RESEARCH ARTICLE   **Open Access**

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

Shawon Ahmmed[1†], Md. Abdullah-Al-Kamran Khan[2†], Md. Mostavi Enan Eshik[1], Nusrat Jahan Punom[1], Abul Bashar Mir Md. Khademul Islam[2*] and Mohammad Shamsur Rahman[1*]

Abstract

Background: Due to its rapid lethal effect in the early development stage of shrimp, acute hepatopancreatic necrosis disease (AHPND) has been 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 a total of 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 III secretion system-1 (T3SS-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 and macrolide groups in both MSR16 and MSR17 strains.

Conclusions: The findings of this research may facilitate the tracking of pathogenic and/or antibiotic-resistant *V. parahaemolyticus* isolates between production sites, and the identification of candidate strains for the 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.

Keywords: *Vibrio parahaemolyticus*, Acute hepatopancreatic necrosis disease, AHPND, Genome sequencing, Virulence, *pir*A, *pir*B, Phylogenetics

Background

Asian shrimp farming industry has encountered enormous production losses because of a *Vibrio* caused disease, known as the early mortality syndrome/acute hepatopancreatic necrosis disease (EMS/AHPND) [1].

AHPND is a shrimp bacterial disease which causes high mortality of cultivated penaeid shrimps commonly occur within the first 30 days after stocking in grow-out ponds [2]. Since 2009, AHPND was first recorded in shrimp farms of southern China [3], in 2010 in Island of Hainan [2], in Vietnam and Malaysia in 2011 [4] and subsequently it spread in the eastern part and other culture areas of Thailand in 2012 [4]. Worldwide the production loss of shrimp farming due to AHPND was estimated at about more than $1 billion per year [5]. In Bangladesh,

* Correspondence: khademul@du.ac.bd; shamsur@du.ac.bd
†Shawon Ahmmed and Md. Abdullah-Al-Kamran Khan contributed equally to this work.
[1]Aquatic Animal Health Group, Department of Fisheries, University of Dhaka, Dhaka 1000, Bangladesh
[2]Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh

AHPND positive *Vibrio parahaemolyticus* were first reported in 2017 [6].

The AHPND affected shrimp shows a pale and atrophied hepatopancreas along with an empty stomach and midgut [3]. The moribund shrimps usually harbor some pathological features like- enlarged hepatopancreatic nuclei, sloughed HP cells-blister-like (B), fibrilla (F), resorptive (R) cells, and the diseased shrimps frequently suffer from secondary bacterial infections [3]. The causative agent of AHPND in shrimp is *Vibrio parahaemolyticus*; a gram-negative rod-shaped bacterium mainly inhabitant in warm marine and estuarine environment, and distributes throughout the world [3, 7]. AHPND causing *V. parahaemolyticus* possesses ~ 69 kb plasmid encoding toxin genes *pir*A and *pir*B [3, 8] which are similar to Photorhabdus insect-related (*pir*) toxin [9] which is one of the major causal factors reported. Moreover, two sets of the type III secretion system (T3SS1 and T3SS2) possessed by *V. parahaemolyticus* are also considered as an important virulence factor of this organism [10]. Though all strains of *V. parahaemolyticus* contain T3SS1, only the human clinical strains possess T3SS2 [10]. AHPND positive *V. parahaemolyticus* strains do not contain TDH, TRH, and T3SS2 which are known virulence factors affecting humans [11]. Amplification of species-specific gene *ldh* (lecithin dependent hemolysin) [12] is utilized to detect *V. parahaemolyticus* isolates whereas AP3 [13] and AP4 [14] primers are commonly used to identify the AHPND positive strains.

Nowadays, whole genome sequencing (WGS) has become a popular tool for the identification and detection of bacterial outbreaks in aquaculture [15]. In whole genome sequencing, all of the single nucleotide polymorphisms (SNPs) are 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 the south-west region 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 (Fig. 1). MSR16 (isolated in 2016) and MSR17

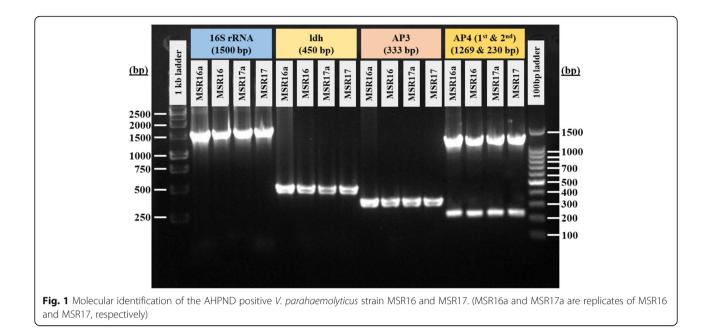(isolated in 2017) strains were finally sequenced for whole genome sequencing.

Features of the assembled genomes

The genomes were assembled into 108 contigs in MSR16 strain and 66 contigs in MSR17 strain. The largest contigs size for MSR16 strain was ~ 1.9 Mbp; and ~ 1.7 Mbp for MSR17 strain. The total GC content was 45.09 and 45.19% for MSR16 and MSR17 strains, respectively. The total genome size of MSR16 was ~ 5.4 Mbp; and ~ 5.2 Mbp for MSR17. MSR16 was found comprised of two circular chromosomes with a length of ~ 3.4 Mbp, ~ 1.8 Mbp while the genome of MSR17 was comprised of similar two circular chromosomes with a length of ~ 3.4 Mbp, ~ 1.7 Mbp. Both MSR16 and MSR17 contain a plasmid with a length of ~ 68 Kbp and ~ 66 Kbp, respectively (Fig. 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.

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 (*pir*A and *pir*B) 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 (*pir*A and *pir*B) 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 (Fig. 3). These particular subsystems are the hallmarks for the pathogenicity and both strains were found to have almost similar amounts of factors across their genomes. The number of genes associated with the general COG functional categories for both strains is provided in (Fig. 4). Both strains are found to possess an 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

Fig. 1 Molecular identification of the AHPND positive *V. parahaemolyticus* strain MSR16 and MSR17. (MSR16a and MSR17a are replicates of MSR16 and MSR17, respectively)

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 has 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 major virulence factors of *V. parahaemolyticus* are thermostable direct hemolysin (*tdh*) [17], TDH-related hemolysin (*trh*) [18] and two type III secretion systems (T3SS1 and T3SS2) [19]. *tdh* and *trh* both genes were not found in these strains but the thermolabile hemolysin (*tlh*) gene was found. Between two types of T3SS, only the T3SS1 type was found in MSR16 and MSR17 strain. Both strains possess the plasmid-borne *pir*A and *pir*B toxins.
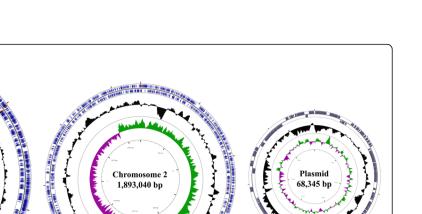
Antibiotic resistance genes were predicted against β-lactam, fluoroquinolone, tetracycline, macrolide and cephalosporin antibiotics in MSR16; and MSR17 strain

has similar resistance genes except for cephalosporin (Additional file 4). Six and two probable prophage regions were found in MSR16 and MSR17 strains, respectively.
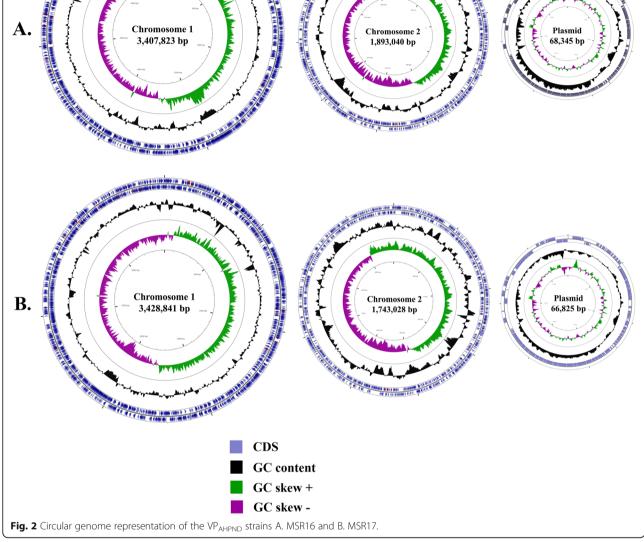
Strains MSR16 and MSR17 have approximately 39 and 27 genomic islands (GI) regions respectively (Additional file 5). In strain MSR16, toxin-antitoxin systems like YoeB-YefM, Doc-Phd; antibiotic resistance proteins like FosA (Fosfomycin resistance protein); components of type-I, type-VI secretion systems, etc. are found in those genomic islands. Genomic islands of strain MSR17 contain 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 [20] 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, as several environmental strains of *V. parahaemolyticus* were found to cause cytotoxicity to human gastrointestinal cells even in the absence of *tdh* and/or *trh* genes [21].

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

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

Ahmmed *et al. BMC Microbiology* (2019) 19:270

Page 4 of 14

Fig. 2 Circular genome representation of the VP$_{AHPND}$ strains A. MSR16 and B. MSR17.

comparison. In this phylogenetic tree, these strains were distributed in 5 major clusters (Fig. 5).

Most Chinese and Thai strains are found in cluster A. Both of our studied strains (MSR16 and MSR17) located at same cluster B and were closely related with one of the Indian strain AP1511 indicating that the mutation and evolutionary pattern of MSR16 and MSR17 strains might be analogous to this Indian strain. The two Spanish *V. parahaemolyticus* strains separately made cluster C. The strains including Vp-4 MK377081.1 China, Ramsar KJ704113.2 Iran belong to separate cluster D. Besides, two AHPND positive *V. owensii* strains were located at separate cluster E.

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 a phylogenetic relationship based on common housekeeping genes (Fig. 6) including (*dna*E, *dtd*S, *gyr*B, *pnt*A, *pyr*C, *rec*A, *tna*A). The 16S rRNA gene was not included because a separate phylogenetic relationship was established based on it. The strains M0605 Mexico, TUMSAT-H10-S6 Thailand, NCKU-TV-3HP Thailand, MSR17 Bangladesh, M1−1 Vietnam, MVP3 Malaysia

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	1,892,806	1,742,619
Total length	5,393,740	5,241,592
GC (%)	45.09	45.19
CDS	4909	4689
Gene	5090	4854
tRNA	119	109
misc_RNA	51	45
rRNA	10	10
tmRNA	1	1

and VP14 India strain located at same cluster B (Fig. 6). The strains 12-009A/1335 Vietnam, MSR16 Bangladesh, 13–028-A2 Vietnam, and NA9 Malaysia strains located at the same cluster C (Fig. 6).

The phylogenetic tree showed that the MSR16 strain was closely related to the 12-009A/1335 Vietnam strain which maintains an antibacterial type VI secretion system with versatile effector repertoires [22] suggesting that MSR16 strain may have originated from Vietnam. MSR17 strain was closely related to the M1–1 Vietnam strain signifying that MSR17 strain might evolve from M1–1 Vietnamese strain. Kumar et al. (2018) reported that M1–1 strain causes a mild form of shrimp AHPND infection [23]. Compared to other virulent strains, the M1–1 genome was reported to have gained a few additional genes and lost several other genes, which may have resulted in the reduced virulence of this strain [23].

The tree also shows that MSR16 strain arises earlier than MSR17 strain. NA7 Malaysia strain belonged to an independent lineage and distantly related to our studied strains (MSR16 and MSR17) signifying that features from this strain might be dispersed to MSR16 and MSR17 strains.

ANI (average nucleotide identity) tree of different AHPND positive *V. parahaemolyticus* strain

A total of 52 genomes of AHPND positive *V. parahaemolyticus* strain including MSR16 and MSR17 were selected for calculating the average nucleotide identity (ANI) (Fig. 7). The ANI tree clearly shows that MSR16 strain belonged to an independent lineage and indicating this strain may have evolved earlier than MSR17. The reason for belonging to an independent lineage might be the presence of an extra ~ 200 kb sequence 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 as well as some Thai strains located in cluster B (Fig. 7).

The strains ND11 Malaysia and ND13 Malaysia belonged to an independent lineage as well as distantly related to our studied strain (MSR16, MSR17) indicating that these two strains' genome sequence might be dispersed to MSR16 and MSR17 strain.
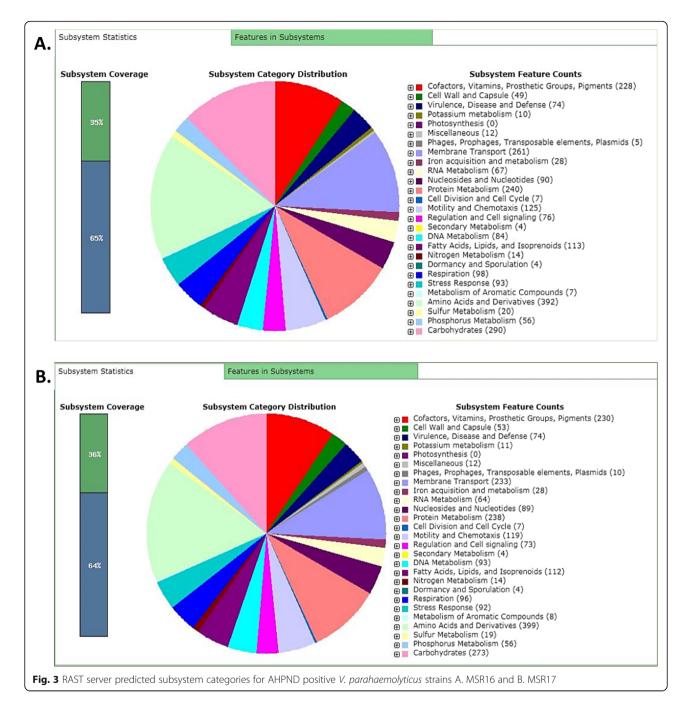
SNP tree of different AHPND positive *V. parahaemolyticus*

A total of 37 genomes of AHPND positive *V. parahaemolyticus* strain including MSR16 and MSR17 were selected for establishing a SNP based relationship (Fig. 8). The strains MSR16 Bangladesh and NA9 Malaysia were closely related and located at the same cluster C (Fig. 8) 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 the same cluster C (Fig. 8) indicating that the mutation and evolutionary pattern of MSR17 might be analogous to the Mexican strain. Five iron acquisition systems (hemin, enterobactin, vibrioferrin, and two TonB), 7 secretion systems (two T2SS, one T3SS, two T2/4SS, and two T6SS) and 14 different toxin genes that are involved in the pathogenicity mechanisms were found in both chromosomes of *V. parahaemolyticus* strain M0605 [11]. Gomez-Gil et al. (2014) also detected four plasmids in the M0605 strain's genome [11]. Strain TUMSAT-H03-S5 Thailand strain belonged to an independent lineage. This strain's mutation and evolutionary pattern might disperse to MSR16 and MSR17.

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 the phylogenetic relationship among the AHPND positive *V. parahaemolyticus* plasmid (Fig. 9). Six plasmids including pMSR16 Bangladesh, pVPA3–1 Vietnam, pMSR17 Bangladesh, pVpR13-71Kb USA, pVPGX1 China, pVPE61a Thailand, were located at Cluster A (Fig. 9).
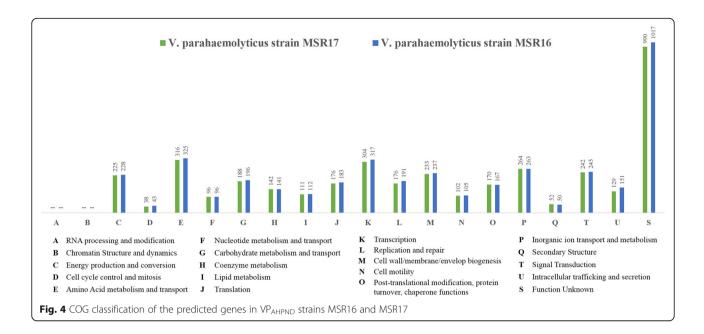
The phylogenetic tree showed that pMSR16 and pMSR17 were closely related to pVPA3–1. The plasmid pVPA3–1 is a Vietnam strain and its accession no. is NC_025152.1. Han et al. (2015) reported that AHPND positive *V. parahaemolyticus* strain 13–028/A3 [9] possess this 69 kb plasmid pVPA3–1 which has 92 open reading frames that encode mobilization proteins, replication enzymes, transposases, virulence-associated proteins, and proteins similar to Photorhabdus insect-related (*Pir*) toxins. The plasmid pV110-KY498540.1

Fig. 3 RAST server predicted subsystem categories for AHPND positive *V. parahaemolyticus* strains A. MSR16 and B. MSR17
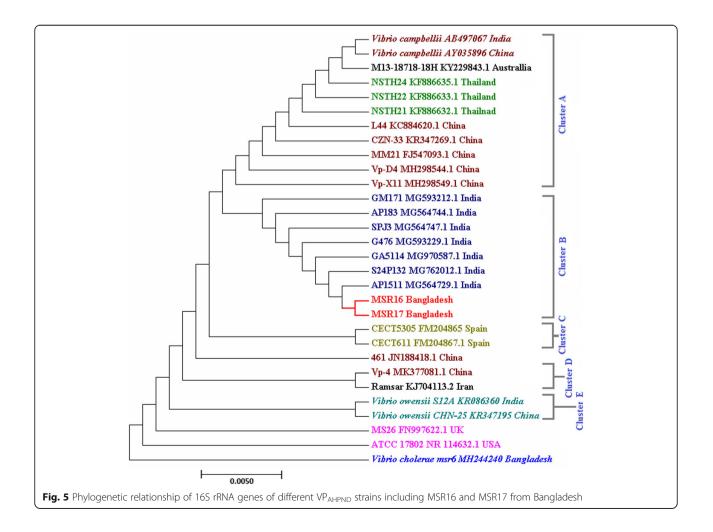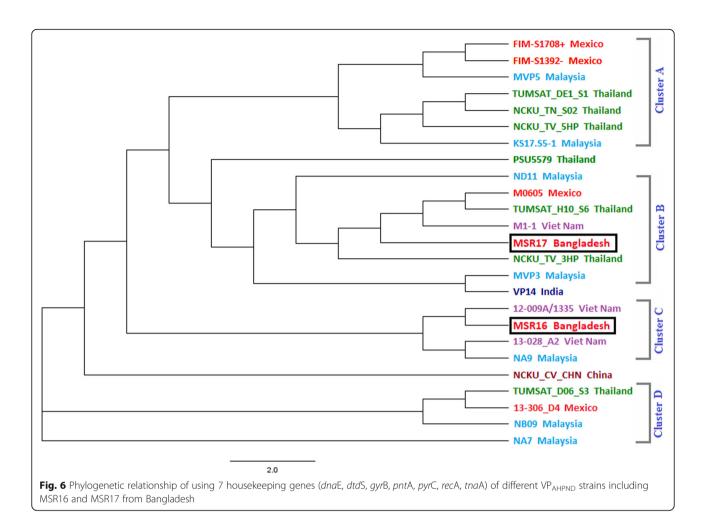
China and p1937–1-NZCP022245.1 China belonged to an independent lineage respectively. These two strains might acquire plasmids from different sources. These two plasmids are also distantly related to our studied plasmids pMSR16 and pMSR17 indicating that they might not have originated from Chinese strains.

Discussion

Penaeus monodon is one of the most important shrimp species that has widely been used for farming in many tropical countries. With the intensification of

shrimp farming worldwide, new pathogens are seen to emerge frequently. A variety of microorganisms, such as the White Spot Syndrome Virus (WSSV), *Vibrio* spp. and Taura Syndrome Virus (TSV) has been constantly posing devastating threats to the sustainability of the shrimp farming industry over the years [24]. Since the very first outbreak occurred in China in 2009, because of its rapid deadly effects in the early developmental stages of shrimp, the AHPND/EMS has spread throughout the globe and caused huge economic losses [25].

Fig. 4 COG classification of the predicted genes in VP_AHPND strains MSR16 and MSR17

Fig. 5 Phylogenetic relationship of 16S rRNA genes of different VP_AHPND strains including MSR16 and MSR17 from Bangladesh

Fig. 6 Phylogenetic relationship of using 7 housekeeping genes (*dna*E, *dtd*S, *gyr*B, *pnt*A, *pyr*C, *rec*A, *tna*A) of different VP$_{AHPND}$ strains including MSR16 and MSR17 from Bangladesh

Recent studies showed that various AHPND positive *V. parahaemolyticus* strains possess the Photorhabdus insect-related (*Pir*) toxin-like genes and these genes (*pir*A and *pir*B like) are likely to be the primary virulence factor in these strains [9]. These gene products are found to be crucial in developing the AHPND in cultured shrimp [26]. The disease has spread and caused major economic losses in Asia as well as in the Americas, and most recently in Texas, USA [27–29]. The *pir*AB region in *V. parahaemolyticus* R13 and R14 strains is encoded on the pVpR13-71Kb and pVpR14-74Kb plasmids, respectively [26, 30]. In *V. parahaemolyticus* R13 strain, the promoter region upstream of *pir*A, the entire open reading frame (ORF) of *pir*A, and part of the 5′ end of the *pir*B ORF were absent and this strain was found avirulent [26]. A particular *V. parahaemolyticus* strain's capability of causing AHPND depends on the presence of plasmid-borne binary toxins *Pir*AVp and *Pir*BVp [8].

V. parahaemolyticus is mostly found in aquatic environments, like- sediments, plankton, and aquatic animals [25]. From the phylogenetic analysis, it has been found

out that all AHPND related isolates could be undoubtedly segregated into distinct clusters, where each cluster is specific for a distinctive region [26].

In this study, the genomes of both MSR16 and MSR17 strains contain a plasmid of ~ 69 Kbp. The plasmid of MSR16 contains a total of 87 genes and MSR17 contains a total of 88 genes. Both of the plasmids carry *Pir*A and *Pir*B genes which are responsible for AHPND disease. The length of *pir*A gene was 336 bp (starts at 64,962 bp and stops at 65,297 bp) and the length of *pir*B gene was 1317 bp (starts at 65,310 bp and stops at 66,626 bp) in MSR16. On the other hand, the length of *pir*A gene was 336 bp (starts at 63,108 bp and stops at 63,443 bp) and the length of *pir*B gene was 1317 bp (starts at 63,456 bp and stops at 64,772 bp) in MSR17.
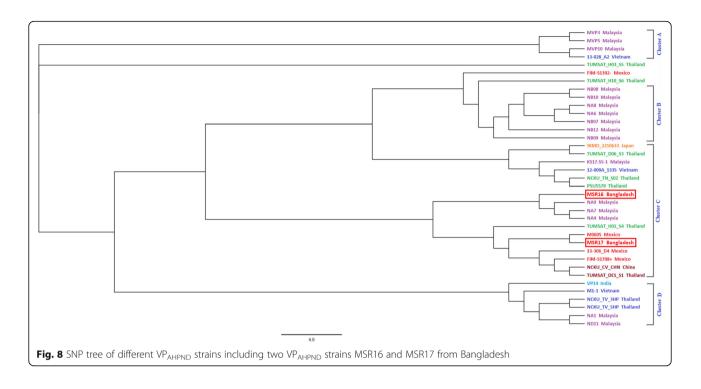
Antibiotic resistance mechanisms can be transmitted from resistant bacteria to other bacteria through the exchange of its naturally occurring resistance genes [31]. It can be observed that both strains possess resistance genes for efflux mechanisms and antibiotic modification which supports our previously reported antibiogram data [6, 32]. While comparing to the

Fig. 7 ANI tree of different VP$_{AHPND}$ strains including two VP$_{AHPND}$ strains MSR16 and MSR17 from Bangladesh

Fig. 8 SNP tree of different VP$_{AHPND}$ strains including two VP$_{AHPND}$ strains MSR16 and MSR17 from Bangladesh

experimental results, the presence of several predicted resistance genes against some antibiotics was found in which they are currently either sensitive or have an intermediate response, which means these strains can gain significant antibiotic resistance in the nearest future.

In the present study, the phylogenetic analysis of MSR16 and MSR17 strain was done in several ways such as the construction of trees based on the 16S rRNA gene, common housekeeping genes excluding 16S rRNA, whole plasmid sequences, SNP and average nucleotide identity (ANI).

Fig. 9 Phylogenetic relationship based on plasmid sequences from VP$_{AHPND}$ isolates including two VP$_{AHPND}$ strains plasmid pMSR16 and pMSR17 from Bangladesh

The phylogenetic tree based on 16S rRNA showed our studied strains MSR16 and MSR17 located at the same cluster and were closely related with one of the 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 the M1–1 Vietnam strain signifying that evolution of both strains might be from Vietnam.

Analyzing the SNP tree, we have found that MSR16 strain is closely related to three Malaysian strains indicating that the mutation and evolutionary pattern of MSR16 might be comparable to these Malaysian strains. On the other hand, MSR17 strains are closely related to the M0605 Mexico strain indicating that the mutation and evolutionary pattern of MSR17 might be analogous to the Mexican strain.

The ANI tree depicts a diversified pattern for strain MSR16 as it was found occupied in an independent lineage, whereas strain MSR17 was found closely related to Mexican strain 13–306-D4.

The plasmid sequence based phylogenetic tree showed that pMSR16 and pMSR17 were closely related to pVPA3–1. The plasmid pVPA3–1 is a Vietnam strain and its accession no is NC-025152.1. Both plasmids carry the causative agent *pir*A and *pir*B gene of AHPND in their sequence. The plasmid of two studied strains might have 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 ~ 5.4 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 the human food chain to prevent potential health hazards.

Methods

Culturing of *V. parahaemolyticus* strains and molecular identification

V. parahaemolyticus strains from the previous study (MSR16 and MSR17 strains were isolated from the infected shrimps collected from Morrelganj and Rampal upazila, Bagerhat district of Bangladesh, respectively) [6, 32] were inoculated in Tryptone soy 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) was 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 [33] using several parameters; ii) Generation of QC reports of trimmed and untrimmed data using FastQC v0.11.7 [34]; iii) Genome Assembly using the trimmed and untrimmed data by SPAdes v3.10 [35] in both general and plasmid mode utilizing different k-mer combinations; iv) Assembly polishing using Pilon v1.22 [36]; v) Determination of the quality and coverage of the assemblies using Quast v5.0.2 [37]; vi) Scaffolding into chromosomes and plasmids by MeDuSa v1.6 [38]. Genomic scaffolds of these two strains were compared using Mauve 2.4.0 [39].

Gene prediction and annotation

Genome annotation was performed by Prokka v1.12 [40], Glimmer v3.02 [41], RASTtk v1.3.0 [42] and tRNA, rRNA annotation was done using Barrnap v0.6, tRNAscan-SE v2.0 [43]. Average nucleotide identity of 50 different AHPND causing *V. parahaemolyticus* strains were calculated using Pyani v0.2.7 [44].

Analysis of virulence, antibiotic resistance genes, and others

Virulence factors were searched using VFanalyzer [45]. ResFinder [46], ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation) [47] and CARD tools [48] were used to search antibiotic resistance genes. Prophage sequences were searched respectively by PHASTER [49]. COG (Clusters of orthologous groups) classification of the genes was achieved by eggNOG-mapper v1 [50]. Genomic islands were predicted using the Islandviewer tool [51].

Phylogenetic analysis and genome comparison

A 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 [52]. The evolutionary history was inferred by using the Maximum Likelihood method with 100 bootstraps based on the best model fit for this dataset analyzed by MEGA, in this case, the Tamura-Nei (TN) [53] model which had the lowest BIC score.

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 with 100 bootstraps based on Kimura two-parameter (K2P) substitution model was constructed using MEGA 7.0 software [52].

SNP based NJ-tree was constructed with genomes of 37 different AHPND positive *V. parahaemolyticus* strains using the Parsnp v1.2 tool [54]. Average nucleotide identity (ANI) based tree was constructed from our generated ANI value matrix (Additional file 1) using PHYLIP package [55]. Twenty-six plasmids from several AHPND positive *V. parahaemolyticus* strains were aligned and a NJ-tree was constructed with 50 bootstraps using MAFFT v7 [56] utilizing the Jukes-Cantor (JK) substitution model.

Quality assurance

16S rRNA genes of *V. parahaemolyticus* strain MSR16 and MSR17 were predicted from the annotation pipeline and also from the BLAST [57] search of the PCR amplified partial sequences of both strains' 16S rRNA genes.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12866-019-1655-8.

Additional file 1. Matrix representation of ANI values of strains MSR16 and MSR17 compared to other AHPND positive *V. parahaemolyticus* strains.

Additional file 2. Unique genes predicted in VP_{AHPND} strains MSR16 and MSR17.

Additional file 3. Virulence property of VP_{AHPND} strains MSR16 and MSR17 predicted by VFAnalyzer.

Additional file 4. Predicted antibiotic resistance genes in MSR16 and MSR17 strains of VP_{AHPND}.

Additional file 5. Number of Genomic islands (GIs) predicted for AHPND positive *V. parahaemolyticus* strains MSR16 and MSR17.

Additional file 6. List of genes found in the Genomic islands (GIs) of VP_{AHPND} strains MSR16 and MSR17.

Author's contribution
MSR and ABMMKI conceived and designed the project. SA and MA-A-KK performed the culture, sequence assembly, annotation, and analyses. MMEE and NJP collected shrimp samples, isolated and identified the AHPND positive *V. parahaemolyticus*. SA, MA-A-KK, MMEE, NJP, MSR, and ABMMKI wrote the manuscript. SA and MA-A-KK contributed equally. All authors read and approved the final 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. J Fish Dis. 2017;40(5):649–59.
2. Lightner DV, Redman R, Pantoja C, Noble B, Tran L. Early mortality syndrome affects shrimp in Asia. Glob Aquaculture Advocate. 2012;15(1):40.
3. Tran L, Nunan L, Redman RM, Mohney LL, Pantoja CR, Fitzsimmons K, et al. Determination of the infectious nature of the agent of acute hepatopancreatic necrosis syndrome affecting penaeid shrimp. Dis Aquat Org. 2013;105(1):45–55.
4. Kondo H, Tinwongger S, Proespraiwong P, Mavichak R, Unajak S, Nozaki R, et al. Draft genome sequences of six strains of *Vibrio parahaemolyticus* isolated from early mortality syndrome/acute hepatopancreatic necrosis disease shrimp in Thailand. Genome Announc. 2014;2(2):e00221–14.

Ahmmed *et al. BMC Microbiology*   (2019) 19:270

Page 13 of 14

5. Fao A. 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. Available: http://www.fao.org/3/i3422e/i3422e00.htm; Accessed: 17 May, 2019.

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. J Bangladesh Acad Sci. 2017;41(2):127–35.

7. McCarter L. The multiple identities of *Vibrio parahaemolyticus*. J Mol Microbiol Biotechnol. 1999;1(1):51–7.

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. Proc Natl Acad Sci. 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. Dis Aquat Org. 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. Infect Immun. 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, Srituyalucksana 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. p. 7–9. Available: https://enaca.org/?id=96; Accessed: 17 May, 2019

14. Dangtip S, Sirikharin R, Sanguanrut P, Thitamadee S, Srituyalucksana K, Taengchaiyaphum S, et al. AP4 method for two-tube nested PCR detection of AHPND isolates of *Vibrio parahaemolyticus*. Aquaculture Rep. 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. Front Microbiol. 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 Genet. 2015;11(3): e1005072.

17. Nishibuchi M, Fasano A, Russell R, Kaper J. Enterotoxigenicity of *Vibrio parahaemolyticus* with and without genes encoding thermostable direct hemolysin. Infect Immun. 1992;60(9):3539–45.

18. Honda T, Ni Y, Miwatani T. Purification and characterization of a hemolysin produced by a clinical isolate of Kanagawa phenomenon-negative *Vibrio parahaemolyticus* and related to the thermostable direct hemolysin. Infect Immun. 1988;56(4):961–5.

19. Makino K, Oshima K, Kurokawa K, Yokoyama K, Uda T, Tagomori K, et al. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of V cholerae. Lancet. 2003;361(9359):743–9.

20. 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.

21. Raghunath P. Roles of thermostable direct hemolysin (TDH) and TDH-related hemolysin (TRH) in *Vibrio parahaemolyticus*. Front Microbiol. 2015;5:805.

22. 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.

23. 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.

24. Flegel TW. Historic emergence, impact and current status of shrimp pathogens in Asia. J Invertebr Pathol. 2012;110(2):166–73.

25. 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.

26. Kanrar S, Dhar AK. Complete genome sequence of a deletion mutant of *Vibrio parahaemolyticus* from Pacific white shrimp (Penaeus vannamei). Genome Announc. 2018;6(25):e00544–18.

27. Dhar A, Piamsomboon P, Aranguren Caro L, Kanrar S. First report of the presence of acute hepatopancreatic necrosis disease (AHPND) in Texas, USA. Las Vegas: Aquaculture America; 2018.

28. Hong X, Lu L, Xu D. Progress in research on acute hepatopancreatic necrosis disease (AHPND). Aquac Int. 2016;24(2):577–93.

29. Nunan L, Lightner D, Pantoja C, Gomez-Jimenez S. Detection of acute hepatopancreatic necrosis disease (AHPND) in Mexico. Dis Aquat Org. 2014; 111(1):81–6.

30. Kanrar S, Dhar AK. Complete genome sequence of a novel mutant strain of *Vibrio parahaemolyticus* from Pacific white shrimp (Penaeus vannamei). Genome Announc. 2018;6(24):e00497–18.

31. Bennett P. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. Br J Pharmacol. 2008;153(S1):S347–S57.

32. Eshik MME, Punom NJ, Begum MK, Khan T, Saha ML, Rahman MS. Molecular identification and characterization of *Vibrio parahaemolyticus* causing acute hepatopancreatic necrosis disease in cultured shrimp. Dhaka Univ J Biol Sci. 2018;27(1):57–68 Available: http://journal.library.du.ac.bd/index.php?journal= dujbs&page=article&op=view&path%5B%5D=1580; Accessed: 17 May, 2019.

33. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

34. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. 2017. Available: http://www.bioinformatics.babraham.ac.uk/ projects/fastqc/; Accessed: 17 May, 2019.

35. 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. J Comput Biol. 2012;19(5):455–77.

36. 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.

37. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072–5.

38. 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.

39. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004; 14(7):1394–403.

40. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9.

41. Aggarwal G, Ramaswamy R. Ab initio gene identification: prokaryote genome annotation with GeneScan and GLIMMER. J Biosci. 2002;27(1):7–14.

42. 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. Sci Rep. 2015;5:8365.

43. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25(5): 955–64.

44. Pritchard L. pyani: Python module for average nucleotide identity analyses. 2017. Available: https://github.com/widdowquinn/pyani; Accessed: 17 May, 2019.

45. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 2018;47(D1):D687–D92.

46. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67(11):2640–4.

47. 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. Antimicrob Agents Chemother. 2014;58(1):212.

48. 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 Res. 2016;45(D1):D566–D73.

49. 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 Res. 2016; 44(W1):W16–21.
50. 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. Mol Biol Evol. 2017;34(8):2115–22.
51. 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.
52. Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform. 2008; 9(4):299–306.
53. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993;10(3):512–26. https://doi.org/10.1093/oxfordjournals.molbev. a040023.
54. 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 Biol. 2014;15(11):524.
55. Felsenstein J. PHYLIP (phylogeny inference package), version 3.5 c: Joseph Felsenstein.; 1993. Available: http://evolution.genetics.washington.edu/ phylip.html; Accessed: 16 May, 2019.
56. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30(4):772–80.
57. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic acids research. 2008;36(Web Server issue):W5-W9. DOI: https://doi.org/10.1093/nar/gkn201

Publisher's Note