

Genomic and evolutionary features of two AHPND positive *Vibrio parahaemolyticus* strains isolated from shrimp (*Penaeus monodon*) of south-west Bangladesh

Shawon Ahmmed

Dhaka University

Md. Abdullah-Al-Kamran Khan

University of Dhaka

Md. Mostavi Enan Eshik

University of Dhaka

Nusrat Jahan Punom

University of Dhaka

Abul B.M.M.K. Islam

University of Dhaka

Mohammad Shamsur Rahman (✉ shamsur@du.ac.bd)

<https://orcid.org/0000-0001-8447-3275>

Research article

Keywords: *Vibrio parahaemolyticus*, Acute hepatopancreatic necrosis disease, AHPND, Genome sequencing, Virulence, *pirA*, *pirB*, Phylogenetics

Posted Date: September 18th, 2019

DOI: <https://doi.org/10.21203/rs.2.14490/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on December 3rd, 2019. See the published version at <https://doi.org/10.1186/s12866-019-1655-8>.

Abstract

Background Due to its rapid lethal effect in the early stage of shrimp, acute hepatopancreatic necrosis disease (AHPND) causing great economic losses, since its first outbreak in southeast China in 2009. *Vibrio parahaemolyticus*, carrying the *pir A* and *pir B* toxin genes is known to cause AHPND in shrimp. The overall objective of this study was to sequence the whole genome of AHPND positive *V. parahaemolyticus* strains isolated from shrimp (*Peneaus monodon*) of the south-west region of Bangladesh in 2016 and 2017 and characterize the genomic features and emergence pattern of this marine pathogen.

Results Two targeted AHPND positive *V. parahaemolyticus* strains were confirmed using PCR with 16S rRNA, *ldh*, AP3 and AP4 primers. The assembled genomes of strain MSR16 and MSR17 were comprised of total 5,393,740 bp and 5,241,592 bp, respectively. From annotation, several virulence genes involved in chemotaxis and motility, EPS type II secretion system, Type three secretion system-1 (TTSS-1) and its secreted effectors, thermolabile hemolysin were found in both strains. Importantly, the ~69 kb plasmid was identified in both MSR16 and MSR17 strains containing the two toxin genes *pir A* and *pir B*. Antibiotic resistance genes were predicted against β -lactam, fluoroquinolone, tetracycline, macrolide and cephalosporin groups in both MSR16 and MSR17 strains.

Conclusions The findings of this research may facilitate the tracking of pathogenic and/or antibiotic resistance *V. parahaemolyticus* isolates between production sites, and the identification of candidate strains for production of vaccines as an aid to control of this devastating disease. Also, the emergence pattern of this pathogen can be highlighted to determine the characteristic differences of other strains found all over the world.

Background

Asian shrimp farming industry has experienced massive production losses due to disease caused by toxins of *Vibrio* bacteria, known as early mortality syndrome/acute hepatopancreatic necrosis disease (EMS/AHPND) (1). AHPND is a shrimp disease which causes high mortality of cultivated penaeid shrimps commonly occur within the first 30 days after stocking in grow-out ponds (2, 3). The early mortality syndrome (EMS), also called AHPND was detected in shrimp farms in southern China as the first record in 2009 and then on the Island of Hainan in 2010 and then in Vietnam and Malaysia in 2011 and consequently in the eastern part of Thailand since 2012 and widely spread in other culture areas of Thailand (4). The loss of shrimp production caused by AHPND in the shrimp farming industry was estimated at more than \$1 billion per year globally (5). In Bangladesh, cases with AHPND were first reported in 2017 (6).

Acute hepatopancreatic necrosis disease (AHPND) causes a pale and atrophied hepatopancreas (HP) together with an empty stomach and midgut of the shrimp. Some pathological features include enlarged hepatopancreatic nuclei, sloughed HP cells-blister-like (B), fibrilla (F), resorptive (R) cells and AHPND is frequently followed by secondary bacterial infections (3). The causative agent of AHPND in shrimp is

Vibrio parahaemolyticus; a gram-negative rod shaped bacterium mainly inhabitant in marine and estuarine environment (3). *V. parahaemolyticus* lives in warm estuarine and marine environments and distributes throughout the world (7). AHPND causing *V. parahaemolyticus* possesses ~69 kb plasmid encoding toxin genes *pirA* and *pirB* (3, 8) which are similar to Photorhabdus insect-related (*pir*) toxin (9) which is one of the major causal factors reported. Two sets of the type III secretion system (T3SS1 and T3SS2) possessed by *V. parahaemolyticus*, which are also proposed to be important virulence factors of this organism. All *V. parahaemolyticus* strains contain T3SS1 whereas T3SS2 is present only in human clinical strains (10). AHPND positive *V. parahaemolyticus* strains do not contain TDH, TRH and T3SS2 which are known virulence factors affecting to humans (11). Detection of *V. parahaemolyticus* isolates is typically based on molecular techniques where the amplification is performed through species-specific gene *ldh* (lecithin dependent hemolysin) (12); as well as to identify the AHPND positive *V. parahaemolyticus* strains, AP3 (13) and AP4 (14) primers are commonly used.

Now-a-days, whole genome sequencing (WGS) has become a popular tools for identification and detection the bacterial outbreaks in aquaculture (15). In whole genome sequencing, all of single nucleotide polymorphisms (SNPs) used to confirm the epidemiological links of outbreak strains with higher typing resolution (16). In this study, we have sequenced two AHPND positive *V. parahaemolyticus* strains (MSR16 and MSR17) which were isolated from shrimp farms of south-west regions of Bangladesh and this is the very first genome sequencing report of AHPND positive *V. parahaemolyticus* strains isolated from shrimps of Bangladesh. Subsequently, we analyzed their genomic features associated with virulence and other factors. Finally, we have performed phylogenetic analyses using several genomic features of this bacteria to find out the relations between the outbreak causing strains around the globe with our sequenced strains.

Results

Identification of the AHPND positive strains

Molecular identification and characterization of suspected AHPND positive *V. parahaemolyticus* isolates were done using 16S rRNA, *ldh*, AP3 and AP4 primers PCR (Figure 1). MSR16 (isolated in 2016) and MSR17 (isolated in 2017) strains were finally sequenced for whole genome sequencing.

Features of the assembled genomes

The genomes were assembled into 35 contigs in MSR16 strain and 53 contigs in MSR17 strain. The largest contigs size for MSR16 strain was 1742619 bp; and 1892806 bp for MSR17 strain. The total GC content was 45.19% and 45.09% for MSR16 and MSR17 strains, respectively. The total genome size of MSR16 was 5232129 bp; and 5377931 bp for MSR17. MSR16 was found comprised of two circular chromosome with length of 3,407,823 bp, 1,893,040 bp while genome of MSR17 was comprised with similar two circular chromosome with length of 3,428,841 bp, 1,743,028 bp. Both MSR16 and MSR17 contain a plasmid with a length of 68,345 bp and 66,825 bp, respectively (Figure 2). Comparing the

genomes, it was observed that chromosome 2 of MSR16 strain has an extra 100kb region. More information about MSR16 and MSR17 genomes are given in Table 1.

Table 1: Summary of the assembled genomes of two strains (MSR16 & MSR17) of AHPND positive *V. parahaemolyticus*.

Features	VP _{AHPND} MSR16	VP _{AHPND} MSR17
Contigs	108	66
Largest contigs	1742619	1892806
Total length	5232129	5377931
GC (%)	45.19	45.09
CDS	4909	4689
Gene	5090	4854
tRNA	119	109
misc_RNA	51	45
rRNA	10	10
tmRNA	1	1

The plasmid of MSR16 contains total 87 genes of which 58 genes are hypothetical protein (67%), 5 repeat regions (6%), 7 conjugative transfer proteins (8%), 3 mobile element protein (3%), 2 antirestriction protein (2%), 2 toxin genes (*pirA* and *pirB*) and 10 other genes (11%). The plasmid of MSR17 contains total 88 genes of which 57 genes are hypothetical protein (65%), 6 repeat regions (7%), 7 conjugative transfer proteins (8%), 3 mobile element protein (3%), 2 antirestriction protein (2%), 2 toxin genes (*pirA* and *pirB*) and 11 other genes (13%).

Out of the RAST server predicted 406 subsystems, MSR16 strain possesses 74 responsible for virulence, disease and defense; five for phages, prophages, transposable elements and plasmids; 28 for iron acquisition and metabolism; and 125 for motility and chemotaxis. While out of the predicted 403 subsystems, MSR17 strain contained 74 responsible for virulence, disease and defense; 10 for phages, prophages, transposable elements and plasmids; 28 for iron acquisition and metabolism; and 119 for motility and chemotaxis (Figure 3). These particular subsystems are the hallmarks for the pathogenicity and both strains were found to have almost similar amount of factors across their genomes. Number of

genes associated with the general COG functional categories for both strains are provided in (Figure 4). Both strains are found to possess equivalent number of genes associated with those categories.

MSR16 and MSR17 strains have average nucleotide identity values of 98.57% with *V. parahaemolyticus* strain M1-1 and 98.65% with *V. parahaemolyticus* strain 13-306D/4 respectively; they also have an average of 95% ANI values with other AHPND positive strains (Additional file 1). Strains MSR16 and MSR17 have 1403 and 1228 hypothetical genes respectively, whose functional prediction can provide more insights into its pathogenicity and other functional pathways. 144 and 94 unique genes were found in strain MSR16 and MSR17 respectively which are uniquely predicted only for one strain (Additional file 2). MSR17 strain contains unique genes for zona occludens toxin, several transposition proteins, integrase, recombinases etc.; whereas MSR16 strain have genes for several conjugative transfer related proteins, bacteriocin immunity proteins etc. Both strains are predicted to have some exclusive genes for diverse metabolic pathways.

Virulence and antimicrobial resistance genes

Most common 9 virulence factor classes involved in- adherence, antiphagocytosis, enzyme, chemotaxis and motility, iron uptake, quorum sensing, secretion system, toxin, immune evasion were found in the MSR16, while MSR17 possess 8 of these such factors except the factors involved in immune evasion; also few genes in these classes of factors were found absent in these strains (Additional file 3). The thermostable direct haemolysin (*tdh*), the TDH-related haemolysin (*trh*) and the two type III secretion systems (T3SS1 and T3SS2) are recognized as major virulence factors in *V. parahaemolyticus* (17). *tdh* and *trh* both genes were not found in these strains but the thermo labile hemolysin (*tlh*) gene was found. Between two types of T3SS only T3SS1 type was found in MSR16 and MSR17 strain. Both strains possess the plasmid borne *pirA* and *pirB* toxins.

Antibiotic resistance genes were predicted against β -lactam, fluoroquinolone, tetracycline, macrolide and cephalosporin antibiotics in MSR16; and MSR17 strain has similar resistance genes except for cephalosporin (Additional file 4). Six and two probable prophage regions were found in MSR16 and MSR17 strains, respectively.

Strains MSR16 and MSR17 have approximately 39 and 27 genomic island (GI) regions respectively (Additional file 5). In strain MSR16, toxin-antitoxin systems like YoeB-YefM, Doc-Phd; antibiotic resistance proteins like FosA (Fosfomycin resistance protein); components of type-I, type-VI secretion systems etc. are found in those genomic islands. Genomic islands of strain MSR17 contains toxin-antitoxin systems like HipA-HipB, YoeB-YefM; type-I, type-III secretion systems; Multidrug resistance efflux pump; several phage and transposon related proteins etc. (Additional file 6).

PathogenFinder tool (18) predicted an overall probability of 0.868 for MSR16 and 0.871 for MSR17 for becoming potential human pathogen, so there is a very high risk of spreading these strains into the human food chain and causing human diseases.

Phylogenetic relationship based on 16S rRNA genes of different AHPND positive *V. parahaemolyticus* strain

A total of 30 strains were selected for establishing phylogenetic relationship based on 16S rRNA gene sequence (Figure 5). The tree includes 25 *V. parahaemolyticus* (including MSR16 and MSR17), two *V. campbellii* and two *V. owensii* strains which were responsible for AHPND outbreak in the recent years in different regions of the world. *V. cholerae* was used for outgroup comparison.

The tree illustrates that *Vibrio campbellii* AB497067 India, *Vibrio campbellii* AY035896 China strains were located at same cluster. The strains including M13-18718-18H KY229843.1 Australia, NSTH24 KF886635.1 Thailand, NSTH22 KF886633.1 Thailand, NSTH21 KF886632.1 Thailand, L44 KC884620.1 China, CZN-33 KR347269.1 China, MM21 FJ547093.1 China, Vp-D4 MH298544.1 China, and Vp-X 11 MH298549.1 China located at same cluster.

Both of our studied strains (MSR16 and MSR17) located at same cluster and were closely related with one of Indian strain AP1511 indicating that the mutation and evolutionary pattern of MSR16 and MSR17 strains might be analogous to the Indian strain. All the Indian *V. parahaemolyticus* strains GM171 MG593212.1 India, AP183 MG564744.1 India, SPJ3 MG564747.1 India, G476 MG593229.1 India, GA5114 MG970587.1 India, S24P132 MG762012.1 India, and AP1511 MG564729.1 India located at same cluster. The two Spanish *V. parahaemolyticus* strains (CECT5305 FM204865 Spain and CECT611 FM204867.1 Spain) separately make a cluster. One of the Chinese *V. parahaemolyticus* strain 461 JN188418.1 China belonged to an independent lineage.

The strains including Vp-4 MK377081.1 China, Ramsar KJ704113.2 Iran belong to separate cluster. Besides, two AHPND positive *V. owensii* strains S12A KR086360 India, CHN-25 KR347195 China were located at separate cluster. *V. cholerae* (*msr6*) strain was distantly related with our studied strains.

Phylogenetic relationship based on housekeeping genes of different AHPND positive *V. parahaemolyticus*

A total of 25 strains were selected for establishing phylogenetic relationship based on common housekeeping genes (Figure 6) including (*dnaE*, *dtdS*, *gyrB*, *pntA*, *pyrC*, *recA*, *tnaA*). The 16S rRNA gene was not included because separate phylogenetic relationship was established based on it. FIM-S1708+ Mexico, FIM-S1392- Mexico and MVP5 Malaysia strains located at same cluster. TUMSAT-DE1-S1 Thailand, NCKU-TN-S02 Thailand, NCKU-TV-5HP Thailand and KS17.S5-1 Malaysia strains located at same cluster. PSU5570 Thailand and ND11 Malaysia strain belong to an independent lineage respectively. The strains M0605 Mexico and TUMSAT-H10-S6 Thailand located at same cluster. NCKU-TV-3HP Thailand, MSR17 Bangladesh, M1-1 Vietnam located at same cluster. MVP3 Malaysia and VP14 India strain located at same cluster. The strains 12-009A/1335 Vietnam, MSR16 Bangladesh located at same cluster. 13-028-A2 Vietnam and NA9 Malaysia strains separately make a cluster. NCKU_CV_CHN_China belong to an independent lineage. TUMSAT-D06-S3 Thailand, 13-306-D4 Mexico and NB09 Malaysia strain located at same cluster.

The phylogenetic tree showed that MSR16 strain was closely related to 12-009A/1335 Vietnam strain which maintain an antibacterial type vi secretion system with versatile effector repertoires (19) suggesting that MSR16 strain is originated from Vietnam. MSR17 strain was closely related to M1-1 Vietnam strain signifying that MSR17 strain might evolves from M1-1 Vietnamese strain. Kumar et al. (2018) reported that M1-1 strain causes a mild form of shrimp acute hepatopancreatic necrosis disease (AHPND) (20). Compared to other virulent strains, the M1-1 genome appeared to express several additional genes, while some genes were missing. These instabilities may be related to the reduced virulence of M1-1.

The tree also illuminates that MSR16 strain arise earlier than MSR17 strain. NA7 Malaysia strain belonged to an independent lineage and distantly related to our studied strains (MSR16 and MSR17) signifying that this strain might be disparate to MSR16 and MSR17.

SNP tree of different AHPND positive *V. parahaemolyticus*

A total of 37 genome of AHPND positive *V. parahaemolyticus* strain including MSR16 and MSR17 were selected for establishing SNP relationship (Figure 7). MVP3 Malaysia and MVP5 Malaysia strain located at same cluster. The strains MVP10 Malaysia and 13-028-A2 Vietnam located at same cluster. TUMSAT-H03-S5 Thailand, FIM-S1392 Mexico and TUMSAT-H10-S6 Thailand strain belongs to an independent lineage respectively. The strains NB08 Malaysia and NB10 Malaysia separately make a cluster. The strain including NA8 Malaysia, NA6 Malaysia and NB07 Malaysia located at same cluster. The strain including NB12 Malaysia and NB09 Malaysia located at same cluster.

The strains MSR16 Bangladesh and NA9 Malaysia were closely related and located at same cluster indicating that the mutation and evolutionary pattern of MSR16 might be comparable to Malaysian strains. NA9 strain was extracted from Malaysian aquaculture pond water which causes AHPND in shrimp and impacting Malaysian shrimp aquaculture. While strains M0605 Mexico and MSR17 Bangladesh were closely related and located at same cluster indicating that the mutation and evolutionary pattern of MSR17 might be analogous to the Mexican strain. Gomez-Gil et al. (2014) reported that, several pathogenicity mechanisms were identified on both chromosomes: five iron acquisition systems (hemin, enterobactin, vibrioferrin, and two TonB) and seven secretion systems (two T2SS, one T3SS, two T2/4SS, and two T6SS). At least 14 different toxin genes were annotated, two of which are large repeats in toxin (RTX), as well as nine hemolysins. Gomez-Gil et al. (2014) also detected four plasmid from M0605 strains genome (11).

Strains NA7 Malaysia and NA4 Malaysia located at same cluster. Strain TUMSAT-H03-S5 Thailand strain belonged to an independent lineage. This strains mutation and evolutionary pattern might disparate to MSR16 and MSR17.

ANI (Average Nucleotide Identity) tree of different AHPND positive *V. parahaemolyticus* strain

A total of 52 genome of AHPND positive *V. parahaemolyticus* strain including MSR16 and MSR17 were selected for calculating the average nucleotide identity (ANI) (Figure 8). The 11 strains including VP14 India, FIMS1392-, 13-028/A2 Vietnam, TUMSAT-S5 Thailand, TUMAT-S4 Thailand, M0605-Mexico, 13-306-D4 Mexico, MSR17, FIMS1708+ Mexico, TUMSATDE1 Thailand, NCKU-CHN China were located at cluster 2. KS17.S5-1 Malaysia, TUMSATD06 Thailand, 12-297/B Vietnam, 12-009A Vietnam, PSU5579 Thailand and NCKU-S02 China located at cluster same cluster.

The ANI tree clearly illustrates that MSR16 strain belonged to an independent lineage and indicating this strain evolve earlier than MSR17. The reasons for belongs to an independent lineage might be the presence of extra 200 kb nucleotide in the genome. The strain MSR17 was closely related to 13-306-D4 Mexico strain signifying that the average nucleotide identity (ANI) of MSR17 is comparable to this Mexican strain.

The strains ND11 Malaysia and ND13 USA belonged to an independent lineage as well as distantly related to our studied strain (MSR16, MSR17) indicating that these 2 strains genome sequence might be disparate to MSR16 and MSR17 strain.

Phylogenetic relationship of identified plasmids found in the AHPND related isolates

A total of 26 *V. parahaemolyticus* isolates plasmid including pMSR16 and pMSR17 were selected for establishing phylogenetic relationship among the AHPND positive *V. parahaemolyticus* plasmid (Figure 9). The tree is unrooted as cannot determine which plasmid arose earliest. 6 plasmid including pMSR16, pVPA3-1 Vietnam, pMSR17, pVpR13-71Kb USA, pVPGX1 China, pVPE61a Thailand, were located at Cluster 1.

The phylogenetic tree showed that pMSR16 and pMSR17 were closely related with pVPA3-1. The plasmid pVPA3-1 is a Vietnam strain and its accession no is NC_025152.1. Han et al. (2015) reported that, 69 kb plasmid pVPA3-1 was identified in *V. parahaemolyticus* strain 13-028/A3 that can cause AHPND (9). It consists of 92 open reading frames that encode mobilization proteins, replication enzymes, transposases, virulence-associated proteins, and proteins similar to Photorhabdus insect-related (*Pir*) toxins. In *V. parahaemolyticus*, these *Pir* toxin-like proteins are encoded by 2 genes (*pirA*- and *pirB*-like) located within a 3.5 kb fragment flanked with inverted repeats of a transposase-coding sequence (~1 kb).

The plasmid p1937-2 China and pFORC4 South Korea: Busan strain were located at cluster 2. The plasmid pFORC14, 04-2192 located in cluster 3. The plasmid pVpR14-74Kb, pVA1, ISF-54-12, HS-22, P2, 04-2551, PMA109-5 located in the cluster 4. The plasmid S372-5, pVPUCMV, pC4602-2, RM-13-3 located at the cluster 5. The plasmid pVpR13-55Kb USA, pVpR14-56Kb USA and pVPGX2 China located at the cluster 6.

The plasmid pV110-KY498540.1 China and p1937-1-NZCP022245.1 China belonged to an independent lineage respectively. This two strains might acquire plasmids form different sources. These two plasmids

are also distantly related with our studied pMSR16 and pMSR17 indicating that they might be not come from China.

Discussion

Penaeus monodon is a very important aquaculture shrimp species in tropical countries. With the intensification of shrimp farming worldwide, new pathogens are seen to be emerged frequently. For years, the sustainability of the shrimp farming industry has been threatened by a variety of microorganism, such as white spot syndrome virus (WSSV), *Vibrio spp.* and Taura Syndrome Virus (TSV) (2). Since it was first found in the outbreaks of China in 2009, acute hepatopancreatic necrosis disease (AHPND) which was previously known as early mortality syndrome (EMS) due to its rapid lethal effect in early stages of shrimp, has spread all over the world, causing great economic loss (21).

Photorhabdus insect-related (*Pir*) toxin-like genes have been recently identified in various AHPND-causing *V. parahaemolyticus* strains and these genes (*pirA* and *pirB* like) were shown to be the primary virulence factor in these strains. *Vibrio parahaemolyticus* carrying the toxin genes *pirA* and *pirB* causes acute hepatopancreatic necrosis disease (AHPND) in cultured shrimp. The binary toxins *pirA* and *pirB* are similar to Photorhabdus insect-related (*Pir*) toxin (22). The disease has spread from Asia to the Americas, and now into Texas in the United States, causing major economic losses throughout its path. The *pirAB*-containing region in *V. parahaemolyticus* R14 strain is encoded on the pVpR14-74Kb plasmid. The toxin genes in the R13 strain are located in the pVpR13-71Kb plasmid. Sequence analysis revealed that the promoter region upstream of *pirA*, the entire open reading frame (ORF) of *pirA*, and part of the 5' end of the *pirB* ORF is deleted in R13. Since the binary toxins are the virulence factor in AHPND causing *V. parahaemolyticus*, deletion of the toxin genes contributed to a virulence in the R13 strain. Devadas et al. (2018) reported that acute hepatopancreatic necrosis disease (AHPND) is a shrimp bacterial disease caused by *Vibrio spp.* carrying plasmid encoding homologues of the *Photorhabdus* insect-related (*Pir*) toxins *pirAvp* and *pirBvp* (23). Kumar et al. (2018) reported that acute hepatopancreatic necrosis disease (AHPND), strains can be described as either virulent or nonvirulent. The ability of a *Vibrio parahaemolyticus* given strain to cause AHPND depends on the presence of the virulence plasmid pVA1, which harbors the binary toxins *PirAVp* and *PirBVp* (20).

In this study, the genomes of both MSR16 and MSR17 strains contain a plasmid of ~69 kbp. The plasmid of MSR16 contains total 87 genes and MSR17 contains total 88 genes. Both of the plasmid carry *PirA* and *PirB* genes which is responsible for AHPND disease. The length of *pirA* gene was 336 bp (starts at 64,962 bp and stops at 65,297 bp) and the length of *pirB* gene was 1,317 bp (starts at 65,310 bp and stops at 66,626 bp) in MSR16. On the other hand the length of *pirA* gene was 336 bp (starts at 63108 bp and stops at 63443 bp) and the length of *pirB* gene was 1,317 bp (starts at 63,456 bp and stops at 64,772 bp) in MSR17.

1. *parahaemolyticus* is found to be widely present in the environments, including sediments, plankton, and aquatic animals (24). Phylogenetic analysis also showed that the all AHPND related isolates could be clearly differentiated into distinct clusters each specific for different regions. The phylogenetic analysis of the core genome of pVA-1 plasmid also showed considerable divergence among the pVA-1 like plasmids with at least three clusters (25). The gene organization of pVA1 showed that two transposes were located in both side of *pirAB*, further suggesting that *pirAB* may be frequently transferred among the *V. parahaemolyticus* or even other *Vibrio* sp. via transposition (26).

Antibiotic resistance-producing bacteria are capable of transmitting naturally occurring resistance genes to other bacteria through genetic exchange, enabling them to neutralize or destroy the antibiotics with which they are challenged (27). It can be observed that both strains possess resistance genes for efflux mechanisms and antibiotic modification which supports our previously reported antibiogram data (6, 28). While comparing to the experimental results, presence of several predicted resistance genes against some antibiotics were found which they are currently either sensitive or have intermediate response, means these strains can gain significant antibiotic resistance in the nearest future.

The phylogenetic analysis of the core genome of pVA-1 plasmid also showed considerable divergence among the pVA-1 like plasmids with at least three clusters (25). The gene organization of pVA1 supports the spreading of *pirAB* toxins into other strains (26). In the present study the phylogenetic analysis of MSR16 and MSR17 strain was done by several ways such as the construction of tree based on 16s rRNA gene, common housekeeping genes excluding 16s rRNA, tree based on plasmid, SNP tree and average nucleotide identity (ANI tree).

The phylogenetic tree based on 16S rRNA showed our studied strains MSR16 and MSR17 located at same cluster and were closely related with one of Indian strain AP1511 indicating that the mutation and evolutionary pattern of MSR16 and MSR17 strains might be analogous to the Indian strain. The phylogenetic tree based on common housekeeping genes showed that MSR16 strain was closely related to 12-009A/1335 Vietnam and MSR17 strain was closely related to M1-1 Vietnam strain signifying that evolution of both strain might be from Vietnam.

The SNP tree illustrates that MSR16 strains is closely related with three Malaysian strain indicating that he mutation and evolutionary pattern of MSR16 might be comparable to these Malaysian strains. On the other hand, MSR17 strains is closely related with M0605 Mexico strain indicating that the mutation and evolutionary pattern of MSR17 might be analogous to the Mexican strain.

The ANI tree clearly illustrates that MSR16 strain belonged to an independent lineage and indicating this strain evolve earlier than MSR17. The reasons for belongs to an independent lineage might be the extra 200 kb nucleotide in the genome. The strain MSR17 was closely related to 13-306-D4 Mexico strain signifying that the average nucleotide identity (ANI) of MSR17 is comparable to this Mexican strain.

The phylogenetic tree showed that pMSR16 and pMSR17 were closely related with pVPA3-1. The plasmid pVPA3-1 is a Vietnam strain and its accession no is NC-025152.1. Both plasmid carry the causative agent *pirA* and *pirB* gene of AHPND in their sequence. The plasmid of two studied strain might evolved from Vietnam. From the above explanations it can be said that the *V. parahaemolyticus* (AHPND outbreaks) have multiple origins.

Conclusion

In this study, we report the 5.3 Mbp and 5.2 Mbp genome sequences of *V. parahaemolyticus* strains MSR16 and MSR17 having distinct virulence factors for causing the outbreaks in Bangladesh. Complete resequencing of these genomes of AHPND causing strains MSR16 and MSR17 should provide genomic insights into the pathogenicity and virulence mechanisms of VP_{AHPND}. Additional comparative genomics and phylogenetic studies of these two strains may provide understandings of their emergence, spreading patterns so that future outbreaks can be predicted. Also, with the help of different genome sequences collected from outbreaks around the world along with our reported sequences, novel vaccines or drug targets can be identified to tackle any future outbreaks in shellfishes and to reduce the chances of getting these strains introduced in human food chain to prevent potential health hazards.

Methods

Revitalization of *V. parahaemolyticus* strains and molecular identification

1. *parahaemolyticus* strains from the previous study (6, 28) were inoculated in Tryptone soya broth (TSB) with 2% salt. The bacteria grow in the TSB were streaked on TCBS agar plate. From TCBS agar plate the bacterial isolates were re-streaked on ChromAgar *Vibrio* medium (CHROMagar, Paris-France). The bacterial isolates were further streaked on Tryptone soya agar (TSA) with 2% salt to obtain pure isolates. To support the vigorous growth of *V. parahaemolyticus* strain, Luria bertani (LB) broth was used with 2% salt. Total genomic DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. The DNA quality was quantified using NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA). Polymerase chain reaction (PCR) were used for the partial amplification of 16S rRNA, *ldh*, AP3 and AP4 genes for the molecular identification of suspected AHPND positive *V. parahaemolyticus* strains.

Sequencing and assembly

A genomic library was constructed and employed for 150 bp paired-end whole-genome sequencing using an Illumina MiSeq platform (Illumina, San Diego, CA, USA). An in house pipeline was built to perform the

whole assembly process which performed i) Adapter and low quality base trimming using Trimmomatic v0.38 (29) using several parameters; ii) Generation of QC reports of trimmed and untrimmed data using FastQC v0.11.7 (30); iii) Genome Assembly using the trimmed and untrimmed data by SPAdes v3.10 (31) in both general and plasmid mode utilizing different k-mer combinations; iv) Assembly polishing using Pilon v1.22 (32); v) Determination of the quality and coverage of the assemblies using Quast v5.0.2 (33); vi) Scaffolding into chromosomes and plasmids by MeDuSa v1.6 (34). Genomic scaffolds of these two strains were compared using Mauve 2.4.0 (35).

Gene prediction and annotation

Genome annotation was performed by Prokka v1.12 (36), Glimmer v3.02 (37), RASTtk v1.3.0 (38) and tRNA, rRNA annotation was done using Barrnap v0.6 , tRNAscan-SE v2.0 (39). Average nucleotide identity of 50 different AHPND causing *V. parahaemolyticus* strains was calculated using Pyani v0.2.7 (40).

Analysis of virulence, antibiotic resistance genes and others

Virulence factors were searched using VFAnalyzer (41). ResFinder (42), ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation) (43) and CARD tools (44) were used to search antibiotic resistance genes. Prophage sequences were searched respectively by PHASTER (45). COG (Clusters of orthologous groups) classification of the genes was achieved by eggNOG-mapper v1 (46). Genomic islands were predicted using the Islandviewer tool (47).

Phylogenetic analysis and genome comparison

Phylogenetic tree was constructed based on 16S rDNA sequences of several AHPND causing *V. parahaemolyticus* strains from around different parts of the world using MEGA 7.0 software (48). The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (49).

Housekeeping genes of 25 different *V. parahaemolyticus* strains were obtained from MLST 2.0 server (<https://cge.cbs.dtu.dk/services/MLST/>) using the assembled genomes of those strains. Gene sequences were extracted and concatenated using in house shell scripts and a Neighbor-joining tree was constructed using MEGA 7.0 software (48).

SNP based NJ-tree was constructed with genomes of 37 different AHPND positive *V. parahaemolyticus* strains using Parsnp v1.2 tool (50). Average nucleotide identity (ANI) based tree was constructed from our generated ANI value matrix (Additional file 1) using PHYLIP package (51). 26 plasmids from several AHPND positive *V. parahaemolyticus* strains were aligned and a NJ-tree was constructed using MAFFT v7 (52).

Quality assurance

16S rRNA genes of *V. parahaemolyticus* strain MSR16 and MSR17 were predicted from the annotation pipeline and also from the BLAST (53) search of the PCR amplified partial sequences of both strains' 16S rRNA genes.

Declarations

List of abbreviations

EMS	Early mortality syndrome
AHPND	Acute hepatopancreatic necrosis disease
T3SS	Type three secretion system
<i>pir</i>	Photorhabdus insect-related
WGS	Whole genome sequencing
SNP	Single nucleotide polymorphism
COG	Clusters of orthologous groups
GI	Genomic Island
ANI	Average Nucleotide Identity

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The draft sequences of both *Vibrio parahaemolyticus* strains MSR16 and MSR17 can be found in NCBI Bioproject ID PRJNA505599.

Competing interests

The authors declare no conflict of interests.

Funding

None.

Authors' contributions

MSR and ABMMKI conceived and designed the project. SA and MAAKK performed the culture, sequence assembly, annotation and analyses. MMEE and NJP collected shrimp samples, isolated and identified the AHPND positive *V. parahaemolyticus*. SA, MAAKK, MMEE, NJP, MSR and ABMMKI wrote the manuscript. SA and MAAKK contributed equally. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank Mst. Khadiza Begum, Lecturer, Department of Fisheries, University of Dhaka for her assistance during strain isolation and helpful criticism of this manuscript.

References

1. Boonyawiwat V, Patanasatienkul T, Kasornchandra J, Poolkhet C, Yaemkasem S, Hammell L, et al. Impact of farm management on expression of early mortality syndrome/acute hepatopancreatic necrosis disease (EMS/AHPND) on penaeid shrimp farms in Thailand. *Journal of fish diseases*. 2017;40(5):649-59.
2. Flegel TW. Historic emergence, impact and current status of shrimp pathogens in Asia. *Journal of invertebrate pathology*. 2012;110(2):166-73.
3. Tran L, Nunan L, Redman RM, Mohny LL, Pantoja CR, Fitzsimmons K, et al. Determination of the infectious nature of the agent of acute hepatopancreatic necrosis syndrome affecting penaeid shrimp. *Diseases of aquatic organisms*. 2013;105(1):45-55.
4. Zorriehzahra M, Banaederakhshan R. Early mortality syndrome (EMS) as new emerging threat in shrimp industry. *Adv Anim Vet Sci*. 2015;3(2S):64-72.

5. FAO. Report of the FAO/MARD technical workshop on early mortality syndrome (EMS) or acute hepatopancreatic necrosis syndrome (AHPNS) of cultured shrimp (under TCP/VIE/3304). FAO Rome; 2013. p. 54.
6. Eshik MME, Abedin MM, Punom NJ, Begum MK, Rahman MS. Molecular Identification of AHPND Positive *Vibrio Parahaemolyticus* causing an outbreak in South-West Shrimp Farming Regions of Bangladesh. Journal of Bangladesh Academy of Sciences. 2017;41(2):127-35.
7. Xie T, Xu X, Wu Q, Zhang J, Cheng J. Prevalence, molecular characterization, and antibiotic susceptibility of *Vibrio parahaemolyticus* from ready-to-eat foods in China. Frontiers in microbiology. 2016;7:549.
8. Lee C-T, Chen I-T, Yang Y-T, Ko T-P, Huang Y-T, Huang J-Y, et al. The opportunistic marine pathogen *Vibrio parahaemolyticus* becomes virulent by acquiring a plasmid that expresses a deadly toxin. Proceedings of the National Academy of Sciences. 2015;112(34):10798-803.
9. Han JE, Tang KF, Tran LH, Lightner DV. Photorhabdus insect-related (*Pir*) toxin-like genes in a plasmid of *Vibrio parahaemolyticus*, the causative agent of acute hepatopancreatic necrosis disease (AHPND) of shrimp. Diseases of aquatic organisms. 2015;113(1):33-40.
10. Okada N, Iida T, Park K-S, Goto N, Yasunaga T, Hiyoshi H, et al. Identification and characterization of a novel type III secretion system in trh-positive *Vibrio parahaemolyticus* strain TH3996 reveal genetic lineage and diversity of pathogenic machinery beyond the species level. Infection and immunity. 2009;77(2):904-13.
11. Gomez-Gil B, Soto-Rodríguez S, Lozano R, Betancourt-Lozano M. Draft genome sequence of *Vibrio parahaemolyticus* strain M0605, which causes severe mortalities of shrimps in Mexico. Genome Announc. 2014;2(2):e00055-14.
12. Joshi J, Srisala J, Truong VH, Chen I-T, Nuangsaeng B, Suthienkul O, et al. Variation in *Vibrio parahaemolyticus* isolates from a single Thai shrimp farm experiencing an outbreak of acute hepatopancreatic necrosis disease (AHPND). Aquaculture. 2014;428:297-302.
13. Sirikharin R, Taengchaiyaphum S, Sritunyalucksana K, Thitamadee S, Flegel TW, Mavichak R, et al. A new and improved PCR method for detection of AHPND bacteria. Network of Aquaculture Centres in Asia-Pacific (NACA). 2014:7-9.
14. Dangtip S, Sirikharin R, Sanguanrut P, Thitamadee S, Sritunyalucksana K, Taengchaiyaphum S, et al. AP4 method for two-tube nested PCR detection of AHPND isolates of *Vibrio parahaemolyticus*. Aquaculture Reports. 2015;2:158-62.
15. Bayliss SC, Verner-Jeffreys DW, Bartie KL, Aanensen DM, Sheppard SK, Adams A, et al. The promise of whole genome pathogen sequencing for the molecular epidemiology of emerging aquaculture pathogens. Frontiers in microbiology. 2017;8:121.
16. Didelot X, Pang B, Zhou Z, McCann A, Ni P, Li D, et al. The role of China in the global spread of the current cholera pandemic. PLoS genetics. 2015;11(3):e1005072.
17. Ceccarelli D, Hasan NA, Huq A, Colwell RR. Distribution and dynamics of epidemic and pandemic *Vibrio parahaemolyticus* virulence factors. Frontiers in cellular and infection microbiology. 2013;3:97.

18. Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O. PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data. PLOS ONE. 2013;8(10):e77302.
19. Li P, Kinch LN, Ray A, Dalia AB, Cong Q, Nunan LM, et al. Acute hepatopancreatic necrosis disease-causing *Vibrio parahaemolyticus* strains maintain an antibacterial type VI secretion system with versatile effector repertoires. Appl Environ Microbiol. 2017;83(13):e00737-17.
20. Kumar R, Chang C-C, Ng TH, Ding J-Y, Tseng T-C, Lo C-F, et al. Draft genome sequence of *Vibrio parahaemolyticus* strain M1-1, which causes acute hepatopancreatic necrosis disease in shrimp in Vietnam. Genome Announc. 2018;6(3):e01468-17.
21. Soto-Rodriguez SA, Gomez-Gil B, Lozano-Olvera R, Betancourt-Lozano M, Morales-Covarrubias MS. Field and experimental evidence of *Vibrio parahaemolyticus* as the causative agent of acute hepatopancreatic necrosis disease of cultured shrimp (*Litopenaeus vannamei*) in Northwestern Mexico. Appl Environ Microbiol. 2015;81(5):1689-99.
22. Kanrar S, Dhar AK. Complete Genome Sequence of a Deletion Mutant of *Vibrio parahaemolyticus* from Pacific White Shrimp (*Penaeus vannamei*). 2018.
23. Devadas S, Bhassu S, Soo TCC, Yusoff FM, Shariff M. Draft Genome Sequence of the Shrimp Pathogen *Vibrio parahaemolyticus* ST17. P5-S1, Isolated in Peninsular Malaysia. Microbiol Resour Announc. 2018;7(11):e01053-18.
24. Su Y-C, Liu C. *Vibrio parahaemolyticus*: a concern of seafood safety. Food microbiology. 2007;24(6):549-58.
25. Fu S, Tian H, Wei D, Zhang X, Liu Y. Delineating the origins of *Vibrio parahaemolyticus* isolated from outbreaks of acute hepatopancreatic necrosis disease in Asia by the use of whole genome sequencing. Frontiers in microbiology. 2017;8:2354.
26. Xiao J, Liu L, Ke Y, Li X, Liu Y, Pan Y, et al. Shrimp AHPND-causing plasmids encoding the PirAB toxins as mediated by pirAB-Tn903 are prevalent in various *Vibrio* species. Scientific reports. 2017;7:42177.
27. Bennett P. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. British journal of pharmacology. 2008;153(S1):S347-S57.
28. Eshik MME, Punom NJ, Begum MK, Tahsin K, Saha ML, Rahman MS. Molecular identification and characterization of *Vibrio parahaemolyticus* causing Acute Hepatopancreatic Necrosis Disease in cultured Shrimp. Dhaka University Journal of Biological Sciences. 2018;27(1):57-68.
29. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
30. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. 2017.
31. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology. 2012;19(5):455-77.
32. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS one.

- 2014;9(11):e112963.
33. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QCAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;29(8):1072-5.
 34. Bosi E, Donati B, Galardini M, Brunetti S, Sagot M-F, Lió P, et al. MeDuSa: a multi-draft based scaffolder. *Bioinformatics*. 2015;31(15):2443-51.
 35. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. *Genome Research*. 2004;14(7):1394-403.
 36. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068-9.
 37. Aggarwal G, Ramaswamy R. Ab initio gene identification: prokaryote genome annotation with GeneScan and GLIMMER. *Journal of biosciences*. 2002;27(1):7-14.
 38. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific reports*. 2015;5:8365.
 39. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic acids research*. 1997;25(5):955-64.
 40. Pritchard L. pyani: Python module for average nucleotide identity analyses. 2017.
 41. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic acids research*. 2018;47(D1):D687-D92.
 42. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*. 2012;67(11):2640-4.
 43. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a New Bioinformatic Tool To Discover Antibiotic Resistance Genes in Bacterial Genomes. *Antimicrobial Agents and Chemotherapy*. 2014;58(1):212.
 44. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research*. 2016;45(D1):D566-D73.
 45. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*. 2016;44(W1):W16-W21.
 46. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, et al. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Molecular Biology and Evolution*. 2017;34(8):2115-22.
 47. Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, et al. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res*. 2017;45(W1):W30-w5.
 48. Kumar S, Nei M, Dudley J, Tamura K. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics*. 2008;9(4):299-306.

49. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*. 1993;10(3):512-26.
50. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biology*. 2014;15(11):524.
51. Felsenstein J. PHYLIP (phylogeny inference package), version 3.5 c: Joseph Felsenstein.; 1993.
52. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*. 2013;30(4):772-80.
53. Johnson M, Zaretskaya I, Raytselis Y, Merezuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *Nucleic Acids Research*. 2008;36(suppl_2):W5-W9.

Figures

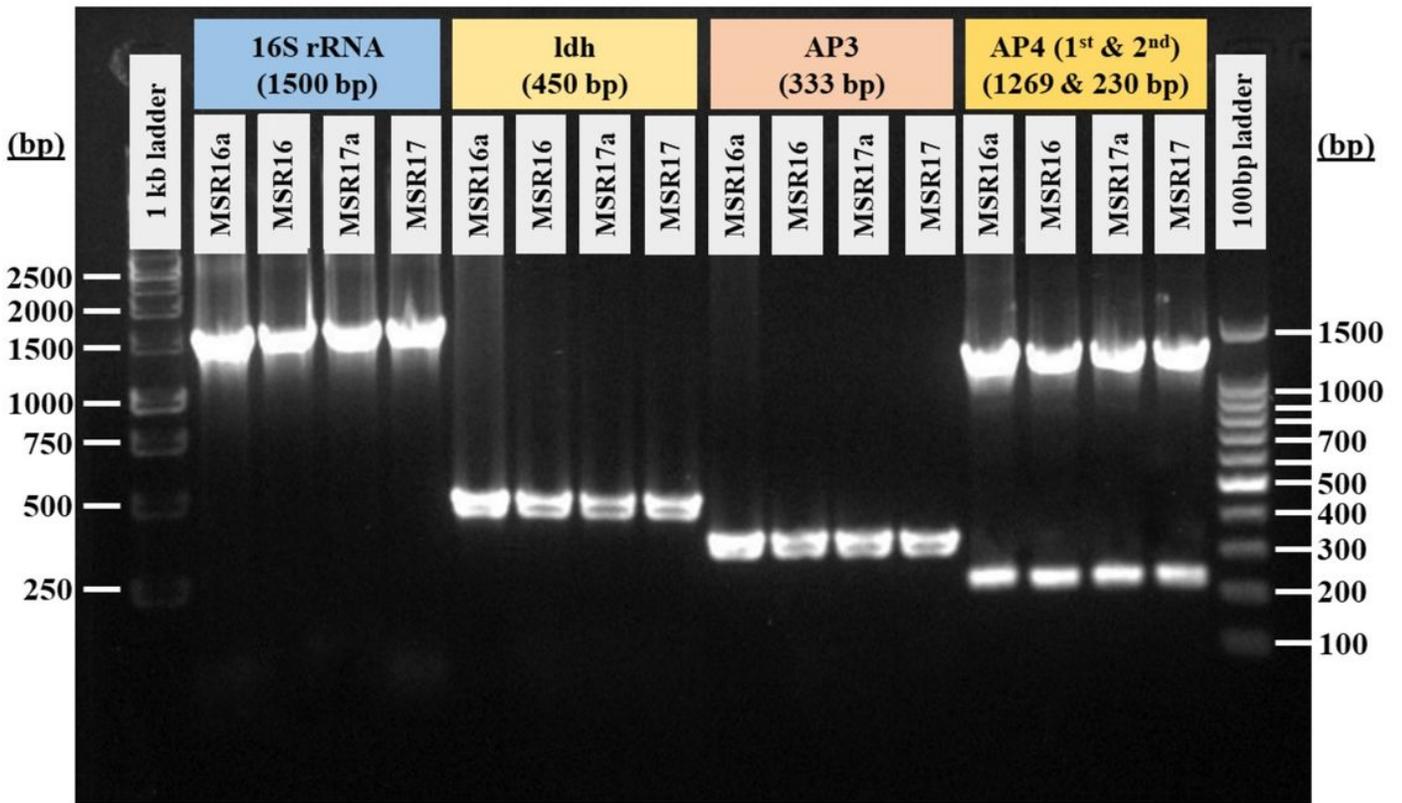


Figure 1

Molecular identification of the AHPND positive *V. parahaemolyticus* strain MSR16 and MSR17. (MSR16a and MSR17a are replicates of MSR16 and MSR17, respectively)

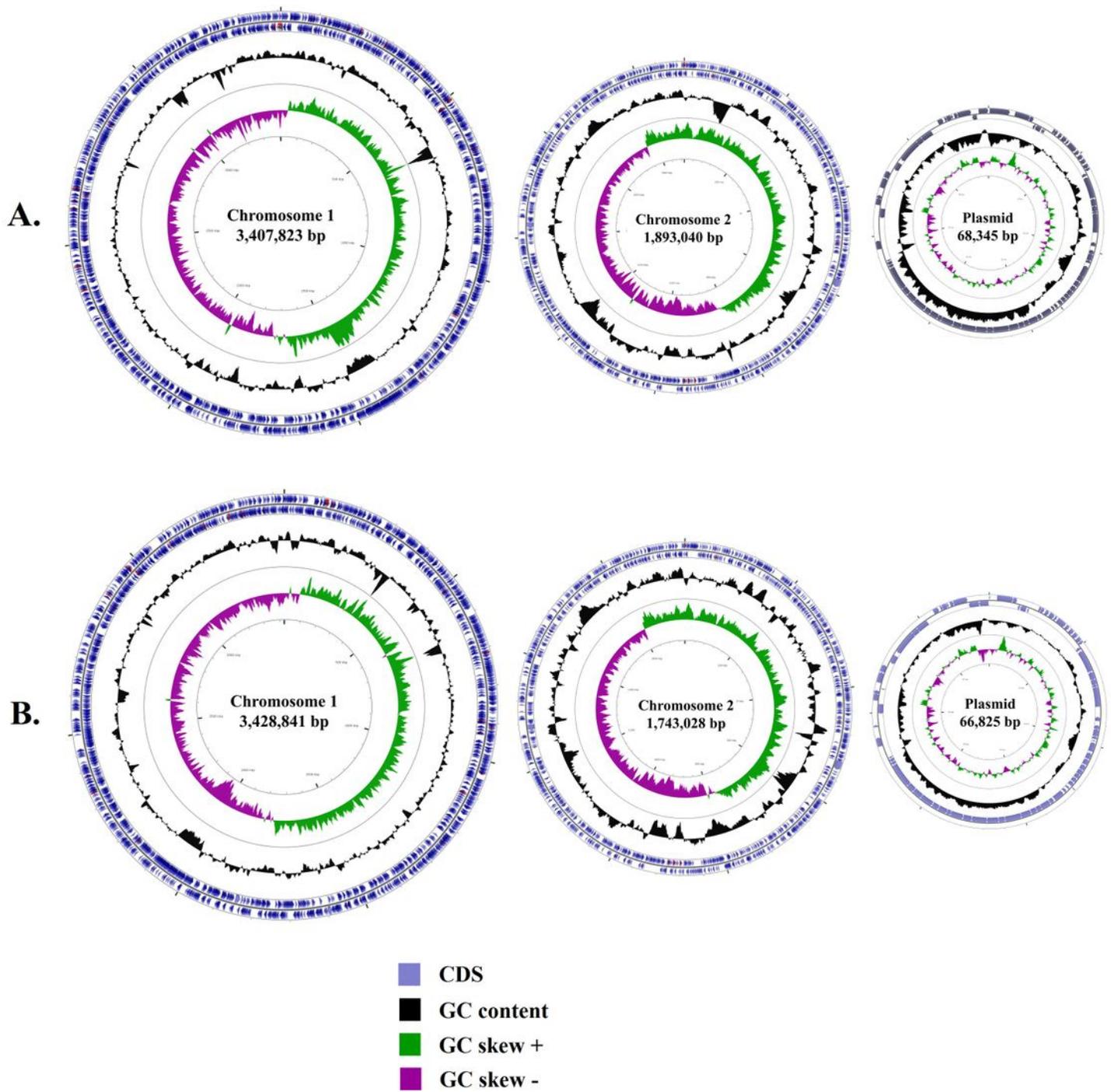


Figure 2

Circular genome representation of the VPAHPND strains A. MSR16 and B. MSR17.

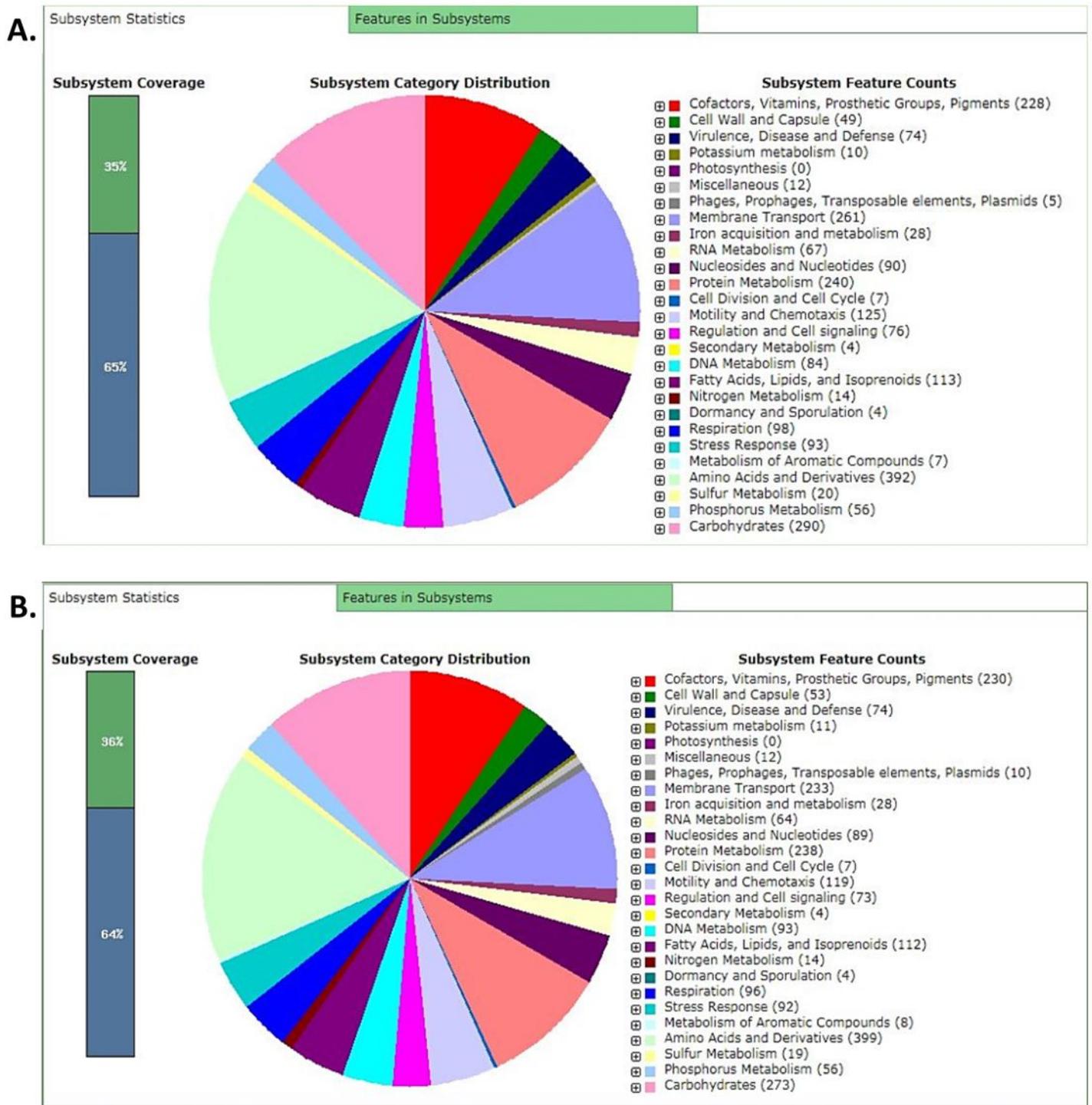


Figure 3

RAST server predicted subsystem categories for AHPND positive *V. parahaemolyticus* strains A. MSR16 and B. MSR17.

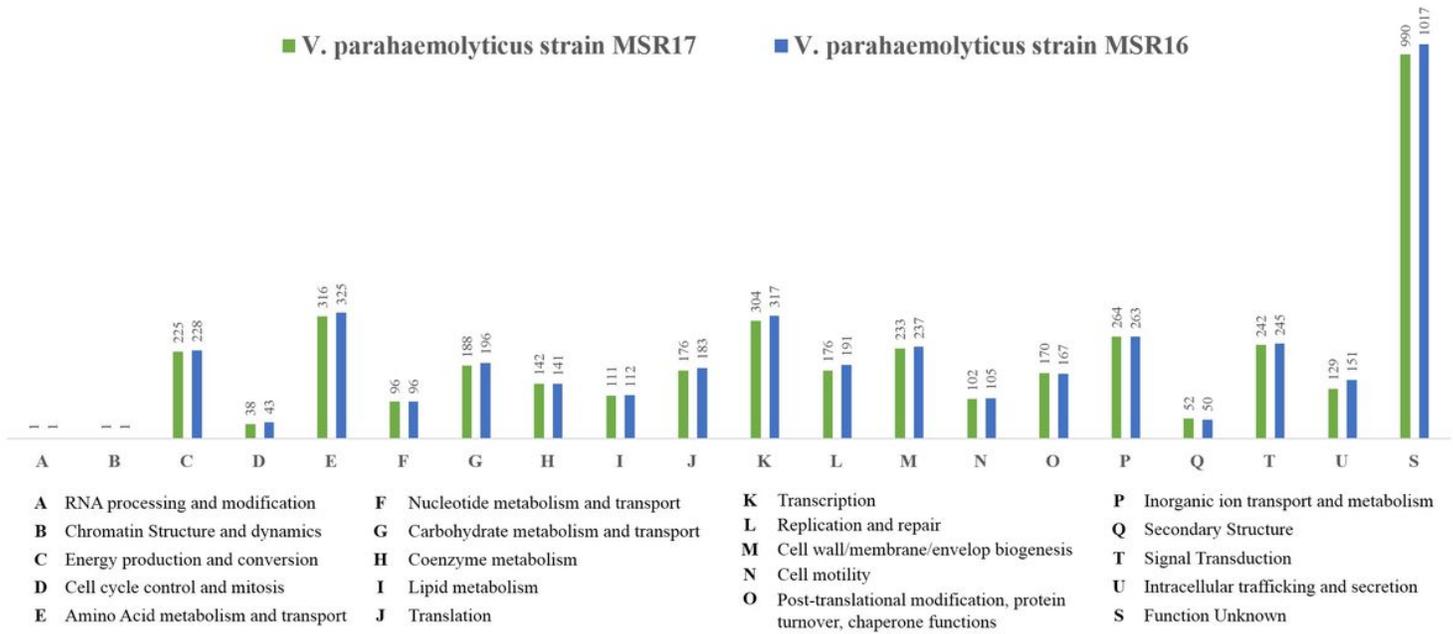


Figure 4

COG classification of the predicted genes in VPAHPND strains MSR16 and MSR17.

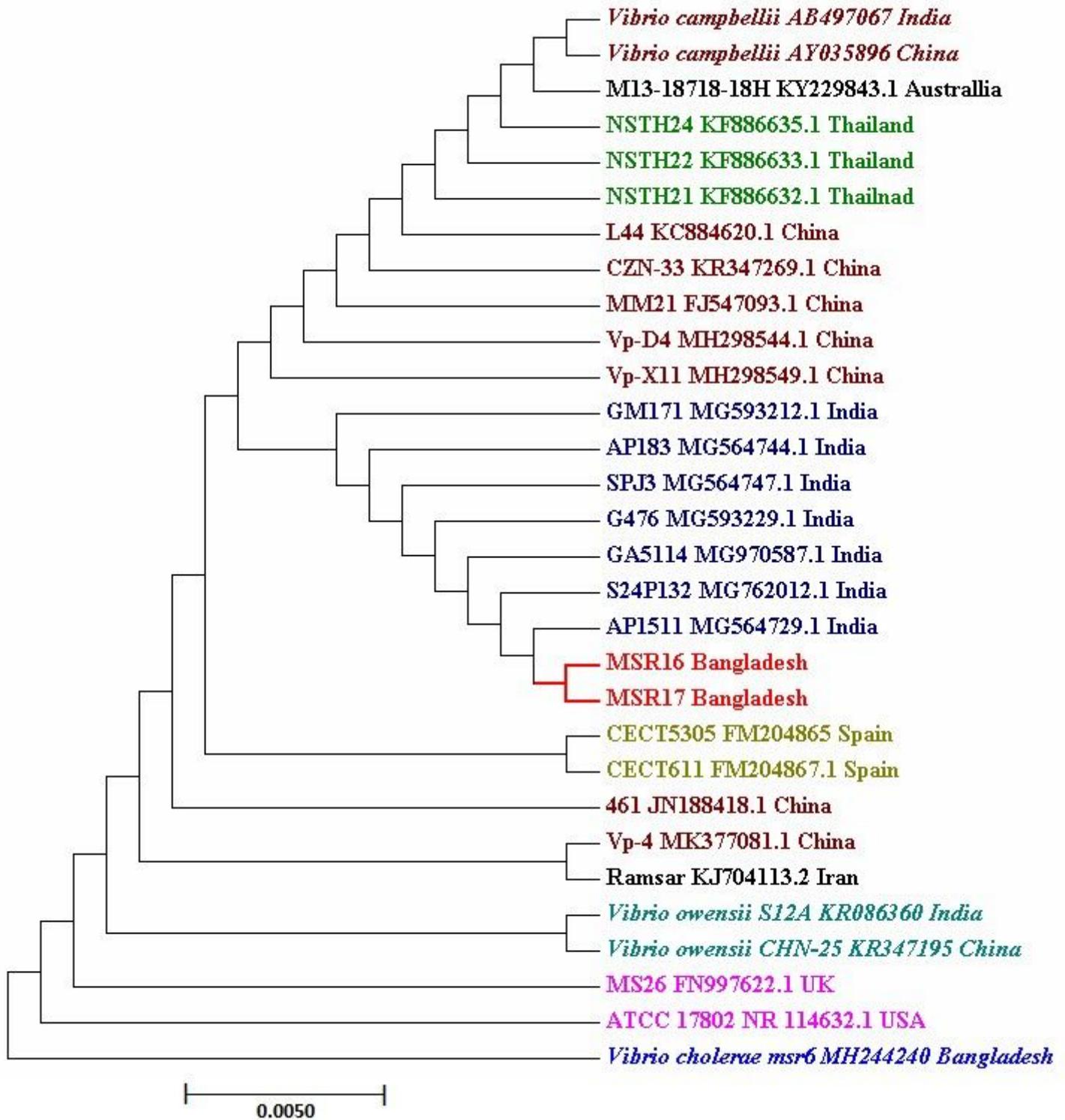


Figure 5

Phylogenetic relationship of 16S rRNA genes of different VPAHPND strains including MSR16 and MSR17 from Bangladesh.

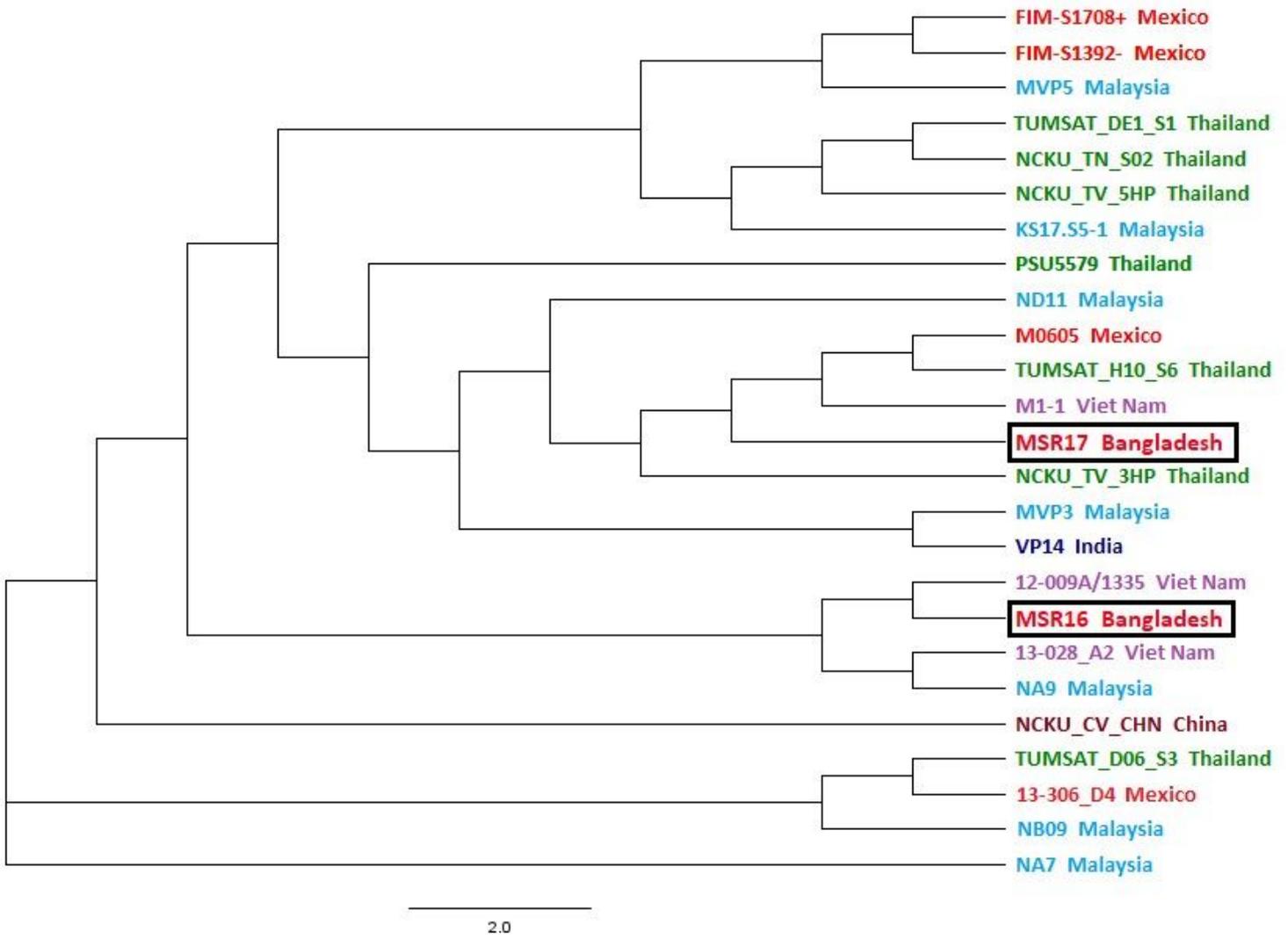


Figure 6

Phylogenetic relationship of using 7 housekeeping genes (dnaE, dtdS, gyrB, pntA, pyrC, recA, tnaA) of different VPAHPND strains including MSR16 and MSR17 from Bangladesh.

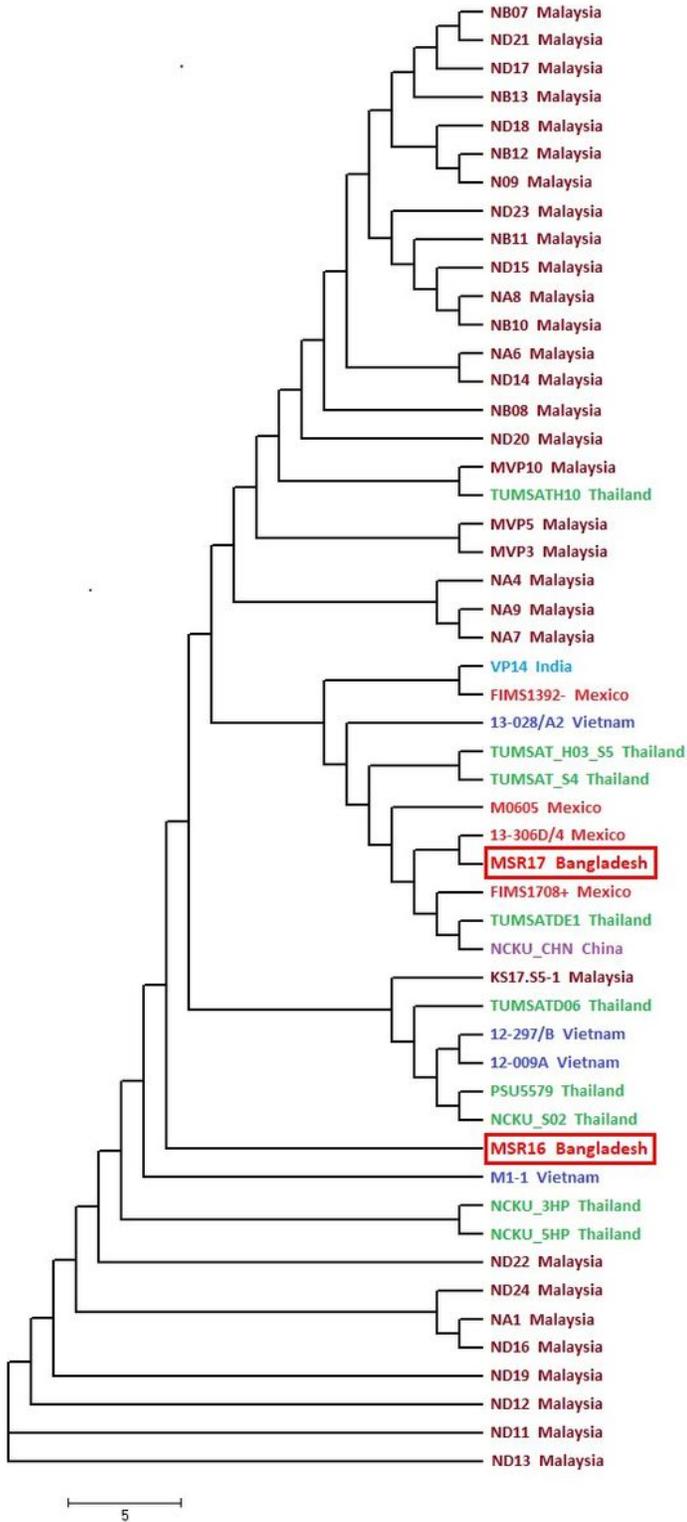


Figure 7

SNP tree of different VPAHPND strains including two VPAHPND strain MSR16 and MSR17 from Bangladesh.

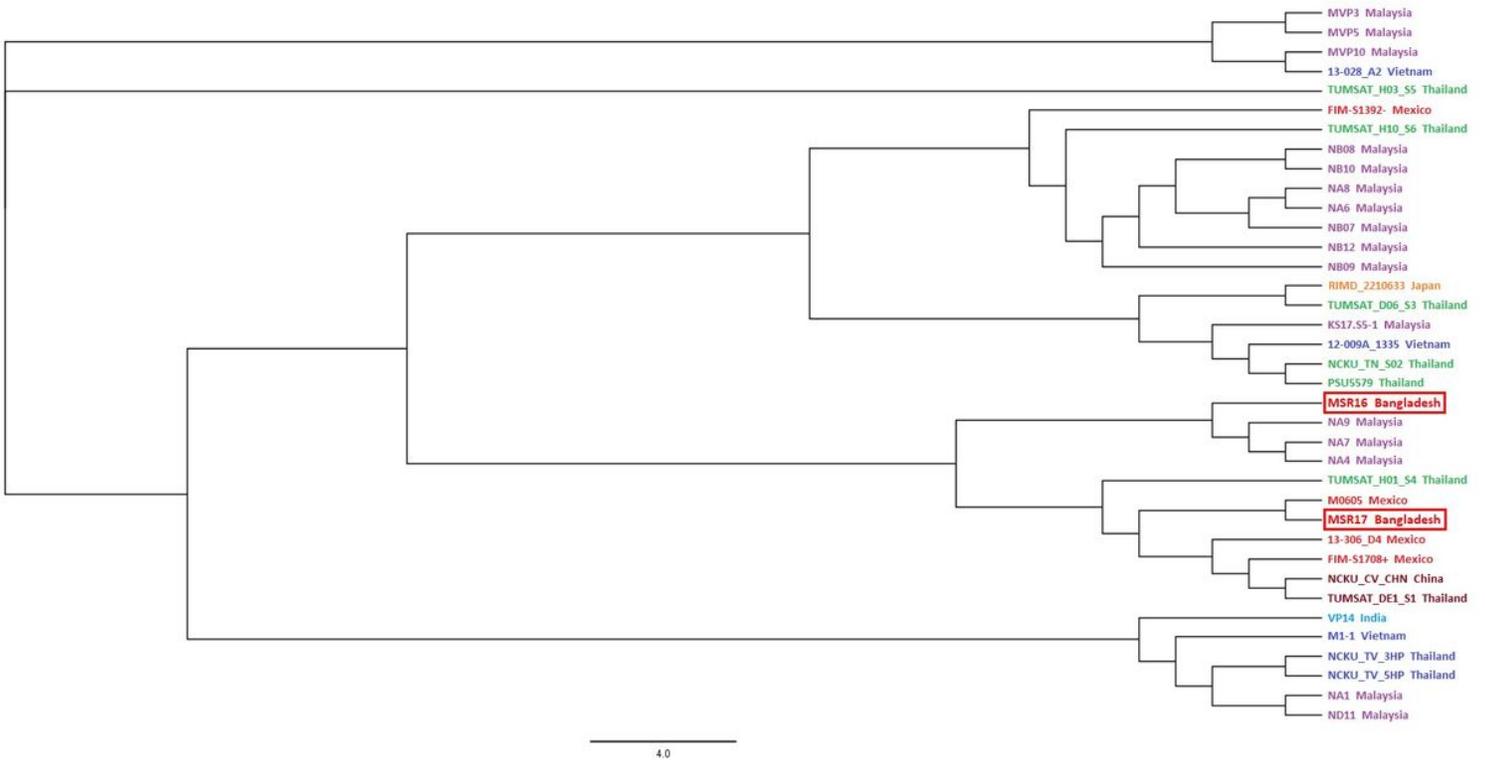


Figure 8

ANI tree of different VPAHPND strains including two VPAHPND strains MSR16 and MSR17 from Bangladesh.

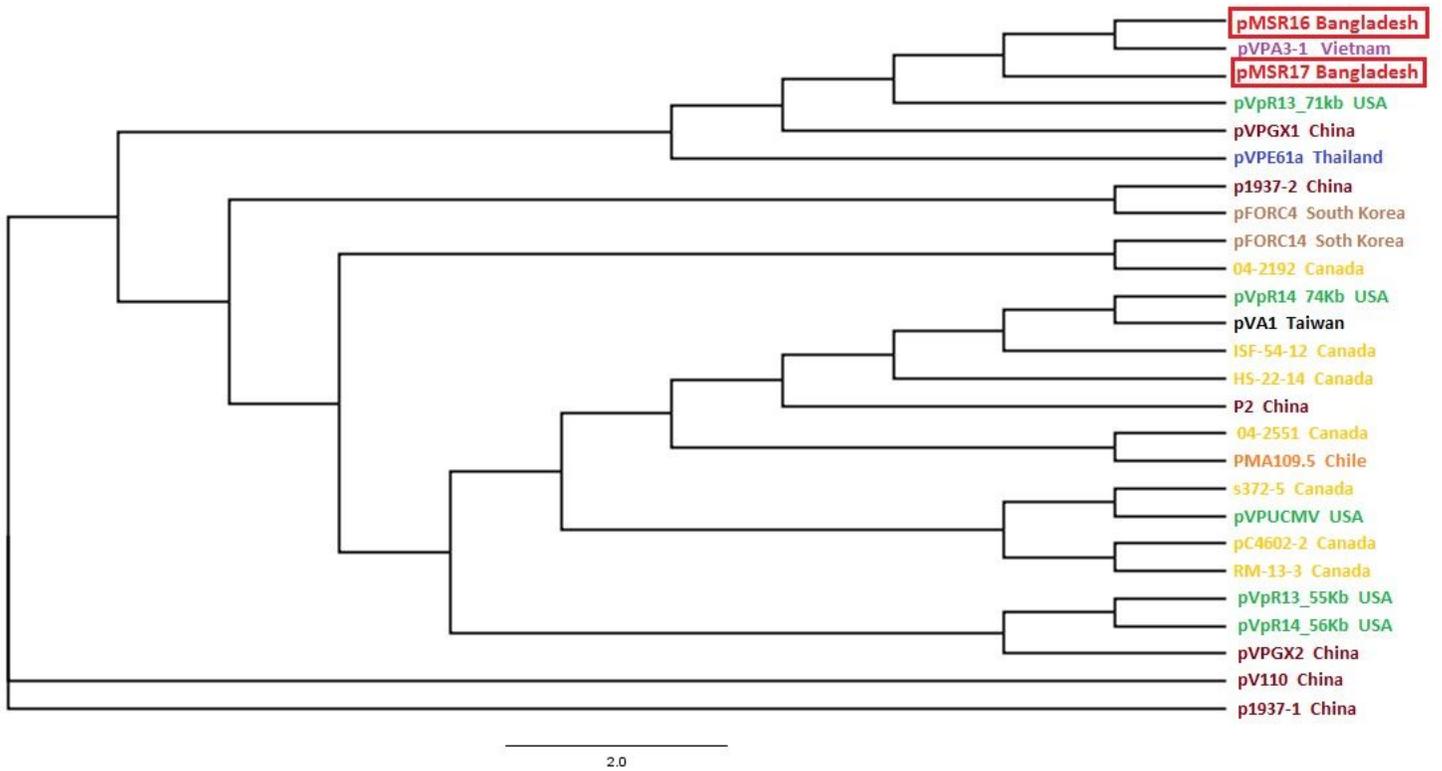


Figure 9

Phylogenetic relationship based on plasmid sequences from VPAHPND isolates including two VPAHPND strains plasmid pMSR16 and pMSR17 from Bangladesh.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile6.xlsx](#)
- [Additionalfile4.xlsx](#)
- [Additionalfile1.xlsx](#)
- [Additionalfile5.docx](#)
- [Additionalfile2.docx](#)
- [Additionalfile3.docx](#)