ASFV pp62 Protein Has 9 Dominant B–T Cell Combined Epitopes

Cao Qianda[1*], Dai Bihong[2], Yang Yang[1], Cao Xuelian[3]

1. Fushun Animal Disease Control Center, Zigong 643200, China; 2. Neijiang Vocational & Technical College, Neijiang 641000, China; 3. Yanjiang Animal Disease Control Center, Ziyang 641300, China

Abstract **[Objective]** The paper was to study the basic characteristics of the pp62 protein of the African swine fever virus strain Pig/HLJ/2018, and to provide more basic data for the study of the virus. **[Method]** We used the ProtParam program to analyze the physical and chemical properties of the pp62 protein. TMHMM-2.0 and SignalP-5.0 were used to analyze protein transmembrane region and signal peptide, and Lasergene 7.0 Protean program was used to study protein antigen index, hydrophilicity, surface accessibility and titration curve. Protein N-glycosylation site and O-glycosylation site came out with NetNGlyc-1.0 and NetOGlyc-4.0 online servers. Then, PSIPRED-4.0, NetSurfP-2.0, and PSRSM were used to analyze protein secondary structure, BepiPred1.0 and IEDB tools were used to analyze protein B-cell epitopes, and NetMHC 4.0 and NetMHCpan tools were used to analyze protein T-cell epitopes. And we used Swiss-Model to analyze the high-level structure of the protein, the EzMol tool to visually analyze B-T cell combined epitopes, and finally, MEGA 7.0 to analyze the genetic evolutionary relationship of the protein. **[Result]** The pp62 protein of African swine fever viral strain Pig/HLJ/2018 had a molecular weight of 60.5 kDa. It was a hydrophilic acid labile protein, and had no transmembrane region and signal peptide. There were 5 N-glycosylation sites and 4 O-glycosylation sites. Analysis of the secondary structure of the protein showed that the proportions of helix, coil and strands were 45.5%, 41.7% and 12.8%, respectively. The study of dominant epitopes revealed that there were 14 dominant B-cell epitopes and 16 dominant T-cell epitopes. And 9 dominant B-T cell combined epitopes located on the surface of the protein molecule were found. The phylogenetic tree constructed with the pp62 protein showed that the evolutionary relationship of Pig/HLJ/2018 strain was the closest to Georgia/2007/1, which belonged to genotype II. **[Conclusion]** The results will provide basic information for pp62 research.

Keywords African swine fever virus; pp62; CP530R; Dominant B-T cell combined epitope; Phylogenetic tree; Visual analysis

The African swine fever virus (ASFV) is the pathogenic microorganism that causes African swine fever (ASF), with virus particles in the shape of a positive twenty-sided body with a diameter of 200 nm and consisting of five layers of structure from the inside out: nucleoid, coreshell, inner membrane, outer capsid, and outer envelope[1]. The viral genome is linear double-stranded DNA (dsDNA) with a similar genomic structure to members of the Rainbow virus and Poxvirus sections, while the replication strategy is the same as that of the Poxvirus[2-3]. An iconic phenomenon of ASFV genomic protein expression is the synthesis of two polyprotein precursors, pp62 and pp220. pp62 is the encoded protein precursor of the *CP530R* gene, belonging to the advanced protein of viral infection, which is hydrolyzed by later S273R protease to form the important structural protein of viral particles, p15 and p35, which make up the main part of the virus nuclear shell, and is the main marker of virus assembly[4–6]. Therefore, pp62 plays an important role in the assembly and maturation of ASFV virus particles. Gallardo and others used pp62 as a diagnostic antigen to show that the ELISA diagnostic specificity of pp62 protein was higher than those of p32 and p54, reaching 99%, and the sensitivity reached 97%, while for serum samples stored for 1 month at 37 ℃, the pp62 specificity and sensitivity reached 100%[7–8]. That suggested that pp62 produced antibodies were much more stable than that of the other two proteins, or more affinity, and it had the potential to act as a high-quality diagnostic antigen. At present, there is little research on pp62 protein, and this paper analyzes the basic characteristics, physical and chemical properties, valuable antigen epitopes, protein structure and the evolutionary relationship of pp62 protein of ASFV strain pig/HLJ/2018 isolated from China in 2018, intending to provide basic information for pp62 research.

Received: 2022–10–12 Accepted: 2022–11–28

*Corresponding author. E-mail: cqd123.2009@163.com

1 Materials and Methods

1.1 Sequence *CP530R* gene of ASFVPig/HLJ/2018 strain is located in the virus genome 117 290 bp to 124 720 bp, and its encoded protein is pp62. The protein sequence (GeneBank ID: QBH90580.1) was obtained in NCBI.

1.2 Methods

1.2.1 Analysis of the physical and chemical properties of pp62 protein. The physical and chemical properties of pp62 protein molecular weight, isoelectric point, amino acid composition, atomic number,

hydrophobic properties, *etc*., were analyzed using ExPASy online analysis tools ProtParam[9] (https: //web.expasy. org/protparam/) and Lasergene software.

1.2.2 Protein transmembrane region and signal peptide analysis. The protein cross-membrane region was analyzed by TMHMM-2.0[10] (https: //services. healthtech. dtu. dk/service.php?TMHMM-2.0) online server, and the protein signal peptide was analyzed by the signalP-5.0[11] (https: //services.healthtech.dtu. dk/service. php?SignalP-5.0).

1.2.3 Protein antigen index, hydrophilicity, surface accessibility, titration curve analysis. The protein antigen index, hydrophilicity, surface accessibility, and titration curve were analyzed using the Lasergene 7.0 Protean module.

1.2.4 Protein N-glycosylation site and O-glycosylation site analysis. N-glycosylation sites and O-glycosylation sites of pp62 protein were planned by NetNGlyc-1.0 online server (https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0) and NetOGlyc-4.0 online server (https://services.healthtech.dtu.dk/service.php?NetOGlyc-4.0).

1.2.5 pp62 protein secondary structure prediction. PSIPRED-4.0[12](http://bioinf.cs.ucl.ac.uk/psipred/), NetSurfP-2.0[13] (https://services.healthtech.dtu.dk/service.php?NetSurfP-2.0) and PSRSM[14-15] (http://210.44.144.20:82/protein_PSRSM/default.aspx) online servers were used to analyze the secondary structure of protein sequence, and the results of the three methods were combined to obtain more accurate protein secondary structure information. Removing areas helix, strand that were difficult to form antigen epitopes, and filtering out areas that were prone to form antigen epitopes, in other words were potential advantage epitope regions.

1.2.6 Dominant B, T cell antigen epitope prediction of pp62. The online analysis tools BepiPred1.0 [16] (http://www.cbs.dtu.dk/services/BepiPred-1.0) and IEDB[17] (http://tools. iedb.org/bcell/) were used to analyze B-cell antigen epitope region, and the results of the two methods were brought together to arrive at the B-cell epitope region. Net-MHC 4.0 online servers [18] (http://www.cbs.dtu.dk/services/NetMHC/) and NetMHCpan [19] (https://services.healthtech.dtu.dk/service.php?NetMHCpan-4.1) were used to analyze sequence T-cell antigen epitope region: species selection, pig; MHC species, SLA-10401, SLA-20401 and SLA-30401; peptide segment length selection, 9aa, other parameters as default. Lower rank value indicated the peptide segment affinity was stronger. The results were combined to get T-cell epitope region. The secondary structure analysis results, protein B and T-cell epitope analysis results, were brought together to result in dominant B, T-cell antigen epitope.

1.2.7 Advanced structure prediction of pp62 protein. The advanced structure of the pp62 protein was predicted using the Swiss-Model online prediction server [20] (https://swissmodel.expasy.org/).

1.2.8 Visual analysis of pp62 protein advantage B-T cell combined epitope. Based on the optimal B, T cell antigen epitopes, B-T cell combined antigen epitopes were comprehensively screened out. By visualizing the filtered epitopes through EzMol[21] (http://www.sbg.bio.ic.ac.uk/ezmol/), the B-T cell combined epitopes exposed to protein surface were then screened.

1.2.9 The evolutionary tree based on pp62 protein of ASFV HLJ/2018 strain. Sequences were got through the NCBI BlastP program. Clustal W was utilized to compare the selected sequences. The evolutionary tree was drawn using MEGA 7.0[22], bootstrap was set to 1 000 times, and Poisson model was selected for validation. The built evolutionary tree was edited and labeled through iTOL[23] (https://itol.embl.de/).

2 Results and Analysis

2.1 Physical and chemical properties of pp62 protein Depending on the analysis results of ProtParam and Lasergene, the pp62 protein contained 530 amino acids. The amino acid with the highest content was leucine (Leu, 10%), followed by threonine (Thr, 7.4%), proline (Pro, 6.8%), lysine (Lys, 6.8%), and isoleucine (Ile, 6.8%). The molecular weight was 60517.26 Da, and the theoretical isoelectric point was 6.53. The molecular formula was $C_{2738}H_{4252}N_{722}O_{793}S_{17}$, which contained 8522 atoms in total. 160 charged amino acids accounted for 30.19%, of which 60 negatively charged amino acids (acidic) accounted for 11.32%, and 56 positively charged amino acids (basic) accounted for 10.57%, that is, theoretically, the protein was an negatively charged acidic protein. The extinction coefficient was 62 800 – 63 175 $M^{-1}cm^{-1}$ (measured at 280 nm in water), dependent on the number of disulfide bonds formed by cysteine pairing. The half-life in mammalian reticulocytes was 30 h, and those in yeast and *E. coli* were more than 20 h and more than 10 h, respectively. Protein instability (II) was 42.41, which was an unstable protein. The fat-soluble index was 86.64: 178 hydrophobic amino acids accounted for 33.58% and 148 polar amino acids accounted for 27.92%. The total average hydrophilicity (GRAVY) of pp62 was -0.361, which was a fat-soluble hydrophilic protein.

2.2 Protein transmembrane region and signal peptide analysis According to the analysis results of online analysis software, there was no transmembrane region (Fig.1) and signal peptide (Fig.2) in the pp62 protein.

2.3 pp62 protein hydrophilicity, antigen index, surface accessibility, flexible area, and titration curve analysis Results are presented in Figs.3–4. The protein hydrophilicity analysis chart showed that most of the proteins were hydrophilic, and were comprehensively assessed as hydrophilic proteins, which was consistent with the ExPASy analysis result. Comprehensive antigen index, surface accessibility and flexibility analysis results displayed that the flexible areas with high antigen index and exposed on the surface of the

protein were the potential dominant epitope areas, mainly distributed in the peptides 17-24aa, 50-70aa, 85-97aa, 102-108aa, 113-116aa, 127-133aa, 141-175aa, 183-190aa, 215-218aa, 239-246aa, 263-266aa, 291-297aa, 304-310aa, 340-348aa, 368-375aa, 384- 395aa, 403-412aa, 424-438aa, 456-478aa, 495-500aa, 514-524aa.

2.4 Protein N-glycosylation site and O-glycosylation site analysis results The analysis found that the pp62 protein had 5 N-glycosylation sites, located at 34aa, 262 aa, 417 aa, 464 aa, and 486 aa, respectively, as showed in Fig. 5. 4 O-glycosylation sites were located at 169 aa, 172 aa, 420 aa, and 470 aa, respectively.

2.5 Results of protein secondary structure analysis At present, the accuracy of each secondary structure analysis method can not reach 100%, and there are a few differences in the results among analysis methods. The secondary structure of pp62 protein was analyzed by three analytical methods. The prediction results showed that the proportions of helix, coil and strands were 45.5%, 41.7% and 12.8%, respectively. The secondary structure map is shown in Fig. 6, and the detailed statistical results are shown in Tab.1. The helix, strand regions were sometimes difficult to form epitopes, and the coiled regions could be used as candidate regions for screening of dominant epitopes.

2.6 Dominant B-cell epitope of pp62 protein By combining the results of BepiPred 1.0 and IEDB online prediction of B-cell epitopes, as well as the results of potential dominant epitope regions of the secondary structure of pp62 protein, 14 dominant B-cell epitope sequences were analyzed. According to the two analyses, the average of the basic scores of the method was ranked from high to low, and the higher the ranking, the higher the score. The statistics of the results are shown in Tab.2.

2.7 Dominant T-cell epitope of pp62 protein According to the analysis results of two online T-cell epitope analysis methods, combining the SLA-10401, SLA-20401, and SLA-30401 epitope Rank values and the expected number of binding sites, the secondary structure of the binding protein was easy to form an epitope. In potential dominant epitope regions, 16 potential dominant T cell epitope sequences were screened, and were arranged in order of potential binding ability, with strong binding first (Tab.3).

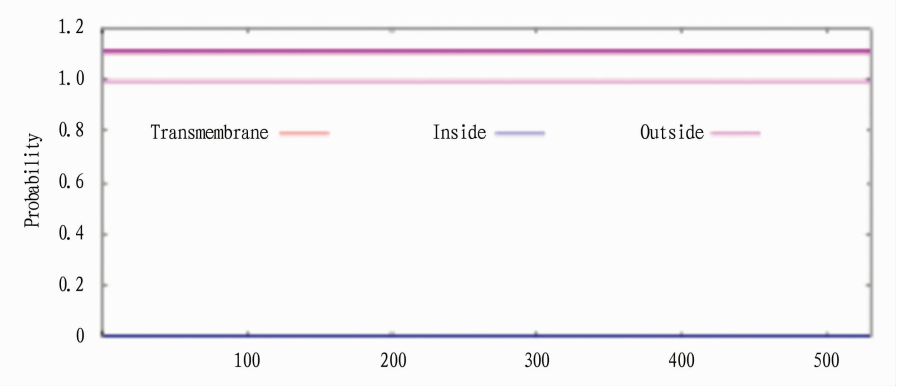


Fig.1 Transmembrane region prediction of pp62

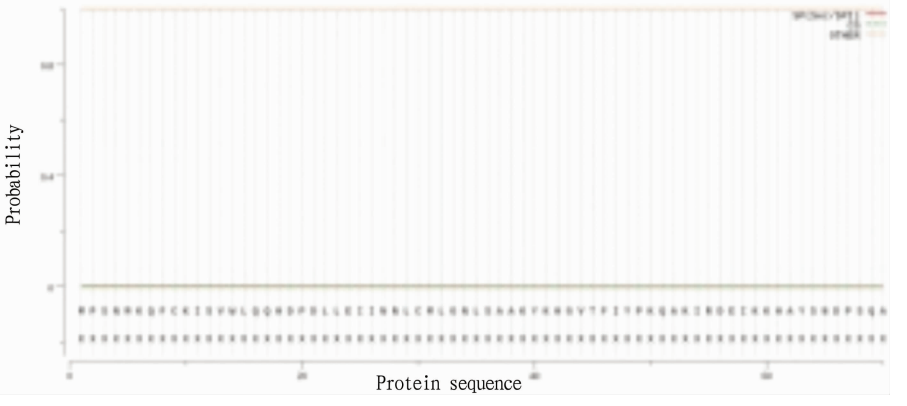


Fig.2 Signal peptide prediction results

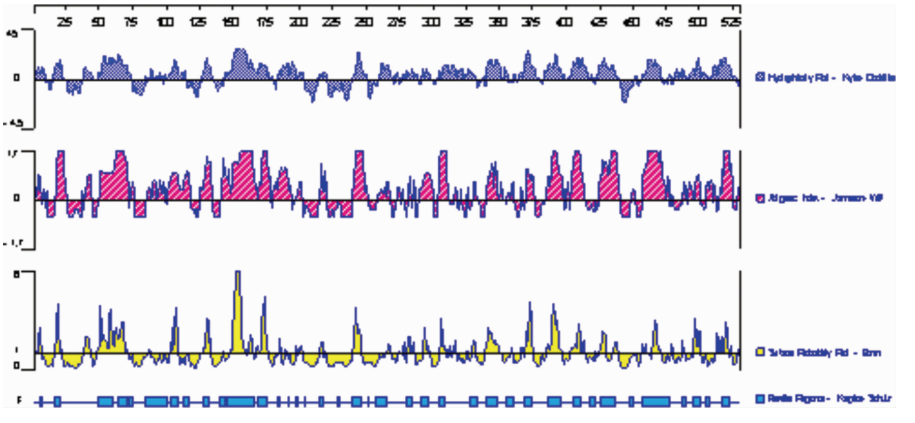


Fig.3 Protein hydrophilicity, antigen index, surface accessibility and flexibility area titration curve analysis by Lasergene 7.0 Protean module

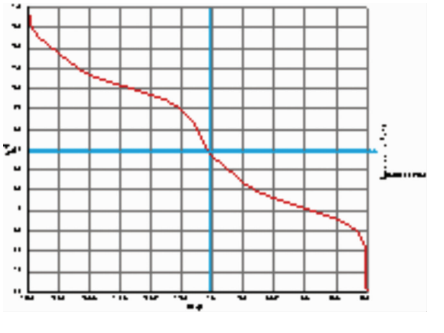


Fig.4 pp62 titration curve analysis results

2.8 Prediction results of high –level protein structure The advanced structure of pp62 was analyzed through the Swiss-model online analysis tool, and the analysis results were verified with the QMEAN 4 tool that came with the system, and the advanced structure of the protein was obtained, as shown in Fig.7.

2.9 Dominant B-T cell combined epitope analysis results Combining the results of epitope analysis of dominant B and T-cell epitopes, 9 combined epitopes were found. The statistical results are presented in Tab.4. Fig.8 shows the location of the 9 B-T cell combined epitopes in the protein. All 9 peptides were lain on the surface of the protein, suggested that the 9 B-T cell combined epitopes screened out had the possibility of developing vaccines.

2.10 The results of phylogenetic tree construction based on pp62 protein Through the BlastP program, 14 full-length protein sequences homologous to the pp62

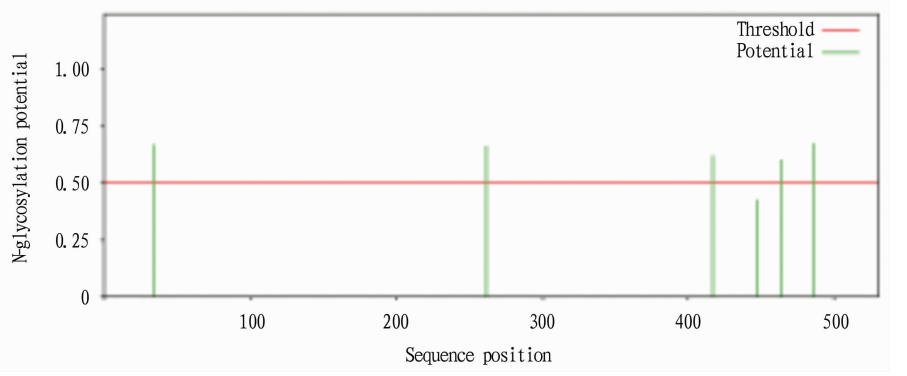
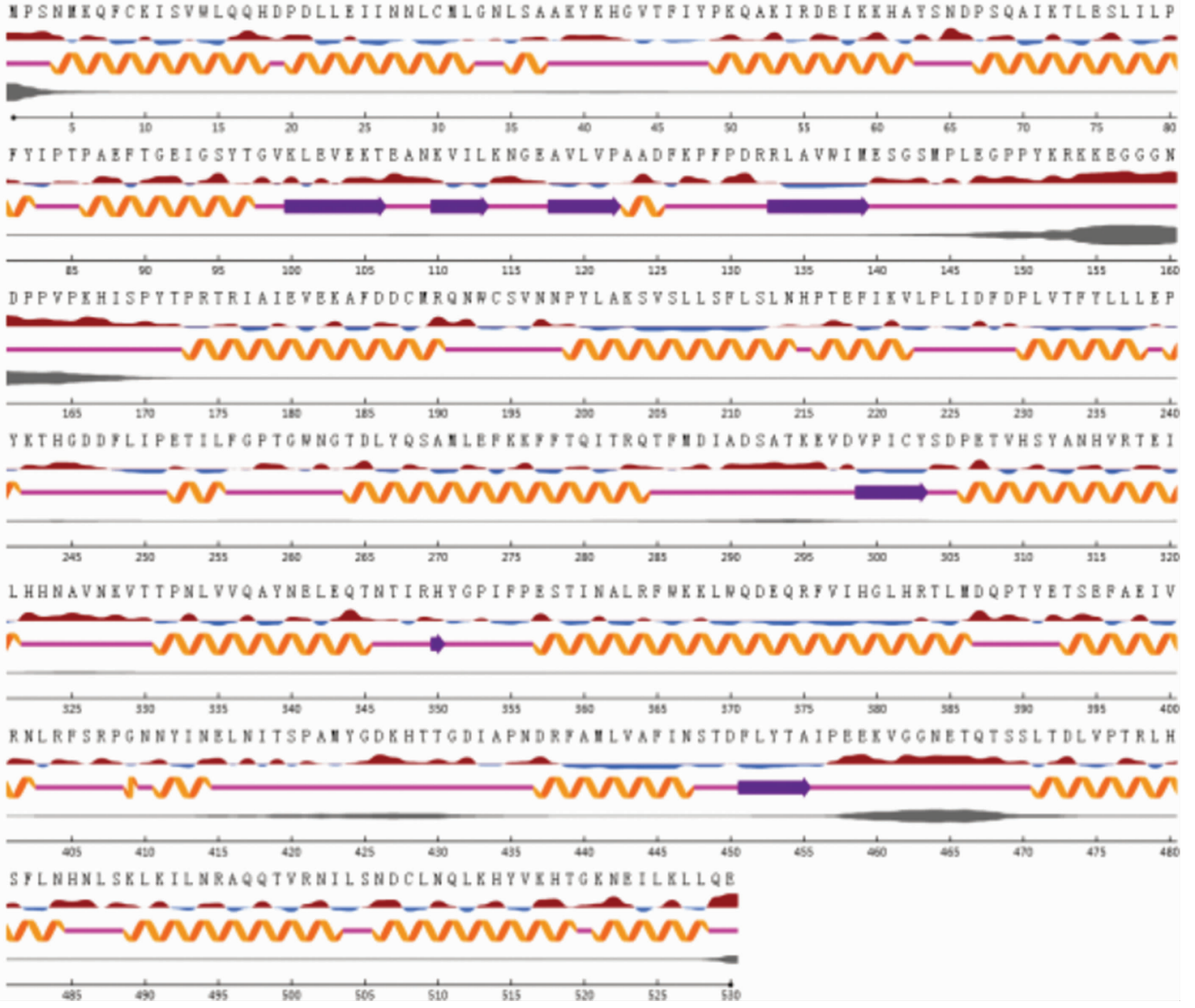


Fig.5 pp62 N-glycosylation site prediction result



Note: Helix; Strand; Coil; Red is exposed and blue is buried, thresholded at 25%; Thickness of line equals probability of disordered residue.

Fig.6 pp62 protein secondary structure analysis

protein were screened. By comparison with the pp62 protein sequence, a phylogenetic tree was successfully constructed, and the constructed phylogenetic tree was further annotated by iTOL, as shown in Fig.9. From the evolutionary tree of ASFV constructed with pp62 protein, it could be seen that the HLJ/2018 strain together with Georgia/2007/1 strain was in the same branch and was closely related, which all belonged to genotype II. The sequence comparison showed that the sequence similarity was 100%.

3 Discussion

Analysis of physicochemical properties of the pp62 protein showed that the protein was an acidic protein with a molecular weight of 60.5kDa. The half-life in the animal reticulum was the longest of 30 h. Analysis of protein stability and fat solubility index showed that the protein was an unsteady hydrophilic protein. The protein had no transmembrane region and signal peptide, suggesting the protein did not get the function of a protein channel and guide protein to subcellular organelles. It was not a secreted protein. Combined with the structure of ASFV, it could be observed the pp62 protein was mainly involved in the assembly of the virus capsid, which was a late structural protein. Through the comprehensive analysis of protein hydrophilicity, antigenic index, and flexible region, it was found that pp62 had good antigenicity, with rather wide distribution area, and 21 regions that may form the dominant epitopes had been found.

Tab.1 pp62 protein secondary structure analysis results

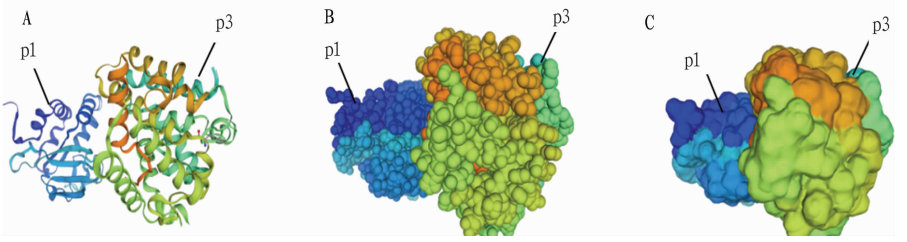
Method	Position//aa		
	Helix	Strand	Coil
PSRSM	6–8, 11–18, 20–35, 52–63, 67–78, 89–91, 123–125, 165, 173–188, 199–212, 218–220, 222–225, 230, 232–237, 262, 264, 267, 270–280, 286–292, 306–324, 329–344, 359–386, 393–406, 421–424, 437–447, 468–469, 478, 480, 484–489, 491–494, 496–519, 521–528	44–47, 100–105, 110–113, 117–120, 133–139, 221, 238–239, 249, 254–256, 301, 303, 416–419, 428–430, 451–455, 481–482	1–5, 9–10, 19, 36–43, 48–51, 64–66, 79–88, 92–99, 106–109, 114–116, 121–122, 126–132, 140–164, 166–172, 189–198, 213–217, 226–229, 231, 240–248, 250–253, 257–261, 263, 265–266, 268–269, 281–285, 293–300, 302, 304–305, 325–328, 345–358, 387–392, 407–415, 420, 425–427, 431–436, 448–450, 456–467, 470–477, 479, 483, 490, 495, 520, 529–530
PSIPRED	4–9, 11–18, 20–32, 35–37, 52–63, 67–77, 86–89, 176–190, 199–213, 230–238, 264–280, 283–293, 306–320, 332–344, 357–387, 393–405, 422–423, 436–447, 458–460, 471–474, 476–483, 487–504, 507–519, 521–529	44–49, 100–105, 110–113, 118–120, 134–140, 219–224, 254–255, 301–303, 452–455	1–3, 10, 19, 33–34, 38–43, 50–51, 64–66, 78–85, 90–99, 106–109, 114–117, 121–133, 141–175, 191–198, 214–218, 225–229, 239–253, 256–263, 281–282, 294–300, 304–305, 321–331, 345–356, 388–392, 406–421, 424–435, 448–451, 456–457, 461–470, 475, 484–486, 505–506, 520, 530
NetSurfP	4–18, 20–32, 35–37, 49–62, 67–82, 86–97, 123–125, 173–190, 199–214, 216–222, 230–238, 240–241, 252–255, 264–284, 306–321, 331–345, 357–386, 393–402, 409, 411–414, 437–447, 471–484, 489–503, 506–519, 521–528	100–106, 110–113, 118–122, 133–139, 299–303, 350, 451–455	1–3, 19, 33–34, 38–48, 63–66, 83–85, 98–99, 107–109, 114–117, 126–132, 140–172, 191–198, 215, 223–229, 239, 242–251, 256–263, 285–298, 304–305, 322–330, 346–349, 351–356, 387–392, 403–408, 410, 415–436, 448–450, 456–470, 485–488, 504–505, 520, 529–530

Tab. 2 pp62 dominant B-cell epitope prediction results

Rank	Position//aa	Sequence	Length
1	140–175	ESGSMPLEGPPYKRKKEGGGNDPPVPKHISPYTPRT	36
2	456–477	IPEEKVGGNETQTSSLTDLVPT	22
3	385–395	LMDQPTYETSE	11
4	256–266	FGPTGWNGTDL	11
5	239–253	EPYKTHGDDFLIPET	15
6	52–72	AKIRDEIKKHAYSNDPSQAIK	21
7	78–99	ILPFYIPTPAEFTGEIGSYTGV	22
8	302–315	CYSDPETVHSYANH	14
9	281–300	ITRQTFMDIADSATKEVDVP	20
10	121–133	VPAADFKPFPDRR	13
11	403–436	LRFSRPGNNYINELNITSPAMYGDKHTTGDIAPN	34
12	103–109	VEKTEAN	7
13	341–358	ELEQTNTIRHYGPIFPES	18
14	321–332	LHHNAVNKVTTP	12

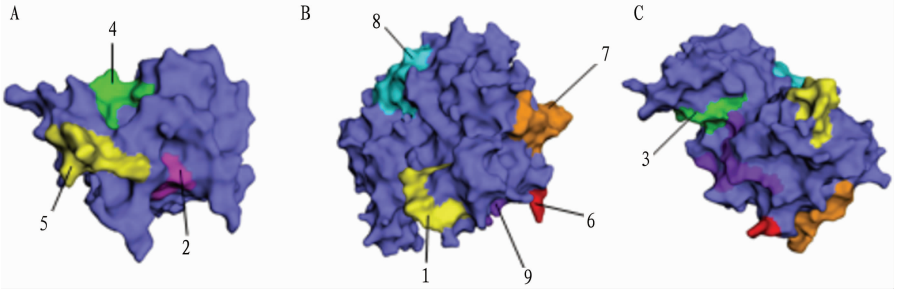
Tab.3 Dominant T-cell epitope of pp62 protein

No.	Position//aa	Peptide
1	383–391	RTLMDQPTY
2	38–46	AKYKHGVTF
3	211–219	LSLNHPTEF
4	227–235	DFDPLVTFY
5	73–81	TLESLILPF
6	143–151	SMPLEGPPY
7	416–424	LNITSPAMY
8	445–453	FINSTDFLY
9	40–48	YKHGVTFIY
10	118–126	AVLVPAADF
11	129–137	FPDRRLAVW
12	163–171	PVPKHISPY
13	253–261	TILFGPTGW
14	259–267	TGWNGTDLY
15	304–312	SDPETVHSY
16	404–412	RFSRPGNNY



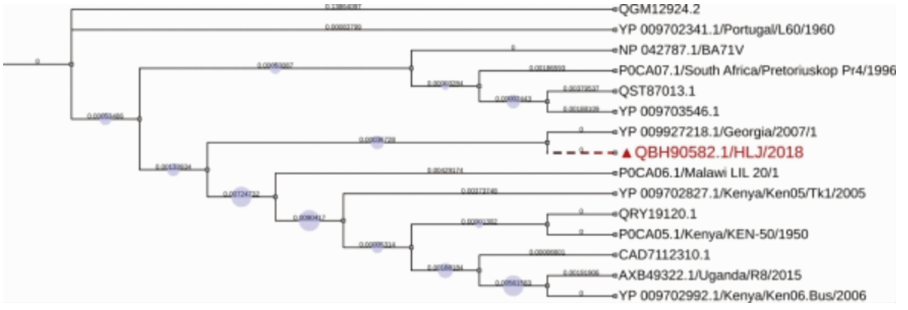
Note: A. Secondary structure of pp62 (blue is p15, others are p35); B. Atomic structure of pp62; C. Protein surface of pp62.

Fig.7 pp62 high level structure



Note: A. 3 dominant B-T cell combined epitopes on p15; B-C. 5 dominant B-T cell combined epitopes on p35; numbers correspond to Tab.4.

Fig.8 Visualization of pp62 dominant B-T cell combined epitopes



Note: HLJ/2018 is displayed in bold red and marked with"▲".

Fig.9 Phylogenetic tree construction based on pp62

Glycosylation is an important post-translational modification as well as a complex modification, it is deemed to affect protein folding, positioning and transport, protein solubility, antigenicity, biological activity, and half-life, as well as cell-to-cell interactions[24-28]. Glycosylation sites play a significant role in protein conformation and change the function and stability of protein biological products, virus replication, and so on[29-32]. In this paper, we got 5 N-glycosylation sites and 4 O-glycosylation sites of pp62 protein through online server analysis, which is useful to study and change the structure and function of the protein.

The secondary structure of the protein is an indicator for the study of protein function[33-35], and residue mutations could lead to the change of the protein structure and function[36-38]. The secondary structure information of pp62 protein was comprehensively analyzed by the three methods of PSIPRED-4.0, NetSurfP-2.0, and PSRSM, and the results of each analysis method were counted and displayed. As a whole, the proportion of helix was high (45.5%), followed by coils (41.7%), and that of strand was the least (about 12.8%). By removing the helix, strand, and other areas that was difficult to form epitopes in the secondary structure, the results of protein B and T cell epitope analysis were combined to obtain the dominant B and T cell

Tab.4 Dominant B-T cell combined epitopes of pp62

No.	Position//aa	Peptide	Length
1	385–391	LMDQPTY	7
2	78–81	ILPF	4
3	416–424	LNITSPAMY	9
4	121–126	VPAADF	6
5	129–133	FPDRR	5
6	163–171	PVPKHISPY	9
7	256–266	FGPTGWNGTDL	11
8	304–312	SDPETVHSY	9
9	404–412	RFSRPGNNY	9

epitopes of pp62 protein. By integrating the protein secondary structure and B-cell epitope analysis results, we got 14 dominant B-cell epitopes. The statistical results were described in detail in Tab.2. Similarly, 16 dominant T-cell epitopes were gained, and the statistical results are shown in Tab.3.

The visual analysis of the high-level structure of the protein could visually display information such as the secondary structure, atomic structure, and surface structure of the protein, which showed the discovery of protein functional sites[39-41]. The high-level structure of pp62 was analyzed by the Swiss-model online analysis tool. The analysis results showed the protein was mainly composed of 2 parts, namely, the two structural proteins, p15 and p35. The visualization results respectively showed the protein secondary structure distribution, atomic conformation, and surface conformation, and the 9 dominant B-T cell combined epitopes selected were visually displayed through EzMol. The 9 combined epitopes were all located on the surface of the protein, suggesting that these epitopes would provide basic data for the following study of the function of pp62.

Building a phylogenetic tree could visually display the genetic evolutionary relationship of ASFV, and an in-depth study of the genetic variation of ASFV could deepen our understanding of the evolutionary mechanism of ASFV, and

provide better suggestions for prevention and purification of ASF[42–44]. The 14 sequences with intensive sequence homology were successfully found through the pp62 protein sequence use BlastP, and the phylogenetic tree was successfully built using MEGA7.0 software based on pp62 protein. The phylogenetic tree was further noted through iTOL. The results showed the ASFV HLJ/2018 strain belonged to the same branch as Georgia/2007/1 in the evolutionary tree, and was closely related, which all belonged to genotype II.

4 Conclusions

The pp62 protein of the ASFV HLJ/2018 strain was a hydrophilic acidic labile protein with a molecular weight of 60.5 kDa. It had no transmembrane region and signal peptide, and had 5 N-glycosylation sites and 4 O-glycosylation sites. The proportions of helix, coil and strands were 45.5%, 41.7% and 12.8%, respectively. Screening and analysis got 14 dominant B-cell epitopes and 16 dominant T-cell epitopes, and found 9 B-T cell combined epitopes on the surface of the protein molecule further. The phylogenetic tree was successfully constructed with pp62 protein of ASFV HLJ/2018 strain, which was in the same branch as Georgia/2007/1 strain by evolutionary relationship and belonged to genotype II. All the data collected from this research will provide basic data for the in-depth study of ASFV.

References

[1] NAN W, DONGMING Z, JIALING W, *et al.* Architecture of African swine fever virus and implications for viral assembly[J]. Science, 2019, 366(6465): 640–644.

[2] WANG H, ZHOU C, DONG N. The first outbreak of African swine fever was occurred in Shenyang, Liaoning province, China[J]. Modern journal of animal husbandry and veterinary medicine, 2018 (10): 40–43.

[3] QUEMBO CJ, JORI F, VOSLOO W, *et al.* Genetic characterization of African swine fever virus isolates from soft ticks at the wildlife/domestic interface in Mozambique and identification of a novel genotype[J]. Transbound Emerg Dis, 2018, 65(2): 420–431.

[4] OU YW, LIU LJ, DAI JF, *et al.* Roles of African swine fever virus structural proteins in viral infection [J]. Biotechnology Bulletin, 2019, 35 (6): 156–163.

[5] SUAREZ C, SALAS ML, RODRIGUEZ JM. African swine fever virus polyprotein pp62 is essential for viral core development[J]. J Virol, 2010, 84(1): 176–187.

[6] ANDRES G, ALEJO A, SALAS J, *et al.* African swine fever virus polyproteins pp220 and pp62 assemble into the core shell [J]. J Virol, 2002, 76(24): 12473–12482.

[7] BAI C, WANG T, ZHAO S, *et al.* Preparation of monoclonal antibodies against recombinant p62 protein of African swine fever virus and its preliminary application[J]. Acta Veterinaria et Zootechnica Sinica, 2020, 51(5): 1074–1082.

[8] GALLARDO C, BLANCO E, RODRIGUEZ JM, *et al*. Antigenic properties and diagnostic potential of African swine fever virus protein pp62 expressed in insect cells [J]. J Clin Microbiol, 2006, 44(3): 950–956.

[9] GASTEIGER E, HOOGLAND C, GATTIKER A, *et al.* Protein identification and analysis tools on the ExPASy server[A]. John M. Walker: The proteomics protocols handbook[C]. Humana Press, 2005: 571–607.

[10] KROGH A, LARSSON B, VON HEIJNE G, *et al.* Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes[J]. J Mol Biol, 2001, 305 (3): 567–580.

[11] ALMAGRO ARMENTEROS JJ, TSIRIGOS KD, SØNDERBY CK, *et al*. SignalP 5.0 improves signal peptide predictions using deep neural networks [J]. Nature Biotechnology, 2019, 37 (4): 420–423.

[12] BUCHAN DWA, JONES DT. The PSIPRED protein analysis workbench: 20 years on [J]. Nucleic Acids Research, 2019, 47(1): 402–407.

[13] KLAUSEN MS, JESPERSEN MC, NIELSEN H, *et al.* NetSurfP-2.0: improved prediction of protein structural features by integrated deep learning[J]. Proteins: Structure, Function, and Bioinformatics, 2019.

[14] KRIEGER S, KECECIOGLU J. Boosting the accuracy of protein secondary structure prediction through nearest neighbor search and method hybridization[J]. Bioinformatics, 2020, 36(S1): 317–325.

[15] MA Y, LIU Y, CHENG J. Protein secondary structure prediction based on data partition and semi-random subspace method[J]. Scientific Reports, 2018, 8(1): 9856.

[16] LARSEN J E, LUND O, NIELSEN M. Improved method for predicting linear B-cell epitopes[J]. Immunome Research, 2006(2): 2.

[17] HASTE ANDERSEN P, NIELSEN M, LUND O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures [J]. Protein Science, 2006, 15(11): 2558–2567.

[18] NIELSEN M, LUNDEGAARD C, WORNING P, *et al.* Reliable prediction of T-cell epitopes using neural networks with novel sequence representations[J]. Protein Science, 2003, 12(5): 1007–1017.

[19] JURTZ V, PAUL S, ANDREATTA M, *et al.* NetMHCpan-4.0: improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data [J]. The Journal of Immunology: Official Journal of the American Association of Immunologists, 2017, 199(9): 3360–3368.

[20] WATERHOUSE A, BERTONI M, BIENERT S, *et al*. SWISS-MODEL: homology modelling of protein structures and complexes[J]. Nucleic Acids Research, 2018, 46(1): 296–303.

[21] REYNOLDS CR, ISLAM SA, STERNBERG MJE. EzMol: A web server wizard for the rapid visualization and image production of protein and nucleic acid structures[J]. Journal of Molecular Biology, 2018, 430 (15): 2244–2248.

[22] KUMAR S, NEI M, DUDLEY J, *et al.* MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences [J]. Briefings in Bioinformatics, 2008, 9(4): 299–306.

[23] LETUNIC I, BORK P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation[J]. Nucleic Acids Research, 2021:

[24] DIANE EW, MERLIN CT. Post-translational modification of plant-made foreign proteins; glycosylation and beyond [J]. Biotechnology Advances, 2012, 30(2): 410–418.

[25] TAN YC, WANG QJ, ZHAO GP, *et al.* Protein post-translational modification in prokaryotes [J]. Progress in Biochemistry & Biophysics, 2011, 38(3): 197–203.

[26] HU J, GUO Y, LI Y. Research progress in protein post-translational modification[J]. Chinese Science Bulletin, 2006, 51(6): 633–645.

[27] LANCE W, STEPHEN AW, GERALD WH. O-GlcNAc: A regulatory post-translational modification[J]. Biochemical & Biophysical Research Communications, 2003, 302 (3): 435–441.

[28] TADASHHI S, KEN K, SADAKO I, *et al*. N-Glycosylation/deglycosylation as a mechanism for the post-translational modification/remodification of proteins [J]. Glycoconjugate Journal, 1995, 12(3): 183–193.

[29] SHENG L, HE Z, CHEN J, *et al*. The impact of N-glycosylation on conformation and stability of immunoglobulin Y from egg yolk[J]. International Journal of Biological Macromolecules, 2016, 96: 129–136.

[30] NAIK NG, WU HN, KIRKEGAARD K. Mutation of putative N-Glycosylation sites on dengue virus NS4B decreases RNA replication[J]. Journal of Virology, 2015, 89(13): 6746–6760.

[31] MENG J, PARROCHE P, GOLENBOCK DT, *et al*. The differential impact of disulfide bonds and N-Linked glycosylation on the stability and function of CD14 [J]. Journal of Biological Chemistry, 2008, 283(6): 3376–3384.

[32] HOUNSELL EF, DAVIES MJ, RENOUF DV. O-linked protein glycosylation structure and function[J]. Glycoconjugate Journal, 1996, 13 (1): 19–26.

[33] RONDA L, BRUNO S, BETTATI S, *et al*. From protein structure to function via single crystal optical spectroscopy [J]. Frontiers in Molecular Biosciences, 2015(2): 1–14.

[34] KLEPEIS JL, LINDORFF-LARSEN K, DROR RO, *et al*. Long-timescale molecular dynamics simulations of protein structure and function[J]. 2009, 19(2): 120–127.

[35] SINHA N, SMITHGILL SJ. Protein structure to function via dynamics [J]. Protein & Peptide Letters, 2002, 9(5): 367–377.

[36] STUDER R, DESSAILLY B, ORENGO C. Residue mutations and their impact on protein structure and function: Detecting beneficial and pathogenic changes[J]. The Biochemical Journal, 2013, 449: 581–594.

[37] SAWAI MV, WARING AJ, KEARNEY WR, *et al*. Impact of single-residue mutations on the structure and function of ovispirin/novispirin antimicrobial peptides[J]. Protein Engineering, 2002, 15(3): 225–232.

[38] GOBEL U, SANDER C, SCHNEIDER R, *et al*. Correlated mutations and residue contacts in proteins [J]. Proteins Structure Function & Bioinformatics, 2010, 18(4): 309–317.

[39] CHRISTIAN S, KENNETH SS, JULIAN H, *et al*. Integrated visual analysis of protein structures, sequences, and feature data [J]. BMC Bioinformatics, 2015, 16(S11): S7.

[40] FRENCH SL, OSHEIM YN, SCHNEIDER DA, *et al*. Visual analysis of the yeast 5S rRNA gene transcriptome: regulation and role of La protein[J]. Molecular & Cellular Biology, 2008, 28(14): 4576.

[41] JANKUNKELLY TJ, LINDEMAN AD, BRIDGES SM. Exploratory visual analysis of conserved domains on multiple sequence alignments[J]. Bmc Bioinformatics, 2009, 10(S11): S7–S7.

[42] ASLANYAN L, AVAGYAN H, KARALYAN Z. Whole-genome-based phylogeny of African swine fever virus[J]. Veterinary World, 2020, 13(10): 2118–2125.

[43] PORTUGAL R, COELHO J, HOPER D, *et al*. Related strains of African swine fever virus with different virulence: genome comparison and analysis [J]. Journal of General Virology, 2015, 96(2): 408–419.

[44] ETIENNE PDV, CARMINA G, MARISA A, *et al*. Phylogenomic analysis of 11 complete African swine fever virus genome sequences [J]. Virology, 2010, 400(1): 128–136.