

Acquisition of Type I methyltransferase via horizontal gene transfer increases the drug resistance of *Aeromonas veronii*

Jiayue Ma, Honghao Zhao, Shuangyi Mo, Juanjuan Li, Xiang Ma, Yanqiong Tang, Hong Li* and Zhu Liu*

Abstract

Aeromonas veronii is an opportunistic pathogen that affects both fish and mammals, including humans, leading to bacteraemia, sepsis, meningitis and even death. The increasing virulence and drug resistance of *A. veronii* are of significant concern and pose a severe risk to public safety. The Type I restriction-modification (RM) system, which functions as a bacterial defence mechanism, can influence gene expression through DNA methylation. However, little research has been conducted to explore its origin, evolutionary path, and relationship to virulence and drug resistance in *A. veronii*. In this study, we analysed the pan-genome of 233 *A. veronii* strains, and the results indicated that it was 'open', meaning that *A. veronii* has acquired additional genes from other species. This suggested that *A. veronii* had the potential to adapt and evolve rapidly, which might have contributed to its drug resistance. One Type I methyltransferase (MTase) and two complete Type I RM systems were identified, namely AveC4I, AveC4II and AveC4III in *A. veronii* strain C4, respectively. Notably, AveC4I was exclusive to *A. veronii* C4. Phylogenetic analysis revealed that AveC4I was derived from horizontal gene transfer from *Thiocystis violascens* and exchanged genes with the human pathogen *Comamonas kerstersii*. Single molecule real-time sequencing was applied to identify the motif methylated by AveC4I, which was unique and not recognized by any reported MTases in the REBASE database. We also annotated the functions and pathways of the genes containing the motif, revealing that AveC4I may control drug resistance in *A. veronii* C4. Our findings provide new insight on the mechanisms underlying drug resistance in pathogenic bacteria. By identifying the specific genes and pathways affected by AveC4I, this study may aid in the development of new therapeutic approaches to combat *A. veronii* infections.

DATA SUMMARY

The authors confirm all supporting data have been provided within the article or through supplementary data files.

INTRODUCTION

Initially considered to be part of the innate immune system in bacteria, recent studies have shown that restriction-modification (RM) systems can also modify specific DNA sites through methylation, shaping bacterial epigenomes and regulating traits such as virulence and adaptability [1, 2]. RM systems are divided into four types (I, II, III and IV) [3]. The Type I RM system is the most intricate, encoded by three *hsd* (host specificity of DNA) genes, namely *hsdR*, *hsdM* and *hsdS* [4]. The genes *hsdM* and *hsdS* are closely located under the control of the same promoter, while *hsdS* is unique to the Type I RM system [5]. A restriction endonuclease ($R_2M_2S_1$) is composed of two restriction (R) subunits encoded by *hsdR*, two modification (M) subunits encoded by *hsdM* and one specificity (S) subunit encoded by *hsdS*. The M_2S_1 complex of the $R_2M_2S_1$ endonuclease solely functions as a DNA methyltransferase (MTase) [6]. The Type I RM system cleaves nonmethylated foreign DNA using $R_2M_2S_1$, while protecting

Received 05 June 2023; Accepted 14 September 2023; Published 27 September 2023

Author affiliations: ¹School of Life Sciences, Hainan University, Haikou, PR China.

*Correspondence: Hong Li, lihongbio@hainanu.edu.cn; Zhu Liu, zhuliu@hainanu.edu.cn

Keywords: *Aeromonas veronii*; drug resistance; horizontal gene transfer; methylome; Type I restriction-modification system.

Abbreviations: BGI, Beijing Genomics Institute; BPGA, bacterial pan genome analysis; CRISPR, clustered regularly interspaced short palindromic repeats; HGT, horizontal gene transfer; KEGG, kyoto encyclopedia of genes and genomes; MGEs, mobile genetic elements.; MTase, methyltransferase; NR, non-redundant protein sequence; RM, restriction-modification; SMRT, single molecule real-time; SSR, simple sequence repeat; TRD, target recognition domain; T1SS, type I secretion system.

This complete genome project has been deposited at GenBank under accession number CP110364.

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Two supplementary figures and eight supplementary tables are available with the online version of this article.

001107 © 2023 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Impact Statement

Numerous studies have reported that methylation of the Type I restriction-modification (RM) system affects gene expression regulation, virulence modulation and drug resistance in several pathogens. In recent years, *Aeromonas veronii* has become a focus of research due to its growing drug resistance, which poses a notable threat to human health and the aquatic economy. However, despite its significance, the function and role of the Type I RM system in *A. veronii* continue to remain elusive. In this investigation, we sequenced the genome of *A. veronii* C4 and conducted comparative genomic analysis with 19 other *A. veronii* strains in the database. Our analysis indicated that *A. veronii* possesses an open pan-genome, and *A. veronii* C4 displayed a distinctive Type I RM system methyltransferase gene in comparison to other strains. Horizontal gene transfer predictions in bacteria and archaea indicated that the methyltransferase gene might have been acquired from *Thiocystis violascens* and *Comamonas kerstesii*. The gene was found to be associated with drug resistance by methylation analysis. Overall, our findings suggest that the epigenetic changes caused by the horizontally acquired methyltransferase gene of the Type I RM system considerably contribute to enhanced drug resistance in *A. veronii*.

endogenous DNA through M₂S₁ methylation [4, 7, 8]. The S subunit is shaped like a dumbbell and contains a simple sequence repeat (SSR) in the centre, as well as various target recognition domains (TRDs) at both ends [6]. The length of the SSR limits the number of nonspecific bases 'N' in the motif, while the TRD determines the specific DNA sequence [9]. Apart from the classical Type I RM system mentioned above, which generates 6mA modifications, the noncanonical system utilizes two different M subunits to form MTase instead of identical copies, modifying m6A on one strand and m4C on the other [10–12].

The Type I RM system possesses the characteristics of mobile genetic elements (MGEs) and facilitates gene exchange between strains possessing homologous Type I RM systems. Theoretically, the presence of a greater number of Type I RM systems within a strain is positively associated with an increased gene abundance and frequency of gene exchange [13]. Furthermore, as observed in various pathogens such as *Neisseria meningitidis* [14], *Pseudomonas aeruginosa* [15], *Staphylococcus aureus* [16], certain species of the genus *Streptococcus* [17–19], *Mycoplasma* [20, 21] and others [22–25], the Type I RM system regulates gene expression through methylation modification, which subsequently affects strain phenotypes. Inactivation of the MTase of *Pseudomonas aeruginosa* reveals an attenuated-virulence phenotype in the *Galleria mellonella* infection model [15]. Variability in the repeat sequences of the Type I RM system of *Mycoplasma pneumoniae* determines its macrolide resistance [21]. However, most recent studies have focused on the function of the Type I RM system, and its origin and evolutionary traits have yet to be traced.

Aeromonas veronii is a rod-shaped, Gram-negative, facultatively anaerobic bacterium with strong motility, virulence and environmental adaptability, making it ubiquitous worldwide [26, 27]. It is primarily isolated from fish lesions and causes symptoms such as haemorrhages, ulceration and visceral enlargement, resulting in significant financial losses in the aquaculture industry [28]. Additionally, it infects amphibians, reptiles and mammals, and spreads through wounds, leading to a variety of diseases and even death [29]. Of particular concern is the increasing virulence and drug resistance of *A. veronii*, which pose a significant threat to public health [27, 30–35]. Bacterial whole-genome analysis serves as an essential method for understanding the genetic variations, functionality and evolution of bacteria, offering basic support and evidence for disease prevention and environmental conservation [36, 37]. Although the genomes of individual *A. veronii* strains have been sequenced, little is known about the key regulatory elements such as the Type I RM systems and their contribution to pathogenesis and drug resistance. For *A. veronii*, analysis of its complete genome, specifically elucidating the presence and distribution patterns of the Type I RM system, bears substantial significance for the prevention and treatment of infections caused by this pathogen. It remains unclear whether *A. veronii* acquired the Type I RM system from other bacteria during its evolution, or whether there was horizontal transfer of this system between *A. veronii* and other bacterial species.

The present study employed genomic sequencing and comparative analysis to investigate *A. veronii* strain C4. In addition, the epigenome of *A. veronii* C4 was examined to identify the specific motif methylated by AveC4I. Subsequently, the obtained data were subject to functional annotation and enrichment analysis. This work holds theoretical and practical value in elucidating the drug-resistant mechanisms of the pathogen, identifying novel drug targets, and exploring new strategies for pathogen control.

METHODS

Genome sequencing and assembly

In 2015, *A. veronii* C4 was isolated from the liver lesions of diseased *Ctenopharyngodon idella* in Shantou, China [38, 39]. Since then, the strain has been stored in our laboratory. Single colonies were picked and cultured in LB liquid medium at 30°C with shaking at 150 r.p.m. for 16 h. The bacterial cells were then collected and stored at -80°C until further analysis. Genomic DNA was extracted using the MgiEasy Microbial DNA Extraction Kit.

For Illumina sequencing, DNA fragments were generated using an ultrasonic DNA breaker (Covaris), followed by end-repair, 3'-adenylation, and ligation to the adaptors. The amplified and purified fragments were used to construct high-quality library, which was sequenced on the Illumina HiSeq 4000 platform as paired-end 2×150 bp reads.

For single molecule real-time (SMRT) sequencing, DNA fragments of 10–15 kb were treated with enzymatic digestion, damage-repair and end-repair, and connected with the barcode sequence linker. The SMRTbell library was obtained and sequenced on the PacBio RS II platform. Both Illumina and SMRT sequencing were performed at Beijing Genomics Institute (BGI).

To ensure the accuracy of the results, Illumina and SMRT reads with low-quality (≤ 20) bases (40% as default) or Ns exceeding a certain threshold (10% as default), adapter and duplication contamination, and SMRT reads shorter than 1000 bp were removed. The SMRT reads were self-corrected and assembled using Pbdagcon (<https://github.com/PacificBiosciences/pbdagcon>) and Celera Assembler (<http://wgs-assembler.sourceforge.net>), respectively. The assembled genome was corrected by GATK (<https://www.broadinstitute.org/gatk/>) and SOAP tools (including SOAP2, SOAPSnp and SOAPindel) based on Illumina reads. The complete genome of *A. veronii* C4 was submitted to the NCBI GenBank database (accession number CP110364). Gene prediction was performed using glimmer3 (<http://ccb.jhu.edu/software/glimmer/index.shtml>) with the Hidden Markov model. tRNAs, rRNAs and sRNAs were identified by comparing with the tRNAscan-SE [40], RNAmmer [41] and Rfam [42] databases. Protein sequence alignments against the Non-Redundant Protein Sequence (NR), UniProt, Kyoto Encyclopedia of Genes and Genomes (KEGG) and Pfam databases were carried out to obtain *hsd* genes in *A. veronii* C4

Pan-genome analysis

The genomes of 19 other *A. veronii* strains were obtained from the NCBI Genome database using the NCBI Datasets command-line tool with the parameters ‘datasets download genome taxon “aeromonas veronii” --assembly-level complete_genome --assembly-source refseq --dehydrated’ in November 2021. RefSeq assembly accession numbers are provided as follows: GCF_000204115.1 [43], GCF_000464515.2 [44], GCF_001593245.1, GCF_002803925.1, GCF_002803945.1, GCF_003491365.1, GCF_003722175.1, GCF_008693705.1 [45], GCF_009755745.1, GCF_009834065.1, GCF_011045495.1, GCF_013415825.1, GCF_014168715.1 [46], GCF_014168995.1 [46], GCF_014169795.1 [46], GCF_014169835.1 [46], GCF_014169875.1 [46], GCF_017310275.1 and GCF_020172765.2 [47]. To ensure the timeliness and comprehensiveness of the data, we also removed the restriction of assembly level and downloaded 232 *A. veronii* genomes on 28 July 2023. Pan-genome analysis was performed using Bacterial Pan Genome Analysis (BPGA) [48]. The results were visualized with gnuplot software and the Tutools platform (<http://www.cloudtutu.com>). OrthoFinder was used to identify homologous genes of the Type I RM systems in *A. veronii* C4 [49]. The Wagner parsimony method implemented in the Count software package was used to infer the orthologous genes of ancestors, with a gain/loss penalty rate of 2 [50].

Phylogenetic analyses of HsdM1, HsdS1 and *Aeromonas* species

The protein annotations of fully assembled reference genomes of 3357 bacteria and 220 archaea were downloaded from the NCBI Genome database using the NCBI Datasets command-line tool with the parameters ‘datasets download genome taxon bacteria --assembly-level complete_genome --reference --dehydrated’ and ‘datasets download genome taxon archaea --assembly-level complete_genome --reference --dehydrated’. Homologous proteins of HsdM1 and HsdS1 in *A. veronii* C4 were identified by aligning their sequences against HsdM1 and HsdS1 using a sequence identity threshold of 30% and coverage threshold of 40%. The program MUSCLE was used to perform multiple sequence alignments [51], and IQ-TREE was executed to reconstruct phylogenetic trees of HsdM1 and HsdS1 based on the maximum likelihood method [52]. Species that contained homologous HsdM1 and HsdS1 were selected, and single-copy core genes from these species were identified using OrthoFinder. Multiple sequence alignments for each orthogroup were carried out using MUSCLE, and spurious sequences or poorly aligned regions were removed using trimAL [53]. High-quality-alignment single-copy core genes were concatenated, and further were used to reconstruct the phylogenetic tree of the species containing homologous HsdM1 or HsdS1 by IQ-TREE. The taxonomy of the species was annotated using Taxonkit [54]. All phylogenetic trees were visualized using iTOL [55]. Gene clusters were visualized using ChiPlot (<https://www.chiplot.online/>). The reliability of the phylogenetic trees was evaluated using 1000 bootstraps.

Inference of evolutionary relationships

The phylogenetic tree of species containing homologous HsdM1 or HsdS1, as well as the phylogenetic trees of HsdM1 or HsdS1 were analysed using Ranger-DTL to infer gene evolution [56]. In total, 1000 reconciliations were calculated for each gene with different random seeds, and the results were aggregated using the AggregateRanger function. Gene duplication or horizontal gene transfer (HGT) events were considered supported if the corresponding support value was at least 50%.

Methylation analysis of the SMRT sequencing data

The genome index of *A. veronii* C4 was generated using samtools, and the SMRT reads were aligned to the genome using pbalign. The software ipdSummary was then utilized to identify 6mA sites, and the methylated motifs containing the 6mA sites were detected using motifMaker. A motif was defined as a DNA sequence pattern that occurred repeatedly. These steps were carried out

in the SMRT Analysis platform (v2.3.0, <http://www.pacb.com/products-and-services/analyticalsoftware/smrt-analysis/analysis-applications/epigenetics/>). The Blast tool in the REBASE database was used to assign motifs to restriction modification systems [57]. A self-written Perl script was used to search the genome of *A. veronii* C4 for the positions of all AveC4I motifs (https://github.com/JiayueMa0529/perl_findMotif/blob/master/1_motif_loc.pl). The genomic distribution of the recognition sequence was visualized using the Gene-Density tool available in Hiplot [58]. The genes harbouring the recognition sequence were then annotated in their corresponding KEGG pathways. Enrichment analysis was employed using OmicShare tools (<https://www.omicshare.com/tools/Home/Soft/pathwaygseasenior>), and the outcomes were visualized through ImageGP [59].

RESULTS AND DISCUSSION

A. veronii has an ‘open’ pan-genome

The genome of *A. veronii* C4 was sequenced, which was found to have a G+C content of 58.50% and a total length of 4926096 bp. This strain harboured 4724 protein-coding genes and 168 noncoding RNA genes. After downloading the genomes of the other 19 *A. veronii* strains, whose assembly levels were complete, from the National Center for Biotechnology Information (NCBI) Genome database, their pan-genome was acquired, and this contained 9174 orthologous gene families (Table S1, available in the online version of this article). Of these, 2668 genes were shared by the 20 strains (defined as core genes), 3459 were unique to individual strains (defined as unique genes), and the remaining were accessory genes, indicating a high degree of genetic variation (Fig. 1a and Table S1). The number of unique genes in different *A. veronii* strains exhibited significant variation, ranging from 74 to 384 (Fig. 1a and Table S2). After removing the restriction of assembly level, the genomes of 233 *A. veronii* strains including *A. veronii* C4 were obtained and the number of unique and accessory genes increased accordingly (Table S3). There were 2179 core genes for the 233 strains (Table S3). This suggested that certain strains may accept essential foreign genes and, in some cases, even discard defence mechanisms such as CRISPR (clustered regularly interspaced short palindromic repeats) in order to facilitate gene integration [13, 60, 61]. Notably, *A. veronii* C4 had a relatively large genome among the compared strains, implying that it had a greater capacity for environmental adaptability and had probably acquired genes through HGT.

The gene accumulation curves depicted in Fig. 1(b) showed that the pan-genome of 20 *A. veronii* strains was ‘open’, as its size continued to increase without showing signs of saturation with the inclusion of more genomes. In contrast, the size of the core genome gradually decreased but did not seem to approach a plateau. To ensure the robustness of the results, the pan-genome of 233 *A. veronii* strains was also analysed and the same phenomenon was still observed (Fig. S1A). This indicated that *A. veronii* strains frequently and readily incorporated genes from other species, resulting in a constant increase in the diversity and functionality of genes. Interestingly, bacteria with ‘closed’ pan-genomes are typically found in stable environments, such as the commensal bacterium *Staphylococcus lugdunensis* [62]. However, most free-living bacteria possess ‘open’ pan-genomes to adapt to variable environments. In fact, nine bacteria have been shown to have ‘open’ pan-genomes [63], and even the bacterial domain as a whole does [64].

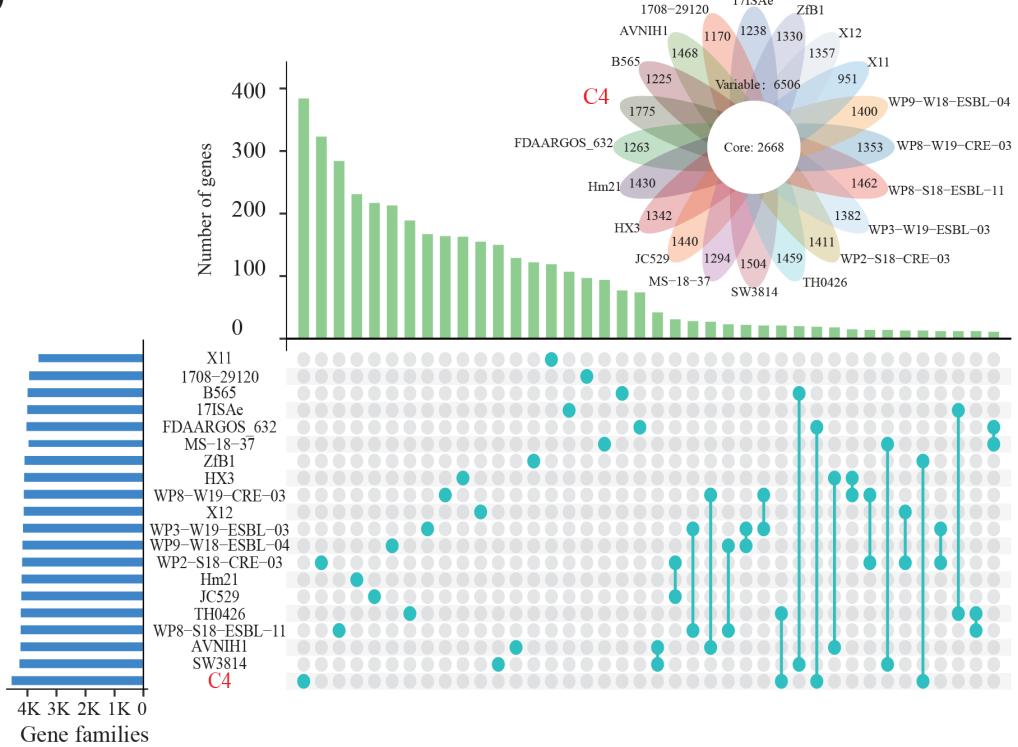
An ‘open’ pan-genome confers a greater adaptive potential for a species to thrive in varying ecological niches. To unravel the functional characteristics of the pan-genome of the 20 *A. veronii* strains, we conducted functional annotation of the core, unique and accessory genes using the COG database (Fig. 1c). A substantial portion of the pan-genome, representing 67.76%, was not well characterized with respect to functional annotation, as indicated by ‘general function prediction only’ or ‘function unknown’. The core genome exhibited diverse functions and played a crucial role in various biological processes, such as ‘amino acid transport and metabolism’ (9.73%), ‘transcription’ (6.74%), ‘energy production and conversion’ (6.45%), ‘signal transduction mechanisms’ (6.26%), and ‘translation ribosomal structure and biogenesis’ (6.15%) (Fig. 1c and Table S3). The accessory genome was primarily involved in ‘transcription’ (9.41%), while the unique genome was enriched for genes related to ‘replication, recombination and repair’ (12.52%) and ‘cell wall/membrane/envelope biogenesis’ (11.19%) (Fig. 1c and Table S4). When analysing the pan-genome of 233 *A. veronii* strains, the functional repertoire of three categories of genes did not show a clear change (Fig. S1B and Table S5).

Analysis of gene gain and loss in the evolution of *A. veronii*

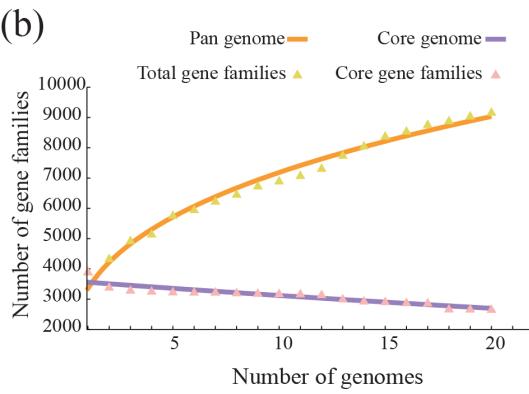
To analyse the dynamics of gene families in the evolution of *A. veronii*, we reconstructed a phylogenetic tree of 20 strains based on 2728 single-copy core genes (Fig. 2a). Based on the topological structure and evolutionary distance, the tree was divided into five clades, with the exception of the independently evolved strains 1708-29120 and WP2-S18-CRE-03 [65]. The strains in Clade I, including WP8-S18-ESBL-11, WP9-W18-ESBL-04 and WP3-W19-ESBL-03, were isolated from water in Tokyo, Japan. Interestingly, strain WP8-W19-CRE-03 was also isolated from the same location but belonged to Clade V (Fig. 2a). Similarly, strains X11 and X12 exhibited distinct core genomes, despite being isolated from the same location. Strains C4 and ZfB1 evolved from a common ancestor and together with strain 17ISAe constituted Clade II (Fig. 2a).

The orthologous genes of ancestors was inferred (Table S6), and the ancestral genome of the 20 *A. veronii* strains was estimated to contain 4152 gene families (Table S7), while the ancestral genome of 29 *Aeromonas* species contained 3821 gene families [66]. The phylogenetic tree of 20 *A. veronii* strains showed evidence of gene gain and loss, indicating a dynamic evolutionary process in *A. veronii* (Fig. 2a and Table S7). Gene gain was more frequent than gene loss in most strains (Fig. 2a and Table S7), suggesting

(a)



(b)



(c)

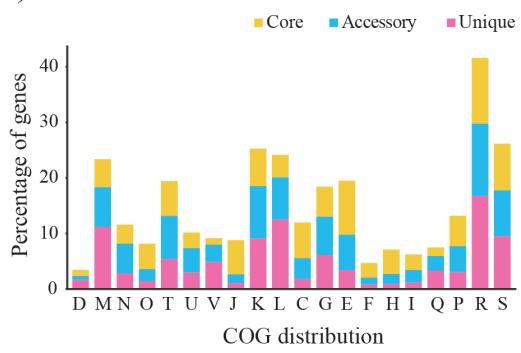


Fig. 1. Pan-genome structure of 20 *Aeromonas veronii* strains. (a) Comparative genomic map of 20 *A. veronii* strains. The blue bar represents the number of gene families in each strain. The green bar on the genomic map denotes the count of either unique genes (glaucous dots) present in individual strains, or accessory genes (glaucous dots connected by the glaucous line) found in at least two strains. The flower plot represents the number of core genes and the remaining genes (defined as variable genes) in each strain. (b) Gene accumulation curves for the pan-genome and core genome of 20 *A. veronii* strains. (c) Distribution of Clusters of Orthologous Groups (COG) functional annotations of 20 *A. veronii* strains. [C] Energy production and conversion, [D] Cell cycle control, cell division, chromosome partitioning, [E] Amino acid transport and metabolism, [F] Nucleotide transport and metabolism, [G] Carbohydrate transport and metabolism, [H] Coenzyme transport and metabolism, [I] Lipid transport and metabolism, [J] Translation, ribosomal structure and biogenesis, [K] Transcription, [L] Replication, recombination and repair, [M] Cell wall/membrane/envelope biogenesis, [N] Cell motility, [O] Posttranslational modification, protein turnover, chaperones, [P] Inorganic ion transport and metabolism, [Q] Secondary metabolite biosynthesis, transport and catabolism, [R] General function prediction only, [S] Function unknown, [T] Signal transduction mechanisms, [U] Intracellular trafficking, secretion and vesicular transport, [V] Defence mechanisms, [Y] Nuclear structure.

that *A. veronii* acquired new genes through HGT or other mechanisms to adapt to diverse lifestyles [66]. In contrast, gene loss was more common in the 29 species of *Aeromonas*, reflecting vertical descent in the species [65]. The patterns of gene gain and loss were not consistent across *A. veronii* strains, even among closely related strains (Fig. 2a). For example, strain C4 gained 245 genes and lost 114 genes, while strain Zfb1 gained 144 genes and lost 196 genes. These results suggested that gene gain and loss had played a major role in driving the evolution and shaping the genetic diversity of *A. veronii*.

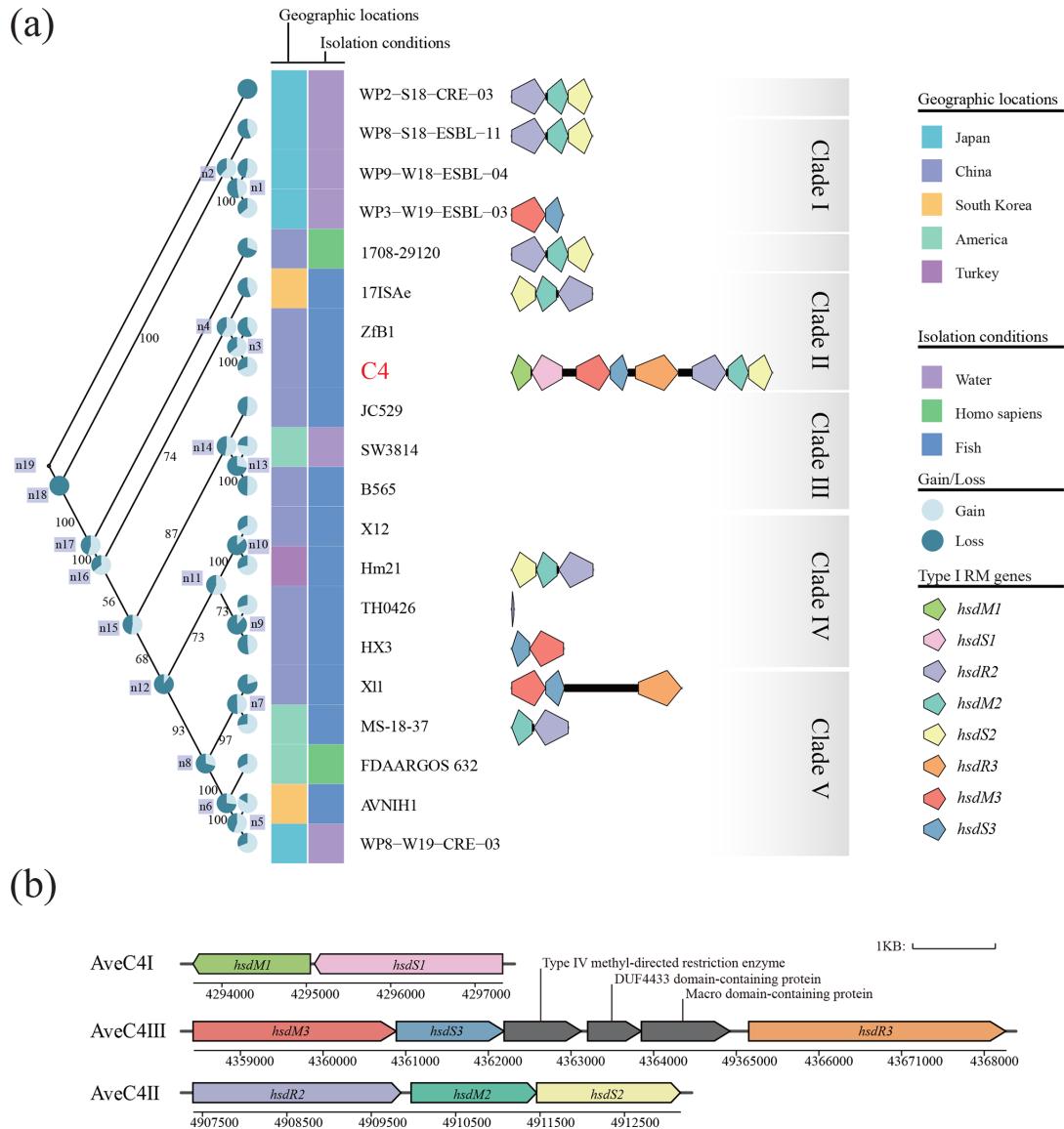


Fig. 2. Distribution of *hsd* genes in 20 *Aeromonas veronii* strains. (a) Phylogenetic analysis of 20 *A. veronii* strains. The phylogenetic tree based on 2728 single-copy core genes of the 20 strains is divided into five clades. Gene gain (light blue) and loss (dark blue) are represented by pie charts. Geographical location and isolation conditions are represented by heat maps. The gene cluster consists of homologous genes of eight *hsd* genes in *A. veronii* C4. Bootstrap values from 1000 replicates are shown. (b) Gene clusters of AveC4I, AveC4II and AveC4III in *A. veronii* C4.

Scattered distribution of *hsd* genes in *A. veronii*

Three *hsdM* genes (Locus tags: ONR73_20345, ONR73_23265 and ONR73_20695), three *hsdS* genes (Locus tags: ONR73_20340, ONR73_23270 and ONR73_20700) and two *hsdR* genes (Locus tags: ONR73_20720 and ONR73_23260) were found. These *hsd* genes constituted one Type I MTase and two Type I RM systems (Fig. 2a). They were all located on defence islands and were designated as AveC4I, AveC4II and AveC4III [67]. AveC4I was composed of *hsdM1* (Locus tag: ONR73_20345) and *hsdS1* (Locus tag: ONR73_20340), while AveC4II was encoded by *hsdR2* (Locus tag: ONR73_23260), *hsdM2* (Locus tag: ONR73_23265) and *hsdS2* (Locus tag: ONR73_23270). AveC4III consisted of *hsdR3* (Locus tag: ONR73_20720), *hsdM3* (Locus tag: ONR73_20695) and *hsdS3* (Locus tag: ONR73_20700). In AveC4I and AveC4III, the *hsdM* and *hsdS* genes were closely located, but *hsdR* was either absent or separated, resulting in the decay of an RM cluster or the potential for conjunct function with other RM systems [68]. Between *hsdS3* and *hsdR3*, three genes were integrated, including a Type IV methyl-directed restriction enzyme, a DUF4433 domain-containing protein and a Macro domain-containing protein (Fig. 2b). These mosaic genes probably arose from evolutionary recombination and may be redundant, gradually fusing or disappearing in the near future, ultimately leading to the development of a tightly continuous Type I RM system [69].

Comparative genomic analysis was conducted to examine homologous genes of the Type I RM systems and MTase in 19 other *A. veronii* strains, in comparison with *A. veronii* C4 (Fig. 2a). Homologous genes of AveC4II or AveC4III were present in ten strains, while the remaining strains lacked these genes. Notably, AveC4I only existed in *A. veronii* C4 when analysing 233 *A. veronii* strains (Table S8). The Type I RM systems and MTase play an important role in enhancing the environmentally adaptive capacity of bacterial strains. Under high adverse pressure, strains tend to acquire Type I RM systems and MTase from outside sources and regulate gene expression to adapt by altering specific motifs. Since the Type I RM system requires ATP consumption, strains tend to discard the system to reduce metabolic burden when the stress is relieved [60, 61, 70–72]. Despite this, AveC4I, AveC4II and AveC4III have been retained in the genome of *A. veronii* C4, probably due to experiencing tremendous pressure during the course of evolution.

Incongruence between the phylogenetic trees hinted at potential HGT of AveC4I

We analysed 3357 bacterial and 220 archaeal reference genomes that were fully assembled in order to identify homologous proteins of *HsdM1* and *HsdS1*. Our analysis identified 109 homologous *HsdM1* in 107 species (Fig. 3a), whereas only 59 homologous *HsdS1* were identified in 58 species (Fig. 3b), probably due to the highly variable TRDs on both ends of the S subunit in different species [4]. A total of 152 species, mostly belonging to the classes *Gammaproteobacteria*, *Flavobacteriia*, *Actinomycetia* and *Betaproteobacteria*, contained either *HsdM1* or *HsdS1*. Of these species, 13 contained both *HsdM1* and *HsdS1*. To explore the evolutionary relationships between these 152 species, a phylogenetic tree was reconstructed using 18 single-copy core genes (Fig. S2A, B). Species of archaea from five classes were clustered into a large clade and evolved independently from bacteria. Similarly, species of bacteria from the same class tended to be clustered together. The species of *Gammaproteobacteria* evolved later and shared a common ancestor with the species of *Betaproteobacteria*. *A. veronii* C4, classified in *Gammaproteobacteria*, was clustered together with the genus *Shewanella* (Fig. S2A, B).

Phylogenetic trees of *HsdM1* and *HsdS1* as well as the phylogenetic trees of the species containing homologous *HsdM1* or *HsdS1* were reconstructed (Fig. 3a, b). Unlike the phylogenetic tree of the species, the phylogenetic tree of *HsdM1* or *HsdS1* separated species belonging to the same class into different clades, suggesting that both *hsdM1* and *hsdS1* have undergone HGT. The phylogenetic tree for *HsdM1* was divided into two large clades (Fig. 3a), each containing several small clades composed of species from the classes *Flavobacteriia*, *Actinomycetia* and *Gammaproteobacteria*. The phylogenetic tree for *HsdS1* was also divided into two clades (Fig. 3b), but the distribution of species was more irregular. Even for two species in the same clade, the total branch length was high, further explaining the high variability of *HsdS1*. In the phylogenetic trees of *HsdM1* and *HsdS1*, *A. veronii* C4 was most closely related to *Comamonas kerstersii* (Fig. 3a, b). Between the two species, the protein sequences of *HsdM1* and *HsdS1* shared 88 and 60% identity, respectively. *Comamonas kerstersii*, like *A. veronii* C4, is a drug-resistant pathogen that causes peritoneal infections, bacteraemia and other diseases [73, 74].

AveC4I has complex HGT routes

Using RANGER-DTL, HGT and gene duplication events were inferred to reconstruct the evolutionary history of AveC4I. The analysis revealed extensive HGTs of *hsdM1* and *hsdS1*, which were not limited to species within the same class but also occurred between bacteria and archaea (Fig. S2). Although several species possessed *hsdM1* or *hsdS1*, only the *hsdM1* of *Candidatus Nitrososphaera evergladensi* was found to be duplicated (Fig. S2A).

The evolutionary history of *hsdM1* was revealed to involve extensive HGT events, as inferred by RANGER-DTL (Fig. S2). HGT of *hsdM1* was not limited to species within the same class, but also occurred between species of different classes, including bacteria and archaea. In total, 29 species and ten putative ancestors were predicted as donors of *hsdM1*, including three species of archaea (*Methanocaldococcus fervens* of *Methanococci*, *Methanosaeta harundinacea* of *Methanomicrobia* and *Thermococcus nautili* of *Thermococci*) (Fig. S2A). The *hsdM1* of these three species was transferred to the classes *Bacteroidia*, *Epsilonproteobacteria*, *Gammaproteobacteria* and *Acidithiobacillia*. Among these donor archaea, two strains were isolated from extreme environments, and only one of them transferred *hsdM1* to the bacteria in the same environment. The remaining recipient bacteria were isolated from lake bottom sediments or healthy animals. The phylogenetic tree of *HsdM1* showed several clades within *Flavobacteriia*, but their topological structures were inconsistent with those in the phylogenetic tree of the species (Fig. 3a). Based on the prediction of HGT routes, we speculated that HGT of *hsdM1* occurred within the class *Flavobacteriia*. Eight routes were obtained, and two of the donor species were isolated from extreme environments, while the majority of recipients displayed pathogenicity (Fig. S2A and Table 1). In *Cyanobacteria*, HGT of *hsdM1* did not occur, suggesting that the *hsdM1* of these species was acquired through vertical gene transfer. Although numerous species of *Gammaproteobacteria* possessed *hsdM1*, there were relatively few HGT events within the class, and the majority of *hsdM1* were derived from ancestors (Fig. S2A).

In contrast to *hsdM1*, the horizontal transfer of *hsdS1* was typically observed in a radial pattern, especially in *Proteobacteria* (Fig. S2B). The transfer of *hsdS1* from *Helicobacter cетorum* (*Epsilonproteobacteria*) to *Terriglobus albidus* (*Acidobacteriia*) and *Desulfovibulus propionicus* (*Deltaproteobacteria*) resulted in the initial appearance of *hsdS1* in these species. *Pseudomonas qingdaonensis* and *Azotobacter chroococcum* in the same clade frequently transferred *hsdS1* to other species, including four species of *Actinomycetia*. Species of *Cyanobacteria* acquired *hsdS1* from other classes, but did not transfer it to other species. The donor

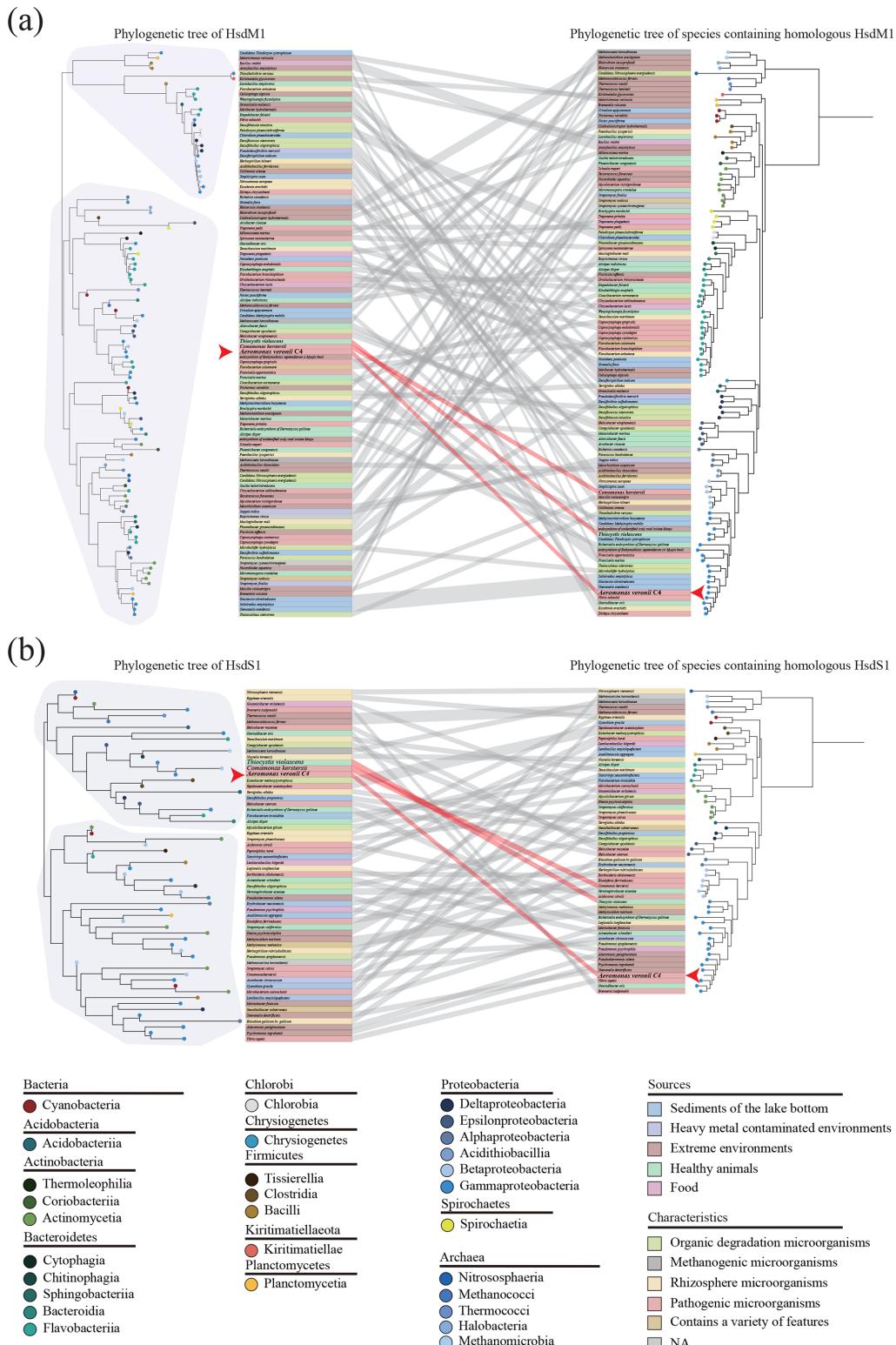


Fig. 3. Comparison between phylogenetic trees. (a) The phylogenetic tree of HsdM1 was reconstructed based on 109 homologous HsdM1 (left), and the phylogenetic tree of the 107 species containing homologous HsdM1 was reconstructed based on 21 single-copy core genes (right). (b) The phylogenetic tree of HsdS1 was reconstructed based on 59 homologous HsdS1 (left), and the phylogenetic tree of the 58 species containing homologous HsdS1 was reconstructed based on 26 single-copy core genes (right). In (a) and (b) dots of various colours symbolize the species belonging to different classes. Different background colours highlight the species that come from diverse sources or have different characteristics. Lines connect the same species. *A. veronii* C4 is marked with a red arrowhead.

Table 1. Horizontal gene transfer of *hsdM1* within *Flavobacteriia*

Donor	Source/characteristic	Recipient(s)	Source/characteristic
<i>Cellulophaga algicola</i>	Extreme environment	<i>Flavobacterium anhuiense</i> <i>Wenyingzhuangia fucanilytica</i>	Rhizosphere microorganism Healthy animal
<i>Chryseobacterium lactis</i>	Food	<i>Ornithobacterium rhinotracheale</i>	Pathogenic microorganism
<i>Elizabethkingia anophelis</i>	Pathogenic microorganism	<i>Capnocytophaga endodontalis</i>	Pathogenic microorganism
<i>Flavobacterium columnare</i>	Contains a variety of features	<i>Capnocytophaga gingivalis</i>	Pathogenic microorganism
<i>Maribacter hydrothermalis</i>	Extreme environment	<i>Empedobacter falsenii</i>	Healthy animal
n38	–	<i>Flavobacterium branchiophilum</i>	Pathogenic microorganism
n28	–	n37	–

strains were pathogens or microbes with organic degradation ability. The recipient bacteria mostly were pathogens. Interestingly, *A. veronii* C4 received *hsdM1* from *Comamonas kerstersii* (Fig. S2A) and donated *hsdS1* to *Comamonas kerstersii* (Fig. S2B), resulting in a win-win situation. They were all isolated from China and were pathogenic bacteria, indicating that there may be genetic exchange among pathogenic bacteria in the same region.

AveC4I is associated with the chemotaxis of *A. veronii* C4

Seven methylated motifs, including 'GTANNNNNNCTTC', 'GAAGNNNNNTAC', 'GCCAYNNNNNDTGA', 'TCAHNNNNN-RTGGC', 'CCAYNNNNNCTG', 'CAGNNNNNRTGG' and 'GATC', were identified as the recognition motifs when analysing the SMRT sequencing data. To search for homologous proteins and their corresponding recognition motifs, the sequences of HsdS1, HsdS2 and HsdS3 were submitted to the REBASE database (<http://tools.neb.com/blast/>) [57]. The protein S.Pfe13482III was found to exhibit 11% sequence identity with HsdS1 and to recognize the conserved motif 'GAAYNNNNNNCTG'. The protein S.SenS922ORF19115P demonstrated 53% sequence identity with HsdS2 and was observed to specifically recognize the conserved motif 'GTANNNNNNNGTTC'. The protein S.Ecos1613ORF5225P exhibited 32% sequence identity with HsdS3 and was found to recognize the conserved motif 'CCAYNNNNNGTTY'. Given that TRDs have been shown to recognize analogous sequences with a minimum of 25% identity in HsdS [4], referring to previous research methods [75], we assigned the sequences 'GCCAYNNNNNDTGA', 'GTANNNNNNNCTTC' and 'CCAYNNNNNCTG' to AveC4I, AveC4II and AveC4III, respectively.

The 14-base novel motif, 'GCCAYNNNNNDTGA/TCAHNNNNNRTGGC', was methylated by AveC4I, and included five arbitrary bases in the middle and specific bases at both ends recognized by the TRDs of HsdS1. A total of 571 sequences matching this motif were found dispersed throughout the genome (Fig. 4a), with 509 of these sequences located in 462 genes. Of these genes, 34 contained at least two matched sequences. Notably, three adjacent genes (ONR73_14145, ONR73_14150 and ONR73_14155) spanning from 2974154 to 2983700 bp each contained more than five matched sequences (Fig. 4a), indicating that they were likely to be important regulatory targets of AveC4I. They encode the Type I secretion C-terminal target domain-containing protein, Type I secretion system (T1SS)-143 repeat domain-containing protein, and RTX toxin-related Ca²⁺-binding protein, respectively. Recognition of the Type I secretion C-terminal target domain-containing protein is triggered by the C-terminal secretion signal, which leads to assembly of the functional transenvelope complex and confers specificity to the recognition by the ABC protein and RTX toxins [76]. The T1SS 143 repeat domain-containing protein tends to share several properties with the RtxA protein of *Vibrio cholerae* [77]. The RTX toxin-related Ca²⁺-binding protein can promote the secretion of RTX, which is a type of bacterial exotoxin and can cause damage to host cells [78]. The expression of genes in regions enriched with methylation is regulated by methylation status [22]. DNA methylation can either activate or repress a gene, depending on the presence of other DNA-binding proteins that compete with DNA MTases [79, 80]. The results implied that AveC4I played a role in the regulation of virulence.

KEGG annotation and enrichment analyses were conducted to investigate the functions of the 462 genes containing the novel motif recognized by AveC4I (Fig. 4b). These genes were primarily involved in several pathways, including 'Carbon metabolism', 'Biosynthesis of amino acids', 'Bacterial chemotaxis' and 'Flagellar assembly'. Among them, the 'Bacterial chemotaxis' pathway showed significant enrichment (hypergeometric test, $P<0.01$). A total of 16 genes were annotated in the pathway, including *cheY*, *cheB* and the gene encoding methyl-accepting chemotaxis protein. The expression of these genes related to chemotaxis is suggested to be affected by methylation of AveC4I. Bacterial chemotaxis is a fundamental process that plays a critical role in a multitude of biological processes. It encompasses the directed movement of bacteria towards favourable chemical gradients, while avoiding toxic chemical gradients [81–83]. Depletion of chemotaxis-related genes has been shown to affect the antimicrobial susceptibilities of *Aeromonas hydrophila* [84]. Chemotaxis not only facilitates the ability of *Shewanella oneidensis* MR-1 to colonize but also maintains metabolic activity in regions where the antibiotic ciprofloxacin exhibits high toxicity [85]. The chemotaxis histidine kinase CheA has the ability to phosphorylate not only its specific response regulator CheY2, but also one of the response regulators

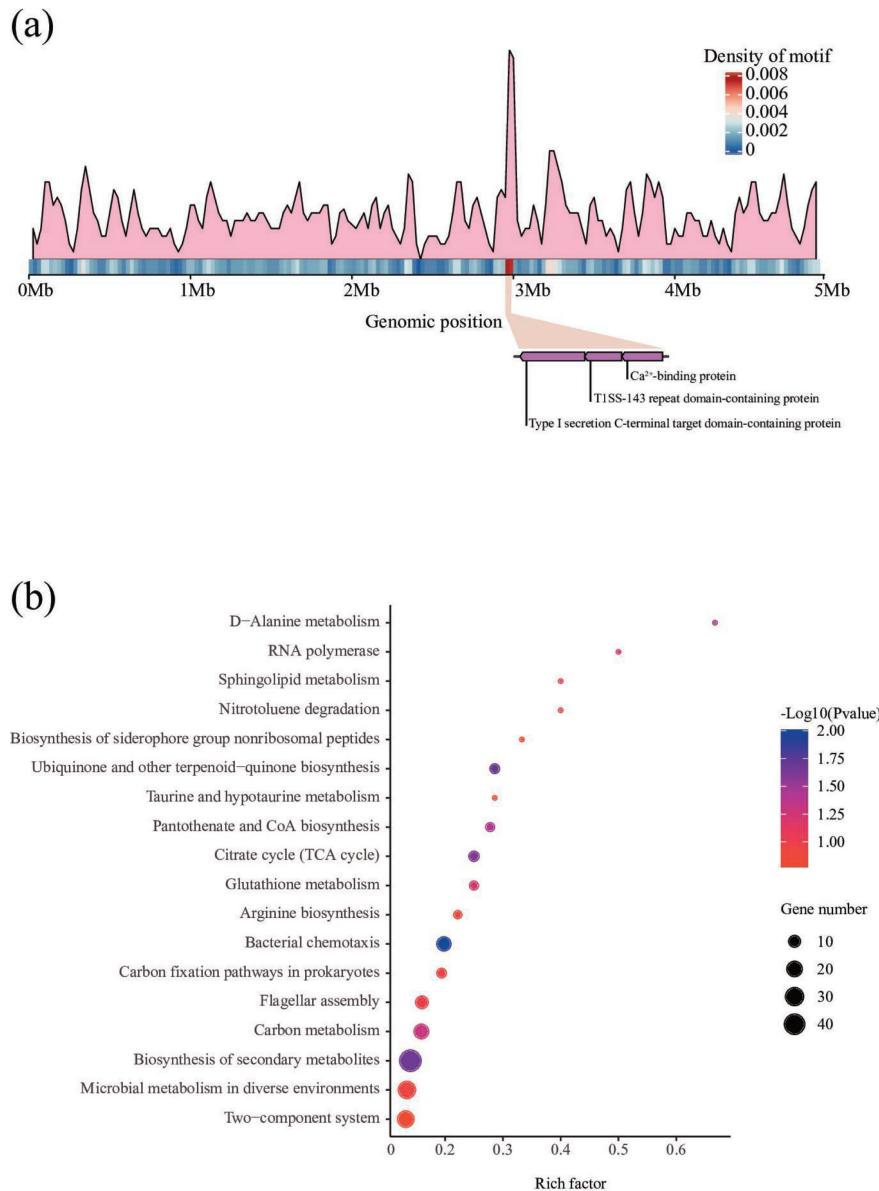


Fig. 4. Methylome analysis of AveC4I. (a) Distribution of the motif methylated by AveC4I in *A. veronii* C4. The motif is enriched in the region (2.9–3.0 Mb), which includes ONR73_14145 (Type I secretion C-terminal target domain-containing protein), ONR73_14150 (T1SS 143 repeat domain-containing protein) and ONR73_14155 (Ca²⁺-binding protein). (b) Scatter plot of KEGG analysis of the genes with the motif. The 'Bacterial chemotaxis' pathway showed the most significant enrichment (hypergeometric test, $P<0.01$).

involved in the pathway that mediates biofilm formation, FlmD [86]. The chemotaxis genes *cheY* and *cheW* in *Helicobacter pylori* also influence its pathogenicity in the stomach [87]. In *Pseudomonas aeruginosa*, *cheB2* assists in adapting to the environment [88]. The observed enrichment of genes within the pathway provided compelling evidence supporting the involvement of AveC4I in the regulation of bacterial chemotaxis and associated biological processes. These findings suggested a potential relationship between AveC4I and drug resistance of *A. veronii*.

CONCLUSIONS

Comparative genomic and phylogenetic analyses have revealed that *A. veronii* possesses an 'open' pan-genome, rendering it susceptible to the acquisition of genes from other species via extensive HGTs. Consequently, Type I RM systems and MTase genes were dispersed throughout the *A. veronii* genome. Of all the *A. veronii* strains examined, *A. veronii* C4 possessed a larger genome and greater number of Type I system genes. The AveC4I gene was unique to *A. veronii* C4 and was determined to have

been horizontally transferred from *Comamonas kerstensii*. SMRT sequencing and methylation analysis revealed that the methylation sites of AveC4I were clustered on three genes associated with virulence. The expression of genes in regions enriched with methylation is regulated by the methylation status [22]. It is plausible that AveC4I may exert an influence on the virulence of *A. veronii* C4. Pathway analyses of genes featuring AveC4I recognition motifs implied that AveC4I may regulate drug resistance pathway. Methylation has been reported to affect virulence and drug resistance in various bacteria, such as *Streptococcus pyogenes* and *Mycoplasma pneumoniae* [18, 21]. Therefore, we speculate that Type I RM systems also could be a promising therapeutic target for the treatment of diseases related to *A. veronii*.

Funding information

This research was supported by the Hainan Provincial Natural Science Foundation of China (322RC589 to Hong Li) and the National Natural Science Foundation of China (32060153 to Hong Li, 32160038 to Zhu Liu, 32000089 and 31960027 to Xiang Ma).

Acknowledgements

We express our gratitude to Zhuo Wang from the Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, for providing valuable suggestions on the analysis of the results. We would also like to thank TopEdit (www.topeditsci.com) for its linguistic assistance during the preparation of the manuscript.

Author contributions

Conceptualization, J.M., H.L. and Z.L.; Data curation, J.M. and H.L.; Formal analysis, J.M., H.Z., S.M., Y.T., X.M., J.L. and H.L.; Funding acquisition, X.M., H.L. and Z.L.; Investigation, Y.T. and X.M.; Methodology, J.M. and H.Z.; Project administration, H.L. and Z.L.; Resources, J.L.; Supervision, H.L. and Z.L.; Visualization, J.M. and S.M.; Writing—original draft, J.M.; Writing—review & editing, H.L. and Z.L.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

- Seib KL, Srikhanta YN, Atack JM, Jennings MP. Epigenetic regulation of virulence and immuno-evasion by phase-variable restriction-modification systems in bacterial pathogens. *Annu Rev Microbiol* 2020;74:655–671.
- Vasu K, Nagaraja V. Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiol Mol Biol Rev* 2013;77:53–72.
- Bickle TA, Krüger DH. Biology of DNA restriction. *Microbiol Rev* 1993;57:434–450.
- Murray NE. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol Mol Biol Rev* 2000;64:412–434.
- Roberts GA, Chen K, Cooper LP, White JH, Blakely GW, et al. Removal of a frameshift between the *hsdM* and *hsdS* genes of the EcoKI Type IA DNA restriction and modification system produces a new type of system and links the different families of type I systems. *Nucleic Acids Res* 2012;40:10916–10924.
- Gao Y, Cao D, Zhu J, Feng H, Luo X, et al. Structural insights into assembly, operation and inhibition of a type I restriction-modification system. *Nat Microbiol* 2020;5:1107–1118.
- Cooper LP, Roberts GA, White JH, Luyten YA, Bower EKM, et al. DNA target recognition domains in the type I restriction and modification systems of *Staphylococcus aureus*. *Nucleic Acids Res* 2017;45:3395–3406.
- De Ste Croix M, Vacca I, Kwun MJ, Ralph JD, Bentley SD, et al. Phase-variable methylation and epigenetic regulation by type I restriction-modification systems. *FEMS Microbiol Rev* 2017;41:S3–S15.
- Price C, Lingner J, Bickle TA, Firman K, Glover SW. Basis for changes in DNA recognition by the EcoR124 and EcoR124/3 type I DNA restriction and modification enzymes. *J Mol Biol* 1989;205:115–125.
- Morgan RD, Luyten YA, Johnson SA, Clough EM, Clark TA, et al. Novel m4C modification in type I restriction-modification systems. *Nucleic Acids Res* 2016;44:9413–9425.
- Zhu J, Gao Y, Wang Y, Zhan Q, Feng H, et al. Molecular insights into DNA recognition and methylation by non-canonical type I restriction-modification systems. *Nat Commun* 2022;13:6391.
- Loenen WAM, Dryden DTF, Raleigh EA, Wilson GG. Type I restriction enzymes and their relatives. *Nucleic Acids Res* 2014;42:20–44.
- Oliveira PH, Touchon M, Rocha EPC. Regulation of genetic flux between bacteria by restriction-modification systems. *Proc Natl Acad Sci U S A* 2016;113:5658–5663.
- Budroni S, Siena E, Dunning Hotopp JC, Seib KL, Serruto D, et al. *Neisseria meningitidis* is structured in clades associated with restriction modification systems that modulate homologous recombination. *Proc Natl Acad Sci U S A* 2011;108:4494–4499.
- Doberenz S, Eckweiler D, Reichert O, Jensen V, Bunk B, et al. Identification of a *Pseudomonas Aeruginosa* Pao1 DNA methyltransferase, its targets, and physiological roles. *mBio* 2017;8.
- Roberts GA, Houston PJ, White JH, Chen K, Stephanou AS, et al. Impact of target site distribution for type I restriction enzymes on the evolution of methicillin-resistant *Staphylococcus aureus* (MRSA) populations. *Nucleic Acids Res* 2013;41:7472–7484.
- DebRoy S, Shropshire WC, Tran CN, Hao H, Gohel M, et al. Characterization of the type I restriction modification system broadly conserved among group A *Streptococci*. *mSphere* 2021;6:e0079921.
- Nye TM, Jacob KM, Holley EK, Nevarez JM, Dawid S, et al. DNA methylation from a type I restriction modification system influences gene expression and virulence in *Streptococcus pyogenes*. *PLoS Pathog* 2019;15:e1007841.
- Oliver MB, Swords WE. Comparative analysis of *Streptococcus pneumoniae* type I restriction-modification loci: variation in *hsdS* gene target recognition domains. *Pathogens* 2020;9:712.
- Dordet-Frisoni E, Vandecasteele C, Contarin R, Sagné E, Baranowski E, et al. Impacts of *Mycoplasma agalactiae* restriction-modification systems on pan-epigenome dynamics and genome plasticity. *Microb Genom* 2022;8:mgen000829.
- Lee J-K, Seong M-W, Yun K-W, Choi E-H. Association of tandem repeat number variabilities in subunit S of the type I restriction-modification system with macrolide resistance in *Mycoplasma pneumoniae*. *J Clin Med* 2022;11:715.
- Oliveira PH, Ribis JW, Garrett EM, Trzilova D, Kim A, et al. Epigenomic characterization of *Clostridioides difficile* finds a conserved DNA methyltransferase that mediates sporulation and pathogenesis. *Nat Microbiol* 2020;5:166–180.
- Blakeway LV, Tan A, Lappan R, Ariff A, Pickering JL, et al. *Moraxella catarrhalis* restriction-modification systems are associated with phylogenetic lineage and disease. *Genome Biol Evol* 2018;10:2932–2946.

24. Casselli T, Tourand Y, Scheidegger A, Arnold WK, Proulx A, et al. DNA methylation by restriction modification systems affects the global transcriptome profile in *Borrelia burgdorferi*. *J Bacteriol* 2018;200:e00395-18.
25. Manso AS, Chai MH, Attack JM, Furi L, De Ste Croix M, et al. A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nat Commun* 2014;5:5055.
26. Hickman-Brenner FW, MacDonald KL, Steigerwalt AG, Fanning GR, Brenner DJ, et al. *Aeromonas veronii*, a new ornithine decarboxylase-positive species that may cause diarrhea. *J Clin Microbiol* 1987;25:900–906.
27. Suzuki T, Franchi L, Toma C, Ashida H, Ogawa M, et al. Differential regulation of caspase-1 activation, pyroptosis, and autophagy via Ipaf and ASC in Shigella-infected macrophages. *PLoS Pathog* 2007;3:e111.
28. Nawaz M, Khan SA, Khan AA, Sung K, Tran Q, et al. Detection and characterization of virulence genes and integrons in *Aeromonas veronii* isolated from catfish. *Food Microbiol* 2010;27:327–331.
29. Fernández-Bravo A, Figueras MJ. An update on the genus *Aeromonas*: taxonomy, epidemiology, and pathogenicity. *Microorganisms* 2020;8:129.
30. Chandrarathna H, Nikapitiya C, Dananjaya SHS, Wijerathne CUB, Wimalasena S, et al. Outcome of co-infection with opportunistic and multidrug resistant *Aeromonas Hydrophila* and *A. Veronii* in Zebrafish: identification, characterization, pathogenicity and immune responses. *Fish Shellfish Immunol* 2018;80:573–581.
31. Sánchez-Céspedes J, Figueras MJ, Aspiroz C, Aldea MJ, Toledo M, et al. Development of imipenem resistance in an *Aeromonas veronii* biovar sobria clinical isolate recovered from a patient with cholangitis. *J Med Microbiol* 2009;58:451–455.
32. Sánchez-Céspedes J, Blasco MD, Martí S, Alba V, Alcalde E, et al. Plasmid-mediated QnrS2 determinant from a clinical *Aeromonas veronii* isolate. *Antimicrob Agents Chemother* 2008;52:2990–2991.
33. Elgendi MY, Shaalan M, Abdelsalam M, Eissa AE, El-Adawy MM, et al. Antibacterial activity of silver nanoparticles against antibiotic-resistant *Aeromonas veronii* infections in Nile Tilapia, *Oreochromis niloticus* (L.), in vitro and in vivo assay. *Aquac Res* 2022;53:901–920.
34. Zhang L, Yu W, Tang Y, Li H, Ma X, et al. RNA chaperone hfq mediates persistence to multiple antibiotics in *Aeromonas veronii*. *Microp Pathog* 2019;132:124–128.
35. Wang D, Li H, Khan WU, Ma X, Tang H, et al. SmpB and tmRNA orchestrate purine pathway for the trimethoprim resistance in *Aeromonas veronii*. *Front Cell Infect Microbiol* 2020;10:239.
36. Kim Y, Gu C, Kim HU, Lee SY. Current status of pan-genome analysis for pathogenic bacteria. *Curr Opin Biotechnol* 2020;63:54–62.
37. Wan Makhtar WRW, Wan Ratmaazila W, Bharudin I, Samsulrizal NH, Yusof NY. Whole genome sequencing analysis of *Salmonella enterica* serovar typhi: history and current approaches. *Microorganisms* 2021;9:2155.
38. Liu Z, Liu P, Liu S, Song H, Tang H, et al. Small protein B upregulates sensor kinase bvgS expression in *Aeromonas veronii*. *Front Microbiol* 2015;6:579.
39. Liu P, Chen Y, Wang D, Tang Y, Tang H, et al. Genetic selection of peptide aptamers that interact and inhibit both small protein B and alternative ribosome-rescue factor A of *Aeromonas veronii* C4. *Front Microbiol* 2016;7:1228.
40. Chan PP, Lin BY, Mak AJ, Lowe TM. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res* 2021;49:9077–9096.
41. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007;35:3100–3108.
42. Kalvari I, Nawrocki EP, Ontiveros-Palacios N, Argasinska J, Lamkiewicz K, et al. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Res* 2021;49:D192–D200.
43. Li Y, Liu Y, Zhou Z, Huang H, Ren Y, et al. Complete genome sequence of *Aeromonas veronii* strain B565. *J Bacteriol* 2011;193:3389–3390.
44. Bomar L, Stephens WZ, Nelson MC, Velle K, Guillemin K, et al. Draft genome sequence of *Aeromonas veronii* Hm21, a symbiotic isolate from the medicinal leech digestive tract. *Genome Announc* 2013;1:e00800-13.
45. Sichtig H, Minogue T, Yan Y, Stefan C, Hall A, et al. FDA-ARGOS is a database with public quality-controlled reference genomes for diagnostic use and regulatory science. *Nat Commun* 2019;10:3313.
46. Sekizuka T, Inamine Y, Segawa T, Hashino M, Yatsu K, et al. Potential KPC-2 carbapenemase reservoir of environmental *Aeromonas hydrophila* and *Aeromonas caviae* isolates from the effluent of an urban wastewater treatment plant in Japan. *Environ Microbiol Rep* 2019;11:589–597.
47. Lovero KG, Mota-Bravo L. Closed genome sequence of an environmental *Aeromonas veronii* strain from California, United States, with an IncA/C plasmid carrying an extended-spectrum β-lactamase gene, *bla*_{VEB-3}. *Microbiol Resour Announc* 2022;11:e0103321.
48. Chaudhari NM, Gupta VK, Dutta C. BPGA- an ultra-fast pan-genome analysis pipeline. *Sci Rep* 2016;6:24373.
49. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019;20:238.
50. Csurös M. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* 2010;26:1910–1912.
51. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res* 2022;50:W276–W279.
52. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, et al. Corrigendum to: IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;37:1530–1534.
53. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;25:1972–1973.
54. Shen W, Ren H. TaxonKit: a practical and efficient NCBI taxonomy toolkit. *J Genet Genomics* 2021;48:844–850.
55. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;49:W293–W296.
56. Bansal MS, Kellis M, Kordi M, Kundu S. RANGER-DTL 2.0: rigorous reconstruction of gene-family evolution by duplication, transfer and loss. *Bioinformatics* 2018;34:3214–3216.
57. Roberts RJ, Vincze T, Posfai J, Macelis D. REBASE--a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res* 2015;43:D298–D299.
58. Li J, Miao B, Wang S, Dong W, Xu H, et al. Hiplot: a comprehensive and easy-to-use web service for boosting publication-ready biomedical data visualization. *Brief Bioinform* 2022;23:bbac261.
59. Chen T, Liu Y-X, Huang L. ImageGP: an easy-to-use data visualization web server for scientific researchers. *iMeta* 2022;1.
60. Pleška M, Qian L, Okura R, Bergmiller T, Wakamoto Y, et al. Bacterial autoimmunity due to a restriction-modification system. *Curr Biol* 2016;26:404–409.
61. Woods LC, Gorrell RJ, Taylor F, Connallon T, Kwok T, et al. Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation. *Proc Natl Acad Sci U S A* 2020;117:26868–26875.
62. Argemi X, Matelska D, Ginalski K, Riegel P, Hansmann Y, et al. Comparative genomic analysis of *Staphylococcus lugdunensis* shows a closed pan-genome and multiple barriers to horizontal gene transfer. *BMC Genomics* 2018;19:621.
63. Costa SS, Guimarães LC, Silva A, Soares SC, Baraúna RA. First steps in the analysis of prokaryotic pan-genomes. *Bioinform Biol Insights* 2020;14:117793220938064.
64. Lapierre P, Gogarten JP. Estimating the size of the bacterial pan-genome. *Trends Genet* 2009;25:107–110.
65. Zhong C, Han M, Yang P, Chen C, Yu H, et al. Comprehensive analysis reveals the evolution and pathogenicity of *Aeromonas*, viewed

- from both single isolated species and microbial communities. *mSystems* 2019;4:e00252-19.
66. Csürös M. Ancestral Reconstruction by Asymmetric Wagner Parsimony over Continuous Characters and Squared Parsimony over Distributions. In: Nelson CE and Vialette S (eds). *Comparative Genomics*. Berlin, Heidelberg: Springer; 2008. pp. 72–86.
 67. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, et al. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res* 2003;31:1805–1812.
 68. Kobayashi I. Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res* 2001;29:3742–3756.
 69. Bower EKM, Cooper LP, Roberts GA, White JH, Luyten Y, et al. A model for the evolution of prokaryotic DNA restriction-modification systems based upon the structural malleability of type I restriction-modification enzymes. *Nucleic Acids Res* 2018;46:9067–9080.
 70. Bernheim A, Sorek R. The pan-immune system of bacteria: antiviral defence as a community resource. *Nat Rev Microbiol* 2020;18:113–119.
 71. Koonin EV, Makarova KS, Wolf YI, Krupovic M. Evolutionary entanglement of mobile genetic elements and host defence systems: guns for hire. *Nat Rev Genet* 2020;21:119–131.
 72. Omer S, Harlow TJ, Gogarten JP. Does sequence conservation provide evidence for biological function? *Trends Microbiol* 2017;25:11–18.
 73. Almuzara MN, Cittadini R, Vera Ocampo C, Bakai R, Traglia G, et al. Intra-abdominal infections due to *Comamonas kerstesii*. *J Clin Microbiol* 2013;51:1998–2000.
 74. Moser AI, Campos-Madueno EL, Keller PM, Endimiani A. Complete genome sequence of a third- and fourth-generation cephalosporin-resistant *Comamonas kerstesii* isolate. *Microbial Resour Announc* 2021;10:e0039121.
 75. Furuta Y, Namba-Fukuyo H, Shibata TF, Nishiyama T, Shigenobu S, et al. Methylome diversification through changes in DNA methyltransferase sequence specificity. *PLoS Genet* 2014;10:e1004272.
 76. Delepelaire P. Type I secretion in gram-negative bacteria. *Biochim Biophys Acta* 2004;1694:149–161.
 77. Sheahan K-L, Cordero CL, Satchell KJF. Identification of a domain within the multifunctional *Vibrio cholerae* RTX toxin that covalently cross-links actin. *Proc Natl Acad Sci U S A* 2004;101:9798–9803.
 78. O'Brien DP, Hernandez B, Durand D, Hourdel V, Sotomayor-Pérez A-C, et al. Structural models of intrinsically disordered and calcium-bound folded states of a protein adapted for secretion. *Sci Rep* 2015;5:14223.
 79. Portnoy DA, Low DA, Weyand NJ, Mahan MJ. Roles of DNA Adenine methylation in regulating bacterial gene expression and virulence. *Infect Immun* 2001;69:7197–7204.
 80. Wion D, Casadesús J. N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nat Rev Microbiol* 2006;4:183–192.
 81. Celani A, Vergassola M. Bacterial strategies for chemotaxis response. *Proc Natl Acad Sci U S A* 2010;107:1391–1396.
 82. Keegstra JM, Carrara F, Stocker R. The ecological roles of bacterial chemotaxis. *Nat Rev Microbiol* 2022;20:491–504.
 83. Wong-Ng J, Celani A, Vergassola M. Exploring the function of bacterial chemotaxis. *Curr Opin Microbiol* 2018;45:16–21.
 84. Li W, Ali F, Cai Q, Yao Z, Sun L, et al. Quantitative proteomic analysis reveals that chemotaxis is involved in chlortetracycline resistance of *Aeromonas hydrophila*. *J Proteomics* 2018;172:143–151.
 85. Alcalde RE, Dundas CM, Dong Y, Sanford RA, Keitz BK, et al. The role of chemotaxis and efflux pumps on nitrate reduction in the toxic regions of a ciprofloxacin concentration gradient. *ISME J* 2021;15:2920–2932.
 86. Huang Z, Wang Y-H, Zhu H-Z, Andrianova EP, Jiang C-Y, et al. Cross talk between chemosensory pathways that modulate chemotaxis and biofilm formation. *mBio* 2019;10:e02876-18.
 87. Williams SM, Chen Y-T, Andermann TM, Carter JE, McGee DJ, et al. *Helicobacter pylori* chemotaxis modulates inflammation and bacterium-gastric epithelium interactions in infected mice. *Infect Immun* 2007;75:3747–3757.
 88. Garvis S, Munder A, Ball G, de Bentzmann S, Wiehlmann L, et al. *Caenorhabditis elegans* semi-automated liquid screen reveals a specialized role for the chemotaxis gene cheB2 in *Pseudomonas aeruginosa* virulence. *PLoS Pathog* 2009;5:e1000540.

Five reasons to publish your next article with a Microbiology Society journal

1. When you submit to our journals, you are supporting Society activities for your community.
2. Experience a fair, transparent process and critical, constructive review.
3. If you are at a Publish and Read institution, you'll enjoy the benefits of Open Access across our journal portfolio.
4. Author feedback says our Editors are 'thorough and fair' and 'patient and caring'.
5. Increase your reach and impact and share your research more widely.

Find out more and submit your article at microbiologyresearch.org.