Defining *Aeromonas* species boundaries with large-scale average nucleotide identity analyses and genomic phylogeny

Alex Lu[a], Ruochen Wu[a], Ruiting Lan[a], Li Zhang[a]#

[a]School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW 2052, Australia

Running title: Redefining *Aeromonas* species boundaries

Address correspondence to Li Zhang l.zhang@unsw.edu.au

Key words: *Aeromonas*, core genome, ANI, dDDH, species delineation, phylogeny

Abstract

Aeromonas taxonomy has long been complicated by overlapping phenotypic, biochemical, and protein profiles. Here, we establish a robust genome-based framework for *Aeromonas* species delineation. We analyzed genome-wide average nucleotide identity (ANI) across 3,782 genomes representing the 33 currently recognized *Aeromonas* species and identified 95.2% ANI as the species boundary. This threshold defined 34 ANI-based species, comprising 30 corresponding to existing *Aeromonas* species and four potential novel species, while three existing species were merged into closely related species. A genus-level core genome of 673 loci was defined using 392 representative genomes, and core-genome phylogenetic analysis of 1,505 genomes showed consistency with ANI-based clustering. Digital DNA–DNA hybridization (dDDH) values between clusters were ≤ 55%, further supporting species separation. We also found that nine *Aeromonas* species had intraspecies dDDH values below 70%, cautioning the use of this cutoff for species delineation. We provide the tool *Aeromonas* ANI species typer (AeromonasANIStyper), an ANI-based species typing tool that assigns query genomes using ANI similarity to medoid genomes and achieved 100% accuracy. This framework supports standardized genome-based *Aeromonas* taxonomy and future surveillance.

Importance

Aeromonas species have gained increased attention as emerging human enteric pathogens. *Aeromonas* taxonomy has long been complicated by overlapping phenotypic, biochemical, and protein profiles. Although an average nucleotide identity (ANI) threshold was proposed previously for the *Aeromonas* species delineation, the growing number and diversity of genomes requires a more rigorous reassessment of species boundaries. By analyzing 3,782 *Aeromonas*

genomes, we identified 95.2 % ANI as an optimal cutoff for defining *Aeromonas* species. Using this threshold, we delineated 34 ANI-defined species, which were further validated by core genome phylogeny and digital DNA-DNA hybridization (dDDH). In addition, we clarified previously misidentified species. Importantly, we provide the *Aeromonas* ANI species typer (AeromonasANIStyper), a tool that accurately assigns *Aeromonas* genomes to ANI-defined species. These findings clarify *Aeromonas* species identification and support improved pathogen surveillance, evolutionary and epidemiological studies.

Introduction

The genus *Aeromonas* consists of highly diverse group of bacteria that occupy a wide range of aquatic, environmental and host-associated niches (1). *Aeromonas hydrophila* was recognized as a highly virulent opportunistic pathogen responsible for severe disease outbreaks in fish, resulting in substantial economic losses in global aquaculture (2). Several other *Aeromonas* species have also been identified as causative agents of disease outbreaks in fish and other farmed animals, highlighting their ecological diversity and pathogenic breadth of this genus (3, 4). Beyond their importance in animal health, *Aeromonas* species are also recognized as human pathogens. *Aeromonas* species are well-documented causes of human soft tissue, wound, and bloodstream infections, and their role in human gastrointestinal diseases is increasingly recognized (5, 6). Human infections are often associated with contaminated water, food or aquatic animals and can be especially severe in immunocompromised individuals, highlighting the public health relevance of the genus (5, 7, 8).

The taxonomy of *Aeromonas* has been challenging due to high conservation of the 16S rRNA gene, overlap of phenotypic, biochemical, and protein profiles, which has led to unstable species assignment and frequent misidentification in routine laboratories (9-11). With decreasing costs of whole genome sequencing, average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) have been widely adopted for defining bacterial species. In general, prokaryotic strains with ANI similarity of 95% to 96% are considered to belong to the same species, a threshold that broadly corresponds to the traditional DDH cutoff of 70% for species delineation (12-15). However, studies evaluating ANI, dDDH and genome-scale phylogeny at the genus level for *Aeromonas* species are limited. Colston et al. in 2014 examined 58

Aeromonas genomes and recommended 96% ANI threshold as a practical species for the genus of *Aeromonas* (16). In recent years, the availability of *Aeromonas* whole genome sequences has increased dramatically, encompassing broader ecological niches and clinical cases.

This expanded genomic dataset makes it essential to reassess whether the current species assignments remain valid within *Aeromonas*. A stable genome-based species framework is not only a fundamental requirement for resolving *Aeromonas* taxonomy, but also critical for accurate disease surveillance. Increasing clinical and epidemiological evidence supports the growing importance of *Aeromonas* as human enteric pathogens, underscoring the need to move beyond genus-level classification. Precise species-level identification is necessary because distinct *Aeromonas* species differ substantially in virulence factors, ecological niches, and pathogenic potential, with direct implications for risk assessment, and outbreak investigation (17-19).

In this study, we defined the optimal ANI boundary for species delineation in *Aeromonas* species through analyses of 3,782 *Aeromonas* genomes. The identified ANI threshold is supported by the genus core genome phylogeny and dDDH values. Furthermore, we developed a tool that accurately assigns query genomes to ANI-defined species. This study establishes an updated framework that supports future genome-based identification and pathogen surveillance within *Aeromonas* genus.

Materials and Methods

Aeromonas genome retrieval and establishment of analysis datasets

Assembled *Aeromonas* genomes were downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/) (20). Additionally, publicly available paired end reads of *Aeromonas* were retrieved from the European Nucleotide Archive and assembled using the Shovill v1.1.0 pipeline (https://github.com/tseemann/shovill). The program processed the raw reads using Trimmomatic, Lighter, Fast Adjustment of Short reads (Flash) and assembled them using strategic K-mer extension for scrupulous assemblies (SKESA) (21-24). The quality of all genomes was then assessed for outlying assembly metrics in total assembly length, contig number, N50 and the checkM metrics of completeness and contamination (25).

ANI-based genome comparison and clustering

ANI values were generated using fastANI v1.33 and compiled into a dissimilarity matrix using an in-house python script (12). Genomes were clustered using the average linkage clustering algorithm from SciPy hierarchical clustering package (26). Clusters formed at ANI cutoffs ranging from 94% to 96.5%, with increments of 0.1%, were examined to identify the optimal ANI cutoff for species delineation.

Medoid genome identification

The pairwise dissimilarity matrix of ANI values and the clusters formed at the different ANI cutoffs were used as input for the Partition Around Medoids (PAM) algorithm. The FasterPAM algorithm in the package Fast k-medoids clustering in Python at default setting was used to identify the medoid genomes of each ANI cluster (27). For each cluster, the medoid genome was identified as the genome with the lowest dissimilarity score to all other members of the cluster.

***Aeromonas* genus core gene identification**

PROKKA was used to locate open-reading frames from the core genome dataset, from which the resultant GFF files were input into Roary v3.13 to define the *Aeromonas* genus core loci (28, 29). With consideration for genus diversity, Roary v3.13 was run at multiple BLASTp percentage identities to identify potential core genes. The presence of core loci in each genome was determined by identifying genes with ≥ 85% minimum BLASTp percentage identity between genomes. Core genes that were paralogous or orthologous but were split into multiple fragments by erroneous assembly and/or annotation were identified and subsequently removed using a previously described Python script (30).

Phylogenetic analysis based on genus core genome alignments

Allele sequences of the *Aeromonas* genus core genes were called from genomes representing all ANI-defined species. The individual alleles were then constructed into core genomes sequences by concatenating alignments of individual core loci. To reduce computational complexity, a maximum of 200 genomes per ANI-defined species was included, resulting in a total of 1505 genomes used for phylogenetic analysis. A phylogenetic tree was constructed using IQTREE v3.0.1 (31) based on the core genome sequences. Branch supports were assessed with non-parametric bootstraps at 1,000 replicates. The resulting phylogenetic trees were visualized using GrapeTree v1.5.0 (32).

dDDH analysis

dDDH values were calculated to further evaluate species boundaries defined by ANI. DDH analysis was through the GGDC web service tool with the recommended values of Formula 2 taken as output (33). For interspecies comparisons, one medoid genome representing each ANI-defined species was selected, and pairwise dDDH values were computed between medoid genomes. These comparisons were used to assess genomic separation between ANI-defined species. To evaluate the reliability of the commonly applied 70% dDDH threshold for species delineation in the genus *Aeromonas*, dDDH values of representative strains within the same species were also examined.

Development of an *Aeromonas* genomic species typing tool

The tool *Aeromonas* Genomic Species typer (AeromonasANIStyper) was developed to assign *Aeromonas* species to query genomes based on their ANI similarity to medoid genomes found at

the 95.2% ANI cutoff. The python program utilizes fastANI v1.33 and can be found at (https://github.com/lizhanglab/aeromonasanistyper) (12).

Results

Aeromonas genome dataset

Aeromonas genomes were downloaded, assembled and quality checked (Supplementary Table 1). A total of 3,782 genomes, consisting of 302 complete genomes, 78 chromosome level genomes and 3402 draft genomes were in the final dataset used in this study, representing 33 currently recognized *Aeromonas* species in the NCBI taxonomy database. The genome list can be viewed in Supplementary Table 2.

Determining optimal ANI cutoff for *Aeromonas* species delineation

To determine an optimal ANI cutoff for species delineation, 3,782 *Aeromonas* genomes were compared pairwise using ANI. The clusters were then generated at increments of 0.1% between 94% to 96.5% ANI. An ANI between 95% to 96% is recommended as a cutoff for species delineation in prokaryotes (13). The analysis showed that the previously recommended 96% ANI cutoff is too stringent option for *Aeromonas* species, as this threshold divided the 3782 *Aeromonas* genomes into 53 clusters (Figure 1). Instead, a stable plateau between ANI 94.5% to 95.2%, consistently defined 34 ANI-based species, approximating the 33 traditionally defined species with some differences as detailed below. Therefore, we selected 95.2%, the highest value within the plateau, as the ANI cutoff for species delineation (Figure 1).

Aeromonas species defined at 95.2% ANI

Using the threshold of 95.2% ANI, 34 species were defined among the 3782 *Aeromonas* genomes, with a representative medoid genome identified for each species (Table 1). Of these, 16 showed direct one to one correspondence with currently recognized *Aeromonas* species,

whereas 14 predominantly contained the genomes from one currently recognised species, demonstrating broad consistency between ANI-based delineation and existing taxonomy. For these 30 ANI defined species, we assigned the name of the sole traditional species or the predominant species in that ANI-defined species to maintain consistency and continuity with traditional species names. However, several ANI-defined species contained genomes annotated as different species in the NCBI database, indicating frequent misidentification or ambiguous classification under previous species identification and assignment practices (Table 1).

Four ANI-based species did not match any existing *Aeromonas* species and likely represent novel species. We provisionally designated them as *Aeromonas redevero*, *Aeromonas ichironi*, *Aeromonas rubrola*, and *Aeromonas olawany*. *A. redevero* contained 11 genomes and the remaining three putative species consisted of a single genome each (Table 1).

In addition, three previously defined species were not supported as distinct ANI-based species, instead clustering into closely related species. Specifically, *Aeromonas lacus* (two genomes) was grouped with *Aeromonas jandaei*, *Aeromonas piscicola* (three genomes) with *Aeromonas bestiarum*, and *Aeromonas mytilicola* (two genomes) with *Aeromonas rivipollensis* (Table 1). Further details for each of the 3782 analysed *Aeromonas* genomes, including those previously misidentified, are provided in Supplementary Table 2.

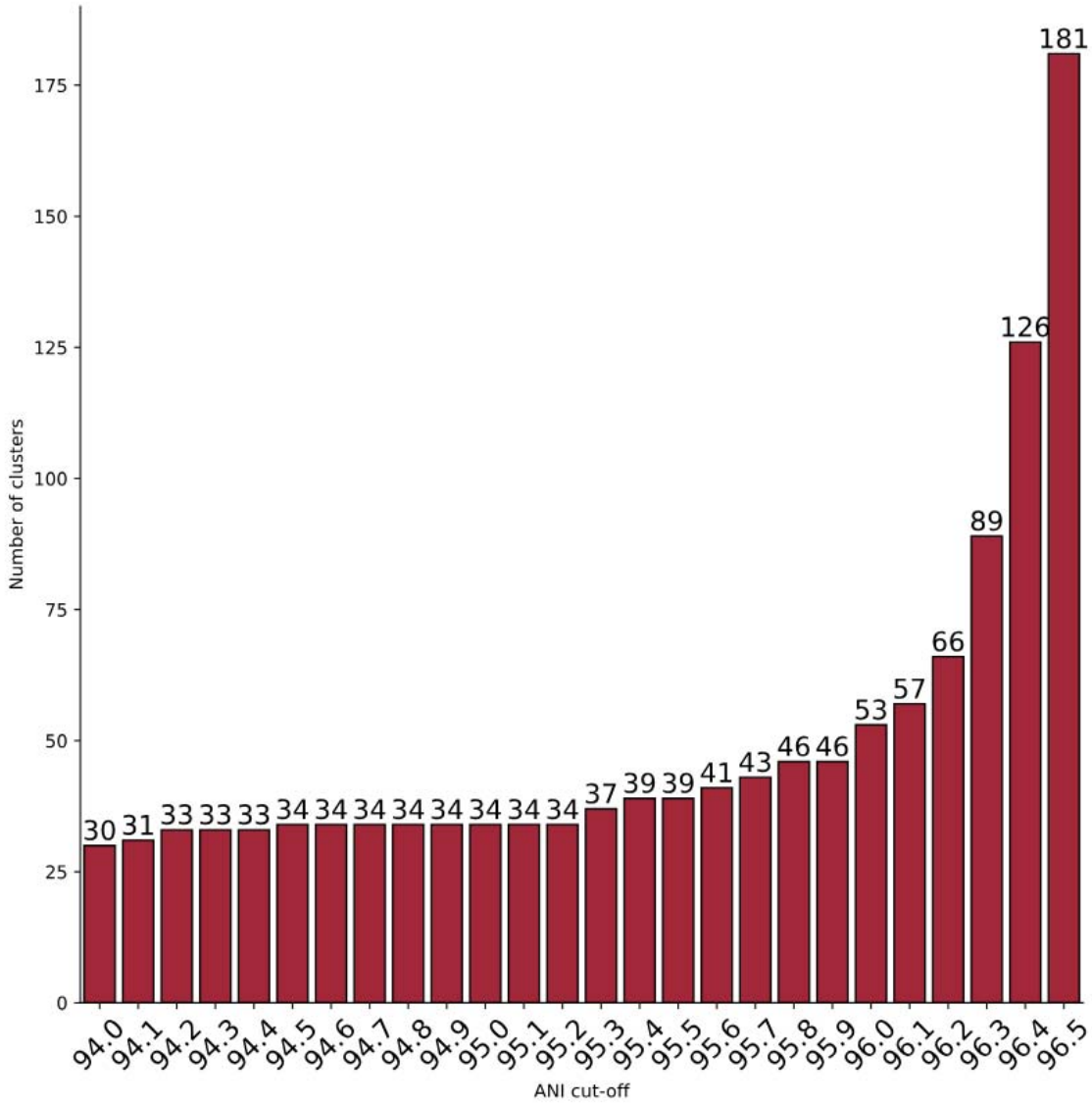

Figure 1. Number of genome clusters defined at different average nucleotide identity (ANI) cutoffs

A total of 3782 *Aeromonas* genomes were analysed, comprising 33 existing NCBI species. Genomes were clustered based on pairwise ANI comparison. The x-axis shows ANI cutoffs ranging from 94% to 96.5%, whilst the y-axis shows the number of ANI clusters. A stable plateau of 34 ANI clusters is observed between 94.5% and 95.2% ANI. Within this plateau, the highest value 95.2% ANI was selected as the threshold for species delineation.

Table 1. Delineation of 34 species from 3,782 *Aeromonas* genomes using a 95.2% average nucleotide identity (ANI) cutoff

ANI cluster	Total genomes	Species currently defined in NCBI database (N)	ANI-defined species assignment	Medoid Genome Accession
1	6	*Aeromonas simiae* (6)	*Aeromonas simiae*	GCF_030514525.1
2	2	*Aeromonas diversa* (2)	*Aeromonas diversa*	GCF_040790265.1
3	11	*Aeromonas schubertii* (11)	*Aeromonas schubertii*	DRR293937
4	13	*Aeromonas bivalvium* (13)	*Aeromonas bivalvium*	GCF_003265565.1
5	4	*Aeromonas molluscorum* (4)	*Aeromonas molluscorum*	SRR32477519
6	4	*Aeromonas rivuli* (4)	*Aeromonas rivuli*	GCF_024809075.1
7	68	*Aeromonas enteropelogenes* (62), *Aeromonas sobria* (1), Unspecified Species (5)	*Aeromonas enteropelogenes*	ERR4559376
8	58	*Aeromonas jandaei* (54), *Aeromonas lacus* (2), *Aeromonas veronii* (1), Unspecified Species (1)	*Aeromonas jandaei*	GCF_034721165.1
9	24	*Aeromonas sobria* (24)	*Aeromonas sobria*	SRR24219967
10	60	*Aeromonas allosaccharophila* (46), *Aeromonas veronii* (11), *Aeromonas hydrophila* (1), Unspecified Species (2)	*Aeromonas allosaccharophila*	SRR29645202
11	11	*Aeromonas veronii* (10), Unspecified Species (1)	*Aeromonas redevero**	SRR15464897
12	754	*Aeromonas veronii* (710), *Aeromonas hydrophila* (2), *Aeromonas sobria* (1), Unspecified Species (41)	*Aeromonas veronii*	SRR11281514
13	1	*Aeromonas veronii* (1)	*Aeromonas olawany**	SRR13288071
14	1	*Aeromonas australiensis* (1)	*Aeromonas australiensis*	ERR660686
15	1	*Aeromonas fluvialis* (1)	*Aeromonas fluvialis*	ERR660681
16	1	*Aeromonas finlandensis* (1)	*Aeromonas finlandensis*	GCF_041929505.1
17	537	*Aeromonas salmonicida* (509), *Aeromonas caviae* (1), Unspecified Species (27)	*Aeromonas salmonicida*	SRR26182261
18	5	*Aeromonas popoffii* (5)	*Aeromonas popoffii*	GCF_018138755.1
19	21	*Aeromonas bestiarum* (17),	*Aeromonas bestiarum*	ERR14335777

		Aeromonas piscicola (3), *Aeromonas hydrophila* (1)		
20	238	*Aeromonas dhakensis* (222), *Aeromonas hydrophila* (11), *Aeromonas bestiarum* (1), *Aeromonas veronii* (1), Unspecified Species (3)	*Aeromonas dhakensis*	ERR12867583
21	412	*Aeromonas hydrophila* (403), *Aeromonas dhakensis* (5), *Aeromonas bestiarum* (2), *Aeromonas enteropelogenes* (1), Unspecified Species (1)	*Aeromonas hydrophila*	SRR3656082
22	71	*Aeromonas media* (66), Unspecified Species (5)	*Aeromonas media*	ERR595939
23	64	*Aeromonas rivipollensis* (33), *Aeromonas media* (25), *Aeromonas mytilicola* (2), *Aeromonas hydrophila* (1), Unspecified Species (3)	*Aeromonas rivipollensis*	GCF_030386695.1
24	1336	*Aeromonas caviae* (1206), *Aeromonas hydrophila* (4), *Aeromonas media* (1), *Aeromonas veronii* (1), Unspecified Species (124)	*Aeromonas caviae*	ERR4559438
25	26	*Aeromonas taiwanensis* (21), *Aeromonas hydrophila* (4), Unspecified Species (1)	*Aeromonas taiwanensis*	GCF_004524025.1
26	36	*Aeromonas sanarellii* (28), *Aeromonas hydrophila* (1), Unspecified Species (7)	*Aeromonas sanarellii*	SRR31059843
27	1	*Aeromonas tecta* (1)	*Aeromonas tecta*	GCF_000820185.1
28	1	*Aeromonas lusitana* (1)	*Aeromonas lusitana*	GCF_002812985.1
29	7	*Aeromonas encheleia* (7)	*Aeromonas encheleia*	GCF_900637545.1
30	2	*Aeromonas eucrenophila* (2)	*Aeromonas eucrenophila*	GCF_000819865.1
31	2	*Aeromonas aquatica* (2)	*Aeromonas aquatica*	GCF_022549315.1
32	1	*Aeromonas veronii* (1)	*Aeromonas ichironi**	GCF_040215105.1
33	1	*Aeromonas cavernicola* (1)	*Aeromonas cavernicola*	GCF_002795305.1
34	1	Unspecified Species (1)	*Aeromonas rubrola**	SRR5194379

* Novel ANI-defined species with proposed new species names. The three previously recognized species *A. lacus*, *A. piscicola*, and *A. mytilicola* were merged into *A. jandaei*, *A.bestiarum*, and *A. rivipollensis*, respectively. N: number of genomes

Identification of the *Aeromonas* genus core genome

We conducted genus-level core genome analysis to further verify the 34 ANI-defined species, using 392 genomes representing the 33 *Aeromonas* species currently recognized in the NCBI database. The 392 genomes comprised of 301 complete genomes, 78 chromosome level genomes and 13 draft genomes. With the *A. hydrophila* strain ATCC 7966 as the reference (4,490 genes), 705 genes initially passed the core genome selection criteria. Subsequently, 28 paralogous genes and 4 genes with overlapping regions were identified and removed. A final *Aeromonas* genus core genome comprising 673 genes was established and used for phylogenetic analysis.

The genus core genome phylogenetic analysis supports ANI-defined *Aeromonas* species

A core-genome phylogenetic tree was constructed using 1505 *Aeromonas* genomes representing the 34 ANI-defined species. As shown in Figure 2, genomes assigned to the same ANI-defined species clustered together at bootstrap values of 100, supporting the ANI-based species delineation. Genomes previously classified as *A. lacus*, *A. piscicola*, and *A. mytilicola* did not form distinct clades in the genus core-genome phylogeny, further supporting their reassignment to closely related species.

The overall features of the core genome phylogenetic tree are consistent with previously reported findings. For example, *A. hydrophila* and *A. dhakensis* formed a common clade, consistent with reclassification of *A. dhakensis* as a separate species from *A. hydrophila* subsp. *dhakensis* (34). *A. media* and *A. rivipollensis* are closely related and are collectively referred to as the *A. media* species complex (35).

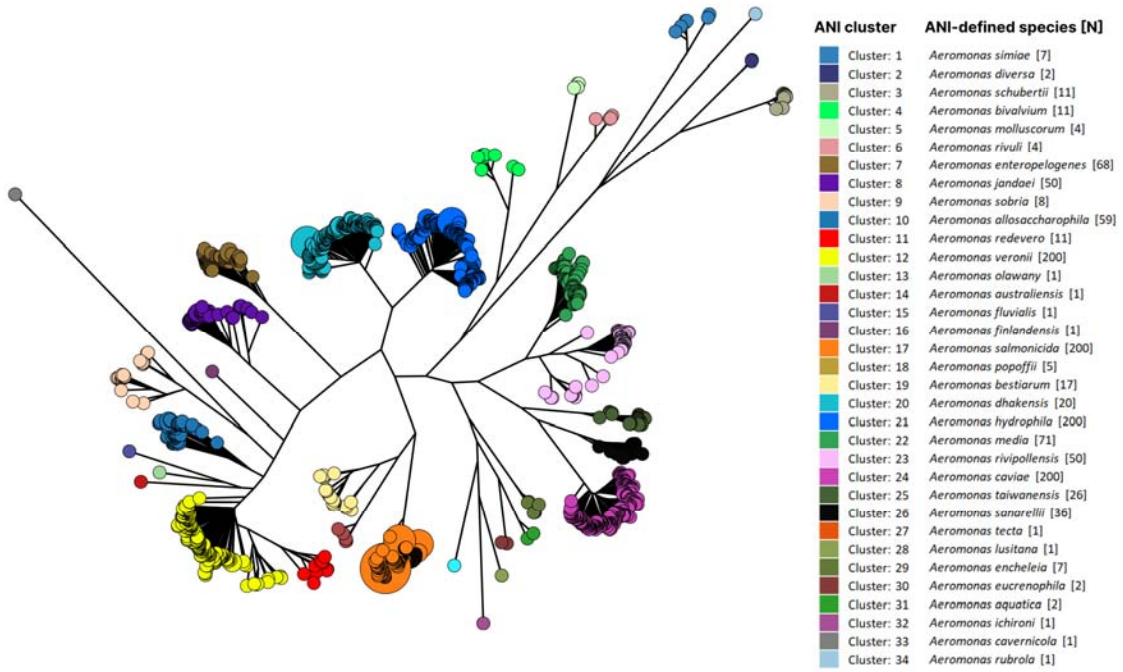

Figure 2. Core genome phylogenetic tree of the 34 *Aeromonas* species defined at a 95.2% average nucleotide identity (ANI) cutoff

A total of 1505 genomes representing the 34 *Aeromonas* species defined at 95.2% ANI cutoff were analysed. The maximum likelihood tree was constructed using concatenated alignments of 673 *Aeromonas* genus core genes as input for IQ-TREE v3.0.1 and visualized with GrapeTree v1.5.0. Branch support was assessed using 1000 ultrafast bootstrap replicates, which showed values of 100 were all higher than at each ANI species node. Each genomes core alignment length was greater than 370 kb. Numbers in brackets following each species name indicate the number of genomes included in the analysis. N: number of genomes.

The dDDH analysis supports ANI-defined species

Using the representative medoid genomes for each species defined at 95.2% ANI, pairwise dDDH results were calculated using the GGDC (Formula 2). Comparisons of dDDH values between ANI-based species were consistently low, with most below 30. The highest dDDH value between ANI species medoid genomes was 55% (Figure 3). These results showed the dDDH values between ANI-defined species were well below the traditional DDH cutoff of 70% for species, providing further support for species delineation based on 95.2% ANI threshold.

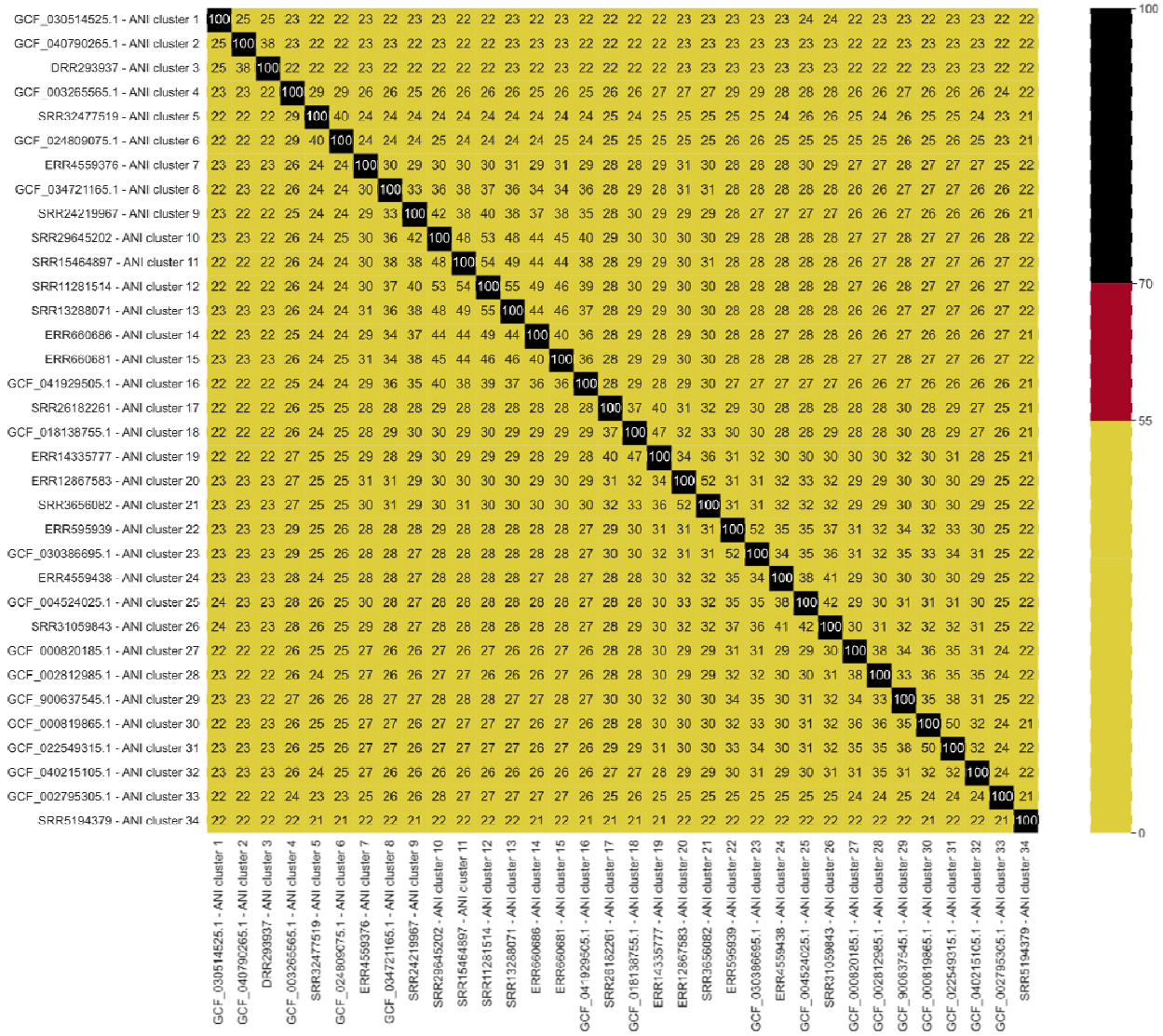

Figure 3. Matrix of digital DNA-DNA hybridization (dDDH) comparisons for the 34 species defined at a 95.2% average nucleotide identity (ANI) cutoff

The representative medoid genome of each species was used for pairwise comparison. Calculations of dDDH values were made using Formula of the GGDC web service and are shown as a colour coded matrix. The interspecies dDDH values are consistently low, with a maximum of 55%.

***Aeromonas* species with intraspecies dDDH values below 70%**

The 70% cutoff for dDDH originates from its proposal for use in traditional DDH to demarcate bacterial species (33). Several recent studies proposed new *Aeromonas* species using the 70% dDDH cutoff as key evidence (36-38). However, these newly proposed species such as *A. lacus* and *A. mytilicola* were not supported by 95.2% ANI and core genome phylogeny in this study. This discrepancy prompted us to investigate the dDDH values within a *Aeromonas* ANI-defined species.

Representative genomes within *Aeromonas* ANI-defined species were compared pairwise using dDDH (GGDC formula 2). Nine species exhibited intraspecies dDDH below 70%, including *A. allosaccharophila*, *A. bestiarum*, *A. bivalvium*, *A. hydrophila*, *A. jandaei*, *A. rivipollensis*, *A. redevero*, *A. veronii*, and *A. sobria* (Supplementary Figure 2). Additionally, all intraspecies dDDH values were above 55%, which was the maximum interspecies dDDH observed (Figure 3). Eight of the nine species were consistently defined by conventional taxonomy and ANI-defined species delineation, whereas *A. redevero* was defined in this study based on 95.2% ANI. These results suggest that the traditional 70% dDDH cutoff may not be universally applicable to *Aeromonas* species delineation and should be interpreted with caution in this genus.

Development of AeromonasANIStyper for identification of *Aeromonas* ANI species

To facilitate the identification of ANI-defined species for newly sequenced genomes, we developed a python program, AeromonasANIStyper. This program utilizes fastANI to compare query *Aeromonas* genomes to the representative medoid genome of each of the 34 ANI-species defined at the 95.2% ANI threshold. Assignment of the query genome to an ANI-based species was based on the highest ANI similarity to a reference medoid genome, provided that this value exceeds 95.2%. Genomes with less than 95.2% ANI to all medoid genomes are flagged as potentially novel species. When compared to the dataset of 3,782 genomes, 100% were assigned correctly to their initial ANI-based species.

Recently described *Aeromonas oralensis* belongs to *A. hydrophila* species

A study in 2025 proposed species *Aeromonas oralensis*, consisting of a single strain AB005 (37). The status of a novel species was proposed because this strain had a dDDH value less than 70% in comparison to a reference *A. hydrophila* genome. When the genome of *A. oralensis* strain AB005 was compared using AeromonasANIStyper, the closest species was *A. hydrophila* with a 96.4% ANI value. Further dDDH comparisons (GGDC Formula 2) against genomes representing *A. hydrophila* intraspecies dDDH clusters (Supplementary Figure 1) revealed a 79.3% dDDH value in comparison to GCF_026651505.1. Both dDDH and ANI values support that the proposed *A. oralensis* is a member of *A. hydrophila* rather than a distinctive species.

Discussion

In this study, through analysis of 3872 *Aeromonas* genomes, we identified an optimal ANI threshold of 95.2% for species delineation within the *Aeromonas* genus. This threshold is supported by genus-level core-genome phylogeny and dDDH values. We have also resolved several previously misidentified species and strains. To facilitate use of ANI, we developed AeromonasANIStyper, a tool for rapid ANI-based *Aeromonas* species identification and delineation.

A 96% ANI cutoff for *Aeromonas* species delineation was proposed in 2014, based on analysis of 58 *Aeromonas* genomes (16). However, our analyses of a more comprehensive dataset comprising of 3872 *Aeromonas* genomes demonstrates that this threshold is not appropriate for species delineation within the genus. Application of the 96% ANI cutoff divided the 3872 *Aeromonas* genomes into 53 clusters (ANI-based species), which divides many of the recognized 33 *Aeromonas* species into multiple species.

We observed a stable plateau in the number of genome clusters across ANI thresholds ranging from 94.5% to 95.2%. Based on this stability, we selected 95.2% ANI as the species delineation threshold, as it represents the most stringent value within the plateau. Choosing the upper bound of this range minimizes the risk of merging genomically distinct lineages while preserving stable species assignments.

Core-genome phylogenetic analysis and dDDH comparisons also support the 95.2% ANI framework. First, the genus core-genome phylogeny recovered well separated clades

corresponding to the ANI-defined species. In contrast, three currently recognised species, *A. lacus*, *A. piscicola*, and *A. mytilicola* did not form their own distinct phylogenetic groups and phylogenetically fell well into their respective ANI-defined species. Second, interspecies dDDH values among medoid genomes were consistently well below the traditional 70% cutoff, reinforcing that the 34 clusters represent separated species. The supportive evidence from both genus core genome and dDDH analysis shows that 95.2% ANI threshold provides an accurate and stable framework for *Aeromonas* species delineation.

In the genus core-genome phylogenetic analysis, a representative subset of 1,505 genomes was used instead of the full dataset of 3,872 genomes to reduce computational complexity. These genomes were selected to represent all 34 ANI-defined species and included all genomes previously identified as potentially misclassified. Despite this subsampling, the resulting phylogeny showed strong bootstrap support (>95%) for all interspecies clades, indicating that the inferred species-level relationships are highly robust. This demonstrates that the reduced dataset retains sufficient phylogenetic signal to reliably support the ANI-based species delineation framework.

The 95.2% ANI threshold clarified a considerable number of previously misidentified genomes across the *Aeromonas* genus (Table 1), demonstrating that misclassification is widespread and affects multiple species, particularly those historically assigned to *A. hydrophila*, *A. veronii*, *A. dhakensis*, and related taxa. These revised species assignments are consistent with both genus-level core-genome phylogeny and dDDH analyses. Recent studies have highlighted the growing importance of *Aeromonas* species as causative agents in human gastrointestinal diseases,

underscoring the need for accurate species-level identification in clinical and epidemiological contexts (39, 40). As closely related *Aeromonas* species can differ in pathogenic potential and epidemiological patterns, surveillance at the species level is likely to become increasingly important (5, 17, 41). Misidentification can obscure species-specific disease associations and hinder reliable surveillance. By resolving historical misclassifications, the 95.2% ANI framework provides a standardized and reproducible basis of species identification for future *Aeromonas* surveillance and for comparative studies of species-associated pathogenicity.

We also observed intraspecies dDDH values can be well below the 70% threshold usings representative genomes in nine *Aeromonas* species. We found that the cutoff of 70% has little value in *Aeromonas* species assignment as too many species would be defined. These dDDH ranges were also found in *Sphingobacterium*, indicating that rigid dDDH cutoffs may not fully capture evolutionary relationships (42). In contrast, a 95.2% ANI threshold provides a more robust framework for species delineation in *Aeromonas*, as it is consistently supported by genus-level core genome phylogeny as well as overall dDDH patterns. The observed inconsistencies among dDDH, ANI, and core-genome phylogenetic analyses likely reflect fundamental differences in how these approaches capture genomic relatedness. dDDH incorporates both sequence divergence of homologous regions and differences in overall genome content, including accessory genes and mobile genetic elements such as plasmids. In contrast, ANI and core-genome phylogenies are primarily driven by nucleotide divergence within shared genomic regions and are therefore not influenced by variation in accessory genome composition (12, 15, 30, 33, 42). Discrepancies between a 70% dDDH cutoff, 95-96% ANI, and genome-scale phylogenetic approaches have also been reported in other bacterial taxa such as *Streptomyces*,

Pseudomonas, and *Burkholderia* (15, 43-45). Collectively, these findings suggest that the 70% dDDH cutoff should be applied with caution for *Aeromonas* species delineation as intraspecies dDDH can be well below the cutoff. Our findings underscore the importance of integrating ANI and core genome phylogenetic analyses.

To facilitate ANI-defined *Aeromonas* species delineation, we developed a tool namely AeromonasANIStyper, which assigns ANI-defined species to query genomes by running the fastANI comparisons against the medoid genomes of each of the 34 ANI-defined *Aeromonas* species. Similar approaches have been established for *Bacillus cereus* group and *Campylobacter* species with high accuracy (30, 46). Our tool has achieved 100% accuracy when applied to the 3872 *Aeromonas* genomes. Currently, multilocus sequence typing (MLST) is also available for *Aeromonas* species identification (47). However, MLST assignments are based on existing species delineation and are not suitable for demarcating new species.

In conclusion, we defined 95.2% ANI as the threshold for *Aeromonas* species delineation and validated the 34 ANI-defined species using core-genome phylogenetic and dDDH analyses. The names of 30 traditionally recognised species were maintained with 3 species removed and 4 new species proposed. We also provide a tool for rapid ANI-based *Aeromonas* species identification. This study establishes an updated species framework that supports genome-based *Aeromonas* species identification and facilitates precision public health surveillance of *Aeromonas* infections.

Author Contributions

Alex Lu: Analysis, Validation, Writing – original draft, review and editing.

Peter Ruochen: Methodology, Writing – review and editing.

Ruiting Lan: Methodology, Validation, Writing – review and editing.

Li Zhang: Conceptualisation, Funding acquisition, Supervision, Writing – review & editing.

Acknowledgements

The authors would also like to thank the UNSW ResTech Technology Services Team for their ongoing assistance with the high-powered computing and data management systems. Alex Lu is thankful for the scholarship support by the Australian government and the University of New South Wales. This project was supported by the BABS Research Grant (IR001/BABS/PS46772) awarded to Associate Professor Li Zhang from the University of New South Wales

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Majeed, S, De Silva, L, Kumarage, PM, Heo, GJ. Occurrence of potential virulence determinants in *Aeromonas* spp. isolated from different aquatic environments. J Appl Microbiol. 2023;134(3):lxad031.

2. 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.

3. Yu, Z, Oh, Y, Kim, S, Han, K, Srikulnath, K, Li, Q, et al. Multilocus sequence typing and antibiotic resistance of *Aeromonas* isolated from freshwater fish in Hebei Province. PLoS One. 2024;19(3):e0298745.

4. Liu, J, Xie, L, Zhao, D, Yang, T, Hu, Y, Sun, Z, et al. A fatal diarrhoea outbreak in farm-raised *Deinagkistrodon acutus* in China is newly linked to potentially zoonotic *Aeromonas hydrophila*. Transbound Emerg Dis. 2019;66(1):287-298.

5. Yuwono, C, Wehrhahn, MC, Liu, F, Zhang, L. Enteric *Aeromonas* infection: A common enteric bacterial infection with a novel infection pattern detected in an Australian population with gastroenteritis. Microbiol Spectr. 2023;11(4):e0028623.

6. Essers, B, Burnens, AP, Lanfranchini, FM, Somaruga, SG, von Vigier, RO, Schaad, UB, et al. Acute community-acquired diarrhea requiring hospital admission in Swiss children. Clin Infect Dis. 2000;31(1):192-6.

7. Yuwono, C, Wehrhahn, MC, Liu, F, Riordan, SM, Zhang, L. The isolation of *Aeromonas* species and other common enteric bacterial pathogens from patients with gastroenteritis in an Australian population. Microorganisms. 2021;9(7):1440.

8. Kilikdar, M, Srinivas, J, Chitrans, P, Siraj, S, Ansari, A. Skin and Soft Tissue Infections due to *Aeromonas* spp.: An Emerging Pathogen. JCDR. 2023;17(4):DC5-DC9.

9. Dubey, S, Maiti, B, Girisha, SK, Das, R, Lamkhannat, M, Mutoloki, S, et al. *Aeromonas* species obtained from different farmed aquatic species in India and Taiwan show high phenotypic relatedness despite species diversity. BMC Res Notes. 2021;14(1):313.

10. Morandi, A, Zhaxybayeva, O, Gogarten, JP, Graf, J. Evolutionary and diagnostic implications of intragenomic heterogeneity in the 16S rRNA gene in *Aeromonas* strains. J Bacteriol. 2005;187(18):6561-4.

11. Roger, F, Lamy, B, Jumas-Bilak, E, Kodjo, A, Marchandin, H. Ribosomal multi-operon diversity: an original perspective on the genus *Aeromonas*. PLoS One. 2012;7(9):e46268.

12. Jain, C, Rodriguez, RL, Phillippy, AM, Konstantinidis, KT, Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9(1):5114.

13. Riesco, R, Trujillo, ME. Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol. 2024;74(3).

14. Konstantinidis, KT, Tiedje, JM. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A. 2005;102(7):2567-72.

15. Richter, M, Rosselló-Móra, R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A. 2009;106(45):19126-31.

16. Colston, SM, Fullmer, MS, Beka, L, Lamy, B, Gogarten, JP, Graf, J. Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. mBio. 2014;5(6):e02136.

17. Singh, A, Liu, F, Yuwono, C, Wehrhahn, MC, Slavich, E, Young, AM, et al. Age-Dependent Variations in the Distribution of *Aeromonas* Species in Human Enteric Infections. Pathogens. 2025;14(2):120.

18. Wu, CJ, Ko, WC, Lee, NY, Su, SL, Li, CW, Li, MC, et al. *Aeromonas* Isolates from Fish and Patients in Tainan City, Taiwan: Genotypic and Phenotypic Characteristics. Appl Environ Microbiol. 2019;85(21):e01360-19.

19. Li, T, Raza, SHA, Yang, B, Sun, Y, Wang, G, Sun, W, et al. *Aeromonas veronii* Infection in Commercial Freshwater Fish: A Potential Threat to Public Health. Animals (Basel). 2020;10(4):608.

20. Sayers, EW, Bolton, EE, Brister, JR, Canese, K, Chan, J, Comeau, DC, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022;50(D1):D20-d26.

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

22. Song, L, Florea, L, Langmead, B. Lighter: fast and memory-efficient sequencing error correction without counting. Genome Biol. 2014;15(11):509.

23. Magoč, T, Salzberg, SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011;27(21):2957-63.

24. Souvorov, A, Agarwala, R, Lipman, DJ. SKESA: strategic k-mer extension for scrupulous assemblies. Genome Biol. 2018;19(1):153.

25. Parks, DH, Imelfort, M, Skennerton, CT, Hugenholtz, P, Tyson, GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25(7):1043.

26. Virtanen, P, Gommers, R, Oliphant, TE, Haberland, M, Reddy, T, Cournapeau, D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261-272.

27. Sander, U, Lubbe, N. The potential of clustering methods to define intersection test scenarios: Assessing real-life performance of AEB. Accid Anal Prev. 2018;113:1-11.

28. Seemann, T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068.

29. Page, AJ, Cummins, CA, Hunt, M, Wong, VK, Reuter, S, Holden, MT, et al. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691.

30. Wu, R, Payne, M, Zhang, L, Lan, R. Uncovering the boundaries of *Campylobacter* species through large-scale phylogenetic and nucleotide identity analyses. mSystems. 2024;9(4):e0121823.

31. Wong, T, Ly-Trong, N, Ren, H, Baños, H, Roger, A, Susko, E, et al. IQ-TREE 3: Phylogenomic Inference Software using Complex Evolutionary Models. EcoEvorxiv 8916 [Preprint]. 2025.

32. Zhou, Z, Alikhan, NF, Sergeant, MJ, Luhmann, N, Vaz, C, Francisco, AP, et al. GrapeTree: Visualization of core genomic relationships among 100,000 bacterial pathogens. Genome Res. 2018;28(9):1395.

33. Meier-Kolthoff, JP, Auch, AF, Klenk, HP, Göker, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics. 2013;14:60.

34. Beaz-Hidalgo, R, Martínez-Murcia, A, Figueras, MJ. Reclassification of *Aeromonas hydrophila* subsp. *dhakensis* Huys et al. 2002 and *Aeromonas aquariorum* Martínez-Murcia et al. 2008 as *Aeromonas dhakensis* sp. nov. comb nov. and emendation of the species *Aeromonas hydrophila*. Syst Appl Microbiol. 2013;36(3):171.

35. Talagrand-Reboul, E, Roger, F, Kimper, JL, Colston, SM, Graf, J, Latif-Eugenín, F, et al. Delineation of Taxonomic Species within Complex of Species: *Aeromonas media* and Related Species as a Test Case. Front Microbiol. 2017;8:621.

36. Ajmi, N, Duman, M, Coskun, B, Esen, C, Sonmez, O, Tasci, G, et al. Unraveling Genomic and Pathogenic Features of *Aeromonas ichthyocola* sp. nov., *Aeromonas mytilicola* sp. nov., and *Aeromonas mytilicola* subsp. *aquatica* subsp. nov. Animals (Basel). 2025;15(7).

37. Mashzhan, A, Smekenov, I, Bakiyev, S, Utegenova, K, Samatkyzy, D, Daniyarov, A, et al. Whole-Genome Sequencing of a Potentially Novel *Aeromonas* Species Isolated from Diseased Siberian Sturgeon (*Acipenser baerii*) Using Oxford Nanopore Sequencing. Microorganisms. 2025;13(7).

38. Beaz-Hidalgo, R, Latif-Eugenín, F, Hossain, MJ, Berg, K, Niemi, RM, Rapala, J, et al. *Aeromonas aquatica* sp. nov., *Aeromonas finlandiensis* sp. nov. and *Aeromonas lacus* sp. nov. isolated from Finnish waters associated with cyanobacterial blooms. Syst Appl Microbiol. 2015;38(3):161-8.

39. Pessoa, RBG, de Oliveira, WF, Correia, MTD, Fontes, A, Coelho, L. *Aeromonas* and Human Health Disorders: Clinical Approaches. Frontiers in Microbiology. 2022;13.

40. Yuwono, C, Wehrhahn, M, Slavich, E, Liu, F, Zhang, L. Detection of *Aeromonas*, *Campylobacter* and *Salmonella* using concurrent bacterial culture and commercial multiplex PCR, Sydney, Australia, 2023. Euro Surveill. 2026;31(5):2500355.

41. Moradi, F, Hadi, N, Bazargani, A, Abdi, F, Ghorbanian, N. *Aeromonas* characteristics in Iran, Southwest Asia; a systematic review and meta-analysis on

epidemiology, reservoirs and antibiotic resistance profile from aquatic environments to human society during 2000-2023. BMC Vet Res. 2025;21(1):107.

42. Li, S, Wang, XR, Zhang, W, Li, WJ. Genome-based reassessment of *Sphingobacterium* reveals the limitation of the fixed dDDH threshold for prokaryotic species delineation. Curr Res Microb Sci. 2026;10:100524.

43. Hu, S, Li, K, Zhang, Y, Wang, Y, Fu, L, Xiao, Y, et al. New Insights Into the Threshold Values of Multi-Locus Sequence Analysis, Average Nucleotide Identity and Digital DNA-DNA Hybridization in Delineating *Streptomyces* Species. Front Microbiol. 2022;13:910277.

44. Garrido-Sanz, D, Meier-Kolthoff, JP, Göker, M, Martín, M, Rivilla, R, Redondo-Nieto, M. Genomic and Genetic Diversity within the *Pseudomonas fluorescens* Complex. PLoS One. 2016;11(2):e0150183.

45. Bach, E, Volpiano, CG, Sant'Anna, FH, Passaglia, LMP. Genome-based taxonomy of *Burkholderia* sensu lato: Distinguishing closely related species. Genet Mol Biol. 2023;46(3 Suppl 1):e20230122.

46. Carroll, LM, Wiedmann, M, Kovac, J. Proposal of a taxonomic nomenclature for the *Bacillus* cereus group which reconciles genomic definitions of bacterial species with clinical and industrial phenotypes. mBio. 2020;11(1).

47. Martinez-Murcia, AJ, Monera, A, Saavedra, MJ, Oncina, R, Lopez-Alvarez, M, Lara, E, et al. Multilocus phylogenetic analysis of the genus *Aeromonas*. Syst Appl Microbiol. 2011;34(3):189-99.