT II E Kill	2 (RRGKVEVE) 1 1 RRGKVEVE 6	BGHL IDL K RVV 70 1 3 K. 3 K. 3 K. 3 K. 3	B c c LDGSVAT 30 LDGSVAT 4. A. A. A. A. A. A. A. A. A. A. A. A. A. A	<pre>cll epitope PITRVSAE0wGR</pre>	00 VR22 01 VR22 01 01 01 01 01 01 01 01 01 01	332 332 -R 1 -R 2 2 2 2 2 2 2 2 2 2 2 4 4 4 4 4 4 4 4 4 4 4 4 4
Excl_TAGCVSQLLSL#CUVPFCFAV + 0 2 EXCLTAGCVSQLLSL#CUVPFCFAV G. Y. T. C. R	PP A Fpirit WASDESS HILL DYMASDESS H	Ectodo	Sk. Sk. G. TN. TN. TN. SK. G. TN. TN. TN. TN. SK. G. TN. TN. TR. TN. G. TN. TR. SK. G. TN. A. TN. A. TN. A. TN. A. TN. B. TN. SDR. TR. DR. TN. SDR. TR. C. TN. K. T. G. G.	SFV IFPVLTHI VSYGA T	TISHFLDTVAL 90 91TISHFLDTVAL 6	VTVSTAGEVE 100 VTVSTAGEVE AYY. A	RYVLSS YAW 110 RYVLSS YAW 110 10 10 10 10 10 10 10 10 1	$\begin{array}{c} T \ cell \ epito \\ CALAALTCFY I \\ 120 \\ CALAALTCFY I \\ 120 \\ CALAALTCFY I \\ . \\ I \\ . \\ . \\ . \\ . \\ . \\ . \\ . \\$	DE 1 130 TRANCAS 1 130 TRA	Endou RYACTRYINFLLD 140 RYACTRYINFLLD S	Iomain T cell epitype 150 160 KKM, KINSEVY II E KKM, KINSEVY II E KKM, KINSEVY II E KKM, KKM, KKM, KKM, KKM, KKM, KKM, KKM,	2 RRGKVEV 1 1 1 RRGKVEV 6	EGHL IDL KRVV 	B c c LDCSVAT 30 10 10 10 10 10 10 10 10 10 1	<pre>vell epitope P1TRVSAE0WGRP 11TRVSAE0WGRP vestment ve</pre>	00 VR22 G111 JAA HuN QYYYY QYYYY QYYYY QYYYY QYYYY QYYYY QYYYY QYYYY QYYYY QYYYY QYYYY QYYYY QYYYYY QYYYY QYYYY QYYYY QYYYYY QYYYYY QYYYYY QYYYYY QYYYYY QYYYYY QYYYYY QYYYYY QYYYYY QYYYYYY QYYYYYY QYYYYYY QYYYYYY QYYYYYYYY	332 332 -R 1 4 2 2 2 2 2 2 2 2 2 2 2 2 2
$ \begin{array}{c} {\rm Epit} \\ \\ {\rm EKCLTAGCYSQLLSLWCUPPECFAV} \\ + & + \\ \\ {\rm EKCLTAGCYSQLLSLWCUPPECFAV} \\ \\ {\rm FKCLTAGCYSQLLSLWCUPPECFAV} \\ \\ {\rm GV, T. C, R. & \\ {\rm G. C, C, C, F, F, VL, . , \\ {\rm G C, C, F, F, STA, \\ {\rm G C, F, F, STA, \\ {\rm G C, F, F, STA, \\ {\rm G C, F, F, STA, \\ {\rm G$	Open A Epitor 50 40 50 40 50 50 50 50 50 50 50 50 50 40 50 50 50	Ectodo pe B IYNI.TLCEING 50 IYNI.TLCEING 80 80 80 80 80 80 80 80 80 80 80 80 80	Sk. Sk. G. TN TR TR SK G. TN TR SK G. TN TN TR TR SK G. TN TN SK G. TN.N TN.N TN TN A. TN A. TN A. TN B. TR SDR. SDR DR. SDR SDR. SDR C. TR G. SDR SDR. SDR C. T. G. G. G. G. G. G. T. G. G. G. G. G.	ESFV IFPVLTHI VSYGA + + + 70 80 ESFV IFPVLTHI VSYGA T T CV GV GV GV G C.	DTSHPLDTVAL 90 UTSHPLDTVAL 6. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5	VTVSTAGEVE 100 VTVSTAGEVE AYY. AYY. AYY. AYY. AH. IC. TY. IA. H. AH. AH. AH. AYY. AY	RYVLSS YAW 110 RYVLSS YAW 110 I. I. I. I. A. A. A.	T cell epito CALAALTCFYI 120 CALAALTCFYI I. I. I. I. I. I. I. I. I. I. I. I. I.	DE 1 RFAROEST 130 GRAROEST L	Endoo RYACTRYINFLLD 140 RYACTRYINFLLD S	Jomain T cell epitype	2 (RRGKUEVE 1 1 (RGKUEVE 1 (RGKUEVE (G	BGHL IDL KRVV 	B c c LDCSVAT 30 LDCSVAT A. A. A. A. A. A. A. A. A. A. A. A. A.	<pre>vell epitope P1TRVSAEQWCEP P1TRVSAEQWCEP vest vest vest vest vest vest vest vest</pre>	VR22 00 01 01 01 01 01 01 01 01 01 01 01 01	332 332 332 332 332 332 332 332 332 31 3 32 31 3 3 3 3
ERITAGUSQUISURCIVPECRAY + 20 ERCLTAGCUSQUISURCIVPECRAY + 20 ERCLTAGUSQUISURCIVPECRAY G C. C. R. F. YLL., G C. C. R. F. YLL., G C. C. R. F. STA, G. S. C. F. F. STA, G. J. M. SF. STA, G L. C. RS, F. STA, G L. C. RS, F. STA, G	Open A Epitor 50 40 50 40 50 40 50 40 50 40 50 40 50 40 50 40 50 50 50 40 50 50 50	Ectodo	main Epitope C DWLANKED#A 60 DRLANKED#A Q Q TN TR SK G J.N TN TN A DR. A DR. A DR. A TN.N. A DR. DR. A A TN. A B DR. Q Q Q Q Q Q C	ESFV IFPVLTH USYGA + + + 70 80 SSV IFPVLTH USYGA T	DTSHFLDTVAL 00 UTSHFLDTVAL G. G. G. S. S. S. S. S. S. S. S. S. S	VTVSTAGFVE 100 VTVSTAGFVE 4	RYVLSS YAW 110 RYVLSS YAW 110 110 100 100 100 100 100 10	$\begin{array}{c} \mbox{T cell epito} \\ \mbox{CALAALTCFT} \\ \mbox{120} \\ \mbox{CALAATCFT} \\ \mbox{120} \\ \mbox{120} \\ \mbox{CALAATCFT} \\ \mbox{120} \\ 120$	DE 1 130 1474/XICMLS 1 130 1474/XICMLS 1 131 1474/XICMLS 1 1474/XICMLS 1	Endov RYACTRYTNFLLJ 140 RYACTRYTNFLJL 3	Iomain T cell episope Kork JYRWSPY II 150 160 NSR JYRWSPY II 250 160 NSR JYRWSPY III 250 160 NSR JYRWSPY II 250	2 KRGK VEVE 1 1 1 RGKVEVE 6	SGHL IDL KRVV 70 1 3 GHL IDL KRVV 3 K. 3	B c c 1105VAT 30 1105VAT 40 1105VAT 40 105VAT 40 105VAT	CII epitope PITRVSAEQUCEP i 90 21 PITRVSAEQUCEP i 00 21 PITRVSAEQUCEP i 0. 21 PITRVSAEQUCEP i 0. 21 PITRVSAEQUCEP i 0. 21 VIXI	VR2 00 VR2 SI 01 02 02 02 02 02 02 02 02 02 02	332 -R -R 1 2 2 2 2 1 4 4 4 4 4 4 4 4 4 4 4 4 4
ERCLTAGCYSQLLSLWCIVPFCFAY 10 20 ERCLTAGCYSQLLSLWCIVPFCFAY 4 4 4 5 CRCLTAGCYSQLLSLWCIVPFCFAY 4 5 CRCLTAGCYSQLLSLWCIVPFCFAY 4 5 CRCLTAGCYSQLLSLWCIVPFCFAY 5 CRCLTAGCYSQLLSLWCIVPFCFAY 6 CRCLC R. F. YLL. 6 C. C. C. R. F. YLL. 6 C. C. C. R. F. STA 6 C. YC. RS. F. STA 6 C. L. C. R. F. STA 7 C. S. C. F. STA 7 C. S. S. STA 7 C. S. S. S. S. STA 7 C. S. S. S. STA 7 C. S. S. S. S. S. STA 7 C. S. S. S. S. STA 7 C. S. S. S. S. S. STA 7 C. S. S. S. S. S. STA 7 C. S. S. S. S. S. STA 7 S. S. S. S. S. S. S. STA 7 S. S. S. S. S. S. S. STA 7 S. S. S. S. S. S. STA 7 S. S. S. S. S. S. S. STA 7 S. S. S. S. S. S. S. STA 7 S. S. S. S. S. S. S. S. STA 7 S. S. S. S. S. S. S. S. S. STA 7 S. S	Open A Epitor 50 40 50 40 50 40 50 40 50 40 50 40 50 40 50 40 50 50 50	Ectodo	main Epitope C DMLANKED#A 0. Q Q TN TR TR SK G J SK G TN A DR A A A DR DR DR DR DR DR C DR DR DR DR C Q Q TR DR R Q Q Q Q Q Q Q Q Q Q Q	ESFVIFPULTHUSYGA T	TISHFLDTVAL 90 ITISHFLDTVAL 6	VTVSTAGFVE + 100 VTVSTAGFVE AYY. AYY. AH. YY. A. H. A. H. YY. I. A. H. YY. I. A. H. A. H. V. A. H. A. H. V. A. H. V. A. H. A.	RYVLSS YAW 110 RYVLSS YAW 110 I. I. I. I. I. I. I. I. I. I.	$\begin{array}{c} T \ cell \ epito \\ CALAALTCFY I \\ 120 \\ CALAATTCFY I \\ 120 \\ . I $	DE 1 130 174 ANCAS 1 130 174 ANCAS 1 13. 174 ANCAS 1 13. 174 ANCAS 1 13. 13. 13. 13. 13. 13. 13. 13	Endou RYACTRYINFLLJ 140 RYACTRYINFLJ S	Iomain T cell epitope Kork I KINSEY II E Kork I KINSEY II E Kork I KINSEY II E KORK I KINSEY II E KIL KIL KIL KIL KIL KIL KIL KIL	2 KRGK VEVE 1 1 1 RGRKVEVE 6 .	3. K. 3. R. 3. <	B c c LJDCSVAT 30 LJDCSVAT A. A. A. A. A. A. A. A. A. A. A. A. A. A. A. A. A	CII epitope PITRVSAE0wGRP 21TRVSAE0wGRP 221TRVSAE0wGRP 221TRVSAE0wGRP 221TRVSAE0wGRP 22. 22. 22. 22. 22. 22. 22. 22	VR2 00 VR2 0111 JJAA HAN QYP SE GO JJAA QYP SE GO JJAA QD JA SE GO SC GO SC GO SU	332 332 -R 1 - - - - - - - - - - - - -
Excl.TAGCVSQLLSLWCUVPFCFAV + 0 EXCLTAGCVSQLLSLWCUVPFCFAV G. Y. T. C. R G C. C. R. F G C. C. R. F G C. C. R. F G C. C. R. F G.	P. A Fpirit MANDESS HILL AND AND STATES HILL AND AND AND STATES HILL AND AND AND AND AND AND AND AND AND AND	Ectodo	main Epitope C DULANKEDVA	ESFVIFPULTIIVSYGA T	TISHFLDTVAL 90 7TSHFJLDTVAL 6	VTVSTAGEVE 100 VTVSTAGEVE AYY. AYY. AYY. AH. IC. T. Y. AH. A. H. A. Y. A. Y. A. Y. A. Y. A. Y. A. Y. A. Y. A. Y. A. H. A. Y. A. H. Y. A. H. Y. Y. A. H. Y. Y. A. H. Y. Y. A. Y. Y. Y. Y. Y. Y. Y. Y. Y. Y.	RYVLSS YAW 110 RYVLSS YAW 110 10 10 10 10 10 10 10 10 1	$\begin{array}{c} T \ cell \ epito \\ CALAALTCFY I \\ & & 120 \\ CALAALTCFY I \\ & 120 \\ CALAALTCFY I \\ & & 1. \\ & I. \\ & $	DE 1 130 174 MOCMS7 130 174 MOCMS7 13. 13. 13. 13. 13. 13. 13. 13.	Endou RYACTRYINFLLD 140 RYACTRYINFLD S	Iomain T cell epitype	2 RRGKVEV 1 1 1 RRGKVEV 6	EGHL IDL KRVV 	B c c LDCSVAT 30 10 10 10 10 10 10 10 10 10 1	<pre>vell epitops PITRVSAE0WGRP ITRVSAE0WGRP L</pre>	00 VR22 01 VR22 01 01 01 01 01 01 01 01 01 01	332 332 -R 1 4 2 2 2 2 2 2 2 2 2 2 2 2 2

FIGURE 3 | Alignment of the deduced amino acid sequences among QYYZ-like PRRSVs. (A) The positions marked in the figure represent positions of the Nsp2 amino acid sequence and refer to the position of VR-2332. Red indicates the QYYZ strain 36-aa characteristic continuous insertion; yellow indicates the NADC30-like PRRSV 131-aa characteristic discontinuous deletion; blue indicates the HP-PRRSV 30-aa characteristic discontinuous deletion; and purple indicates some additional a deletions in QYYZ-like PRRSVs. (B) Alignment of the deduced amino acid sequence based on the ORF5 gene. The signal peptide and transmembrane (TM) regions are shown in blue and black boxes, respectively. The linear antigenic epitopes and cellular epitopes are indicated in yellow and blue, respectively. The seven amino acids characteristic of QYYZ-like strains are shown in gray.

D strains contained the discontinuous 131-aa deletion in Nsp2 identical to that in NADC30 (111 + 1 + 19 aa) (Figure 3A). In addition, 13 of the 37 Chinese lineage 3 PRRSV strains (XJzx12015, HiNZWQ, GXXNF78-1806, SH/CH/2016, GZgv17, JX/CH/2016, GD-KP, FJFS, GXXNF10-1803, GXXNF74-1806, GDXNF229-1811, GDsg, and SDWH27-1710) also have special amino acid deletions in the Nsp2 region, and they all have different deletion patterns (Figure 3A). Interestingly, 12 out of 13 strains with special amino acid deletions showed an amino acid deletion at positions 468-518 (Figure 3A). QYYZ-like PRRSV strains appeared to have more amino acid insertions or deletions in the Nsp2 region than other PRRSV strains. They seem to be more prone to amino acid deficiencies at position 468-518 of Nsp2. The Nsp2 hypervariable region (323-521) not only plays an important regulatory role in maintaining the balance of different viral mRNA species but also regulates PRRSV tropism to primary porcine alveolar macrophages (PAMs) (35). Therefore, these amino acid deletions in the Nsp2 region may alter the cellular tropism of the strains.

GP5 is the major envelope protein of PRRSV and is responsible for the lack of immunological cross-protection among different PRRSV strains due to its hypervariability (36-38). Comparison analyses of the amino acid sequence of QYYZ-like GP5 with those of the other representative strains showed that most strains encode 201 aa, but GDXNF229-1811, GDYDZZZ, and GXTZJ2325-2112 have a 1-aa deletion at residue 33 in VR2332, which is identical to the mutation in SD53-1603 (Figure 3B). Seven unique amino acid substitutions, namely, $F^{25} \rightarrow S^{25}, A^{26} \rightarrow I^{26}, H^{38} \rightarrow Y^{38}, L^{39} \rightarrow S^{39}, T^{66} \rightarrow C^{66}, A^{92} \rightarrow S^{92},$ and $L^{152} \rightarrow I^{152}$, compared with the VR2332 strain. These amino acid mutations were identified only in all QYYZ-like strains but no other representative PRRSV strains (Figure 3B). Interestingly, although the QYYZ-like strains have low homology among them, they still showed consistent molecular features, which may be used as molecular markers to distinguish QYYZ-like strains from other type 2 PRRSV strains in China. GP5 is an envelope protein essential for viral infection, and at least three B-cell and two Tcell epitopes were identified for GP5 (39, 40). Two characteristic mutations of QYYZ-like strains, at positions 38(Y³⁸) and 39(S³⁹), are in epitope B (37-45 aa), which is a highly conserved peptide sequence that presumably functions as a major target for broadly neutralizing antibodies (37, 41). At the same time, there was one amino acid substitution in T-cell epitope 2 (I¹⁵²), compared to PRRSV strains in other lineages (42). We then analyzed the functional domains in GP5 of QYYZ-like strains, including the signal peptide and transmembrane (TM) domain (43). Two residues (S^{25} and I^{26}) resided within the signal peptide (aa 1-31). Moreover, there was a unique amino acid mutation at position 66 in the TM region of the QYYZ-like strains, substituting T⁶⁶ to C⁶⁶. Briefly, we found 7 unique amino acid mutations in the GP5 protein, and some mutation sites were located in cell epitopes, the signal peptide and the TM region. These amino acid substitutions might lead to the failure of receptor recognition and thus result in the failure of vaccines.

Recombination is a pervasive phenomenon among PRRSV isolates, and there are an increasing number of reports about the recombination of QYYZ-like PRRSVs (14, 16, 44). In recent

years, many recombinant QYYZ-like PRRSVs have reportedly reemerged with increased pathogenicity (12, 18, 34, 45, 46). To identify possible recombination events of the new QYYZlike PRRSVs in China, RDP4 and SimPlot software were used to assess possible recombinant events (Table 4 and Supplementary Figure 1). This analysis demonstrated that all nine new QYYZ-like isolates were recombinant viruses. Six recombinant isolates (GXXNF10-1803, GXXNF53-1805, GDXNF60-1805, GXXNF74-1806, GXXNF78-1806, and GDXNF229-1811) emerged from recombination events between HP-PRRSV isolates and QYYZ-like virus; three recombinant isolates (HNLCL15-1903, HNTZJ1714-2011, and GXTZJ2325-2112) were also derived from NADC30-like PRRSV, HP-PRRSV isolates, and QYYZ-like PRRSV (Table 4 and Supplementary Figure 1). These putative recombination events were further supported by statistically incongruent phylogenetic trees (Supplementary Figure 1). According to a similarity plot, these nine strains have an extremely complex recombination pattern (Supplementary Figure 1). Interestingly, the HNLCL15-1903 strain has almost the same recombination pattern as the previous strain SCya18 (MK144543.1) and high homology (BLAST analysis: 97.99%). SCya18 was isolated in Sichuan Province in 2018 (29), while HNLCL15-1903 was isolated in Henan Province in 2019. The two provinces are not adjacent, and under the background of low homology and complex recombination of QYYZ-like strains, the emergence of two strains with high similarity is noteworthy.

To better explore the recombination characteristics of QYYZ-like strains in China (Table 3), we summarized the reported genome-wide recombination of all QYYZ-like strains (ORF5 classified into sublineage 3.5). Interestingly, all 37 QYYZ-like PRRSV strains except for QYYZ (the QYYZ-like original strain) have undergone recombination (Table 3 and Supplementary Figure 2B). The distribution of recombination breakpoints is dispersed and complex, there is no obvious recombination hotspot, and there are relatively many recombination breakpoints located in Nsp12(7), Nsp2(13) and GP2(13) (Supplementary Figure 2B). In a previous study in China, it was found that between 2014 and 2018, the high-frequency interlineage (mainly lineage 1 and lineage 8) recombination regions were located in Nsp9 and GP2 to GP3 (20, 31). Obviously, the recombination hot spots of QYYZ-like strains (sublineage 3.5) are not exactly the same as those of lineage 1 and lineage 8 strains. At the same time, it can be clearly seen that QYYZ-like strains mainly provide fragments of structural protein regions (GP2-N) for recombinant strains (Supplementary Figure 2A), while in the Nsp1 region of the sequence, only GM2 is provided by the QYYZ strain (Table 3). We used the Nsp1 region of QYYZ-like strains to construct a phylogenetic tree for verification (Figure 2D). The majority (33/37) of the QYYZ-like strains were grouped into sublineage 8.7 (JXA1-like), and FJLIUY-2017 and HNTZJ1714-2011 were grouped into sublineage 1.8 (NADC30-like). Only GM2 and QYYZ were classified as sublineage 3.5 (QYYZ-like). To obtain better information about the QYYZ-like strains, we suggest that researchers add a pair of primers to sequence the Nsp1

Strains	Breakp	oints	Parental se	duence			Detect	ion methods (<i>p</i> -va	lue)		
	Beginning	Ending	Minor	Major	RDP	GENECONV	BootScan	MaxChi	Chimera	SiScan	3Seq
GXXNF10-1803	-	3,507	JXA1	QYYZ	2.448×10^{-36}		1.505×10^{-27}	1.677×10^{-28}	1.081×10^{-23}	9.166×10^{-82}	4.440×10^{-16}
	7,781	12,294	JXA1	QYYZ	2.352×10^{-20}	4.191×10^{-10}	1.829×10^{-19}	8.522×10^{-11}	6.504×10^{-5}	ı	4.440×10^{-16}
HNLCL15-1903	-	1,659	SH/CH/2016	NADC30	5.474×10^{-47}	1.583×10^{-16}	6.869×10^{-38}	1.302×10^{-37}	1.196×10^{-17}	ı	1.110×10^{-16}
	10,531	14,132	SH/CH/2016	NADC30	4.798×10^{-13}	4.699×10^{-8}	3.515×10^{-13}	2.702×10^{-10}	9.689×10^{-3}	ı	1.221×10^{-15}
GXXNF53-1805	12,685	15,314	awz	JXA1	4.068×10^{-37}	ı	4.114×10^{-34}	1.467×10^{-13}	3.866×10^{-12}	ı	3.330×10^{-15}
GDXNF60-1805	12,541	15,316	awz	JXA1	1.153×10^{-26}	1.115×10^{-14}	5.567×10^{-26}	9.478×10^{-09}	1.709×10^{-09}	ı	4.440×10^{-16}
GXXNF74-1806	-	2,339	JXA1	QYYZ	3.589×10^{-85}	5.042×10^{-53}	2.133×10^{-82}	1.171×10^{-31}	7.599×10^{-35}	ı	4.440×10^{-16}
	6,924	8,969	JXA1	QYYZ	1.197×10^{-73}	7.929×10^{-30}	3.125×10^{-73}	3.349×10^{-23}	3.736×10^{-25}	ı	1.110×10^{-16}
GXXNF78-1806	11,987	12,851	awz	JXA1	2.804×10^{-68}	1.001×10^{-35}	1.478×10^{-67}	2.161×10^{-19}	1.175×10^{-21}	ı	4.440×10^{-16}
	13,673	15,266	awz	JXA1	4.474×10^{-83}	6.225×10^{-53}	1.542×10^{-80}	1.379×10^{-28}	6.139×10^{-30}	ı	1.099×10^{-14}
GDXNF229-1811	-	2,285	JXA1	QYYZ	5.898×10^{-65}	1.549×10^{-52}	3.919×10^{-70}	8.204×10^{-25}	7.417×10^{-21}	ı	3.330×10^{-16}
	6,317	8,812	JXA1	QYYZ	6.152×10^{-86}	2.128×10^{-72}	1.950×10^{-86}	8.256×10^{-27}	2.368×10^{-21}	ı	3.330×10^{-16}
HNTZJ1714-2011	488	3,548	NADC30	JXA1	8.808×10^{-125}	2.849×10^{-88}	7.364×10^{-54}	2.100×10^{-50}	8.792×10^{-58}	1.554×10^{-50}	4.440×10^{-15}
	11,668	15,011	awz	JXA1	1.119×10^{-3}	4.470×10^{-19}	1.405×10^{-12}	4.438×10^{-8}	9.950×10^{-6}	8.177×10^{-40}	ı
GXTZJ2325-2112		2,004	JXA1	QYYZ	1.870×10^{-12}		1.777×10^{-13}	1.089×10^{-9}	1.218×10^{-14}	2.968×10^{-11}	
	2,004	5,268	NADC30	QYYZ	2.595×10^{-95}	1.055×10^{-67}	3.404×10^{-92}	7.725×10^{-30}	8.322×10^{-38}	1.696×10^{-51}	1.332×10^{-15}

region of the virus when it is identified as a QYYZ-like strain by ORF5 sequencing.

Now, an overwhelming majority of the PRRSV-2 strains in China can be classified into JXA1-like/CH-1a-like (sublineage 8.7), VR2332-like (sublineage 5.1), NADC30-like/NADC34-like (lineage 1), and QYYZ-like (sublineage 3.5). The CH-1a-like strains first appeared in China and then gradually evolved into JXA1-like (HP-PRRSV) strains (45). The JXA1-like strains are also widespread in Cambodia, Thailand, Vietnam and many other Asian countries (6, 47-49). VR2332-like strains containing the Ingelvac PRRS MLV vaccine sequence are the most widespread, with viruses introduced to more than 10 countries (5). Lineage 1 (NADC30-like and NADC34-like) strains are also globally epidemic and have spread to South America, North America and Asia (7, 33, 50-52). The lineage 3 strains originated in Taiwan. Interestingly, after 20 years of transmission and evolution, these strains are prevalent only in greater China (mainland China, Taiwan, and Hong Kong) (18). Moreover, the QYYZ-like strains (sublineage 3.5) were prevalent only in mainland China. According to the results of this study and previous studies, QYYZ-like strains have been occasionally reported in northern China (19, 26, 28, 53) but are mainly prevalent in southern and central China (12-16, 23-25, 27, 29, 30, 32). Although recombination of PRRSV is very common, many non-recombination strains of JXA1like (54), VR2332-like (28) and NADC30-like/NADC34-like PRRSV (7, 55) have been reported after long-term evolution in China. All QYYZ-like strains were recombined with other lineages of PRRSV except the prototype strain QYYZ. Thus, we speculated that the non-recombination QYYZ-like strains are no longer circulating and that they played a role as a provider of recombination fragments in the PRRSV epidemic in China.

CONCLUSION

In summary, QYYZ-like PRRSV strains did not have a largescale epidemic status but persisted in central and southern China during 2018–2021. QYYZ-like strains have low homology and extremely complex amino acid insertion and deletion patterns in the Nsp2 region. However, they have seven identical amino acid mutations in the GP5 protein. These strains all underwent complex recombination except the prototype strain QYYZ and mainly provided structural protein fragments (GP2-N) for the recombinant strains. These results will help us to understand the overall genomic characteristics of QYYZlike PRRSV, which is useful for the prevention and control of this virus.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study included article/Supplementary Material, are in the further inquiries be directed to can the corresponding author/s.

TABLE 4 | Information on recombination events of QYYZ-like PRRSV isolates detected by RDP4 software

ETHICS STATEMENT

The animal study was reviewed and approved by Sampling procedures were performed in accordance with the guidelines of the Animal Ethics Committee of the School of Harbin Veterinary Research Institute of the Chinese Academy of Agricultural Sciences. The Animal Ethics Committee Approval Number was SYXK(Hei) 2011022.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: MS, Y-DT, HZ, and Z-JT. Performed the experiments: HX and LX. Contributed reagents or materials and assisted in some experiments: CLi, JZ, BG, QS, JP, QW, GZ, TA, and XC. Analyzed the data: LX, CLe, and Z-JT. Contributed to the writing of the manuscript: HX, LX, and HZ. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Parasitic Resources Center (NPRC-2019-194-30), the National

REFERENCES

- Nathues H, Alarcon P, Rushton J, Jolie R, Fiebig K, Jimenez M, et al. Cost of porcine reproductive and respiratory syndrome virus at individual farm level - An economic disease model. *Prev Vet Med.* (2017) 142:16– 29. doi: 10.1016/j.prevetmed.2017.04.006
- Adams MJ, Lefkowitz EJ, King AMQ, Harrach B, Harrison RL, Knowles NJ, et al. Changes to taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses. *Arch Virol.* (2017) 162:2505–38. doi: 10.1007/s00705-017-3358-5
- Kuhn JH, Lauck M, Bailey AL, Shchetinin AM, Vishnevskaya TV, Bao Y, et al. Reorganization and expansion of the nidoviral family Arteriviridae. *Arch Virol.* (2016) 161:755–68. doi: 10.1007/s00705-015-2672-z
- An TQ, Zhou YJ, Liu GQ, Tian ZJ, Li J, Qiu HJ, et al. Genetic diversity and phylogenetic analysis of glycoprotein 5 of PRRSV isolates in mainland China from 1996 to 2006: coexistence of two NA-subgenotypes with great diversity. *Vet Microbiol.* (2007) 123:43–52. doi: 10.1016/j.vetmic.2007.02.025
- Shi M, Lam TT, Hon CC, Hui RK, Faaberg KS, Wennblom T, et al. Molecular epidemiology of PRRSV: a phylogenetic perspective. *Virus Res.* (2010) 154:7– 17. doi: 10.1016/j.virusres.2010.08.014
- Shi M, Lam TT, Hon CC, Murtaugh MP, Davies PR, Hui RK, et al. Phylogenybased evolutionary, demographical, and geographical dissection of North American type 2 porcine reproductive and respiratory syndrome viruses. J Virol. (2010) 84:8700–11. doi: 10.1128/JVI.02551-09
- Xu H, Li C, Li W, Zhao J, Gong B, Sun Q, et al. Novel characteristics of Chinese NADC34-like PRRSV during 2020-2021. *Transbound Emerg Dis.* (2022). doi: 10.1111/tbed.14485. [Epub ahead of print].
- Leng CL, Tian ZJ, Zhang WC, Zhang HL, Zhai HY, An TQ, et al. Characterization of two newly emerged isolates of porcine reproductive and respiratory syndrome virus from Northeast China in 2013. *Vet Microbiol.* (2014) 171:41–52. doi: 10.1016/j.vetmic.2014.03.005
- An TQ, Tian ZJ, Xiao Y, Li R, Peng JM, Wei TC, et al. Origin of highly pathogenic porcine reproductive and respiratory syndrome virus, China. *Emerg Infect Dis.* (2010) 16:365–7. doi: 10.3201/eid1602.090005
- 10. Guo Z, Chen XX, Li X, Qiao S, Deng R, Zhang G. Prevalence and genetic characteristics of porcine reproductive and respiratory syndrome virus in

Natural Science Foundation of China (Grant no. 32172890 and 32002315), and the China Postdoctoral Fund (Grant no. 2020M680788).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fvets. 2022.945381/full#supplementary-material

Supplementary Figure 1 | Recombination analysis of GXXNF10-1803, GXXNF53-1805, GDXNF60-1805, GXXNF74-1806, GXXNF78-1806, GDXNF229-1811, HNLCL15-1903, HNTZJ1714-2011, and GXTZJ2325-2112. Phylogenic trees were constructed based on different parent regions. Blue represents HP-PRRSV, yellow represents NADC30-like PRRSV, and red represents QYYZ-like PRRSV.

Supplementary Figure 2 | Summary of recombination breakpoints of all reported QYYZ strains in China. (A) The number of recombinant fragments provided by QYYZ in each region of 37 recombinant strains. The ORF5 of all 37 recombinant strains was provided by the QYYZ strain, while the Nsp1 region was provided by the QYYZ strain in only two strains. (B) Recombinants are listed in the key, with the red triangles matching each corresponding breakpoint. Most breakpoints were localized in Nsp2 and GP2, and the backgrounds of the two regions are highlighted in green.

central China during 2016-2017: NADC30-like PRRSVs are predominant. *Microb Pathog.* (2019) 135:103657. doi: 10.1016/j.micpath.2019.103657

- Zhang HL, Zhang WL, Xiang LR, Leng CL, Tian ZJ, Tang YD, et al. Emergence of novel porcine reproductive and respiratory syndrome viruses (ORF5 RFLP 1-7-4 viruses) in China. *Vet Microbiol.* (2018) 222:105– 8. doi: 10.1016/j.vetmic.2018.06.017
- 12. Lu WH, Tun HM, Sun BL, Mo J, Zhou QF, Deng YX, et al. Re-emerging of porcine respiratory and reproductive syndrome virus (lineage 3) and increased pathogenicity after genomic recombination with vaccine variant. *Vet Microbiol.* (2015) 175:332–40. doi: 10.1016/j.vetmic.2014.11.016
- Sun YK, Li Q, Yu ZQ, Han XL, Wei YF, Ji CH, et al. Emergence of novel recombination lineage 3 of porcine reproductive and respiratory syndrome viruses in Southern China. *Transbound Emerg Dis.* (2019) 66:578– 87. doi: 10.1111/tbed.13067
- Zhou L, Kang R, Zhang Y, Ding M, Xie B, Tian Y, et al. Whole genome analysis of two novel Type 2 porcine reproductive and respiratory syndrome viruses with complex genome recombination between lineage 8, 3, and 1 strains identified in Southwestern China. *Viruses.* (2018) 10:60328. doi: 10.3390/v10060328
- Han G, Xu H, Wang K, He F. Emergence of Two different recombinant PRRSV strains with low neutralizing antibody susceptibility in China. *Sci Rep.* (2019) 9:2490. doi: 10.1038/s41598-019-39059-8
- Liu J, Wei C, Lin Z, Fan J, Xia W, Dai A, et al. Recombination in lineage 1, 3, 5 and 8 of porcine reproductive and respiratory syndrome viruses in China. *Infect Genet Evol.* (2019) 68:119–26. doi: 10.1016/j.meegid.2018.12.006
- Zhao J, Xu L, Xu Z, Deng H, Li F, Sun X, et al. Emergence and spread of NADC34-like PRRSV in Southwest China. *Transbound Emerg Dis.* (2022) 69:254–67. doi: 10.1111/tbed.14463
- Sun YK, Han XL, Wei YF, Yu ZQ, Ji CH, Li Q, et al. Phylogeography, phylodynamics and the recent outbreak of lineage 3 porcine reproductive and respiratory syndrome viruses in China. *Transbound Emerg Dis.* (2019) 66:2152–62. doi: 10.1111/tbed.13269
- Zhang WL, Zhang HL, Xu H, Tang YD, Leng CL, Peng JM, et al. Two novel recombinant porcine reproductive and respiratory syndrome viruses belong to sublineage 3.5 originating from sublineage 3.2. *Transbound Emerg Dis.* (2019) 66:2592–600. doi: 10.1111/tbed.13320

- Zhang H, Leng C, Ding Y, Zhai H, Li Z, Xiang L, et al. Characterization of newly emerged NADC30-like strains of porcine reproductive and respiratory syndrome virus in China. *Arch Virol.* (2019) 164:401–11. doi: 10.1007/s00705-018-4080-7
- Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* (2016) 33:1870– 4. doi: 10.1093/molbev/msw054
- Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* (2021) 49:W293– 6. doi: 10.1093/nar/gkab301
- Deng Y, Pan Y, Wang D, Zhou Q, Bi Y, Chen F, et al. Complete genome sequence of porcine reproductive and respiratory syndrome virus strain QY2010 reveals a novel subgroup emerging in China. J Virol. (2012) 86:7719– 20. doi: 10.1128/JVI.00977-12
- Fan B, Wang H, Bai J, Zhang L, Jiang P. A novel isolate with deletion in GP3 gene of porcine reproductive and respiratory syndrome virus from mideastern China. *Biomed Res Int.* (2014) 2014:306130. doi: 10.1155/2014/306130
- Zhang Q, Bai J, Hou H, Song Z, Zhao Y, Jiang P. A novel recombinant porcine reproductive and respiratory syndrome virus with significant variation in cell adaption and pathogenicity. *Vet Microbiol.* (2017) 208:150– 8. doi: 10.1016/j.vetmic.2017.07.028
- 26. Zhang X, Li Y, Xiao S, Yang X, Chen X, Wu P, et al. High-frequency mutation and recombination are responsible for the emergence of novel porcine reproductive and respiratory syndrome virus in northwest China. *Arch Virol.* (2019) 164:2725–33. doi: 10.1007/s00705-019-04373-z
- 27. Dong JG, Yu LY, Wang PP, Zhang LY, Liu YL, Liang PS, et al. A new recombined porcine reproductive and respiratory syndrome virus virulent strain in China. *J Vet Sci.* (2018) 19:89–98. doi: 10.4142/jvs.2018.19.1.89
- Zhang Z, Qu X, Zhang H, Tang X, Bian T, Sun Y, et al. Evolutionary and recombination analysis of porcine reproductive and respiratory syndrome isolates in China. *Virus Genes.* (2020) 56:354–60. doi: 10.1007/s11262-020-01751-7
- Zhou L, Kang R, Zhang Y, Yu J, Xie B, Chen C, et al. Emergence of two novel recombinant porcine reproductive and respiratory syndrome viruses 2 (lineage 3) in Southwestern China. *Vet Microbiol.* (2019) 232:30– 41. doi: 10.1016/j.vetmic.2019.01.026
- Zhang F, Song D, Ye Y, Guo N, Zhang M, Li H, et al. Complete genome sequence of a novel subgenotype of porcine reproductive and respiratory syndrome virus strain JX/CH/2016, isolated in Jiangxi, China. *Genome Announc.* (2017) 5:e00980-17. doi: 10.1128/genomeA.00980-17
- Yu F, Yan Y, Shi M, Liu HZ, Zhang HL, Yang YB, et al. Phylogenetics, genomic recombination, and NSP2 Polymorphic patterns of porcine reproductive and respiratory syndrome virus in China and the United States in 2014-2018. J Virol. (2020) 94:19. doi: 10.1128/JVI.01813-19
- 32. Liu J, Xu Y, Lin Z, Fan J, Dai A, Deng X, et al. Epidemiology investigation of PRRSV discharged by faecal and genetic variation of ORF5. *Transbound Emerg Dis.* (2021) 68:2334–44. doi: 10.1111/tbed.13894
- 33. Kim SC, Moon SH, Jeong CG, Park GS, Park JY, Jeoung HY, et al. Wholegenome sequencing and genetic characteristics of representative porcine reproductive and respiratory syndrome virus (PRRSV) isolates in Korea. *Virol J.* (2022) 19:66. doi: 10.1186/s12985-022-01790-6
- 34. Han G, Lei K, Xu H, He F. Genetic characterization of a novel recombinant PRRSV2 from lineage 8, 1 and 3 in China with significant variation in replication efficiency and cytopathic effects. *Transbound Emerg Dis.* (2020) 67:1574–84. doi: 10.1111/tbed.13491
- 35. Song J, Gao P, Kong C, Zhou L, Ge X, Guo X, et al. The nsp2 hypervariable region of porcine reproductive and respiratory syndrome virus strain JXwn06 is associated with viral cellular tropism to primary porcine alveolar macrophages. J Virol. (2019) 93:e01436-19. doi: 10.1128/JVI.01436-19
- Veit M, Matczuk AK, Sinhadri BC, Krause E, Thaa B. Membrane proteins of arterivirus particles: structure, topology, processing and function. *Virus Res.* (2014) 194:16–36. doi: 10.1016/j.virusres.2014.09.010
- Popescu LN, Trible BR, Chen N, Rowland RRR. GP5 of porcine reproductive and respiratory syndrome virus (PRRSV) as a target for homologous and broadly neutralizing antibodies. *Vet Microbiol.* (2017) 209:90–96. doi: 10.1016/j.vetmic.2017.04.016
- 38. Sui X, Guo X, Jia H, Wang X, Lin W, Li M, et al. Genomic sequence and virulence of a novel NADC30-like porcine reproductive and respiratory

syndrome virus isolate from the Hebei province of China. *Microb Pathog.* (2018) 125:349-60. doi: 10.1016/j.micpath.2018.08.048

- 39. de Lima M, Pattnaik AK, Flores EF, Osorio FA. Serologic marker candidates identified among B-cell linear epitopes of Nsp2 and structural proteins of a North American strain of porcine reproductive and respiratory syndrome virus. *Virology*. (2006) 353:410–21. doi: 10.1016/j.virol.2006.05.036
- Vashisht K, Goldberg TL, Husmann RJ, Schnitzlein W, Zuckermann FA. Identification of immunodominant T-cell epitopes present in glycoprotein 5 of the North American genotype of porcine reproductive and respiratory syndrome virus. *Vaccine.* (2008) 26:4747–53. doi: 10.1016/j.vaccine.2008.06.047
- Ostrowski M, Galeota JA, Jar AM, Platt KB, Osorio FA, Lopez OJ. Identification of neutralizing and nonneutralizing epitopes in the porcine reproductive and respiratory syndrome virus GP5 ectodomain. *J Virol.* (2002) 76:4241–50. doi: 10.1128/jvi.76.9.4241-4250.2002
- 42. Zhao K, Ye C, Chang XB, Jiang CG, Wang SJ, Cai XH, et al. Importation and recombination are responsible for the latest emergence of highly pathogenic porcine reproductive and respiratory syndrome virus in China. *J Virol.* (2015) 89:10712–6. doi: 10.1128/JVI.01446-15
- Zhu Z, Yuan L, Hu D, Lian Z, Yao X, Liu P, et al. Isolation and genomic characterization of a Chinese NADC34-like PRRSV isolated from Jiangsu province. *Transbound Emerg Dis.* (2021) 69:254–267. doi: 10.1111/tbed.14392
- Kappes MA, Faaberg KS. PRRSV structure, replication and recombination: Origin of phenotype and genotype diversity. *Virology*. (2015) 479–480:475– 86. doi: 10.1016/j.virol.2015.02.012
- 45. Guo Z, Chen XX, Li R, Qiao S, Zhang G. The prevalent status and genetic diversity of porcine reproductive and respiratory syndrome virus in China: a molecular epidemiological perspective. *Virol J.* (2018) 15:2. doi: 10.1186/s12985-017-0910-6
- 46. Hou FH, Lee WC, Liao JW, Chien MS, Kuo CJ, Chung HP, et al. Evaluation of a type 2 modified live porcine reproductive and respiratory syndrome vaccine against heterologous challenge of a lineage 3 highly virulent isolate in pigs. *PeerJ.* (2020) 8:e8840. doi: 10.7717/peerj.8840
- Lager KM, Schlink SN, Brockmeier SL, Miller LC, Henningson JN, Kappes MA, et al. Efficacy of Type 2 PRRSV vaccine against Chinese and Vietnamese HP-PRRSV challenge in pigs. *Vaccine*. (2014) 32:6457– 62. doi: 10.1016/j.vaccine.2014.09.046
- Jantafong T, Sangtong P, Saenglub W, Mungkundar C, Romlamduan N, Lekchareonsuk C, et al. Genetic diversity of porcine reproductive and respiratory syndrome virus in Thailand and Southeast Asia from 2008 to 2013. *Vet Microbiol.* (2015) 176:229–38. doi: 10.1016/j.vetmic.2015.01.017
- Fukunaga W, Hayakawa-Sugaya Y, Koike F, Van Diep N, Kojima I, Yoshida Y, et al. Newly-designed primer pairs for the detection of type 2 porcine reproductive and respiratory syndrome virus genes. J Virol Methods. (2021) 291:114071. doi: 10.1016/j.jviromet.2021.114071
- Ramirez M, Bauermann FV, Navarro D, Rojas M, Manchego A, Nelson EA, et al. Detection of porcine reproductive and respiratory syndrome virus (PRRSV) 1-7-4-type strains in Peru. *Transbound Emerg Dis.* (2019) 66:1107– 13. doi: 10.1111/tbed.13134
- Kim SC, Jeong CG, Park GS, Park JY, Jeoung HY, Shin GE, et al. Temporal lineage dynamics of the ORF5 gene of porcine reproductive and respiratory syndrome virus in Korea in 2014-2019. *Arch Virol.* (2021) 166:2803– 15. doi: 10.1007/s00705-021-05169-w
- Paploski IAD, Pamornchainavakul N, Makau DN, Rovira A, Corzo CA, Schroeder DC, et al. Phylogenetic structure and sequential dominance of sublineages of PRRSV type-2 lineage 1 in the United States. *Vaccines.* (2021) 9:9060608. doi: 10.3390/vaccines9060608
- 53. Sun YF, Liu Y, Yang J, Li WZ, Yu XX, Wang SY, et al. Recombination between NADC34-like and QYYZ-like strain of porcine reproductive and respiratory syndrome virus with high pathogenicity for piglets in China. *Transbound Emerg Dis.* (2022). doi: 10.1111/tbed.14471. [Epub ahead of print].
- 54. Ding Y, Wubshet AK, Ding X, Zhang Z, Li Q, Dai J, et al. Evaluation of four commercial vaccines for the protection of piglets against the highly pathogenic porcine reproductive and respiratory syndrome virus (hp-PRRSV) QH-08 strain. *Vaccines*. (2021) 9:1020. doi: 10.3390/vaccines90 91020
- 55. Li L, Chen J, Cao Z, Cao Y, Guo Z, Tong W, et al. Recombinant bivalent live vectored vaccine against classical swine fever and HP-PRRS revealed adequate

heterogeneous protection against NADC30-like strain. Front Microbiol. (2021) 12:822749. doi: 10.3389/fmicb.2021.822749

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xu, Xiang, Tang, Li, Zhao, Gong, Sun, Leng, Peng, Wang, Zhou, An, Cai, Tian, Zhang and Song. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.