E-ISSN: 2347-5129
P-ISSN: 2394-0506
(ICV-Poland) Impact Value: 5.62
(GIF) Impact Factor: 0.549
IJFAS 2020; 8(2): 203-210
© 2020 IJFAS
www.fisheriesjournal.com
Received: 15-01-2020
Accepted: 19-02-2020

Mehedi Mahmudul Hasan
Associate Professor, Department of Fisheries and Marine Science Noakhali Science and Technology University Noakhali-3814, Bangladesh

Corresponding Author:
Mehedi Mahmudul Hasan
Associate Professor, Department of Fisheries and Marine Science Noakhali Science and Technology University Noakhali-3814, Bangladesh

Comparison between envelope protein VP19 and tegument protein VP26 of white spot syndrome virus in crustaceans: An *in-silico* approach

Mehedi Mahmudul Hasan

Abstract
This study considered previously published sequences of VP19 and VP26 structural proteins of White Spot Syndrome Virus. The sequences were retrieved from National Center for Biotechnology Information (NCBI). VP 19 sequences were from the isolates of India, China, Singapore, Mexico and the Netherlands. VP26 sequences were from the isolates of India, China, Brazil and Mexico. These were aligned using UNIPROT and the phylogenetic trees were drawn through MEGA 7. VP19 sequence of the Singapore isolate (accession no. AAK96179) and VP26 sequence of the Chinese isolate were reference proteins in this study. Mutations were observed in the amino acid position of 9, 15, 66, 71, 73 and 95 in the VP19 sequences from other geographical locations than the reference one. In the VP26 sequences, mutations started from the 127th amino acid position. Amino acid substitutions were observed in the hydrophobic, polar and charged regions of both the proteins. Although mutations were observed, the sequences of the individual proteins show a conserved nature in different geographical locations for the sequence homology.

Keywords: Sequence, protein, amino acids

1. Introduction
White Spot Syndrome Virus (WSSV) appeared in shrimp farms in 1992 in Taiwan. This virus caused huge mortality followed by an outbreak in Japan in 1994 and to several other countries later [1]. It was thought that crustaceans of Asia had spread to other continents as Australia's shrimp farms got affected by this deadly virus [2]. WSSV is a double-stranded DNA virus which is very large in size varying from 281 kbp to 312 kbp with high levels of homology among the published genomes [3]. Chen *et al* [4] mentioned that the difference in size is due to small insertions and a large deletion with a very little variation. Isolates from different geographic locations do not show much variation in proteome and morphology belong to the family Whispovirus [5].
Structural proteins VP28 and VP19 are placed in the envelope and VP26, VP24 and VP15 are in the nucleocapsid [6]. Later the classification of structural proteins changed and VP26 were included as a tegument protein. VP28, VP26 and VP24 could be the result of gene duplication because of evolution [7]. None of these major structural proteins are glycosylated which is a very exceptional feature among the enveloped viruses [8]. This current study compares the amino acid sequences of the major envelope protein VP19 and nucleocapsid protein VP26.

2. Materials and methods
2.1 Sequence dataset
VP19 and VP26 Sequences were downloaded from NCBI. 8 VP19 sequences and 7 VP26 sequences from different geographic locations of the world were selected for study.

2.2 Multiple sequence alignment and drawing phylogenetic tree
All the sequences were aligned using MEGA 7 [9]. Phylogenetic trees for both the protein sequences were also drawn using MEGA 7. Maximum likelihood method was chosen for generating the phylogenetic trees.

2.3 Analysis of amino acid properties

Annotations and the amino acid properties of both the alignment were done through UNIPROT [10-11]. Similarity, hydrophobic, negative, positive, aliphatic, tiny, aromatic, charged, small, polar, big, serine-threonine were considered for the analysis.

3. Results and Discussion

Table 1 and 2 contain a list of sequences that were downloaded from NCBI, the name of the country from which these isolates were collected including their size and date of submission. Sequences were selected from different geographical locations.

Table 1: List of downloaded sequences of VP19 from NCBI.

Accession no.	Country	Size (aa)	Date of submission
YP_009220610 [12]	China	121	2015
AAY27881 [13]	China	66	2005
ABI74630 [14]	India	121	2006
AAR12964 [15]	India	121	2003
AF369029_182 [5]	The Netherlands	121	2004
AAO69661 [16]	China	121	2003
CAI79042 [17]	Mexico	120	2005
AAK96179 [18]	Singapore	121	2001

Table 2: Downloaded sequences of VP26 from NCBI

Accession no.	Country	Size (aa)	Date of submission
AUJ79481 [19]	Brazil	204	2017
AAO69663 [16]	China	204	2003
CAI79043 [20]	Mexico	189	2005
ABI74632 [14]	India	205	2006
AWQ60442 [21]	Mexico	184	2017
AWQ63411 [21]	Mexico	144	2017
AWQ63811 [21]	Mexico	131	2017

Figure 1 and 2 show the membrane spanning and the alpha-helical regions and of VP19. Multiple sequence alignment is colored to show the regions. Figure 3 describes the similarity in amino acids in the selected sequences. There are mutations at the position 9 with serine in two sequences while others have Proline. There are mutations observed in position 15 of those two same sequences with Valine instead of Alanine in rest of the sequences. There is Proline at the 66th position of sequence from instead of Serine. Only one sequence has been found to have Valine at the 71st position while others are having Aspartic acid. In the position 73, all have Aspertic acid without one having Valine. There is a mutation at position 95 also having Threonine in one sequence while other sequences contained Methionine.

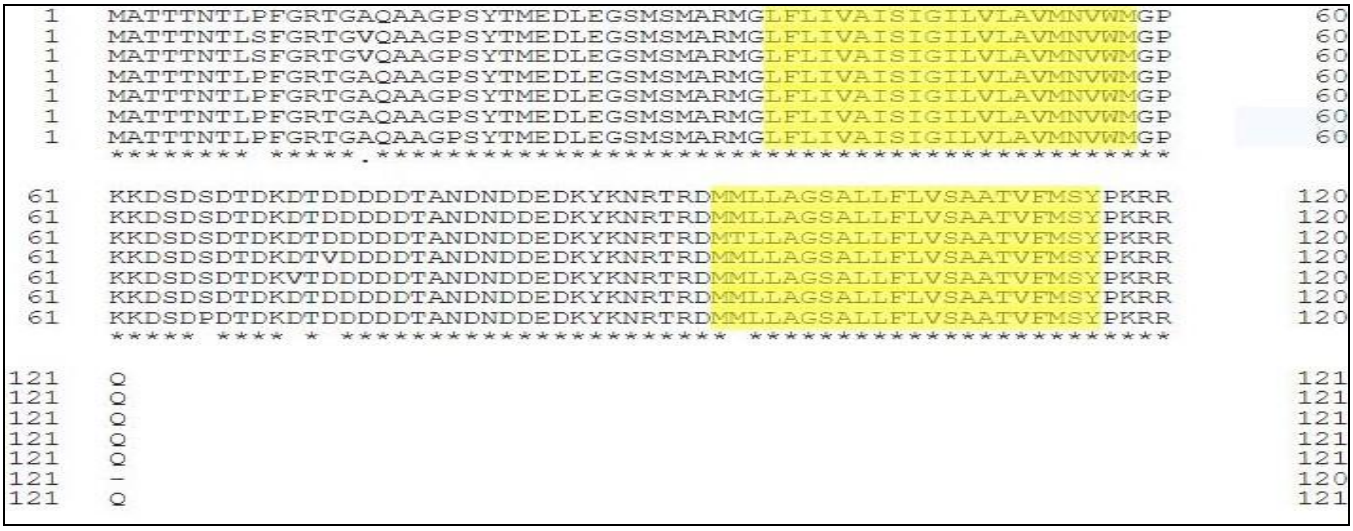

Fig 1: VP 19 transmembrane (Row 1 to 7: Accession no. YP_009220610, AAR12964, ABI74630, AF369029_182, AAO69661, CAI79042, AAK96179)

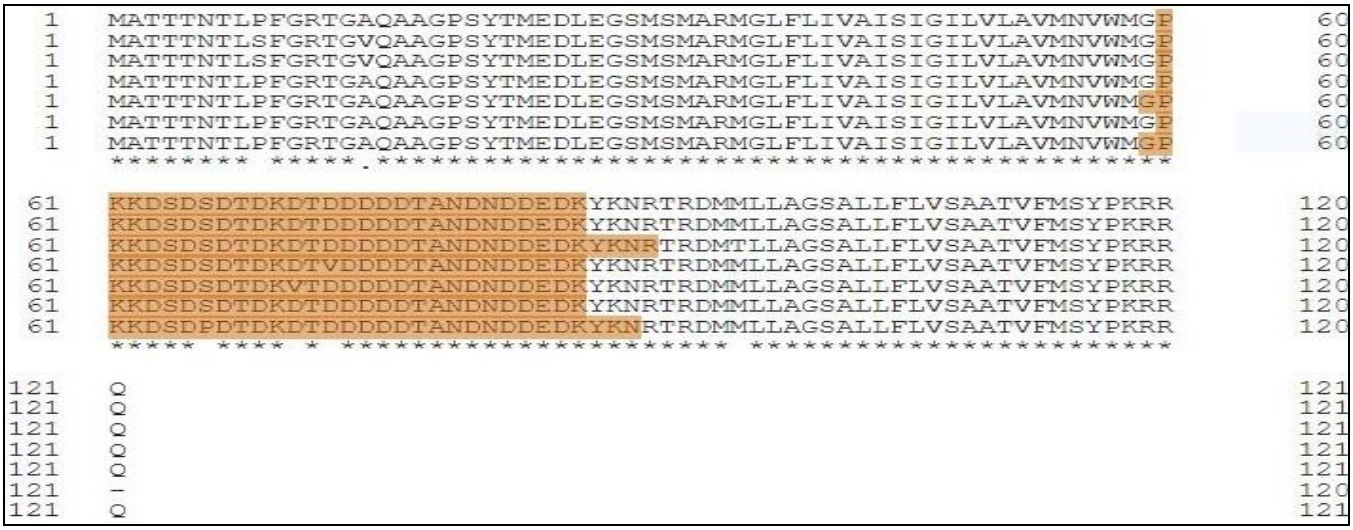

Fig 2: VP19 region (Row 1 to 7: Accession no. YP_009220610, AAR12964, ABI74630, AF369029_182, AAO69661, CAI79042, AAK96179)

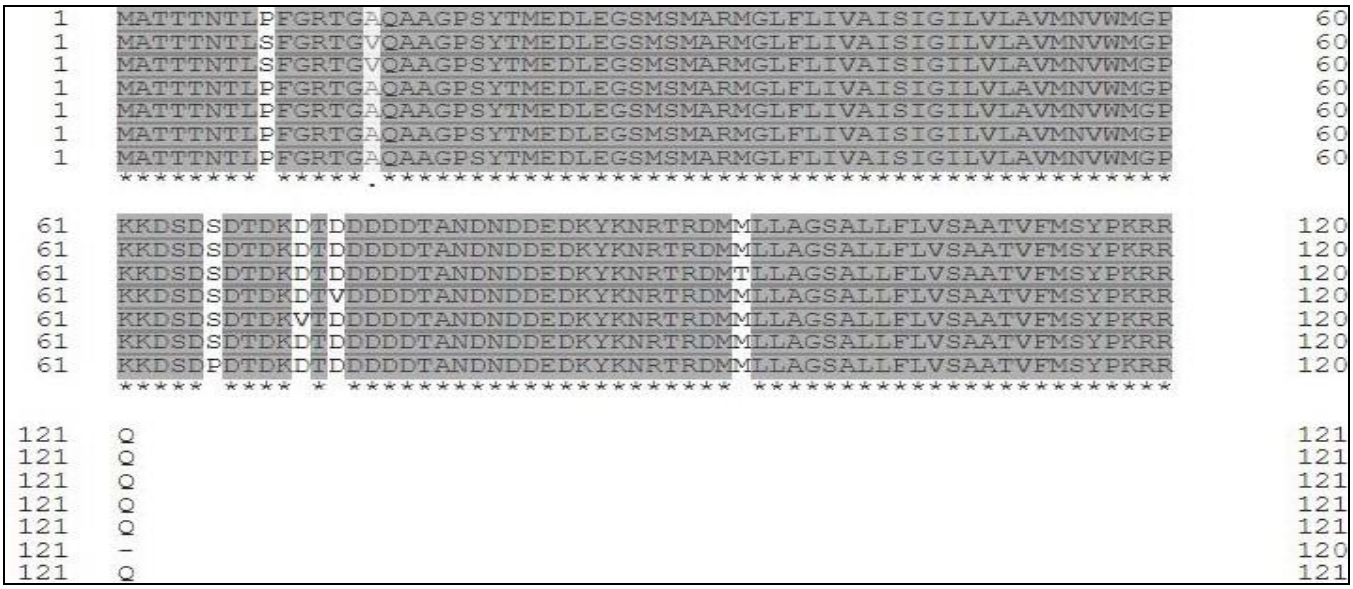

Fig 3: VP19 similarity (Row 1 to 7: Accession no. YP_009220610, AAR12964, ABI74630, AF369029_182, AAO69661, CAI79042, AAK96179)

There are amino acid substitutions in VP19 sequences all over the world in the hydrophobic, polar, charged regions (Fig 4, 5 and 6).

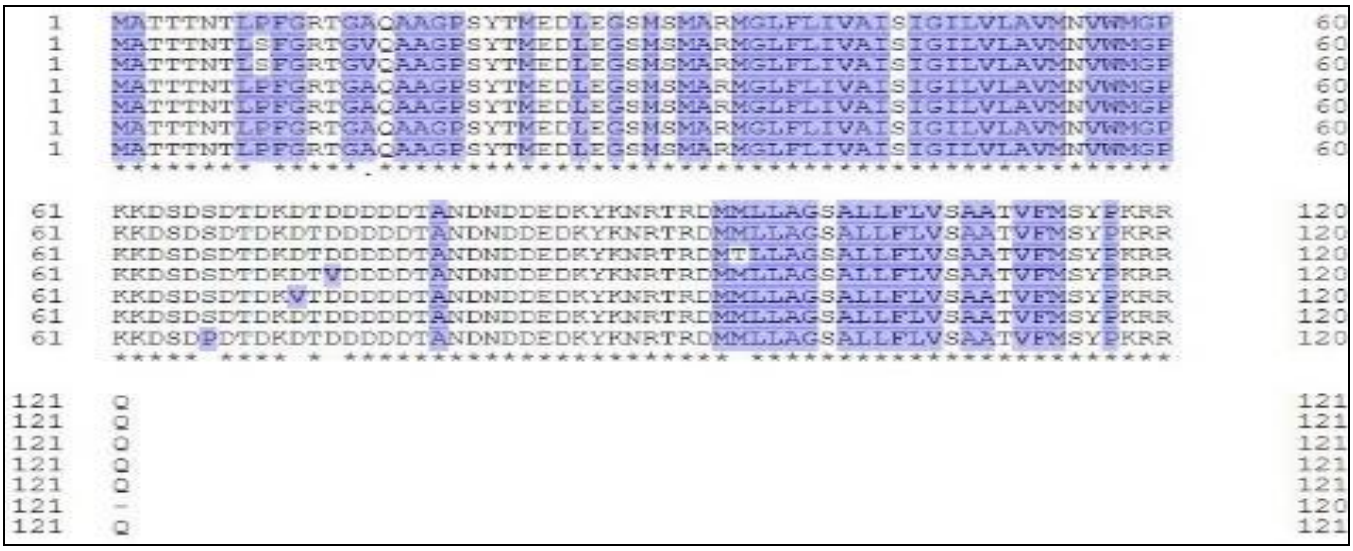

Fig 4: VP19 hydrophobic (Row 1 to 7: Accession no. YP_009220610, AAR12964, ABI74630, AF369029_182, AAO69661, CAI79042, AAK96179)

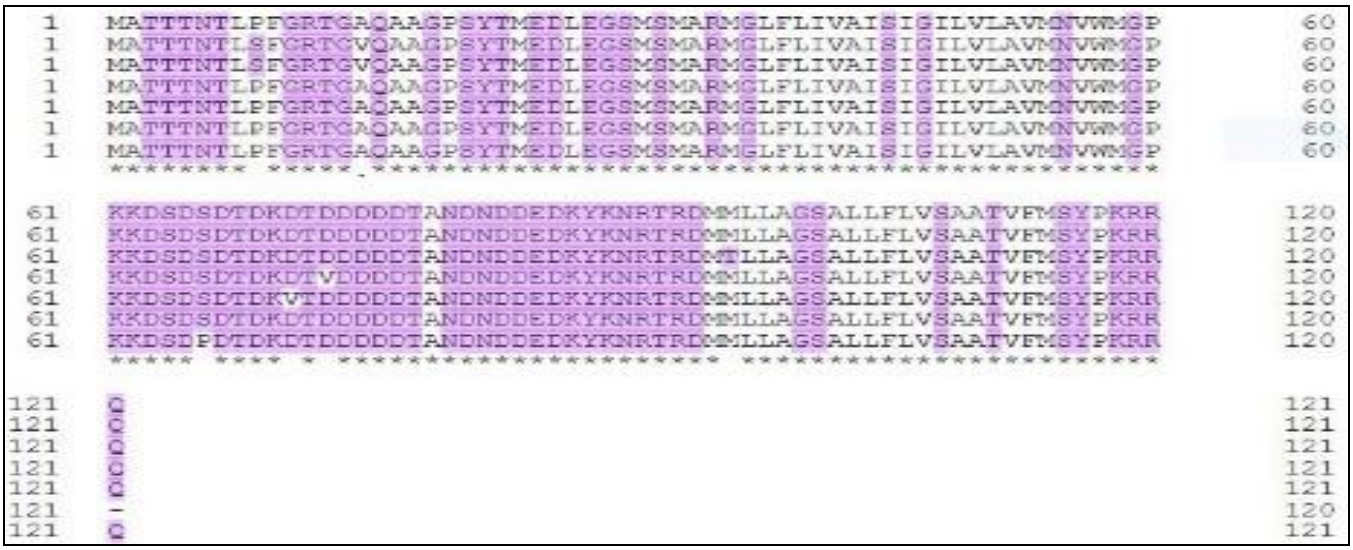

Fig 5: VP19 polar (Row 1 to 7: Accession no. YP_009220610, AAR12964, ABI74630, AF369029_182, AAO69661, CAI79042, AAK96179)

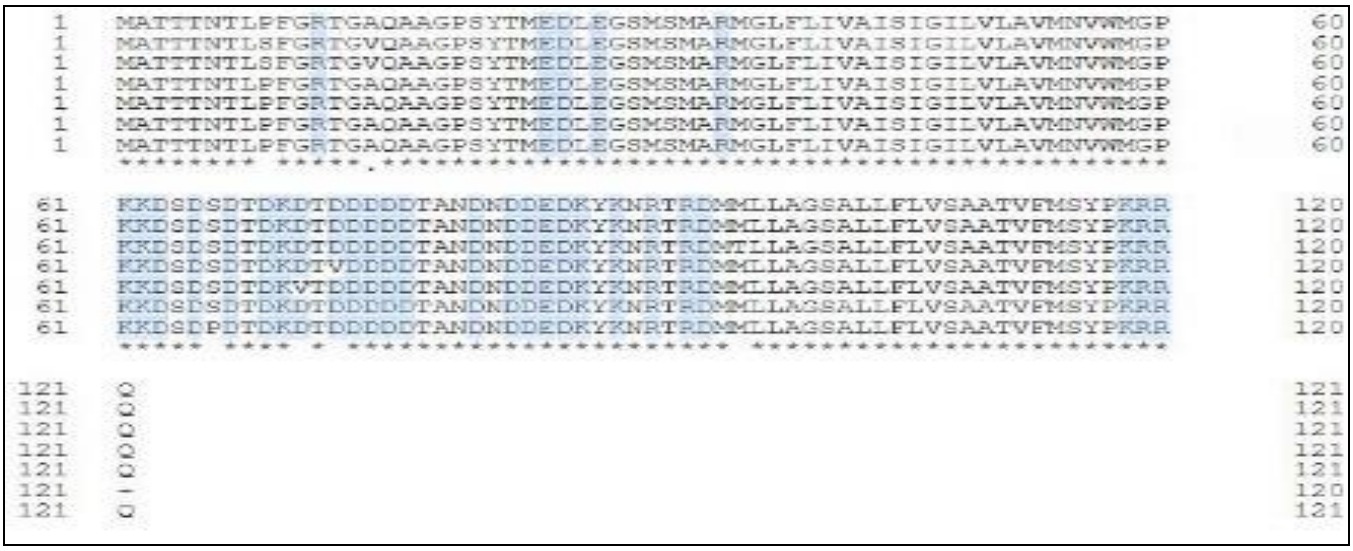

Fig 6: VP19 charged (Row 1 to 7: Accession no. YP_009220610, AAR12964, ABI74630, AF369029_182, AAO69661, CAI79042, AAK96179)

Transmembrane region and domain of the VP26 among the sequences are shown in Figure 7 and 8. The domain of the VP26 includes the regions including the transmembrane and non-terminal residues. All the mutations in the sequences have started after 126 amino acids (Fig 9). Amino acid substitutions are there in hydrophobic, polar and charged regions for the mutations. (Fig 10, 11 and 12).

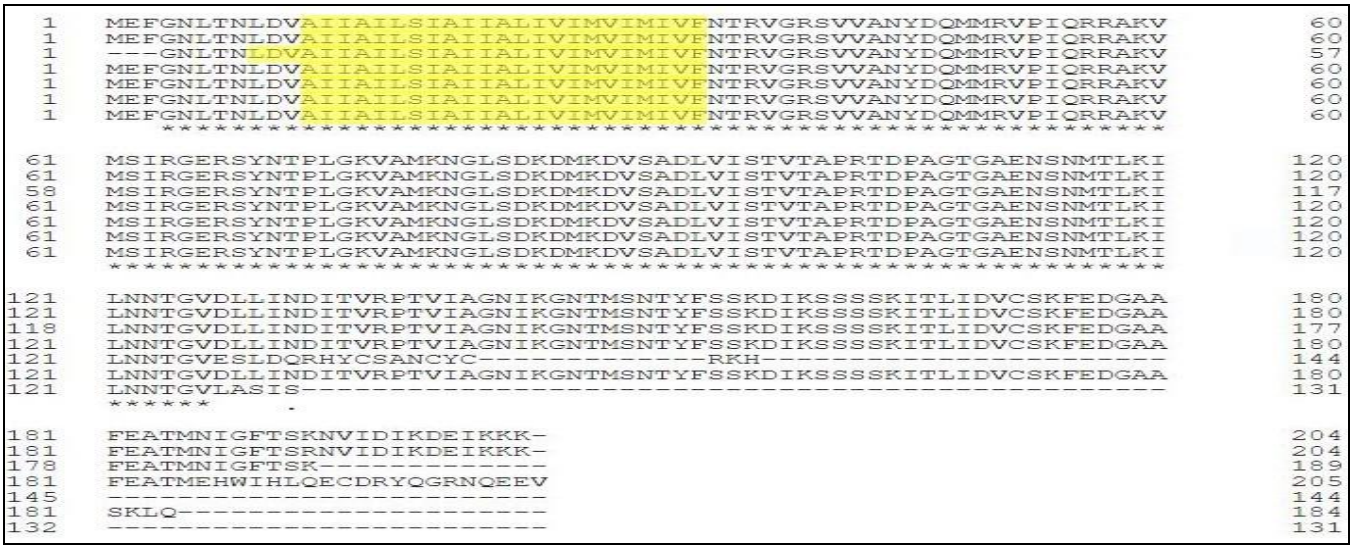

Fig 7: VP26 transmembrane (Row 1 to 7: Accession no. AUJ79481, AAO69663, CAI79043, ABI74632, AWQ63411, AWQ60442, AWQ63811)

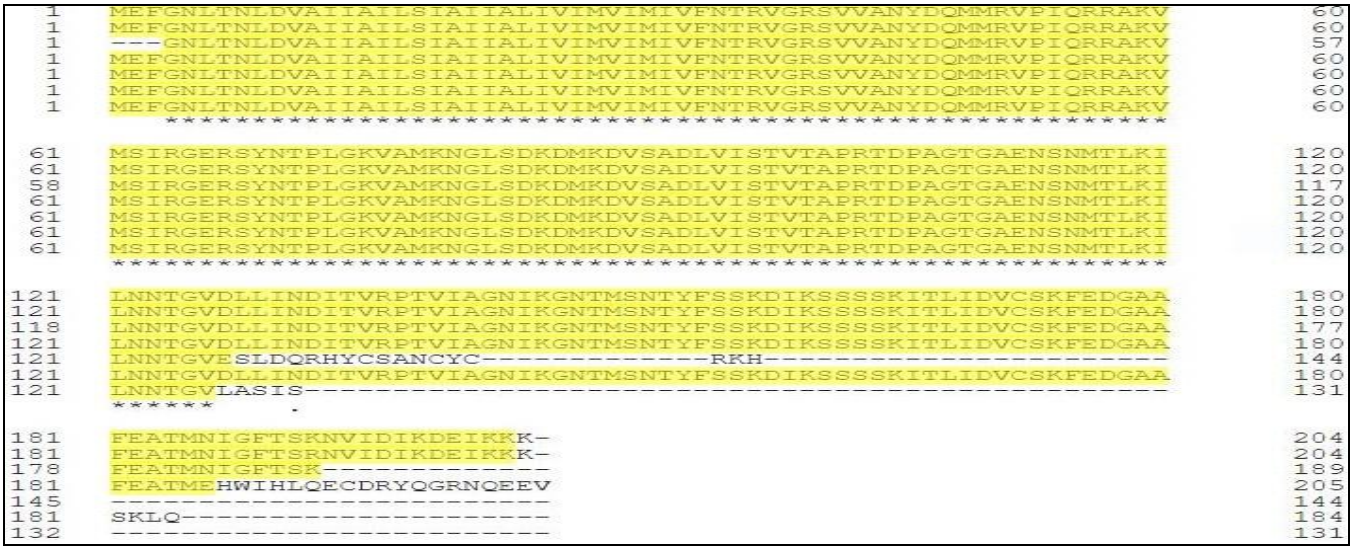

Fig 8: VP26 domain (Row 1 to 7: Accession no. AUJ79481, AAO69663, CAI79043, ABI74632, AWQ63411, AWQ60442, AWQ63811)

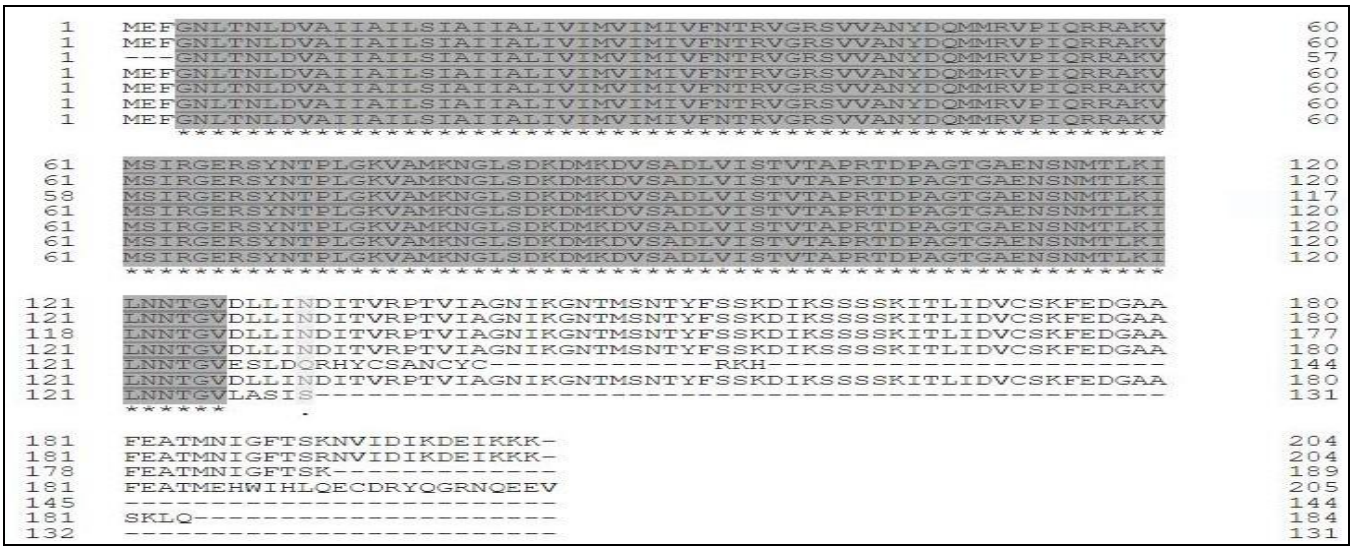

Fig 9: VP26 similarity (Row 1 to 7: Accession no. AUJ79481, AAO69663, CAI79043, ABI74632, AWQ63411, AWQ60442, AWQ63811)

Fig 10: VP26 Hydrophobic (Row 1 to 7: Accession no. AUJ79481, AAO69663, CAI79043, ABI74632, AWQ63411, AWQ60442, AWQ63811)

Fig 11: VP26 Polar (Row 1 to 7: Accession no. AUJ79481, AAO69663, CAI79043, ABI74632, AWQ63411, AWQ60442, AWQ63811)

Fig 12: VP26 Charged (Row 1 to 7: Accession no. AUJ79481, AAO69663, CAI79043, ABI74632, AWQ63411, AWQ60442, AWQ63811)

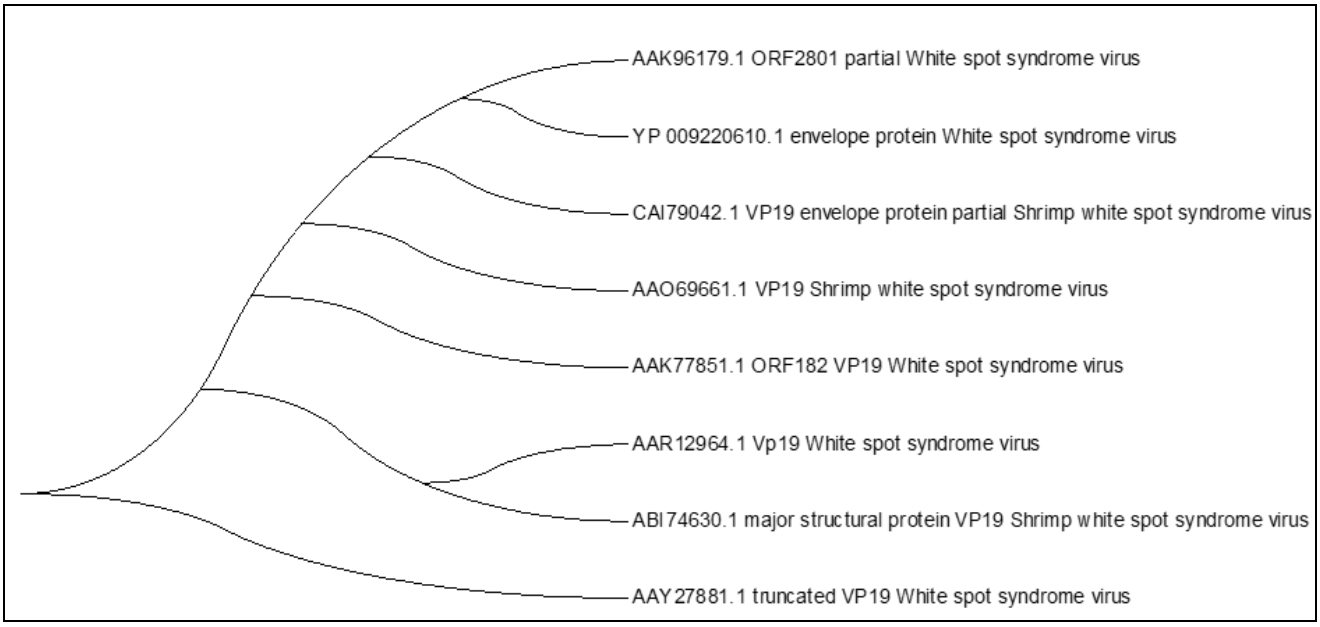

Fig 13: Molecular phylogenetic analysis by Maximum Likelihood method.

The evolutionary history was made on the basis of the JTT matrix-based model [22]. The tree with the maximum log likelihood (-330.45) is publicised. Preliminary tree for the heuristic search was attained automatically by using ‘‘Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances’’ projected applying a JTT model, and then choosing the topology by ‘‘superior log likelihood value’’. The investigation involved 8 amino acid sequences of VP19. All positions with gaps and missing data were removed. Final dataset contained 66 positions. Evolutionary investigation was conducted in MEGA7 [23].

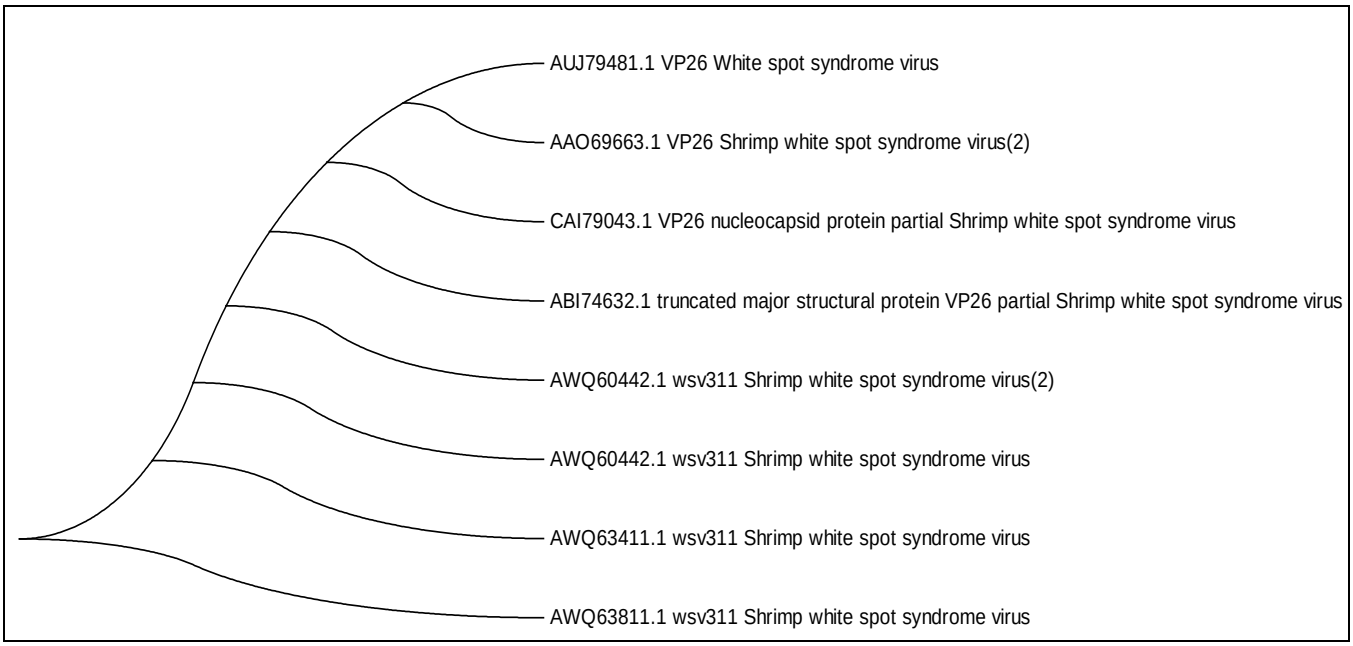

Fig 14: Molecular phylogenetic analysis by Maximum Likelihood method.

The evolutionary history was made on the basis of the JTT matrix-based model [22]. The tree with the maximum log likelihood (-413.01) is publicised. Preliminary tree for the heuristic search was attained automatically by using ''Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances'' projected applying a JTT model, and then choosing the topology by ''superior log likelihood value''. The investigation involved 8 amino acid sequences of VP26. All positions with gaps and missing data were removed. Final dataset contained 127 positions. Evolutionary investigation was conducted in MEGA7 [23].

Since this virus contains a DNA which is double stranded, it is more stable [3]. There is not much variation through its evolution. Since WSSV is found in shrimps, crabs and some other crustaceans, it can transfer from one location to another through the host. The selected sequences were from different geographical locations.

The VP19 is placed with VP28 at the envelope which is involved with systematic infection [24]. Alignment of the sequences of VP19 has shown 94.215% similarity with 114 identical positions (Figure 3). The sequences were chosen from different countries where the culture water temperatures differ from each other. Temperature is a very important parameter that could play a role in outbreak [25]. Among the chosen sequences, the Singapore one was first submitted to NCBI in 2001 [18]. It had proline at the amino acid position 9, alanine at 15, proline at 66, aspartic acid at 71 and 73 and methionine at 95 which were substituted later on in some other sequences. At position 9, proline was substituted by Serine in two of the Indian isolates of 2003 and 2006. At position 15, alanine was substituted by valine in both the Indian isolates. Proline at the position 66 of the Singapore isolate was substituted by serine in all other sequences from different countries. Aspartic acids at the position 71 and 73 were mutated to valine in the Chinese isolate of 2003 and in the Dutch isolate of 2004. Singapore isolate's methione of position 95 which falls into the transmembrane region (Fig 1) was substituted by the one Indian isolate (2006). It is observed in Figure 4 that hydrophobic region has two mutations by serine and one by threonine in Indian isolates. There are changes in amino acid properties of polar and charged regions (Fig 5 and Fig 6) though it cannot be said that these are playing a very important role to change the tertiary structure of the protein since Mishra and Shekhar (2005) mentioned that WSSV of Indian isolate and isolates from other geographical locations show very much similarity in external morphology [26].

In case of the tegument protein VP26, the Chinese isolate [16] is the earliest one among the selected sequences (Table 2). There is no mutation in the transmembrane region of the protein (Fig 7). The domain consists of 203 amino acids of the 204 amino acids which is observed in the sequences of China in 2003 and Brazil in 2017 (Fig 8). Amino acid substitutions are found after 126 amino acid positions in three of the sequences from Mexico which were submitted in 2017 (Fig 9). Mutations were observed in the hydrophobic and polar regions of the sequences of Mexico, accession numbers AWQ63411, AWQ60442 and AWQ63811 and India (ABI74632) (Fig 10 and 11). In the charged region of amino acids, mutations were observed in all the sequences (Fig 12). Both the phylogenetic trees based on VP19 and VP26 sequences display that the sequences did not have much variations and are highly homologous (Fig 13 and 14). Shekhar and Ravichandran (2007) mentioned in their study that these sequences of structural proteins show minor variation in sequences which supports the results of our study [27].

4. Conclusion

The sequences show highly conserved nature of VP19 and VP26. VP19 is an envelope protein that may have important role in binding with the host protein. The mutations could be important for immunogenic response of the host which can only be confirmed after challenging the crustaceans with different phylotypes of these proteins. WSSV structural proteins can be used in search for remedy from white spot disease in crustaceans. These are structural proteins and may have evolved from gene duplication [7]. It is important to know if there is any effect of temperature on mutation and the virulence of isolates from different geographical locations.

5. References

1. Bonilla CME, Sanz VA, Wille M, Sorgeloos P, Pensaert MB, Nauwynck HJ. A review on the morphology, molecular characterization, morphogenesis and pathogenesis of white spot syndrome virus. Journal of Fish Diseases. 2008; 31:1-18
2. Claydon K, Cullen B, Owens L. OIE white spot syndrome virus PCR gives false-positive results in Cheraxquadricarinatus. Diseases of Aquatic Organisms. 2004; 62:265-268.
3. Oakey J, Smith CS. Complete genome sequence of a white spot syndrome virus associated with a disease incursion in Australia. Aquaculture. 2018; 484:152-159.
4. Chen LL, Leu JH, Huang CJ, Chou CM, Chen SM, Wang CH *et al*. Identification of a nucleocapsid protein (VP35) gene of shrimp white spot syndrome virus and characterization of the motif important for targeting vp35 to the nuclei of transfected insect cells. Virology. 2002; 293:44-53.
5. Marks H, van Duijse JJ, Zuidema D, van Hulten MC, Vlak JM. Fitness and virulence of an ancestral White Spot Syndrome Virus isolate from shrimp. Virus Research. 2005; 110(1, 2):9-20.
6. van Hulten MC, Witteveldt J, Peters S, Kloosterboer N, Tarchini R, Fiers M *et al*. The white spot syndrome virus DNA genome sequence. Virology. 2001a; 286(1):7-22.
7. van Hulten MC, Meng-Feng T, Christel AS, Chu-Fang L, Guang-Hsiung K, Just MV. Analysis of a genomic segment of white spot syndrome virus of shrimp containing ribonucleotide reductase genes and repeat regions. Journal of General Virology. 2000; 81:307-316.
8. Van Hulten MC, Martin R, Angela MG, Vermeesch FZ, Vlak M. Identification of VP19 and VP15 of white spot syndrome virus (WSSV) and glycosylation status of the WSSV major structural proteins. Journal of General Virology. 2002; 83:257-265.
9. Gouy M, Guindon S, Gascuel O. Sea View version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evolution. 2009; 27:221-224.
10. Watkins X, Garcia LJ, Pundir S, Martin MJ. UniProt Consortium Prot Vista: visualization of protein sequence annotations. Bioinformatics. 2017; 33:2040-2041.
11. Pundir S, Martin MJ, O'Donovan C. Uni Prot Consortium UniProt Tools Curr. Protoc. Bioinformatics. 2016; 53:1.29.1-1.29.15.
12. Gao M, Yang F, Xu L, LiF. Comparative analysis of genome sequences of three different virulent isolates of white spot syndrome virus. https://www.ncbi.nlm.nih.gov. 1 December, 2019.
13. Jha RK, Xu ZR, Shen J, Bai SJ, Sun JY, Li WF. The efficacy of recombinant vaccines against white spot syndrome virus in Procambarus clarkia. Immunol. Lett. 2006; 105(1):68-76.
14. Musthaq S, Hameed S. Cloning and characterization of Indian WSSV structural genes of Penaeus monodon. https://www.ncbi.nlm.nih.gov. 1 December, 2019.
15. Sathish S, Musthaq S, Sahul H, Narayanan RB. Production of recombinant structural proteins from the Indian WSSV isolate. Aquaculture 2004; 242(1-4):69-80.
16. Xu YX, Xu ZR, Sun JY, Li WF, Lv WP, Gu SH *et al*. https://www.ncbi.nlm.nih.gov. 1 December, 2019.
17. Molina GZJ. Clonacion y caracterizacion molecular de las proteinas estructurales mayores (VP15, VP19, VP26 y VP28) del virus del sindrome de la mancha blanca (Whispovirus, Nimaviridae) patogeno de Litopenaeus sp. cultivado en granjas del pacifico mexicano. Thesis Department of Centro Nacional de Sanidad Acuicola, Universidad Autonoma de Nuevo Leon, FCB, San Nicolas de los Garza, NL, Mexico, 2004.
18. Zhang X, Huang C, Hew CL. A gene from white spot syndrome virus. https://www.ncbi.nlm.nih.gov. 1 December, 2019.
19. Dantas MDA, Teixeira DG, Da silva RB, Soares PET, Agnez-Lima LF, Lima JMS *et al*. Direct sequencing of white spot syndrome virus from Brazil. https://www.ncbi.nlm.nih.gov. 1 December, 2019.
20. Galaviz SL, Molina ZJ, Rosales EJL, Alcocer JM. White spot syndrome virus genetic variants detected in Mexico by a new multiplex PCR method. Aquaculture 2004; 242(1-4):53-68
21. Parrilla TDP, Vibanco-Perez N, Duran MJ, Gomez-Gil B, Llera-Herrera R, Vazquez-Juarez R. Sequencing and comparison of viral genomes of white spot syndrome virus strains from Sinaloa and Nayarit, Mexico. https://www.ncbi.nlm.nih.gov. 1 December, 2019.
22. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Computer Applications in the Biosciences. 1992; 8:275-282.
23. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution. 2016; 33:1870-1874.
24. Van Hulten MCW, Witteveldt J, Snippe M, Vlak JM. White spot syndrome virus envelope protein VP28 is involved in the systemic infection of shrimp. Virology. 2001b, 285.
25. Overstreet RM, Matthews T. Effects of temperature on WSSV and TSV in Litopenaeus vannamei in controlled experiments. ISAAH, 2002.
26. Mishra SS, Shekhar MS. White spot syndrome virus isolates of tiger shrimp *Penaeus monodon* (Fabricius) in India are similar to exotic isolates as revealed by polymerase chain reaction and electron microscopy. Indian Journal of Experimental Biology. 2005; 43:654-661.
27. Shekhar MS, Ravichandran P. Comparison of white spot syndrome virus structural gene sequences from India with those at GenBank. Aquaculture Research. 2007; 38:321-324.