JEPIBIO
Journal of Epidemiology and Infection Biology

Research Article

Pangenome of *Aeromonas hydrophila* Reveals Noticeable Heterogeneity in Antimicrobial Resistance

Afsah Pervaiz*[1], Zabeehullah[2], Muhammad Osama Sajjad[3], Hafiz Waleed Zafar[4]

[1] Institute of Microbiology, University of Veterinary and Animal Sciences, Lahore, Pakistan
[2] Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences, Lahore, Pakistan
[3] College of Medicine and Public Health, Flinders University, Bedford Park, South Australia 5042, Australia
[4] Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences, Lahore, Pakistan

*Corresponding author:

Ms. Afsah Pervaiz
(e-mail: afsahpervaiz@gmail.com)

Date of Receiving: 11/04/2025
Date of Acceptance 22/05/2025
Date of Publishing 10/06/2025

ABSTRACT *Aeromonas hydrophila* is a Gram-negative bacterium found in freshwater and marine environments. It is an opportunistic pathogen capable of infecting fish, amphibians, and humans. The species causes gastrointestinal illness and septicemia in humans and is a major concern in aquaculture due to motile Aeromonas septicemia (MAS). Increasing antibiotic resistance in *A. hydrophila* has raised global health concerns. This resistance is linked to plasmids, insertion sequences, and other mobile elements that facilitate horizontal gene transfer. This study aimed to investigate the genomic diversity of *A. hydrophila* using pangenome analysis. This study analyzed 31 complete or high-quality genomes retrieved from NCBI databases. Genomes were re-annotated using Prokka. Average Nucleotide Identity (ANI) confirmed species identity (>97%) for 27 genomes. PIRATE was used to classify genes into core and accessory clusters. Snippy and IQ-TREE were used for core genome SNP detection and phylogenetic tree construction. Population structure was inferred using hierBAPS. Mobile genetic elements including AMR genes, CRISPRs, insertion sequences, plasmids, and prophages were identified using specialized tools. PCA and correlation analysis were performed to evaluate associations between genomic features. ANI analysis confirmed species-level similarity among strains. Genome size ranged from 4.3 to 5.3 Mb with stable GC content (~61%). The pangenome was open, with high accessory gene diversity. Several strains displayed unique accessory gene profiles. Phylogenetic analysis revealed distinct clades, some showing high divergence. Core SNP-based phylogeny provided strong resolution without strict host or geographic clustering. AMR genes varied across countries, with Chinese strains showing the highest burden. IS elements and plasmids were more frequent in AMR-rich strains. CRISPRs were rare, while prophages were common. PCA grouped strains based on AMR and genomic features. BAPS analysis identified three major SNP-based clusters. This study highlights the genomic plasticity and heterogeneity of *A. hydrophila*. Core and accessory genome analysis revealed evolutionary trends. Associations between AMR and mobile elements were evident. These findings support the role of horizontal gene transfer in shaping resistance patterns. This work enhances understanding of *A. hydrophila* evolution and informs future surveillance and control strategies.

KEYWORDS Aeromonas, Pangenome, Antimicrobial resistance, Genomic diversity, Mobile genetic elements

Introduction

Aeromonas hydrophila is a Gram-negative, facultatively anaerobic bacterium commonly found in aquatic environments (Awan *et al*, 2018). It is widely distributed in freshwater, brackish water, and wastewater. The organism is known for its ability to infect both cold-blooded and warm-blooded hosts. It acts as an opportunistic pathogen in fish, amphibians, reptiles, and mammals. In humans, *A. hydrophila* infects gastrointestinal tract, wound infections, and septicemia (Ali *et al*, 2015). It also occurs as a major pathogen in fish aquaculture where it causes motile Aeromonas septicemia (MAS). The organism has multiple virulence factors which include hemolysins, aerolysins, and enterotoxins. These virulence factors add to its pathogenicity and adaptability. Environmental stress and host immune suppression also increases the pathogenicity. *A. hydrophila* can withstand adverse conditions and cleaning supplies (Abdella *et al*, 2023). This pathogen can grow as biofilms which increase its survival in water bodies. Although, this pathogen is less frequent among human and animal populations since it is widely found and can cause pathogenicity (Jin *et al*, 2020). Over the last few years, *A.*

To cite this article: Pervaiz, A., Zabeehullah, and M. O. Sajjad. (2025). *Pangenome of Aeromonas hydrophila reveals noticeable heterogeneity in antimicrobial resistance*. Journal of Epidemiology and Infection Biology 2(1):14-23.

hydrophila is known to induce genomic variability and has shown high levels of plasticity. This genomic plasticity helps in further adaptability, survive and disease spread.

Globally, antimicrobial resistance (AMR) is becoming a challenge in terms of community wellbeing. *A. hydrophila* has been found with increased antibiotic resistance towards beta-lactams, tetracycline, sulfonamides, and quinolones. Clinical and environmental strains of *A. hydrophila* have also been isolated as multidrug-resistant (MDR) strains which could lead to treatment problems (Lawal *et al*, 2024). Mobile genetic elements (MGE) such as plasmids, insertion sequences, and integrons has enhanced horizontal gene transfer. Such factors play potential role in spread of resistance genes (Ma *et al*, 2023). Natural water resources are prone to act as repositories of AMR determinants. The resistant strains of Aeromonas can get into the food chain or into the water system (Abdella *et al*, 2023). AMR monitoring in *A. hydrophila* is crucial to the community health. Understanding of genomics is vital in understanding the pathogenicity and resistance of this pathogen. Such aspects of the genome as the mobile genetic elements and the resistance islands are of particular interest (Soares *et al*, 2013). The investigation of AMR profiles in *A. hydrophila* allows evaluating risks and planning control measures.

This research aims to explore the genomic diversity of *Aeromonas hydrophila* using pangenome analysis. Pangenome analysis determines the all the aspects of genomes such as virulence factors, core genes, accessory genes etc (Soares *et al*, 2013). Core genes are shared by all strains, while accessory genes vary across isolates. The accessory genome includes virulence and resistance genes. Studying the pangenome reveals patterns of gene gain and loss. It also highlights genomic plasticity and evolutionary trends (Abram *et al*, 2022). These tools allow clustering, variant detection, and population structure analysis. The goal is to link these features with genomic subtypes and SNP-based population structure. This approach provides a holistic view of genomic diversity. It supports surveillance and informs interventions for *A. hydrophila* control.

Materials and Methods

Genomes and Associated Data

A total of 31 genomes of *Aeromonas hydrophila* was retrieved from NCBI RefSeq and GenBank (as of 03/02/2021). Only complete or high-quality draft assemblies were selected. Poor-quality and incomplete genomes were excluded. Metadata included genome size, source, country, and year of isolation. These were extracted from GenBank and BioSample records. Genomes were reannotated using Prokka (version 1.13.4) to ensure consistency. Assembly statistics, Genome size and GC content were calculated using in-house scripts. Duplicates and erroneous entries were removed. Final datasets were used for downstream comparative analysis.

Pangenome estimation

Pangenome analysis was performed using the PIRATE pipeline (version 1.0.5). Genes were clustered based on amino acid similarity. A minimum identity threshold of 95% was used for clustering. PIRATE grouped genes into core, soft-core, shell, and cloud categories. Classification was based on gene presence across genomes. Output matrices were generated for accessory genome analysis. Gene presence/absence data were visualized using in-house script of R-statistical language. Pangenome size and openness were estimated. A gene accumulation curve was plotted.

Core-genome building

Core-genome alignment was built using Snippy (version 4.6.0). A reference genome ATCC 7966 (Accession Number: GCF000014805.1) was used for mapping. SNPs were called under default parameters. Low-quality regions and ambiguous sites were filtered. The final alignment contained high-confidence core SNPs. A phylogenetic tree was constructed using IQ-TREE with 1,000 bootstrap replicates. The best-fit model was selected automatically. Tree robustness was assessed by bootstrap values. The tree was visualized using iTOL (version 7.2). SNP data from Snippy were used for polymorphism analysis. SNP density per gene was calculated. High-variation regions were flagged and annotated. Genomic distribution of polymorphisms was visualized. Plots were generated using custom R scripts. Trends across isolates were compared. Findings suggested variability in genomic hotspots.

Determination of various genomic features

Following core genome building, various genomic features such as AMR, VF, IS-seq, Prophage, Plasmids, MLST, and CRISPR were determined. AMR genes were identified using NCBI AMR Finder Plus (version 4.0.23). The NCBI AMR database was used for annotation. Virulence genes were predicted using ABRicate with the VFDB. Insertion sequences were detected using IS Finder. Prophages were identified using PHASTER. Plasmids were detected using Plasmid Finder. MLST was performed using the mlst tool and PubMLST database. CRISPR elements were annotated with CRISPR CasFinder. Results were compiled in a tabular format for each genome. These features were used in correlation and diversity analyses.

Association of AMR with Other Features

In this study, associations between AMR genes and genomic features were also evaluated. Plasmid count, IS elements, and genome size were included. Significant results ($p < 0.05$) were highlighted. Principal Component Analysis (PCA) was conducted in R using the prcomp() function. Input data included AMR presence, IS count, plasmid count, genome size, and prophage number. All variables were standardized before analysis. Scree plots were used to assess variance explained. Principal components were plotted using ggplot2. Heatmaps and correlation matrices were generated in R. Associations were validated using subset comparisons. Key trends were observed across multiple genomes. These associations helped explain AMR variability. Clusters and outliers were visualized in 2D plots. PCA helped visualize genomic variation. Variance contribution by each feature was analyzed. Feature loadings were also extracted for interpretation.

BAPS Analysis

Population structure was inferred using hierBAPS (Cheng *et al*, 2013). Input was the filtered core SNP alignment. The model was run with 6 hierarchical levels. Genomes were grouped into BAPS levels 1 and 2. Each cluster represented distinct subpopulations. Cluster membership was recorded for each genome. These groups were used in comparative analyses. BAPS results were mapped to the phylogenetic tree. AMR patterns were compared across BAPS clusters.s.

Results

A total of 31 genomes were downloaded from NCBI in the study. All the associated meta data were extracted from NCBI which included countries of origin, sampled animal species, type of sample etc. These genomes were renamed to strain IDs submitted on NCBI as per requirement of downstream software utilized.

Genomes and associated data

The genomic dataset showed high overall similarity among *A. hydrophila* strains based on ANI analysis (Fig. 1a). Most genome pairs shared over 97% ANI, confirming species-level relatedness. A few genomes, such as GCA_014220715.1, had slightly lower ANI, around 95%. These lower values suggest possible sub-lineages or greater diversity. The ANI matrix showed consistent clustering among closely related genomes. In total, the dataset captured both conserved and variable strains. Genomes (n=4) having values lower than < 95% were removed from downstream analyses. Isolates originated mainly from China and the USA (Fig. 1b). China had the highest number of strains, with fish and water as common sources. Host diversity was broader in the USA, including humans and animals. Some isolates lacked host metadata. Genome size ranged from approximately 4.3 Mb to 5.3 Mb (Fig. 1c). GC content varied between 60.8% and 61.6%, remaining relatively stable. No strong correlation was observed between genome size and GC content. PCA based on genome features explained over 97% of variance on the first two axes (Fig. 1d). Genome size and GC% were major contributors to variability. Host-based clusters overlapped, indicating a continuum of genomic diversity across environmental and clinical sources.

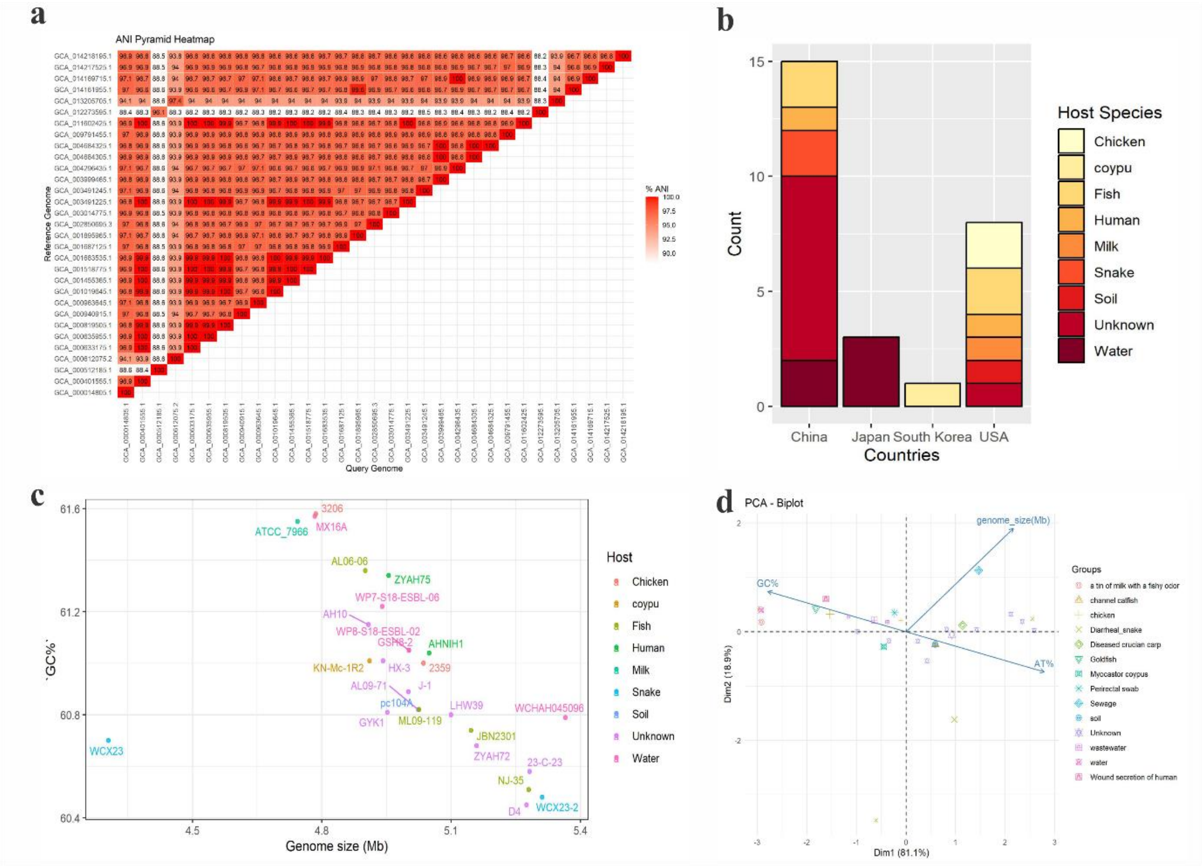

Fig. 1: Genomic dataset of *Aeromonas hydrophila* downloaded from online sources. (a) ANI values determined using FASTANI tool with >95% reference values. (b) Bar plot showing distribution of number of genomes isolated from different sources and countries. (c) Point/Jitter plot showing distribution of Genome size (Mbs) with reference to GC% and Host. (d) PCA biplot showing the relationship of Isolation sources, GC%, AT% and Genome Size (Mb).

Pangenome

The pangenome analysis revealed extensive diversity among *Aeromonas hydrophila* genomes (Fig. 2a). The presence/absence matrix showed a clear division between core and accessory genes. Core genes were consistently present across all isolates while accessory and unique genes were scattered and highlighted both shared and strain-specific genes. Several genomes, such as WCX23 and 23_C_23, showed a distinct accessory gene profile. The distribution of gene clusters across isolates was quantified (Fig. 2b). Most gene clusters were present in only 0–10% or 95–100% of isolates. This confirmed the presence of a large accessory genome and a stable core genome. Fewer genes fell into the mid-frequency range (25–75% of isolates). The overall pattern supports an open pangenome with high genomic plasticity.

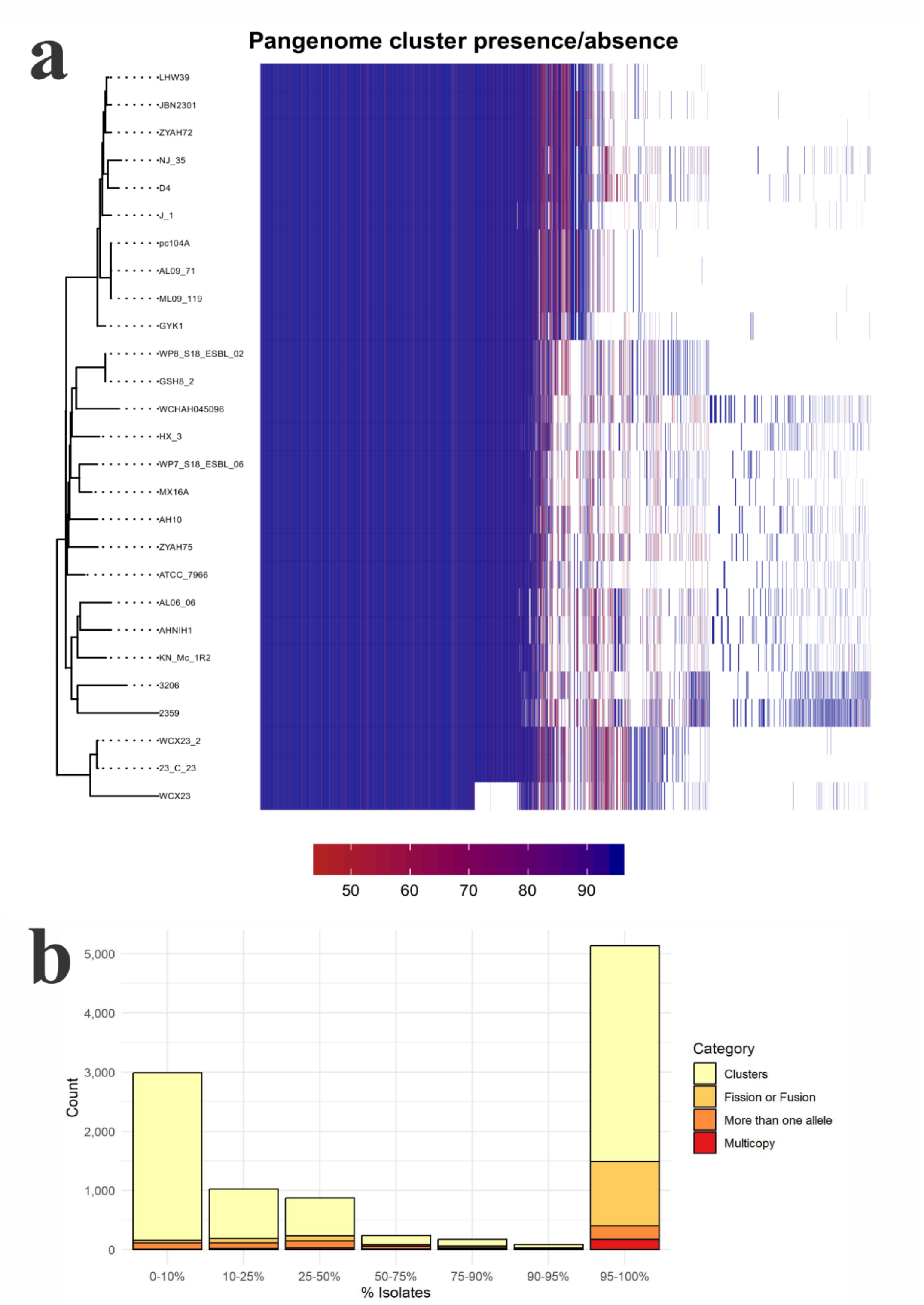

Fig. 2: Pangenome estimation and gene structures was determined using heatmap and bar plot. (a) Heatmap showed the presence or absence of core, accessory, or unique genes in the range of 45% to 95% similarity. (b) Gene structures (intact, fusion, fission, and multicopy) were determined.

Core-genome phylogeny

The core genome phylogeny shows the evolutionary relationships among *A.hydrophila* isolates (Fig. 3). The tree was constructed using core SNPs from all genomes. Most isolates grouped into distinct clades. Several closely related strains clustered together, indicating shared ancestry. Examples include JBN2301 and D4, and also 2359 and 3206. Some strains like ZYAH75 and WCX23 formed long branches, suggesting higher divergence. The reference strain ATCC_7966 clustered with WCHAHA05096 and AL06_06. Certain groups, such as WP97_S18_ESBL_06 and MX16A, were also closely related. The circular layout highlights overall genomic diversity. Isolates were not clustered strictly by host or source. The tree topology suggests multiple evolutionary lineages. Some isolates showed basal branching, indicating ancestral positions. Overall, the core genome was conserved but allowed for phylogenetic resolution. The tree provides a framework for linking genomic variation with phenotype or origin. This phylogeny complements the pangenome and AMR analyses.

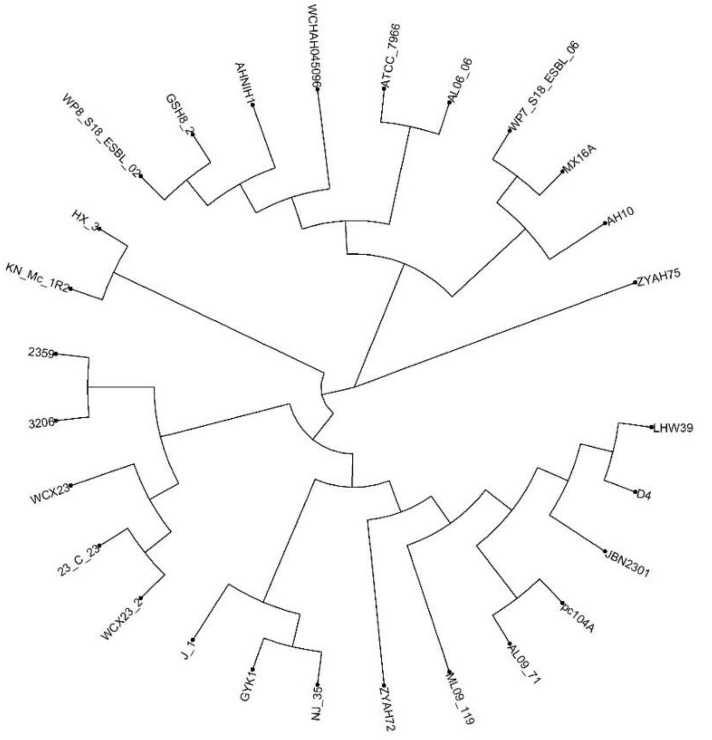

Fig. 3: Neighborhood joining tree for core genome of *Aeromonas hydrophila* showed sub-divisions in phylogeny

Genomic features – (IS-sequences, Prophage, Plasmids, MLST, CRISPR)

The genomic features of the studied strains revealed notable variability (Fig. 4). CRISPR elements were detected in only two strains such as strains 3206 and ZYAH75. Most strains lacked CRISPR-associated sequences entirely. Insertion sequences elements showed substantial variation across strains. WCHAH045096 exhibited the highest IS sequence count, exceeding 150 elements. Other strains such as AHNIH1, WP7-S18-ESBL-06, WP7-S18-ESBL-06, HX-3, GSH8-2 also showed higher IS sequences. The reference strain ATCC7966 showed least count of insertion sequences. Several strains, including 3206, had fewer than 20 IS sequences. Plasmid content ranged from zero to six plasmids per strain. D4 and WCHAH045096 had the highest plasmid count. Strains such as GSH8-2 and AHNIH1 showed only single plasmid presence. Some strains, including AL09-71 and J-1, lacked plasmids entirely. Prophage elements were found in all genomes. Most strains carried 5 to 11 prophage regions. WCHAH045096, HX-3, 23-C-23, and WCX23-2, had the highest prophage content. Strains 3206 and MX16A had relatively lower prophage counts. These features contribute to the genetic diversity observed in the pangenome analysis.

Association of AMR with other features

The distribution of antimicrobial resistance (AMR) genes varied across countries (Fig. 5). Chinese strains showed the highest AMR burden overall (Fig. 5a). WCHAH045096 and ZYAH75 from China exhibited notably high AMR gene counts. Strains from the USA displayed moderate AMR levels, with AHNIH1 showing the highest among them. Japanese strains had comparatively lower AMR gene abundance. South Korean strains showed minimal AMR gene presence. The PCA plot grouped strains based on MLST types. MLST group 251 formed a distinct cluster, indicating close genetic relatedness. Group 1 strains also clustered tightly, suggesting shared genomic features. Untypeable strains were scattered across the PCA space. This indicates higher genetic diversity in untypable genomes. Group 352 and 516 showed isolated positioning in the plot. Dim1 and Dim2 explained 49.8% and 22.1% of total variance, respectively. The PCA revealed both country-specific and MLST-specific clustering patterns. These findings support geographic and genotypic influence on AMR gene distribution.

Trends of polymorphisms in genomic features

The genomic variation among strains was assessed through SNPs, insertions, and deletions (Fig. 6). SNP counts were high across all strains, exceeding 100,000 in most cases. The highest number of SNPs was observed in ZYAH72 and 2359. Strains such as KN-Mc-1R2 and WCX23-2 had slightly lower SNP counts. Insertions were consistently distributed across strains, ranging from 1,400 to 1,700. The highest insertion count appeared in pol0346A and NX18-3A. Strains AL09-79 and ML09-119 showed relatively fewer insertions. Deletions showed more variability compared to insertions. Strain 2359 exhibited the highest number of deletions, exceeding 6,000. Several strains such as GSH8-2 and ATCC_7966 had fewer than 2,000 deletions. Strains from the same geographic region showed similar deletion patterns. Most strains showed a balanced count of insertions and deletions. Overall, SNPs were the most frequent variation type across all genomes. The results indicate significant genomic plasticity among the analyzed strains. These variations contribute to strain-specific genome diversity in the pangenome.

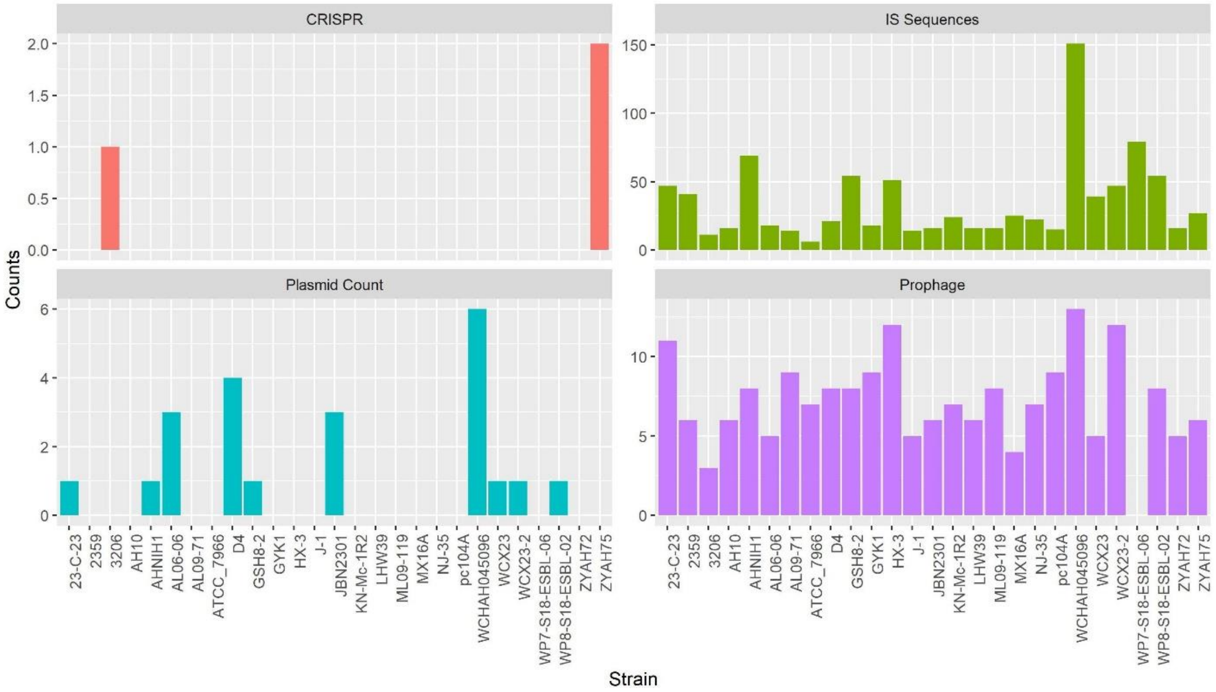

Fig. 4: Number of genomic features determined in *Aeromonas hydrophila* strains.

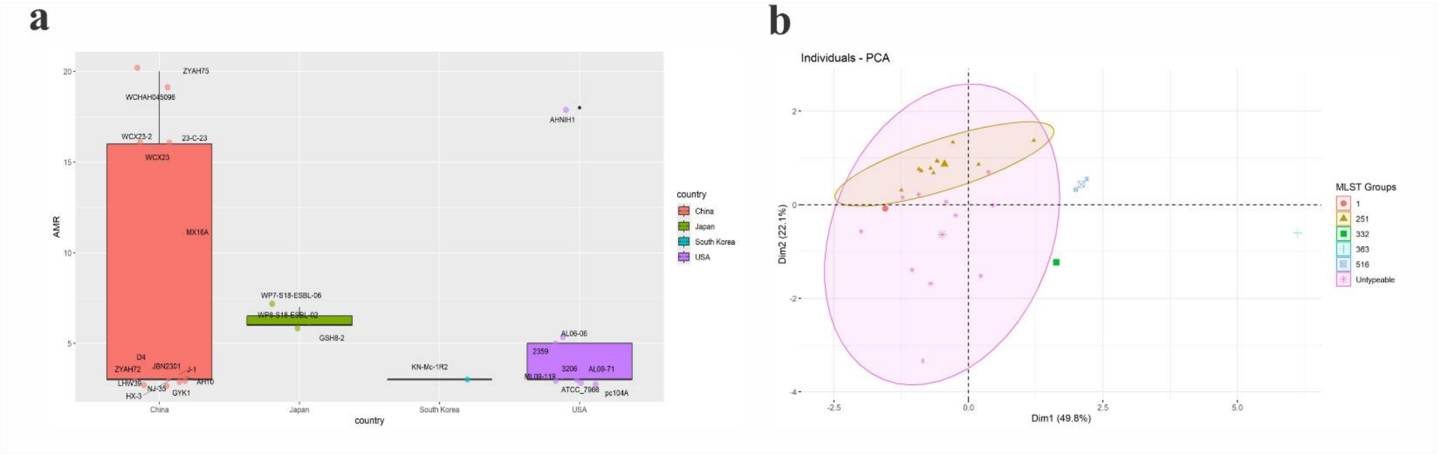

Fig. 5: Relationship of AMR and other genomic features in the *Aeromonas hydrophila* pangenome. (a) AMR gene count across different countries. (b) PCA biplot shows MLST groups were showing diversity.

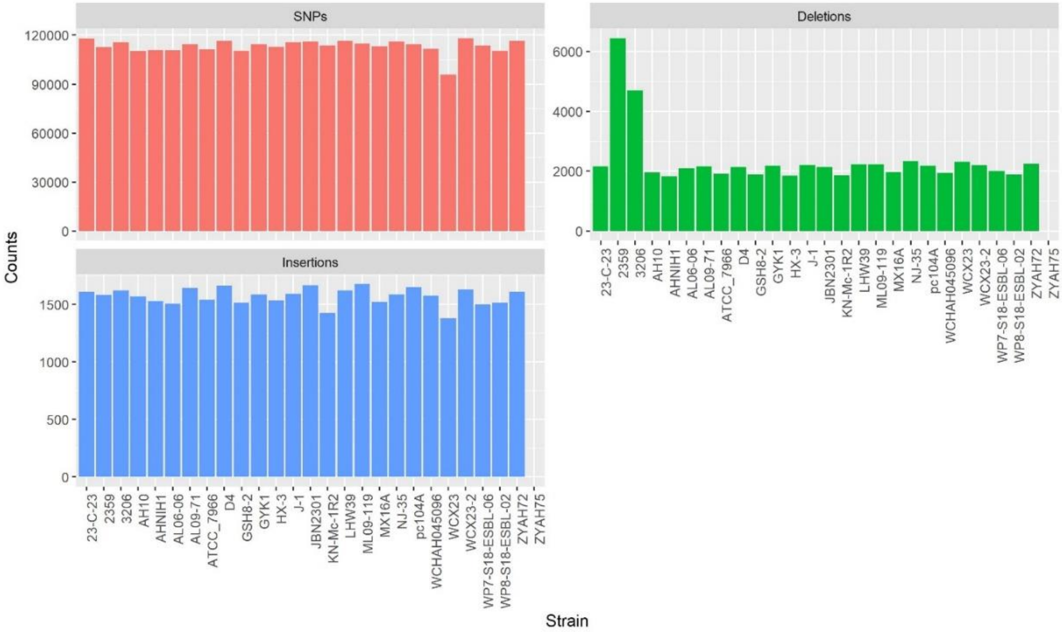

Fig. 6: Comparison of SNPs, insertions, and deletions in *Aeromonas hydrophila* pangenome.

BAPS analysis

The hierBAPS output shows SNP-based clustering of *A. hydrophila* genomes (Fig. 7). The alignment matrix displayed base calls across SNP positions. Each color represented a nucleotide: blue (A), yellow (C), red (G), green (T), and white (indels). Genomes were grouped into hierarchical BAPS clusters. Three main clusters were separated by bold black lines. Each group showed consistent SNP patterns across large genomic regions. Cluster 1 genomes had mostly blue (A) and yellow (C) bases. Cluster 2 contained dominant red (G) and green (T) patterns. Cluster 3 showed mix but distinct transitions in SNP types. The separation between clusters reflects deep genetic divergence. Within each cluster, minor SNP variations are also visible. These may represent sub-lineages or recent mutations. The consistency of color within blocks suggests shared ancestry. Inter-cluster variability supports population structure. Fig. 7 confirmed that *A. hydrophila* genomes were not homogeneous. This SNP-based stratification informed further analysis of AMR and gene content differences.

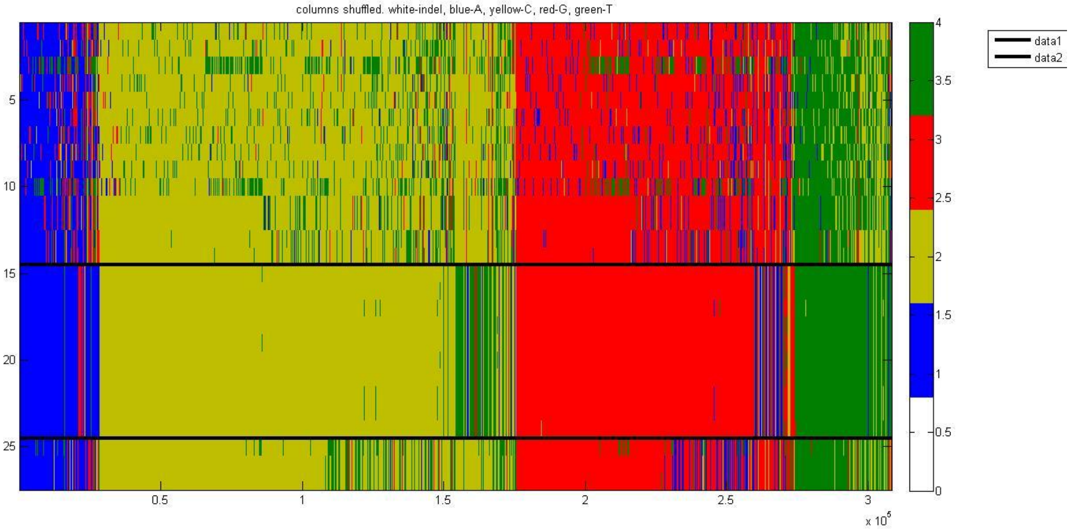

Fig. 7: HeirBAPS clustered revealed the heterogeneity among *Aeromonas hydrophila* genomes.

Discussion

The selection of high-quality genomes ensured robust downstream analysis. The ANI analysis confirmed species-level identity, with most strains sharing >97% identity. This finding aligns with previous studies where ANI >95% was used as a threshold to define species boundaries in Aeromonas (Beaz-Hidalgo *et al*, 2015). A few genomes with borderline ANI values indicate potential sub-speciation or misclassification, which has also been observed in earlier taxonomic revisions of Aeromonas (Martinez-Murcia *et al*, 2011). The genomic dataset covered a broad geographic and host range, similar to the diversity reported in earlier surveillance studies (Yang *et al*, 2019). Genome size and GC content remained relatively stable across strains, consistent with previous reports that describe *A. hydrophila* genomes ranging between 4.3–5.1 Mb and around 61% GC (Colston *et al*, 2014). PCA results showed that genome size and GC% were primary sources of variability, suggesting environmental and evolutionary pressures. Overlapping host-based clusters indicate that *A. hydrophila* lacks strong host specificity, supporting its role as a generalist pathogen. This observation supports earlier findings that describe *A. hydrophila* as a broad host-range pathogen capable of infecting fish, reptiles, and humans (Janda and Abbott Sharon, 2010). These genomic patterns reflect both environmental adaptation and anthropogenic selection pressure. The removal of low-ANI genomes helped maintain phylogenetic integrity. Overall, the genomic dataset provided a reliable foundation for comparative analyses.

Pangenome analysis revealed a clear division between conserved core genes and variable accessory genes. The presence/absence matrix showed that while a core set of genes is conserved, a large proportion of genes varied among strains. This observation supports the open pangenome model previously described in Aeromonas and other aquatic bacteria (Gauthier *et al*, 2017; Sayyaf Dezfuli *et al*, 2023). The gene accumulation curve further confirmed that new genes continue to emerge with additional genomes, consistent with the findings of Piotrowska and Popowska (2015), who reported horizontal gene transfer events in Aeromonas. Accessory genes often include mobile genetic elements, virulence factors, and resistance genes. This variation can add to adaptive phenotypic plasticity and suitability in heterogeneous niches. The clustering pattern showed the genome specific gene repertoires, which was a sign of ecological or evolutionary divergence. Gene structures are dynamic, which entails multicopy and split genes. Similar structural variability has been observed in previous studies related to Vibrio and Pseudomonas (Rouli *et al*, 2015). The open nature of the pangenome highlights the importance of continuous surveillance program (Rasko *et al*, 2008). The variability of accessory genes can play either a role in adaptation to a niche, resistance to antimicrobials or virulence adjustment. This shows again the need to take account of gene content during epidemiological analysis. Comprehensively, the pangenome analysis highlights the flexibility due to DNA in *A. hydrophila*.

Core SNPs were used to find that the phylogenetic tree was quite resolved to strain relationships. Strains closely related to each other were again grouped together, the implication being that they share evolutionary history. These clades were not thought to be specific to hosts or geography in line with previous findings that *A. hydrophila* is an opportunist which is globally widespread (Miranda-Lopez *et al*, 2024). Long branches were observed with some isolates indicating high divergence in these isolates e.g. ZYAH75 and WCX23. This observation is in line with core-genome analysis in other organisms whose divergent branches are usually indicative of horizontal gene transfer or independent evolution (Guo *et al*, 2022). Ancestral Lineages or early divergence events could be suggested by the presence of basal branches. Environmental Aeromonas strains have also been observed to be placed in similar basal location (Nokhwal *et al*, 2025). The pangenome analysis is supplemented by the tree topology in which strains with different accessory gene profiles were also found to be more phylogenetically distant. The phylogenies using only core genes are more useful than traditional MLST or 16S rRNA. High bootstrap support was indicated with the IQ-TREE tool, and it demonstrated that there was stability in the branches. *A. hydrophila* can cross host barrier also indicated by the absence of strict clustering on basis of host. The problem of the relationship between content of the genome and phenotypic traits (e.g. AMR or virulence) can be addressed based on the phylogenetic structuring. Such an approach is imperative to the monitoring of zoonotic or epidemic potential lineages. In general, it is the complicated evolution of this species with its tree.

Genomic feature analysis revealed marked variation in AMR, insertion sequences, plasmids, and CRISPR elements. Chinese strains had more AMR genes, which is in line with the overall trend of AMR worldwide since use of antibiotics in fish farming may lead to selection pressure (Truong *et al*, 2024). Cohort association between plasmid content and AMR favors plasmids as the inheritor of AMR (Zhong *et al*, 2019). High insertion sequence counts in strains such as WCHAH045096 suggest genome rearrangements and horizontal gene transfer (Zhong *et al*, 2019). CRISPRs were rare, which may indicate reduced immune defense against phages or increased reliance on horizontal gene uptake. This aligns with previous reports where *A. hydrophila* showed low CRISPR prevalence (Sakurai *et al*, 2024). The presence of prophages was also prevalent in all the strains and this confirmed the part played by phages in genome evolution. Many engorged genomes were observed in high prophage-containing strain, similar to other Pseudomonas and Vibrio strain results (Rouli *et al*, 2015). The variety of these genomic characteristics introduces the aspects of evolutionary process in the shaping up of the *A. hydrophila* populations. This genomic variation is capable of modifying pathogenicity, resistance and host adaptation. The inclusion of all these characteristics can be used to describe the heterogeneity that is observed in relation to phenotype and epidemiology.

The fact that AMR was linked to genomic characteristics showed revealing tendencies. PCA separate strains by AMR gene content, plasmids and IS elements as well as size of genomes. These associations support the hypothesis that AMR acquisition is linked to mobile genetic elements, as previously reported (Piotrowska et al., 2017). The separation of Chinese strains in PCA plots reflects regional AMR pressure. MLST groups formed distinct clusters, consistent with studies showing MLST type can correlate with resistance and virulence profiles (Godoy *et al*, 2023). Untypeable strains were scattered, suggesting genomic diversity not captured by conventional MLST. Variability in IS elements likely contributes to differential AMR gene mobilization. Plasmid-rich strains exhibited higher AMR, supporting their role in resistance spread. The PCA results align with a model of convergent AMR evolution, where unrelated strains acquire similar resistance via mobile elements. These findings highlight the need for integrated genomic surveillance, especially in aquaculture-dense regions. The results support targeted control strategies focusing on high-risk clones. Overall, AMR patterns in *A. hydrophila* appear to be shaped by both vertical inheritance and horizontal gene transfer.

The analysis of polymorphisms revealed extensive SNP diversity among strains. SNPs were the most frequent type of variation, consistent with previous genomic studies (Fernández-Bravo and Figueras, 2020). High SNP counts in ZYAH72 and 2359 suggest accelerated evolution or adaptation. Balanced insertion and deletion patterns across strains imply stable genomic architecture with localized hotspots. SNP density patterns aligned with known variable regions and mobile elements. A Population structure was noticed with the BAPS clustering: the three significant groups of strains were identified. This observation is aligned with previous works on population genomics of Aeromonas (Rouli *et al*, 2015). Strong SNP signatures in each BAPS group demonstrated that, probably, clonal populations grew or individual groups adapted to their environment, or both. In the case of in-group SNP homogeneity, it suggests that there is recent shared descent whereas between-group disparity suggests divergence that took place over a long period of time. These differences were visually confirmed in the matrix of SNPs color-coded. BAPS analysis completes the picture on phylogenetic and pangenome analysis complementing them with the population-level structure. Genotypes can be linked with resistance, virulence or a trait pertaining to epidemiology by using such clusters (Cheng *et al*, 2013). Together, these data provide strong evidence for heterogeneity and sub structuring within *A. hydrophila*. The combination of SNP and gene-content-based methods enhances resolution. This multi-layered approach is critical for understanding the evolutionary dynamics of emerging aquatic pathogens.

Declaration of Competing Interest

The authors declare that they have no competing or conflict of interests.

Author Contributions

AP: Conceptualization, Methodology, Writing—original draft preparation. **Z:** Methodology, Formal analysis, Writing—original draft preparation. **MOS:** Methodology, Formal analysis, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

References

1. Abdella, B., N. A. Abozahra, N. M. Shokrak, R. A. Mohamed, and E. R. El-Helow. (2023). *Whole spectrum of Aeromonas hydrophila virulence determinants and the identification of novel SNPs using comparative pathogenomics*. Sci Rep 13(1):7712. doi: 10.1038/s41598-023-34887-1
2. Abram, K. Z., S. R. Jun, and Z. Udaondo. (2022). *Pseudomonas aeruginosa Pangenome: Core and Accessory Genes of a Highly Resourceful Opportunistic Pathogen*. Adv Exp Med Biol 1386:3-28. doi: 10.1007/978-3-031-08491-1_1
3. Ali, A., A. Naz, S. C. Soares, M. Bakhtiar, S. Tiwari, S. S. Hassan, F. Hanan, R. Ramos, U. Pereira, D. Barh, H. C. Figueiredo, D. W. Ussery, A. Miyoshi, A. Silva, and V. Azevedo. (2015). *Pan-genome analysis of human gastric pathogen H. pylori: comparative genomics and pathogenomics approaches to identify regions associated with pathogenicity and prediction of potential core therapeutic targets*. Biomed Res Int 2015:139580. doi: 10.1155/2015/139580
4. Awan, F., Y. Dong, J. Liu, N. Wang, M. H. Mushtaq, C. Lu, and Y. Liu. (2018). *Comparative genome analysis provides deep insights into Aeromonas hydrophila taxonomy and virulence-related factors*. BMC Genomics 19(1):712. doi: 10.1186/s12864-018-5100-4
5. Beaz-Hidalgo, R., F. Latif-Eugenín, M. J. Hossain, K. Berg, R. M. Niemi, J. Rapala, C. Lyra, M. R. Liles, and M. J. Figueras. (2015). *Aeromonas aquatica sp. nov., Aeromonas finlandiensis sp. nov. and Aeromonas lacus sp. nov. isolated from Finnish waters associated with cyanobacterial blooms*. Syst Appl Microbiol 38(3):161-168. doi: 10.1016/j.syapm.2015.02.005
6. Cheng, L., T. R. Connor, J. Sirén, D. M. Aanensen, and J. Corander. (2013). *Hierarchical and Spatially Explicit Clustering of DNA Sequences with BAPS Software*. Molecular Biology and Evolution 30(5):1224-1228. doi: 10.1093/molbev/mst028 %J Molecular Biology and Evolution
7. Colston, S. M., M. S. Fullmer, L. Beka, B. Lamy, J. P. Gogarten, and J. Graf. (2014). *Bioinformatic Genome Comparisons for Taxonomic and Phylogenetic Assignments Using Aeromonas as a Test Case*. mBio 5(6):e02136. doi: doi:10.1128/mbio.02136-14
8. Fernández-Bravo, A., and M. J. Figueras. (2020). *An Update on the Genus Aeromonas: Taxonomy, Epidemiology, and Pathogenicity*. Microorganisms 8(1)doi: 10.3390/microorganisms8010129
9. Gauthier, J., A. T. Vincent, S. J. Charette, and N. Derome. (2017). *Strong Genomic and Phenotypic Heterogeneity in the Aeromonas sobria Species Complex*. Front. Microbiol. 8:2434. doi: 10.3389/fmicb.2017.02434
10. Godoy, M., M. Montes de Oca, R. Suarez, A. Martinez, J. P. Pontigo, D. Caro, K. Kusch, Y. Coca, H. Bohle, S. Bayliss, M. Kibenge, and F. Kibenge. (2023). *Genomics of Re-Emergent Aeromonas salmonicida in Atlantic Salmon Outbreaks*. Microorganisms 12(1)doi: 10.3390/microorganisms12010064
11. Guo, Y., C. Zeng, C. Ma, H. Cai, X. Jiang, S. Zhai, X. Xu, and M. Lin. (2022). *Comparative genomics analysis of the multidrug-resistant Aeromonas hydrophila MX16A providing insights into antibiotic resistance genes*. Front Cell Infect Microbiol 12:1042350. doi: 10.3389/fcimb.2022.1042350
12. Janda, J. M., and L. Abbott Sharon. (2010). *The Genus Aeromonas: Taxonomy, Pathogenicity, and Infection*. Clinical Microbiology Reviews 23(1):35-73. doi: 10.1128/cmr.00039-09
13. Jin, L., Y. Chen, W. Yang, Z. Qiao, and X. Zhang. (2020). *Complete genome sequence of fish-pathogenic Aeromonas hydrophila HX-3 and a comparative analysis: insights into virulence factors and quorum sensing*. Sci Rep 10(1):15479. doi: 10.1038/s41598-020-72484-8
14. Lawal, O. U., N. Bryan, V. R. Parreira, R. Anderson, Y. Chen, M. Precious, and L. Goodridge. (2024). *Phylogenomics of novel clones of Aeromonas veronii recovered from a freshwater lake reveals unique biosynthetic gene clusters*. Microbiol Spectr 12(12):e0117124. doi: 10.1128/spectrum.01171-24
15. Ma, J., H. Zhao, S. Mo, J. Li, X. Ma, Y. Tang, H. Li, and Z. Liu. (2023). *Acquisition of Type I methyltransferase via horizontal gene transfer increases the drug resistance of Aeromonas veronii*. Microb Genom 9(9)doi: 10.1099/mgen.0.001107
16. Martinez-Murcia, A. J., A. Monera, M. J. Saavedra, R. Oncina, M. Lopez-Alvarez, E. Lara, and M. J. Figueras. (2011). *Multilocus phylogenetic analysis of the genus Aeromonas*. Syst Appl Microbiol 34(3):189-199. doi: 10.1016/j.syapm.2010.11.014
17. Miranda-Lopez, D. C., E. Perez-Rueda, J. Rojas-Vargas, C. H. Cortez, A. Saldana-Padilla, H. G. Castelan-Sanchez, and G. Castro-Escarpulli. (2024). *Comprehensive comparative analysis of the periodontal pathogen Porphyromonas gingivalis: exploring the pan-genome, the reconstruction of the gene regulatory network and genome-scale metabolic network*. Lett Appl Microbiol 77(5)doi: 10.1093/lambio/ovae048
18. Nokhwal, A., R. K. Vaid, T. Anand, R. Verma, and R. Gulati. (2025). *Aeromonas Species Diversity, Virulence Characteristics, and Antimicrobial Susceptibility Patterns in Village Freshwater Aquaculture Ponds in North India*. Antibiotics (Basel) 14(3)doi: 10.3390/antibiotics14030294
19. Piotrowska, M., and M. Popowska. (2015). *Insight into the mobilome of Aeromonas strains*. Front Microbiol 6:494. doi: 10.3389/fmicb.2015.00494
20. Rasko, D. A., M. Rosovitz, G. S. Myers, E. F. Mongodin, W. F. Fricke, P. Gajer, J. Crabtree, M. Sebaihia, N. R. Thomson, and R. Chaudhuri. (2008). *The pangenome structure of Escherichia coli: comparative genomic analysis of E. coli commensal and pathogenic isolates*. Journal of bacteriology 190(20):6881-6893.

21. Rouli, L., V. Merhej, P. E. Fournier, and D. Raoult. (2015). *The bacterial pangenome as a new tool for analysing pathogenic bacteria*. New Microbes and New Infections 7:72-85. doi: https://doi.org/10.1016/j.nmni.2015.06.005
22. Sakurai, A., M. Suzuki, K. Hayashi, and Y. Doi. (2024). *Taxonomic classification of genus Aeromonas using open reading frame-based binarized structure network analysis*. Fujita Med J 10(1):8-15. doi: 10.20407/fmj.2023-007
23. Sayyaf Dezfuli, B., M. Lorenzoni, A. Carosi, L. Giari, and G. Bosi. (2023). *Teleost innate immunity, an intricate game between immune cells and parasites of fish organs: who wins, who loses*. Front. Immunol. 14:1250835. doi: 10.3389/fimmu.2023.1250835
24. Soares, S. C., A. Silva, E. Trost, J. Blom, R. Ramos, A. Carneiro, A. Ali, A. R. Santos, A. C. Pinto, C. Diniz, E. G. Barbosa, F. A. Dorella, F. Aburjaile, F. S. Rocha, K. K. Nascimento, L. C. Guimaraes, S. Almeida, S. S. Hassan, S. M. Bakhtiar, U. P. Pereira, V. A. Abreu, M. P. Schneider, A. Miyoshi, A. Tauch, and V. Azevedo. (2013). *The pan-genome of the animal pathogen Corynebacterium pseudotuberculosis reveals differences in genome plasticity between the biovar ovis and equi strains*. PLoS One 8(1):e53818. doi: 10.1371/journal.pone.0053818
25. Truong, N. H. M., Q. Nguyen, P. V. Voong, V. Chau, N. H. T. Nguyen, T. H. M. Nguyen, P. H. Vo, L. T. Nguyen, T. T. P. Ha, L. P. H. Nguyen, P. H. Le, D. P. Thanh, and H. D. Nguyen. (2024). *Genomic characterization of Aeromonas spp. isolates from striped catfish with motile Aeromonas septicemia and human bloodstream infections in Vietnam*. Microb Genom 10(5)doi: 10.1099/mgen.0.001248
26. Yang, S., T. He, J. Sun, and S. Sun. (2019). *Distinct Antimicrobial Resistance Profiling Of Clinically Important Aeromonas Spp. In Southwest China: A Seven-Year Surveillance Study*. Infection and Drug Resistance 12(null):2971-2978. doi: 10.2147/IDR.S216926
27. Zhong, C., M. Han, P. Yang, C. Chen, H. Yu, L. Wang, and K. Ning. (2019). *Comprehensive Analysis Reveals the Evolution and Pathogenicity of Aeromonas, Viewed from Both Single Isolated Species and Microbial Communities*. mSystems 4(5)doi: 10.1128/mSystems.00252-19