RESEARCH ARTICLE

Open Access

CrossMark

Comparative genome analysis provides deep insights into *Aeromonas hydrophila* taxonomy and virulence-related factors

Furqan Awan[1], Yuhao Dong[1], Jin Liu[1], Nannan Wang[1], Muhammad Hassan Mushtaq[2], Chengping Lu[1] and Yongjie Liu[1*]

Abstract

Background: *Aeromonas hydrophila* is a potential zoonotic pathogen and primary fish pathogen. With overlapping characteristics, multiple isolates are often mislabelled and misclassified. Moreover, the potential pathogenic factors among the publicly available genomes in *A. hydrophila* strains of different origins have not yet been investigated.

Results: To identify the valid strains of *A. hydrophila* and their pathogenic factors, we performed a pan-genomic study. It revealed that there were 13 mislabelled strains and 49 valid strains that were further verified by Average nucleotide identity (ANI), digital DNA-DNA hybridization (dDDH) and in silico multiple locus strain typing (MLST). Multiple numbers of phages were detected among the strains and among them Aeromonas phi 018 was frequently present. The diversity in type III secretion system (T3SS) and conservation of type II and type VI secretion systems (T2SS and T6SS, respectively) among all the strains are important to study for designing future strategies. The most prevalent antibiotic resistances were found to be beta-lactamase, polymyxin and colistin resistances. The comparative analyses of sequence type (ST) 251 and other ST groups revealed that there were higher numbers of virulence factors in ST-251 than in other STs group.

Conclusion: Publicly available genomes have 13 mislabelled organisms, and there are only 49 valid *A. hydrophila* strains. This valid pan-genome identifies multiple prophages that can be further utilized. Different *A. hydrophila* strains harbour multiple virulence factors and antibiotic resistance genes. Identification of such factors is important for designing future treatment regimes.

Keywords: *Aeromonas*, Pan-genome, Mislabelled strains, Prophages, Antibiotic resistance patterns, Virulence factors

Background

Aeromonas hydrophila is an important emerging zoonotic pathogen of aquatic origin [1]. In humans, *A. hydrophila* can cause gastroenteritis, necrotizing fasciitis, septicaemia and meningitis [2]. It has multifactorial virulence factors, such as type III, type IV and type VI secretion systems (T3SS, T4SS and T6SS, respectively), exotoxins and endotoxins [3]. Moreover, its ability to produce biofilms also threatens the food industry [4]. Although once considered to be susceptible to quinolones, *A. hydrophila* is developing resistance against these

antibiotics according to recent surveys [5]. Because of its ubiquitous nature, the strain-based pathogenicity and virulence of *A. hydrophila* are highly variable [6, 7]. Additionally, characteristics of various strains isolated from humans have not been extensively explored and categorized.

Aeromonas species have dynamic characteristics and are hard to classify into defined taxonomic groups [1]. This overlapping classification has caused a great deal of ambiguity among *A. hydrophila* strains [3]. For example, the strains SSU, Ah-3 and BWH65 were first characterized as *A. hydrophila*, but later bioinformatics analyses differentiated them and they were assigned to different taxa [2, 8, 9]. Traditional classification of *A. hydrophila* is based on multiple hybridization groups (HGs), 16S

---

* Correspondence: liuyongjie@njau.edu.cn
[1]Joint International Research Laboratory of Animal Health and Food Safety, College of Veterinary Medicine, Nanjing Agricultural University, Nanjing, China
Full list of author information is available at the end of the article

Awan *et al. BMC Genomics* (2018) 19:712

Page 2 of 18

rRNA sequencing and multiple locus sequence type (MLST) [1]. Further, classification of *A. hydrophila* has also been proposed on the basis of the presence of specific virulence factors, but this character is highly strain-specific and may not be reliable [3]. Whole genome sequencing (WGS) is now being employed for routine surveillance and for detection of possible outbreaks due to its low cost, less cumbersome protocols and reduced time investment [10]. Based on WGS, core genome MLST (cgMLST), genome-to-genome distance calculations based on digital DNA-DNA hybridization (dDDH) and average nucleotide identity (ANI) have been introduced to classify and identify the isolates [11, 12]. Recently, the pan-genome and core genome based on WGS have been utilized to understand the typing of isolates [13]. This technique is also being used to identify mutations and microevolution among the constantly evolving genes [14]. Identification and distribution of virulence factors, bacteriophages, and antibiotic resistance genes among geographically isolated strains are important features of WGS [15]. Application of WGS to epidemiology has created an opportunity to correctly diagnose disease caused by *A. hydrophila* and has been helpful in outbreak investigation [10].

The aim of this study was to identify the core genome and to estimate the variation within the pan-genome of publicly available *A. hydrophila* genomes from NCBI. Further, these results could be helpful in identifying mislabelled or wrongly classified strains. Additionally, patterns of virulence factors, antibiotic resistance genes and bacteriophages were identified. The results form the basis for classifying *A. hydrophila* along with determining the epidemiological significance of virulence factors.

Results

### Core and pangenome analysis of 62 strains

It was revealed that there were 13 mislabelled strains included in the pan-genome of *A. hydrophila* (Fig. 1a). The remaining 49 strains were considered as valid members of the *A. hydrophila* pan-genome. A core gene tree was constructed based on the core genome (Fig. 1a and b). Among all the strains, several recently reported distinct strains, such as 4AK4, SSU and BWH65, were also included to observe their behaviours. Two of these strains, SSU and BWH65, were previously included in *A. hydrophila* but were later classified into *Aeromonas dhakensis* and *Aeromonas caviae*, respectively. These two strains were found to be distinct compared to the other strains, hence forming a distinct phylogenetic group as seen in Fig. 1a and b. To identify and verify this core genome-based tree, in silico MLST against the *Aeromonas* pubMLST database was performed by uploading the genomes to the CGE webserver. However, in silico MLST could identify some ST groups, while most of the

genomes remained as unknown or nontypeable. For further identification, the ANI percentage and digital DNA-DNA hybridization (dDDH) were determined in comparison to the *A. hydrophila* ATCC7966 reference genome. Strains with greater than 95% ANI were considered to be the same species, while strains with dDDH values of more than 70% were considered to be same species. Figure 1a shows that ANI and dDDH results were consistent with each other and identified the strains closely related to SSU and BWH65 as distinct strains. Hence, the core genome phylogeny verified the presence of other strains in this pan-genome.

A pan-genome phylogenetic tree based on the presence or absence of genes was also constructed (Fig. 2). All the strains that were closely related to SSU and BWH65 showed the same phylogeny, as expected. There was not much difference when the pan-genome tree was compared with the core genome tree. All the distinct strains were closely related to SSU and BWH65, as they were grouped in the core genome phylogeny tree.

Characteristics of the valid 49 *A. hydrophila* genomes

After identification of the 13 mislabelled strains, the remaining 49 strains were considered as valid members of the *A. hydrophila* pan-genome (Fig. 3). These 49 strains were included in further comparative genomic analyses. A total of 217,923 genes were predicted by Prodigal software (Oak Ridge National Laboratory, USA) across all 49 genomes. There was an average of 4448 genes present in each genome. The low quality of the draft genomes could be the cause of this small overestimation. These genes were clustered into 9560 homologue gene clusters (HGCs), which constitute the pan-genome of *A. hydrophila*. A core genome of 2942 HGCs was predicted (Additional file 1). The remaining 49 strains were clustered into different clades. In the reanalysis, a pan-genome matrix was generated on the basis of the absence or presence of genes across all the strains. The phylogenetic relationship of strains did not differ much in comparison to the first analysis (Fig. 3). A comparison of the pan-genome and the core genome based on the progress of the clustering algorithm is shown in Fig. 4. With the addition of genomes, the pan-genome, found to be an open genome, continued to increase, while the core genome decreased initially but stabilized after the addition of a few genomes. Initially, complete genomes were added to avoid any unusual decrease or increase that could occur with the addition of draft genomes. Across all 49 genomes, 1254 ± 144 accessary genes were present, on average. The maximum number of accessary genes was shared by the NJ-35 genome, i.e., 1472 genes. Unique genes were found to be highly variable among the strains as there was no clear pattern. There was a total of 2981 unique

Awan *et al. BMC Genomics* (2018) 19:712

Page 3 of 18
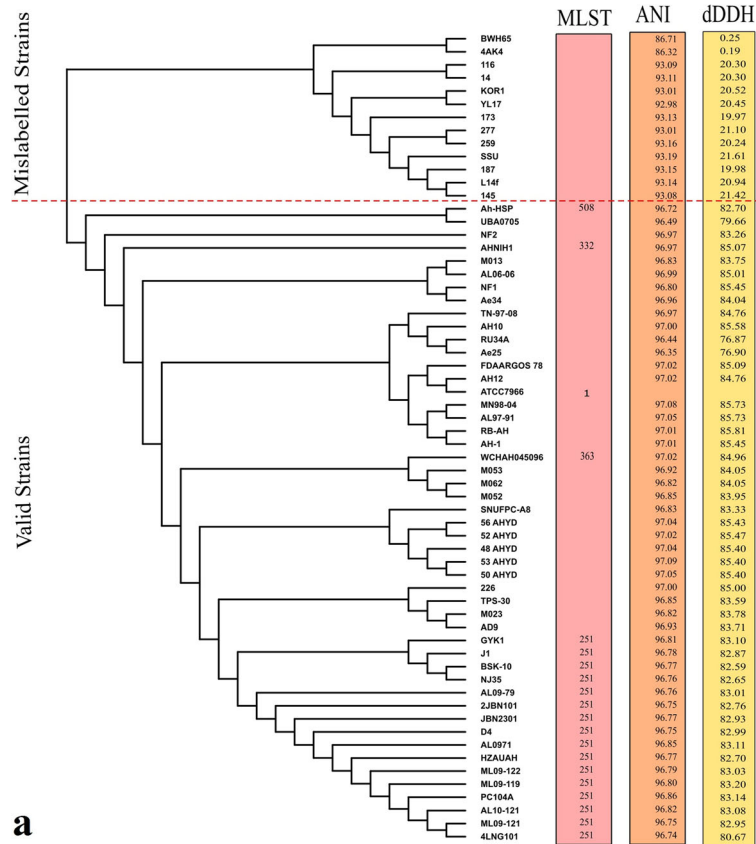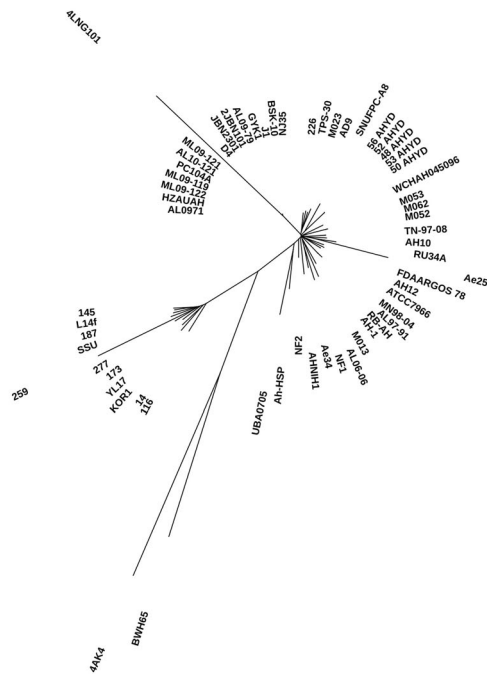
Fig. 1 (See legend on next page.)

Awan *et al. BMC Genomics* (2018) 19:712

Page 4 of 18

(See figure on previous page.)

Fig. 1 Core genome-based phylogeny: **a** The core genome-based phylogeny was verified by in silico MLST, ANI and dDDH. The red dotted line represents the cut-off line that separates the 13 mislabelled strains based on ANI and dDDH results. The remaining 49 strains are considered to be valid strains. **b** Core genome-based un-rooted phylogenetic tree. All the mislabelled strains are shown to be distinct from other valid strains. The remaining 49 strains are closely related to each other

genes present in this pan-genome. The highest number of unique genes ($n = 298$) was found in the WCHAH045096 strain.

Functional annotation

The datasets of core and dispensable genomes were extracted and locally aligned against the Cluster of Orthologous Group (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases using the Usearch algorithm. Identification of COG functional categories revealed that the core genome was enriched in the metabolism and post-transcriptional modification classes (Fig. 5), while the accessary genes were highly enriched in secondary metabolite biosynthesis and catabolism. The unique proteins were mostly found in the replication, defence mechanisms, cell wall biogenesis and cell cycle control classes (Fig. 5). In the KEGG functional annotation (Fig. 6), core genes were found to be higher among the classes of substance dependence, amino acid metabolism, cell cycle and endocrine system. Accessary genes were enriched among the transcription and metabolic diseases classes. Among the unique proteins, more proteins were found to be involved in cell signalling, cellular community and infectious diseases.

Comparison of *A. hydrophila* strains with ST-251 and other STs

The genomes of *A. hydrophila* strains with ST-251 and other STs were also compared and analysed to gain further insights. Among strains in the ST-251 group ($n = 16$), the core genome contained 4052 genes with 272 ± 48 and 289 genes in accessory and unique fractions, respectively (Additional file 2). In contrast, among strains from the other STs group ($n = 33$), the core genome was composed of 2968 genes with 1165 ± 137 and 2936 genes in accessory and unique fractions, respectively (Additional file 3). Based on these results, it is predicted that the ST-251 group has a closed pan-genome whereas the other, more diversified ST groups have an open pan-genome (Fig. 7).

COG superfamily functional categories revealed that unique fractions in ST-251 were higher in information storage processing and metabolism-related functional families than the other STs, whereas cellular processing and poorly characterized gene functional families were higher in the other STs group. The accessory gene fraction was the reverse of the above mentioned unique fraction in both categories for ST-251 and the other STs group. Within both groups, the core genome was

comparatively similar (Fig. 8). KEGG-based functionally annotated superfamilies revealed results very similar to those inferred from the COG database. The ST-251 group had a higher number of unique genes among the cellular processes in comparison to other superfamilies compared to the other STs group. The accessory gene fraction in ST-251 had more genes in environmental information processing, whereas all the remaining superfamilies, such as cellular processes, genetic information processing, metabolism, diseases and organismal systems, were dominate in the other STs group. Similar to the COG results, the core genomes of both groups were equal in superfamilies.

The comparative analysis of genome elements
Prophages

Prophages were detected using the PHAST server across all the genomes. A total of 143 prophages were detected (Fig. 9). Among these, there were 57 intact phages (completeness score was above 90), while 39 were questionable (completeness score was 60–90), and 47 were incomplete phages (completeness score was less than 60). On average, there were approximately 2 prophages present in each strain. The maximum number of phages ($n = 7$) was detected in strain 2JBN101, whereas no prophages were found in the ATCC7966 strain. Among these 143 prophages, there were 38 types of prophages present (Fig. 10). Aeromonas phi O18P was highly distributed among the *A. hydrophila* strains regardless of their ST groups. As expected, there was much more variety of phages among the other STs group compared to ST-251 group. Of greater concern is the presence of shiga toxin-converting bacteriophages, such as 933 W and Stx2. An intact enterobacter 933 W prophage was detected in Ah-10 (other STs group), whereas Stx2 vB EcoP 24B phage was found in the strains 2JBN101, D4 and HZAUAH (ST-251). All the strains carrying shiga toxin bacteriophages were isolated from fish.

Antibiotic resistance genes

More than two antibiotic resistance genes were found across all the genomes in this pan-genome (Fig. 11). Almost all the strains have one or more genes related to beta-lactamases, aminoglycosides, colistin and polymyxins, tetracycline, and chloramphenicol. Among the various types of multidrug efflux pumps found in microbial genomes, resistance nodulation cell division (RND) pumps are very important. Here, this study found that
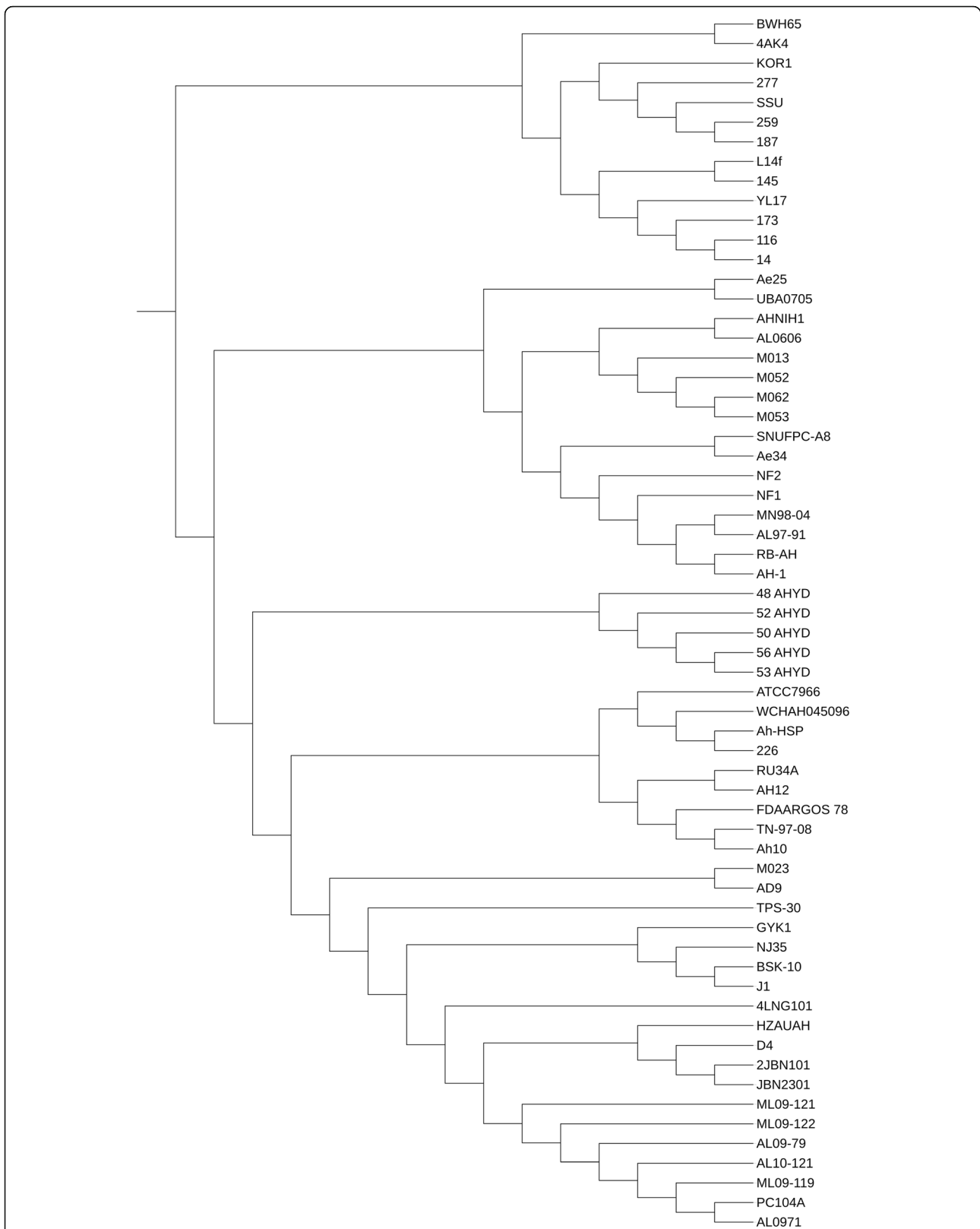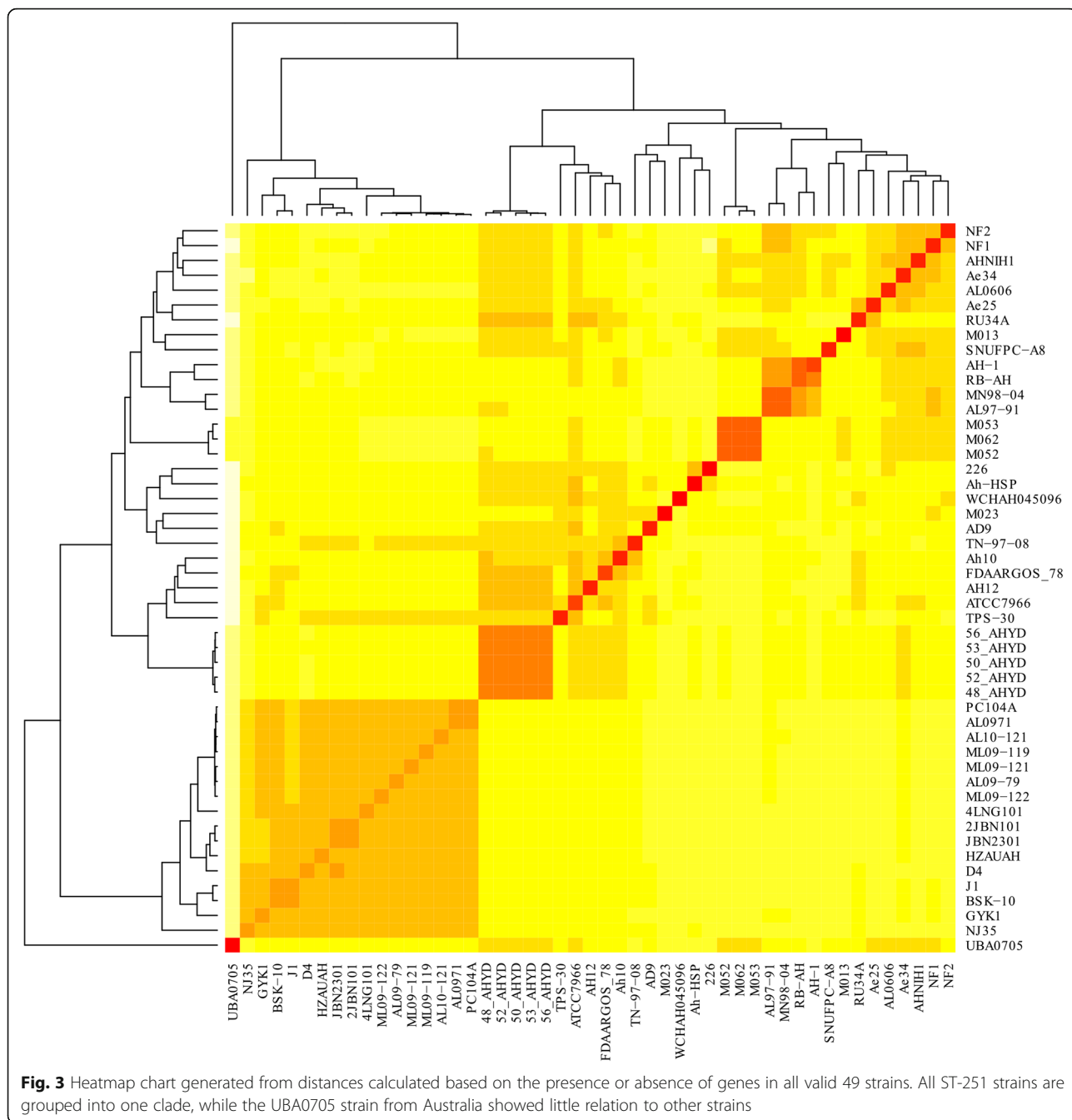
Awan *et al. BMC Genomics* (2018) 19:712

Page 5 of 18

Fig. 2 Pan-genome tree of 62 strains based on the presence or absence of genes with the mislabelled 13 strains represented as a separate group

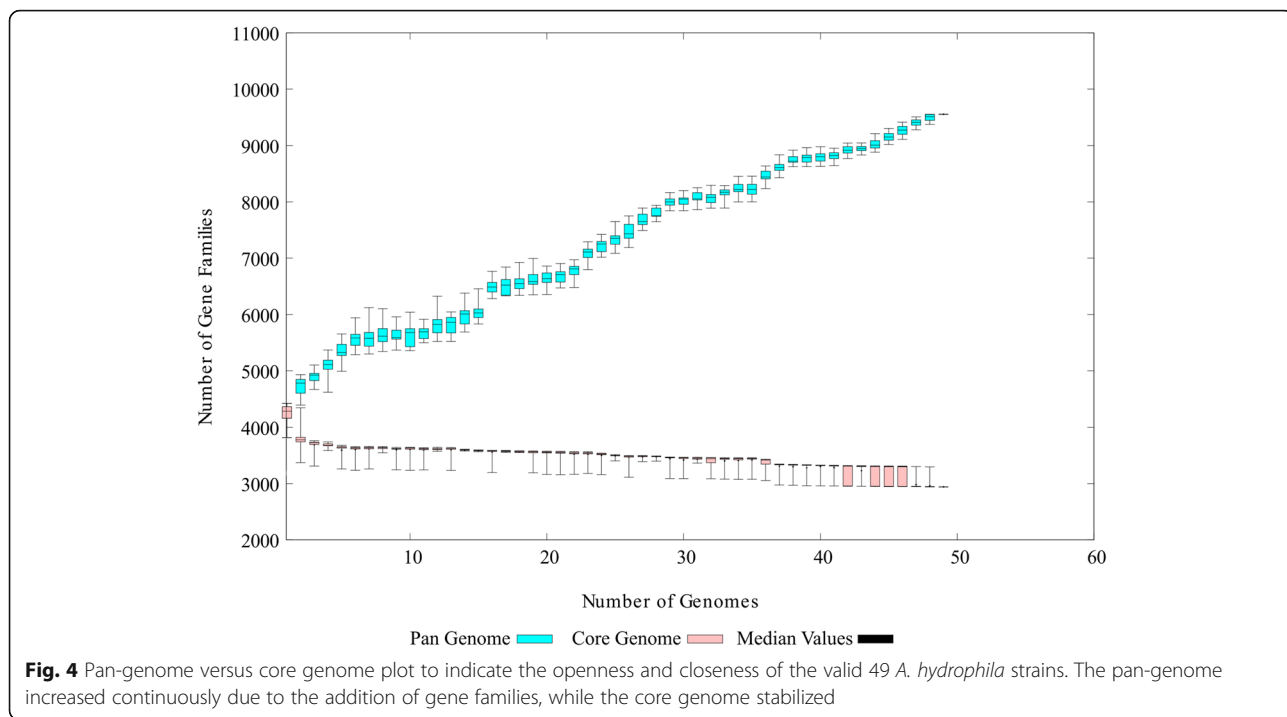Awan *et al. BMC Genomics* (2018) 19:712

Page 6 of 18

Fig. 3 Heatmap chart generated from distances calculated based on the presence or absence of genes in all valid 49 strains. All ST-251 strains are grouped into one clade, while the UBA0705 strain from Australia showed little relation to other strains

the presence of outer membrane proteins, such as OprM, OprJ, OmpK and tolC, was functionally related to antibiotic resistance. In addition, the quinolone resistance gene qnrA4 was found in almost all the strains. All the ST-251 strains had almost the same antibiotic resistance patterns, except that strains NJ35, GYK1, BSK-10 and J1 had oprM, while the other members lacked this outer membrane protein, indicating the same origin and propagating environments. In contrast, the diverse other STs group had different ranges of antibiotic resistances. Among these strains, WCHAH045096

had resistance genes from almost all the classes. This particular strain was isolated from a hospital waste drain, exhibiting clear evidence of horizontal gene transfer.
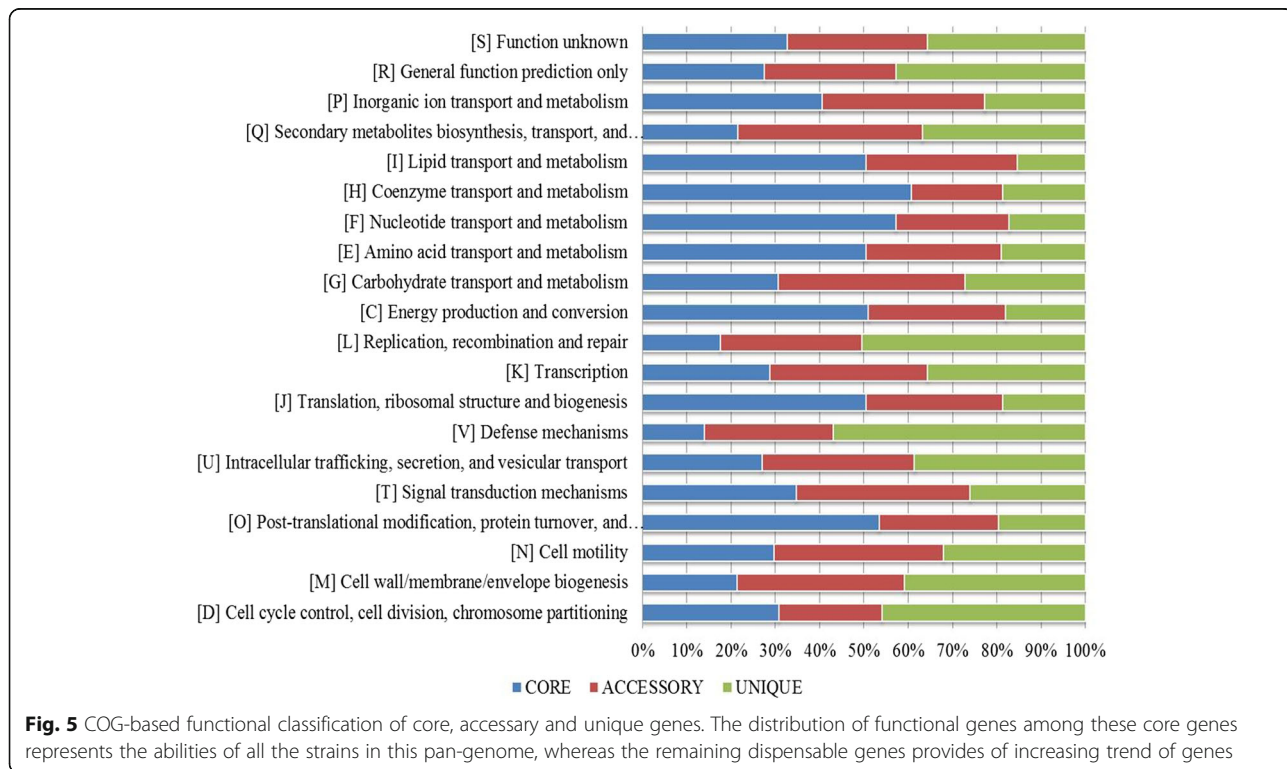
Virulence factors

The Virulence Factors database was screened in this study to determine the distribution of various virulence factors among these 49 strains (Fig. 12). The major virulence factors found among all the strains were secretion systems, the motility and adhesion system, the quorum sensing system and toxins. For secretion systems, T2SS

Awan *et al. BMC Genomics* (2018) 19:712

Page 7 of 18

Fig. 4 Pan-genome versus core genome plot to indicate the openness and closeness of the valid 49 *A. hydrophila* strains. The pan-genome increased continuously due to the addition of gene families, while the core genome stabilized

and T6SS were commonly found, while T3SS and T4SS were observed less in *A. hydrophila* for clinical isolates of both human and fish origin. The RTX toxin-producing genes were found in all those strains that were not harbouring the T3SS genes. There was a good difference in the virulence factors among the ST-251 and other ST groups. With the exceptions of T3SS, T4SS and exotoxin A, all the virulence factors investigated were present in the ST-251 group, whereas the virulence factors were more varied in the other STs group. Among the other
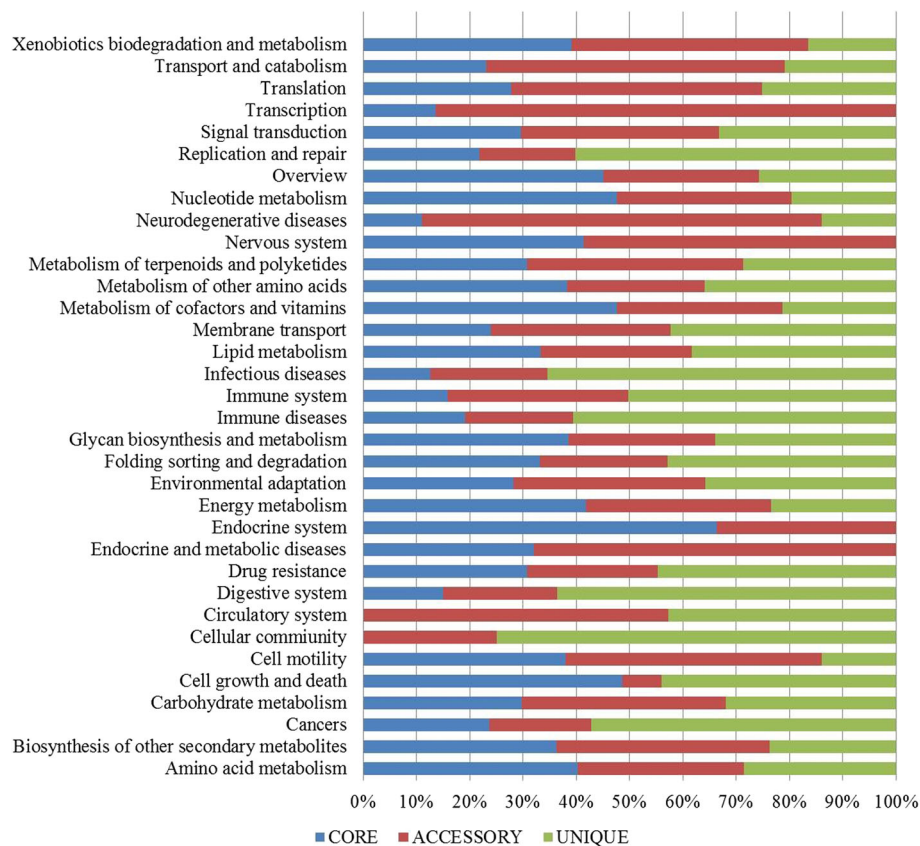

Fig. 5 COG-based functional classification of core, accessary and unique genes. The distribution of functional genes among these core genes represents the abilities of all the strains in this pan-genome, whereas the remaining dispensable genes provides of increasing trend of genes

Awan et al. BMC Genomics (2018) 19:712

Page 8 of 18

Fig. 6 KEGG-based functional annotation of all the protein datasets of the core, accessary and unique genes. Different classes of genes are found among the strains
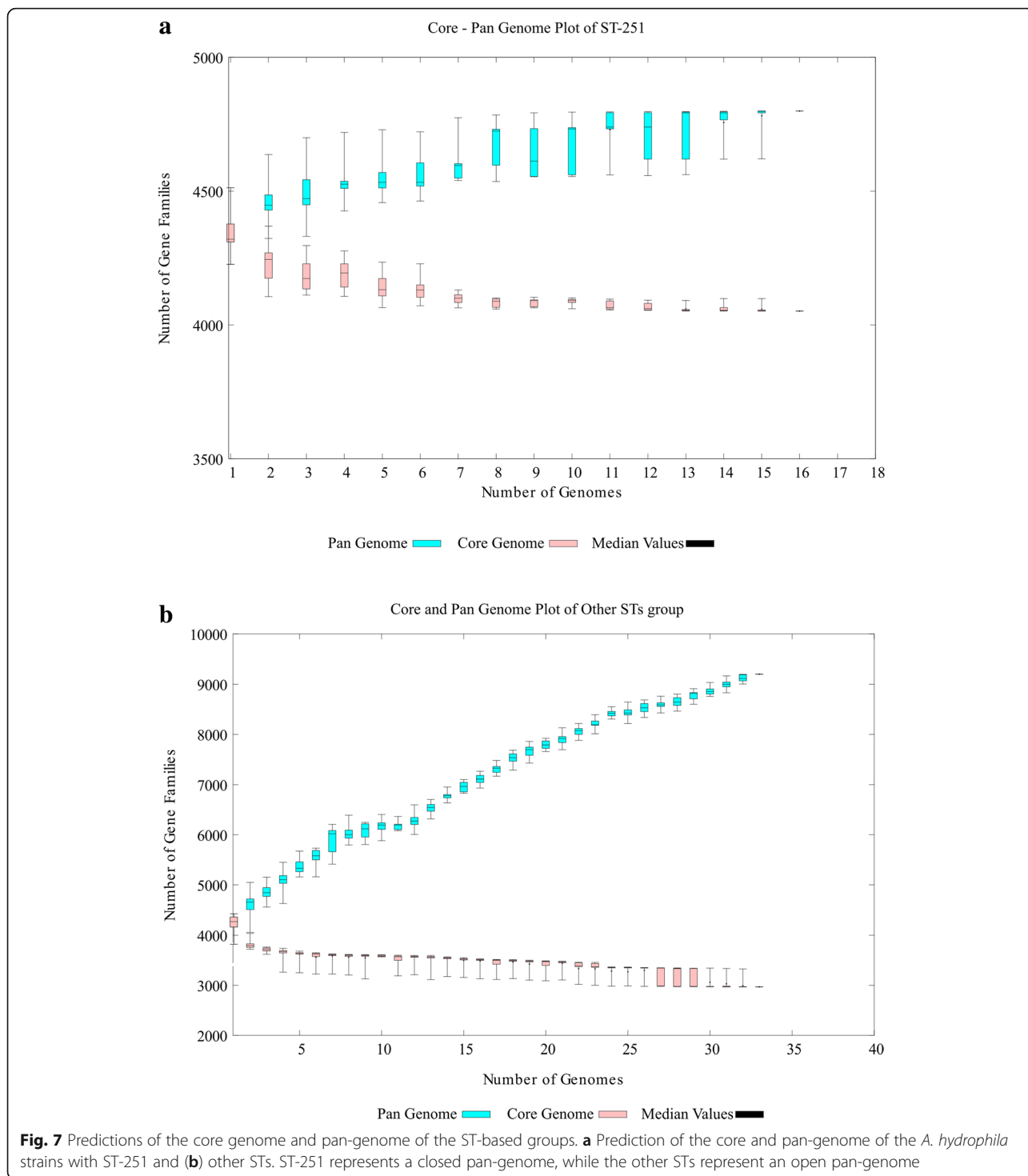
STs' strains, Flp type IV pili and aerolysin (AerA) were mostly found, while the presence of T3SS and exotoxin A showed the dynamic behaviour of this group.

Discussion

Aeromonas species are highly dynamic in nature [3]. Therefore, their classification has remained an unsolved riddle. Pathogenically important *A. hydrophila* is widely present in aquatic sources and causes disease in both human and fish [16]. Because of its global presence, there is a high number of strains reported. This distribution also led to the intraspecific variation and diverse behaviour in strains that could be due to horizontal gene transfer and mutations. Hence, ambiguous classification can lead to misclassification [8]. In this confusing situation, WGS has been helpful for allowing us to investigate the whole genome and classify them accordingly. Additionally, in silico tools have provided the opportunity to determine the differences between strains and to identify the epidemiologically important strains [10]. In this study, identification of core genes and estimation of genetic variation were performed among the 62 publicly

available *A. hydrophila* genomes. By utilizing in silico tools to compare these genomes, how these genomes fit into the presented data was discussed. Prominent results that may form the basis of future classification of *A. hydrophila* were also observed.

In this study, phylogenomic analysis demonstrated that there was wide diversification among the 62 genomes. Moreover, there are many possible distinct species that are confused with *A. hydrophila* and incorrectly classified. This study identified such discrepancies and marked the distinction between valid *A. hydrophila* and incorrectly classified species. In silico tools, such as core genome-based phylogeny, ANI variation and genome-to-genome distance calculations have been proved helpful in identifying the correct classification. The results of this study are validated, as two of the previously misclassified genomes, SSU and BWH65, were also included, and those genomes were classified separately from valid *A. hydrophila* species. In a previous study performed by Beaz-Hidalgo et al. [8], ANI has been used to discriminate the mislabelled strains, but the distinction was not clear and robust.

Awan *et al. BMC Genomics* (2018) 19:712

Page 9 of 18

Fig. 7 Predictions of the core genome and pan-genome of the ST-based groups. **a** Prediction of the core and pan-genome of the *A. hydrophila* strains with ST-251 and (**b**) other STs. ST-251 represents a closed pan-genome, while the other STs represent an open pan-genome

Core genome-based phylogeny is a relatively new proposed method of classification of genomes, which is consistent with MLST in other studies [10, 14]. As previously reported, ANI and dDDH calculations could confirm these variations [15]. In ANI variation and dDDH analyses, when compared with the reference genome ATCC7966, all the core genomes showing variation were already classified into different phylogenetic clusters.

Bacterial genomes are much more flexible in comparison to eukaryotes. Pan-genomic studies are always performed to examine the varying genetic structures present among different individual bacteria strains of a species [17]. In our study, closely related strains of the same species are also
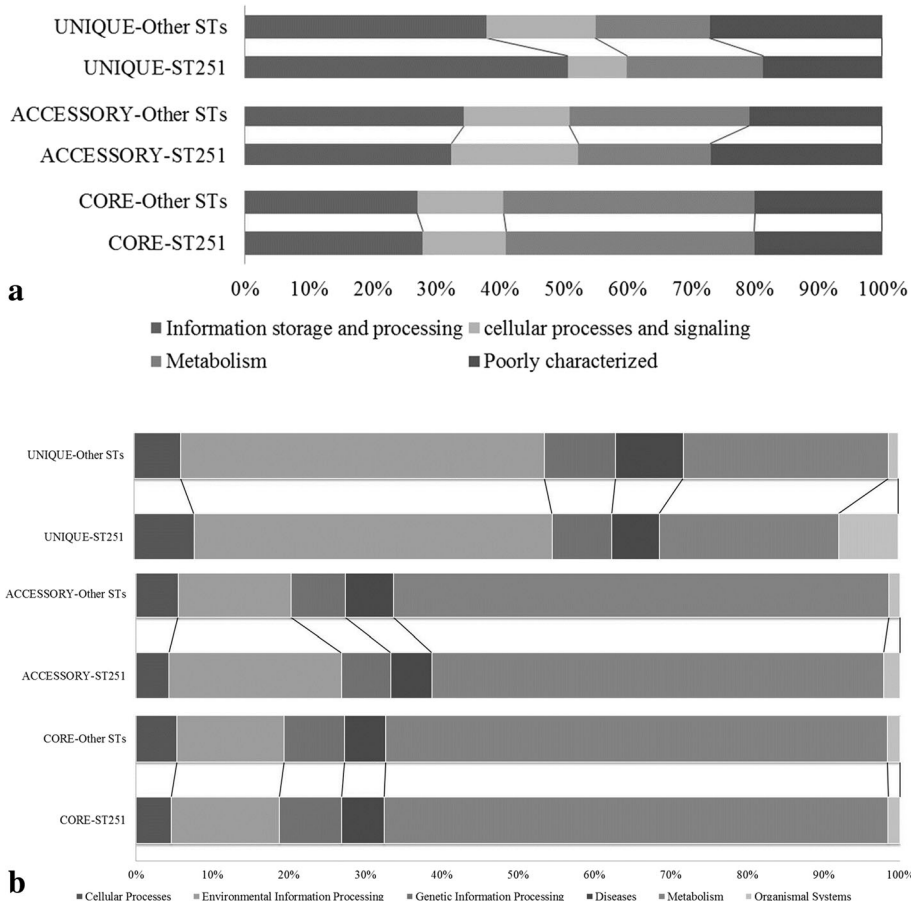
Awan *et al. BMC Genomics* (2018) 19:712

Page 10 of 18

Fig. 8 Comparison of functionally annotated proteins among the ST-251 and other ST groups. **a** Comparison of *A. hydrophila* strains with ST-251 and other STs based on COG functional categories. Unique and accessary classes are different, while there is little difference between the core genes. **b** Distribution of KEGG superfamilies among the various gene fractions of ST-251 and the other STs
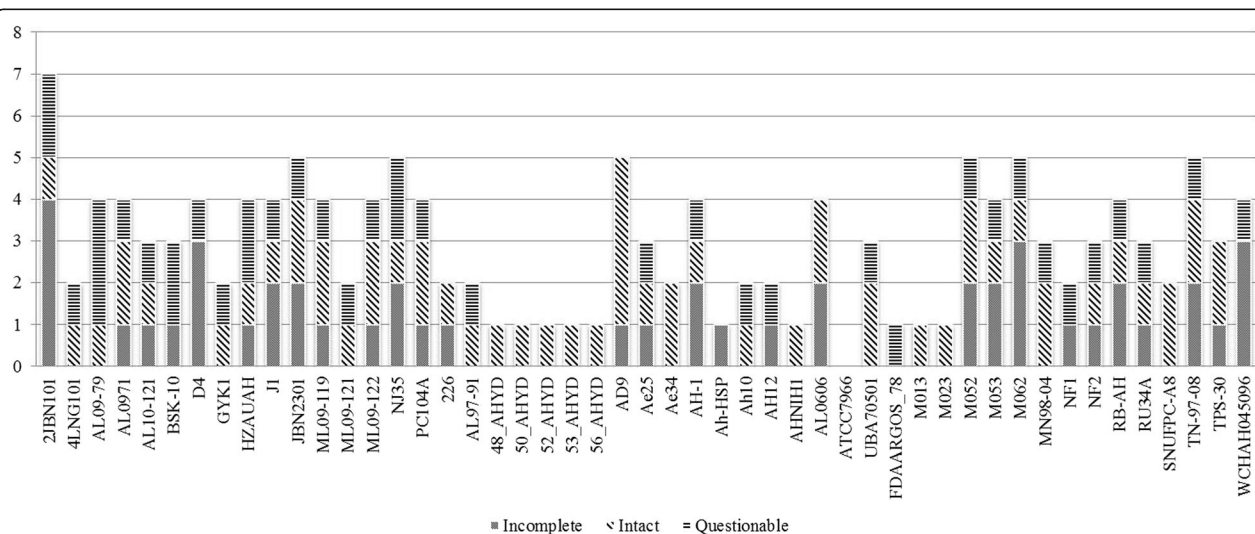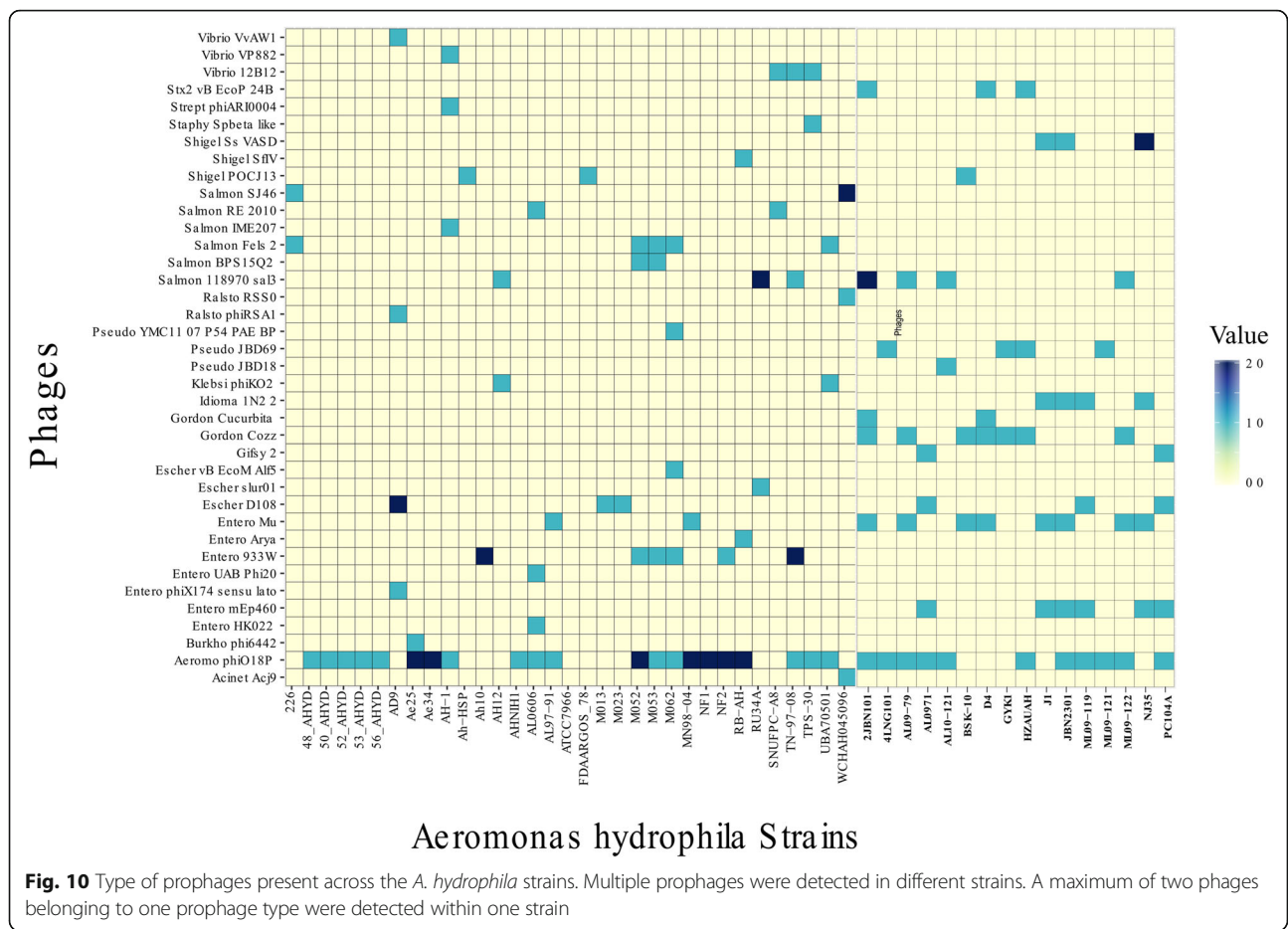

Fig. 9 Number of prophages with their completeness profiles in *A. hydrophila* strains. Only the ATCC-7966 strain did not have any prophages, while the 2JBN101 strain had 7 prophages

Awan *et al. BMC Genomics* (2018) 19:712

Page 11 of 18

Fig. 10 Type of prophages present across the *A. hydrophila* strains. Multiple prophages were detected in different strains. A maximum of two phages belonging to one prophage type were detected within one strain
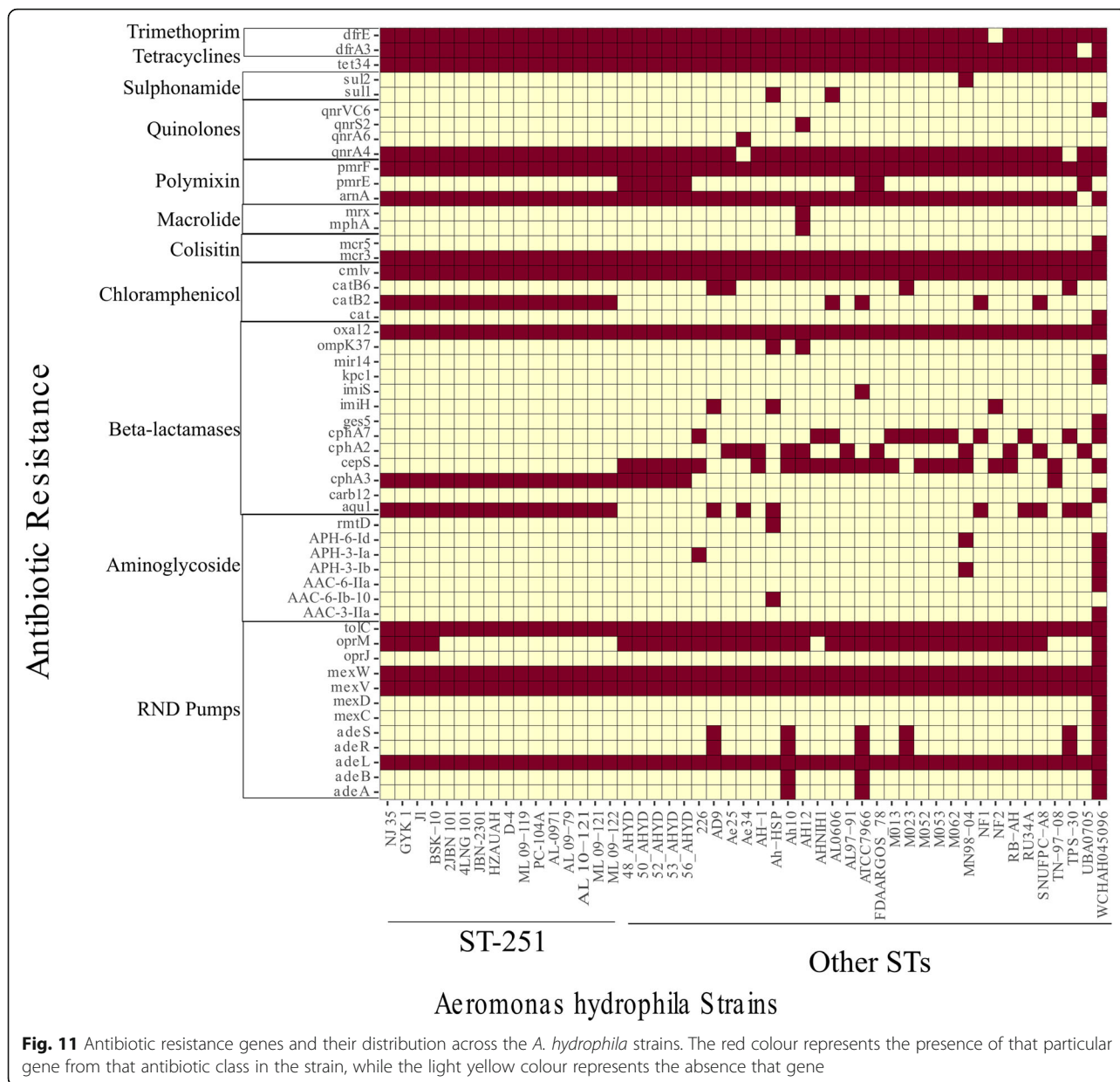
diverse in nature. The variation in terms of encoded proteins among the strains is possibly correlated to geographical isolation and distinct lifestyles, as has been reported in previous studies [18]. Genome comparison further revealed that the dispensable genome truly explained the interstrain variation. These varying fractions include a wide range of genes that could explain the impact of exposure to the environment on the various strains [19]. For instance, environments bearing antibiotic resistant bacteria may eventually transfer the resistance genes to non-resistant bacteria [20]. This scenario could explain why the strain WCHAH045096 harbours the highest number of unique genes, most of which are antibiotic resistance genes. Higher numbers of accessary genes, as in the case of strain NJ-35 in this study, also inferred the presence of shared genes that are prevalent and important for that strain [5]. Based on these fractions, varying patterns and introductions of genes could be estimated, which can be helpful in designing epidemiological strategies and in understanding the changing behaviour of *A. hydrophila*.

Codon usage and amino acid usage patterns may also show interstrain variation and can reveal the molecular basis of evolution of *A. hydrophila* [21]. Exploring these parameters may also reveal the type of natural selection pressure on genes, as found in previous studies [22]. An initial analysis of amino acid frequency of all the genes was performed using BioEdit and CodonW software packages [23, 24]. It revealed significant differences in certain amino acids, such as methionine (Met), leucine (Leu), and alanine (Ala), among the core and dispensable genes (Additional file 4), while different amino acids, such as cysteine (Cys), proline (Pro), arginine (Arg), serine (Ser) and tryptophan (Trp), were found to be significantly frequent in the ST-251 group compared to the other STs group. Although significant differences among the amino acids of different groups exist, we do not know the implications of this difference in the *A. hydrophila* pan-genome. In the future, a thorough codon usage analysis followed by multiple correspondence analyses (MCA) among members of the *A. hydrophila* pan-genome may broaden our understanding regarding selection pressure, point mutations and amino acid conservation.

The functional annotation of COG and KEGG databases to infer the functions of proteins is also of great concern [15, 25, 26], as they categorize the varying fractions into families and superfamilies. Based on both COG and KEGG database assignment results, the core

Awan *et al. BMC Genomics* (2018) 19:712

Page 12 of 18

Fig. 11 Antibiotic resistance genes and their distribution across the *A. hydrophila* strains. The red colour represents the presence of that particular gene from that antibiotic class in the strain, while the light yellow colour represents the absence that gene

genome was enriched in metabolism-related genes, while dispensable genes were found to be higher among the motility, cell communication and defence mechanisms. Surprisingly, both databases produced almost similar results. In previous studies, Bhardwaj and Somanshi [27] found the same results in pan-genomic studies, while in some studies, both databases did not show similar results because KEGG has more metabolism-related genes [25, 28]. This functional annotation can help in the overview of the distribution of proteins and their relation to interstrain variation.

In this study, all of these genomes were also locally aligned against different databases including phages,

antibiotic resistance genes and virulence factors. Phages are of prime importance in impacting bacterial evolution as they are responsible for loci rearrangements and deletions. Approximately 70% of aquatic bacteria are infected with prophages [29]. A higher number of phages are of great concern as they can turn an avirulent strain into a virulent one [30]. As an aquatic bacterium, *A. hydrophila* also harbours multiple phages in its genome. Like other bacteria, the distribution of prophages in *A. hydrophila* is not homogenous. Our study documented the presence of bacteriophages across all the genomes and quantified them. Moreover, the presence of the prophages carrying the shiga toxin gene showed that there
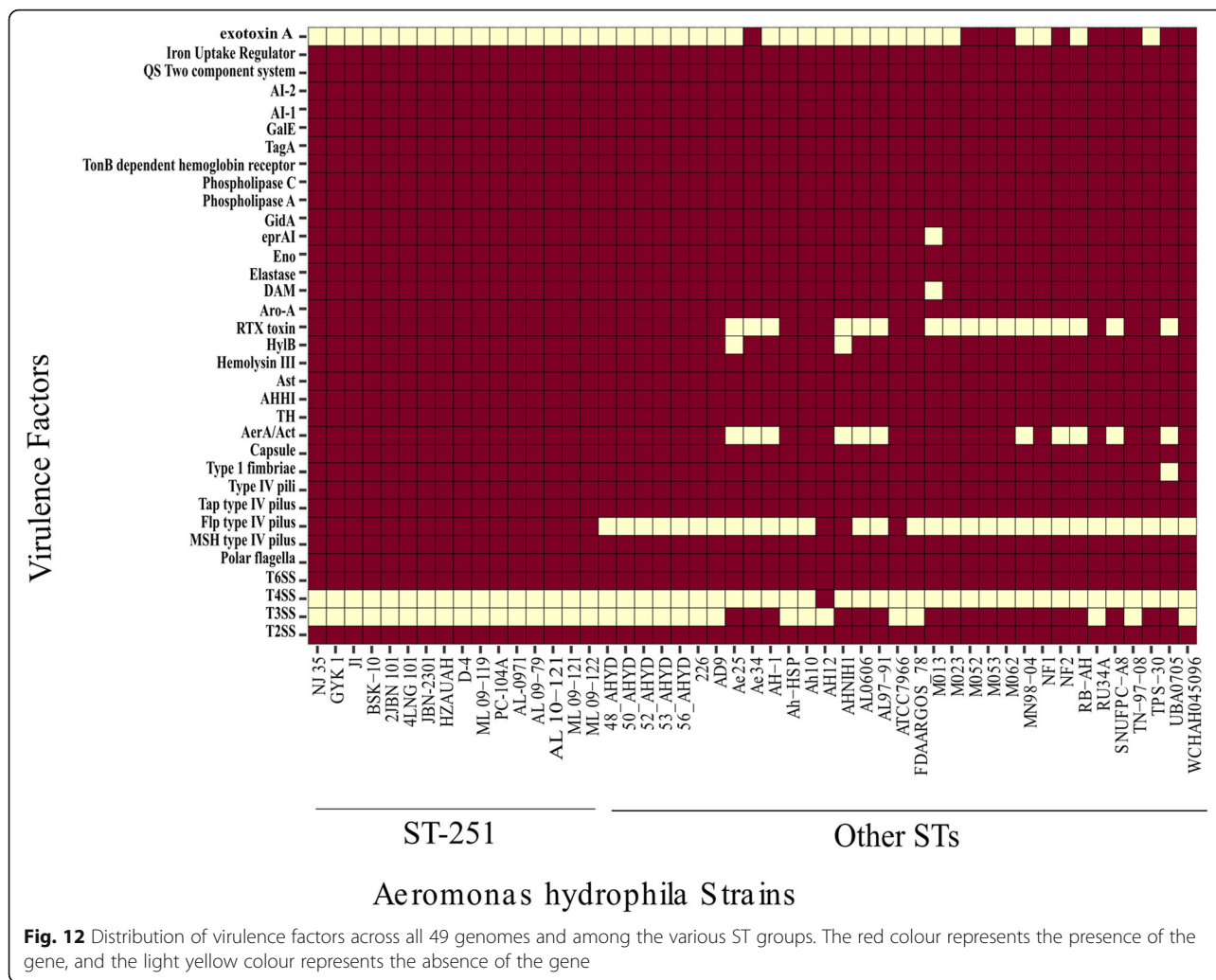
Fig. 12 Distribution of virulence factors across all 49 genomes and among the various ST groups. The red colour represents the presence of the gene, and the light yellow colour represents the absence of the gene

is a chance of possible outbreaks related to severe gastrointestinal disorders. Investigation of prophage function is beyond the scope of this study. However, it would be fascinating to elucidate a specific pattern between phage occurrence and bacterial virulence.

Mobilome-based features, such as virulence factors and antibiotic resistance genes, are very important in interstrain variation as well as pathogenicity. Moreover, virulence factors have also been proposed as the bases of classification systems. As similar bacterial species have conserved virulence factors [3], this study also showed that T2SS and T6SS were conserved among the *A. hydrophila* strains, whereas the occurrence of T3SS was variable among the strains. T3SS has a history of virulence in humans, and the data also showed that strains with T3SS were potentially zoonotic [31]. RTX toxin is also important for human gastrointestinal diseases as it can increase cell rounding and apoptosis [32]. This study also detected the presence of all the important virulence factors regardless of their genetic components. In

previous studies, the virulence factors investigated here have been found to be important for bacterial pathogenicity [3, 33].

Antibiotic resistance was extensively investigated in *A. hydrophila* strains worldwide. Beta-lactamase resistance has already been widely reported [34]. Colistin and polymyxin resistances have also been established for *Aeromonas* spp. [35, 36]. In this study, the presence of multiple resistance genes associated with the three antibiotics supported the previous reports. Additionally, RND pumps are key players in multiple drug resistance and have different types of components that efflux the several antibiotic molecules [37]. We found that several genes encoding RND efflux pumps, including *tolC*, *mexW*, *mexV* and *adeL*, are present in the chromosomes of all the strains. This implies the possibility that the *A. hydrophila* strains investigated here have acquired multiple resistances to many antimicrobials. Moreover, it should be noted that almost all the strains have quinolone resistance gene *qnrA4*, whereas the

other *qnr* genes are rare. We cannot determine whether an association between *qnrA4* and quinolone resistance exists, but it might be interesting to evaluate whether the gene contributes to quinolone resistance. There were no resistance genes found against sulphonamide and aminoglycosides. It is speculated that sulphonamides and aminoglycosides could be a possible component of treatment regimens, although further investigation is needed.

As for the ST group-based comparisons, there was more diversity between the two groups. However, the least variation was in ST-251, although the reported strains were from different continents and countries. This could be possibly owing to the movement of diseased hosts (either fish or human) or loose quarantine practices between countries, as has been previously described [38]. Among the other STs group, substantial variation in virulence factors, antibiotic resistance and phages was found. This diversity is the evidence of the changing behaviour of the bacteria across time and environmental conditions.

Conclusion

Application of comparative genomics to *A. hydrophila* provides insights into taxonomic relationships among these strains. There are also some strains that are mislabelled and should not be regarded as *A. hydrophila* for future analysis. Further, downstream analysis indicated that due to environmental conditions, this bacterium is an emerging potential zoonotic pathogen. Based on the presence or absence of genes, the calculated interstrain relationships are considerable for future analyses. Regardless of their strain typing category, a large number of prophages were found in the pan-genome, especially Aeromonas phi 018 present in multiple strains. Various virulence factors, including T3SS and T6SS, were found in this pan-genome. However, a higher number of virulence factors were detected in ST-251 compared to other STs group. Similarly, various antibiotic resistance genes from different antibiotic classes were found in this pan-genome, such as beta-lactamases, colistin, and polymyxin, indicating the multiple drug resistance ability of this bacterium. Moreover, a specific combination of virulence factors and antibiotic resistance genes might increase the bacterial virulence. Downstream analyses involving proteome analyses could reveal more insights into these factors.

Methods

Genomes, genome properties and gene predictions

All publicly available *A. hydrophila* genome sequences (65 in total: 14 full genomes, 51 draft genomes) at the time of this study were obtained from the National Center for Biotechnology Information (NCBI Resource Coordinators, 2016). The properties of the complete and draft genomes included in this study are summarized in Table 1. Due to odd behaviour in protein prediction and alignment, draft sequences of 3 strains (ARKANSAS 2010, ATCC7966–2, and CIP-107985) were not included. To differentiate two genomes with the same name, AH-1, one of the genomes was renamed to AH-12. G + C % calculations were performed using BioEdit v 7.6.2 [23]. Genome alignment in MAUVE v.20150226 was performed to refine the genome assemblies and genome scaffolding [39]. Protein predictions were performed using Prodigal software (Oak Ridge National Laboratory, USA) [40].

In silico typing, average nucleotide identity and genome-to-genome distance calculation

To identify the strain typing based on loci of housekeeping genes, in silico MLST was performed utilizing the CGE webserver (https://cge.cbs.dtu.dk/services/MLST-1.8/) [41]. As in silico MLST had very few registered loci profiles, for this purpose, bi-directional similarity estimation with the *A. hydrophila* ATCC7966 reference genome was performed utilizing the OrthoANIu tool v 1.2 to identify two-way average nucleotide identity (ANI) values [11]. To further confirm the results of MLST and ANI, dDDH was also performed utilizing the genome-to-genome calculator (http://ggdc.dsmz.de) [42]. The results of genome-to-genome distances were recorded according to recommended formula 2.

Comparative genomic analysis

All the strains based on the ANI and dDDH acceptable results were included in the pangenome (core genome and dispensable genome) analysis using the BPGA algorithm [25]. Usearch was employed for clustering the homologous genes present in the pan genome [43]. Genes shared by all the strains were considered as the core genome, while the dispensable genes either present in two or more strains (accessary proteins) or present in only one strain (unique proteins) were also identified. The core genome plot was based on plotting the total number of shared genes with each subsequent addition of a genome against the number of genomes. The pangenome plot was based on plotting the total number of distinct gene families identified with the addition of each genome and number of genomes. Obtained representative sequences of core and pan genomes were used for functional annotations and further downstream analyses.

The core genome phylogeny was constructed on the basis of conserved genes among all the strains. Core genes were aligned using MUSCLE [44]. To generate a neighbour joining tree, concatenated aligned sequences were utilized in Phylip, and MEGA v 6 was utilized to visualize the phylogenetic tree [44]. Phylogenetic trees were further smoothed using the iTOL tree website

Awan *et al. BMC Genomics* (2018) 19:712

Page 15 of 18

Table 1 All the valid 49 strains of *A. hydrophila* included in this study

Strain	Source	Country	Accession	Base pairs (Mbp)	GC%	Reference
ATCC7966	Fishy milk	USA	NC_008570	4.74	61.50	[50]
AH10	Grasscarp	China	NZ_CP011100	4.91	61.10	
AL0606	Goldfish	USA	NZ_CP010947.1	4.90	61.37	[51]
AL0971	Channel catfish	USA	NZ_CP007566	5.02	60.80	[52]
D4	Fish	China	NZ_CP013965	5.28	60.46	
GYK1	Missing	China	NZ_CP016392	4.95	60.80	
J1	Diseased carp	China	NZ_CP006883	5.00	60.90	[5]
JBN2301	Crucian Caro	China	NZ_CP013178	5.15	60.78	[53]
ML09–119	Diseased catfish	USA	NC_021290	5.02	60.80	[54]
NJ35	Diseased carp	China	NZ_CP006870	5.28	60.50	[5]
PC104A	Pond soil	USA	NZ_CP007576	5.02	60.80	[55]
AHNIH1	Human	USA	NZ_CP016380	5.05	61.07	[56]
2JBN101	Crucian Carp	China	LXME00000000	5.09	60.80	
WCHAH045096	Sewage	China	PDWA00000000	5.12	61.00	
4LNG101	Fish	China	MJGY00000000	4.99	60.10	
UBA705	Environment	Australia	DBMF00000000	4.25	61.60	[57]
48_AHYD	Human	USA	JVFM00000000	4.70	61.60	
50_AHYD	Human	USA	JVES00000000	4.67	61.60	
52_AHYD	Human	USA	JVDW00000000	4.68	61.60	
53_AHYD	Human	USA	JVDL00000000	4.67	61.60	
56_AHYD	Human	USA	JVCD00000000	4.68	61.60	
226	Human Urine	Malaysia	JEML00000000	5.11	60.90	
AD9	Wetland Sediment	USA	JFJO00000000	4.91	61.30	[58]
Ae25	Koi Carp	Sri Lanka	BEYT00000000	4.76	61.30	[59]
Ae34	Koi Carp	Sri Lanka	BAXY00000000	4.71	61.60	[60]
AH-1	Fish	Canada	LYXN00000000	4.76	61.40	
AH-1(2)	Patient	USA	LSZC00000000	5.12	60.90	
Ah-HSP	Human Blood	Brazil	MTPO00000000	5.03	61.18	
AL09–79	Catfish	USA	LRRV00000000	4.97	60.90	[61]
AL10–121	Catfish	USA	LRRW00000000	4.97	60.90	[61]
AL97–91	Tilapia	USA	LYZF00000000	4.83	61.19	[62]
BSK-10	Fish	China	NBOV00000000	4.96	61.00	[63]
FDAARGOS_78	Human Stool	USA	JTBD00000000	4.93	61.00	
HZAUAH	Missing	China	MRDF00000000	5.04	60.90	[64]
M013	WaterFall	Malaysia	JRWS00000000	4.97	61.00	
M023	WaterFall	Malaysia	JSWA00000000	4.91	60.90	[65]
M052	Water	Malaysia	MAKI00000000	4.97	61.10	
M053	Water	Malaysia	MAKJ00000000	4.96	61.10	
M062	WaterFall	Malaysia	JSXE00000000	4.97	61.10	
ML09–121	Catfish	USA	LRRX00000000	4.97	60.90	[61]
ML09–122	Catfish	USA	LRRY00000000	4.97	60.90	[61]
MN98–04	Tilapia	USA	LYZG00000000	4.88	61.10	[62]
NF1	Human	USA	JDWC00000000	4.81	61.10	[66]
NF2	Human	USA	JDWB00000000	4.79	61.30	[66]

Table 1 All the valid 49 strains of *A. hydrophila* included in this study *(Continued)*

Strain	Source	Country	Accession	Base pairs (Mbp)	GC%	Reference
RB-AH	Soil	Malaysia	JPEH00000000	5.09	60.80	
RU34A	Missing	USA	FTME00000000	4.83	61.20	
SNUFPC-A8	Salmon	South Korea	AMQA00000000	4.97	60.80	[67]
TN-97-08	Diseased Bluegill	USA	LNUR00000000	5.09	60.80	[68]
TPS-30	Fish	China	NBWY01000000	4.93	61.20	

(http://itol.embl.de/) [45]. The phylogeny based on the presence/absence matrix was generated by the R statistical language base functions and heatmap function.

Functional annotations

Identified proteins were further functionally annotated against Cluster of Orthologous Group (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases [46]. Integrated prophages were identified using the PHAST server (http://phast.wishartlab.com/index.html) [47]. To identify the prevailing antibiotic resistance genes and virulence factors, genomes were aligned utilizing Ublast with an e-value of 1×10^{-6} and alignment length of 80% [43] against the Comprehensive Antibiotic Resistance Database (CARD) (https://card.mcmaster.ca/) [48] and the Virulence Factors Database (VFDB, www.mgc.ac.cn/VFs/main.htm) [49]. Further, aligned sequences were manually checked for the annotated features to the database and NCBI. All the graphical figures and heat-map charts were generated by R package ggplot2.

Comparison of the genome sequences of *A. hydrophila* with different sequence types (STs)

Our previous study demonstrated that *A. hydrophila* ST-251 is a high-risk type in fish farms [5]. To determine whether there exists a difference in genome size and composition between ST-251 and other STs, a comparative genome analysis was re-run on these categories separately to compare the core and dispensable genes prevalent among both groups (Additional files 2 and 3). Similarly, the functional annotations based on COG and KEGG databases were reanalysed based on these categories.

Additional files

Additional file 1: Number of core, accessory and unique genes among the valid 49 strains of *A. hydrophila*. (DOCX 19 kb)

Additional file 2: Number of core, accessory and unique genes among the 16 strains of the ST-251 group. (DOCX 15 kb)

Additional file 3: Number of core, accessory and unique genes among the 33 strains of the other ST groups. (DOCX 17 kb

Additional file 4: Amino acid frequency of all 49 strains belonging to *Aeromonas hydrophila*. Comparison of core and dispensable genomes on

the basis of all 20 amino acids found in the *A. hydrophila* pan-genome. (DOCX 16 kb)

Awan *et al. BMC Genomics* (2018) 19:712

Page 17 of 18

Author details
[1]Joint International Research Laboratory of Animal Health and Food Safety, College of Veterinary Medicine, Nanjing Agricultural University, Nanjing, China. [2]Department of Epidemiology and Public Health, University of Veterinary and Animal Sciences, Lahore, Pakistan.

References
1. Janda JM, Abbott SL. The genus Aeromonas: taxonomy, pathogenicity, and infection. Clin Microbiol Rev. 2010;23(1):35–73.
2. Grim CJ, Kozlova EV, Sha J, Fitts EC, van Lier CJ, Kirtley ML, et al. Characterization of Aeromonas hydrophila wound pathotypes by comparative genomic and functional analyses of virulence genes. MBio. 2013;4(2):e00064–13.
3. Rasmussen-Ivey CR, Figueras MJ, McGarey D, Liles MR. Virulence factors of Aeromonas hydrophila: in the wake of reclassification. Front Microbiol. 2016;7:1337.
4. Stratev D, Vashin I, Rusev V. Prevalence and survival of Aeromonas spp. in foods - a review. Revue Med Vet. 2012;163:486–94.
5. Pang M, Jiang J, Xie X, Wu Y, Dong Y, Kwok AH, et al. Novel insights into the pathogenicity of epidemic Aeromonas hydrophila ST251 clones from comparative genomics. Sci Rep. 2015;5:9833.
6. Awan F, Dong Y, Wang N, Liu J, Ma K, Liu YJ. The fight for invincibility: environmental stress response mechanisms and Aeromonas hydrophila. Microb Pathog. 2018;116:135–45.
7. Rasmussen-Ivey CR, Hossain MJ, Odom SE, Terhune JS, Hemstreet WG, Shoemaker CA, et al. Classification of a hypervirulent Aeromonas hydrophila pathotype responsible for epidemic outbreaks in warm-water fishes. Front Microbiol. 2016;7:1615.
8. Beaz-Hidalgo R, Hossain MJ, Liles MR, Figueras MJ. Strategies to avoid wrongly labelled genomes using as example the detected wrong taxonomic affiliation for Aeromonas genomes in the GenBank database. PLoS One. 2015;10(1):e0115813.
9. Forn-Cuni G, Tomas JM, Merino S. Genome sequence of Aeromonas hydrophila strain AH-3 (serotype O34). Genome Announc. 2016;4(5): e00919–6.
10. Kaas RS, Friis C, Ussery DW, Aarestrup FM. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse Escherichia coli genomes. BMC Genomics. 2012;13(1):577.
11. Yoon SH, Ha SM, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek. 2017;110(10):1281–6.
12. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol. 2007;57(Pt 1):81–91.
13. De Maayer P, Aliyu H, Vikram S, Blom J, Duffy B, Cowan DA, et al. Phylogenomic, pan-genomic, pathogenomic and evolutionary genomic insights into the agronomically relevant enterobacteria Pantoea ananatis and Pantoea stewartii. Front Microbiol. 2017;8:1755.
14. Breurec S, Criscuolo A, Diancourt L, Rendueles O, Vandenbogaert M, Passet V, et al. Genomic epidemiology and global diversity of the emerging bacterial pathogen Elizabethkingia anophelis. Sci Rep. 2016;6:30379.
15. Tripathi C, Mishra H, Khurana H, Dwivedi V, Kamra K, Negi RK, et al. Complete genome analysis of Thermus parvatiensis and comparative genomics of Thermus spp provide insights into genetic variability and evolution of natural competence as strategic survival attributes. Front Microbiol. 2017;8:1410.
16. Daskalov H. The importance of Aeromonas hydrophila in food safety. Food Control. 2006;17:474–83.
17. Caputo A, Merhej V, Georgiades K, Fournier PE, Croce O, Robert C, et al. Pan-genomic analysis to redefine species and subspecies based on quantum discontinuous variation: the Klebsiella paradigm. Biol Direct. 2015;10:55.
18. Francis F, Kim J, Ramaraj T, Farmer A, Rush MC, Ham JH. Comparative genomic analysis of two Burkholderia glumae strains from different geographic origins reveals a high degree of plasticity in genome structure associated with genomic islands. Mol Gen Genomics. 2013; 288(3–4):195–203.

19. Del Castillo CS, Hikima J, Jang HB, Nho SW, Jung TS, Wongtavatchai J, et al. Comparative sequence analysis of a multidrug-resistant plasmid from Aeromonas hydrophila. Antimicrob Agents Chemother. 2013;57(1): 120–9.
20. Piotrowska M, Popowska M. The prevalence of antibiotic resistance genes among Aeromonas species in aquatic environments. Ann Microbiol. 2014; 64(3):921–34.
21. Wen Y, Zou Z, Li H, Xiang Z, He N. Analysis of codon usage patterns in Morus notabilis based on genome and transcriptome data. Genome. 2017; 60(6):473–84.
22. Berg OG, Kurland CG. Growth rate-optimised tRNA abundance and codon usage. J Mol Biol. 1997;270:544–50.
23. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucl Acids Symp Ser. 1999;41:95–8.
24. John FP. Analysis of codon usage. UK: The University of Nottingham; 1999.
25. Chaudhari NM, Gupta VK, Dutta C. BPGA- an ultra-fast pan-genome analysis pipeline. Sci Rep. 2016;6:24373.
26. Do J, Zafar H, Saier MH. Comparative genomics of transport proteins in probiotic and pathogenic Escherichia coli and Salmonella enterica strains. Microb Pathog. 2017;107:106–15.
27. Bhardwaj T, Somvanshi P. Pan-genome analysis of Clostridium botulinum reveals unique targets for drug development. Gene. 2017;623:48–62.
28. Kiu R, Caim S, Alexander S, Pachori P, Hall LJ. Probing genomic aspects of the multi-host pathogen Clostridium perfringens reveals significant pangenome diversity, and a diverse array of virulence factors. Front Microbiol. 2017;8:2485.
29. Chen F, Wang K, Stewart J, Belas R. Induction of multiple prophages from a marine bacterium: a genomic approach. Appl Environ Microbiol. 2006;72(7): 4995–5001.
30. Canchaya C, Fournous G, Brussow H. The impact of prophages on bacterial chromosomes. Mol Microbiol. 2004;53(1):9–18.
31. Coburn B, Sekirov I, Finlay BB. Type III secretion systems and disease. Clin Microbiol Rev. 2007;20(4):535–49.
32. Suarez G, Khajanchi BK, Sierra JC, Erova TE, Sha J, Chopra AK. Actin cross-linking domain of Aeromonas hydrophila repeat in toxin A (RtxA) induces host cell rounding and apoptosis. Gene. 2012;506(2):369–76.
33. Ghatak S, Blom J, Das S, Sanjukta R, Puro K, Mawlong M, et al. Pan-genome analysis of Aeromonas hydrophila, Aeromonas veronii and Aeromonas caviae indicates phylogenomic diversity and greater pathogenic potential for Aeromonas hydrophila. Antonie Van Leeuwenhoek. 2016;109(7):945–56.
34. Odeyemi OA, Ahmad A. Antibiotic resistance profiling and phenotyping of Aeromonas species isolated from aquatic sources. Saudi J Biol Sci. 2017; 24(1):65–70.
35. Fosse T, Giraud-Morin C, Madinier I. Induced colistin resistance as an identifying marker for Aeromonas phenospecies groups. Lett Appl Microbiol. 2003;36:25–9.
36. Aravena-Roman M, Inglis TJ, Henderson B, Riley TV, Chang BJ. Antimicrobial susceptibilities of Aeromonas strains isolated from clinical and environmental sources to 26 antimicrobial agents. Antimicrob Agents Chemother. 2012;56(2):1110–2.
37. Anes J, McCusker MP, Fanning S, Martins M. The ins and outs of RND efflux pumps in Escherichia coli. Front Microbiol. 2015;6:587.
38. Hossain MJ, Sun D, McGarey DJ, Wrenn S, Alexander LM, Martino ME, et al. An Asian origin of virulent Aeromonas hydrophila responsible for disease epidemics in United States-farmed catfish. MBio. 2014;5(3):e00848–14.
39. Darling AE, Mau B. Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 2010; 5(6):e11147.
40. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11(1):119.
41. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, et al. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol. 2012;50(4):1355–61.
42. Meier-Kolthoff JP, Auch AF, Klenk H-P, Goker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMCBioinformatics. 2013;14:60.
43. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460–1.
44. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.

Awan *et al. BMC Genomics* (2018) 19:712

Page 18 of 18

45. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242–5.

46. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 2012;40(Database issue):D109–14.

47. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. Nucleic Acids Res. 2011;39:W347–W52.

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. 2017;45(D1):D566–d73.

49. Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis--10 years on. Nucleic Acids Res. 2016; 44(D1):D694–7.

50. Seshadri R, Joseph SW, Chopra AK, Sha J, Shaw J, Graf J, et al. Genome sequence of Aeromonas hydrophila ATCC 7966T: jack of all trades. J Bacteriol. 2006;188(23):8272–82.

51. Tekedar HC, Karsi A, Akgul A, Kalindamar S, Waldbieser GC, Sonstegard T, et al. Complete genome sequence of fish pathogen Aeromonas hydrophila AL06-06. Genome Announc. 2015;3(2):e00368–15.

52. Pridgeon JW, Zhang D, Zhang L. Complete genome sequence of the highly virulent Aeromonas hydrophila AL09-71 isolated from diseased channel catfish in West Alabama. Genome Announc. 2014;2(3):e00450–14.

53. Yang W, Li N, Li M, Zhang D, An G. Complete genome sequence of fish pathogen Aeromonas hydrophila JBN2301. Genome Announc. 2016; 4(1):e01615.

54. Tekedar HC, Waldbieser GC, Karsi A, Liles MR, Griffin MJ, Vamenta S, et al. Complete genome sequence of a channel catfish epidemic isolate, Aeromonas hydrophila strain ML09-119. Genome Announc. 2013; 1(5):e00755–13.

55. Pridgeon JW, Zhang D, Zhang L. Complete genome sequence of a moderately virulent Aeromonas hydrophila strain, pc104A, isolated from soil of a catfish pond in West Alabama. Genome Announc. 2014;2(3): e00554–14.

56. Hughes HY, Conlan SP, Lau AF, Dekker JP, Michelin AV, Youn JH, et al. Detection and whole-genome sequencing of carbapenemase-producing Aeromonas hydrophila isolates from routine peri-rectal surveillance culture. J Clin Microbiol. 2016;54(4):1167–70.

57. Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft BJ, Evans PN, et al. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat Microbiol. 2017;2(11):1533–42.

58. Lenneman EM, Barney BM. Draft genome sequences of the alga-degrading bacteria Aeromonas hydrophila strain AD9 and Pseudomonas pseudoalcaligenes strain AD6. Genome Announc. 2014;2(4):e00709–14.

59. Honein K, Jagoda S, Arulkanthan A, Ushio H, Asakawa S. Draft genome sequence of Aeromonas hydrophila strain Ae25, isolated from a septicemic moribund koi carp (Cyprinus carpio) in Sri Lanka. Genome Announc. 2018; 6(5):e01523–17.

60. Jagoda SS, Tan E, Arulkanthan A, Kinoshita S, Watabe S, Asakawa S. Draft genome sequence of Aeromonas hydrophila strain Ae34, isolated from a septicemic and moribund koi carp (Cyprinus carpio koi), a freshwater aquarium fish. Genome Announc. 2014;2(3):e00572–14.

61. Tekedar HC, Kumru S, Karsi A, Waldbieser GC, Sonstegard T, Schroeder SG, et al. Draft genome sequences of four virulent Aeromonas hydrophila strains from catfish aquaculture. Genome Announc. 2016;4(4):e00860–16.

62. Tekedar HC, Kumru S, Kalindamar S, Karsi A, Waldbieser GC, Sonstegard T, et al. Draft genome sequences of three Aeromonas hydrophila isolates from catfish and tilapia. Genome Announc. 2017;5(3):e01509–16.

63. Pan X, Lin L, Xu Y, Yuan X, Yao J, Yin W, et al. Draft genome sequence of Aeromonas hydrophila strain BSK-10 (serotype O97), isolated from Carassius carassius with motile aeromonad septicemia in China. Genome Announc. 2017;5(24):e00497–17.

64. Teng L, Deng L, Dong X, Wei S, Li J, Li N, et al. Genome sequence of hypervirulent Aeromonas hydrophila strain HZAUAH. Genome Announc. 2017;5(11):e00012–7.

65. Tan WS, Yin WF, Chang CY, Chan KG. Whole-genome sequencing analysis of quorum-sensing Aeromonas hydrophila strain M023 from freshwater. Genome Announc. 2015;3(1):e01548–14.

66. Ponnusamy D, Kozlova EV, Sha J, Erova TE, Azar SR, Fitts EC, et al. Cross-talk among flesh-eating Aeromonas hydrophila strains in mixed infection leading to necrotizing fasciitis. Proc Natl Acad Sci U S A. 2016;113:722–7.

67. Han JE, Kim JH, Choresca C, Shin SP, Jun JW, Park SC. Draft genome sequence of a clinical isolate, Aeromonas hydrophila SNUFPC-A8, from a moribund cherry salmon (Oncorhynchus masou masou). Genome Announc. 2013;1(1):e00133–12.

68. Tekedar HC, Kumru S, Karsi A, Waldbieser GC, Sonstegard T, Schroeder SG, et al. Draft genome sequence of Aeromonas hydrophila TN97-08. Genome Announc. 2016;4(3):e00436–16.