

Emergence of *Vibrio cholerae* O1 Sequence Type 75 in Taiwan

Appendix

Experimental Methods

Cholera information and bacterial isolates

We obtained the statistical, demographic, and epidemiologic information on cholera cases in Taiwan from the databases of Taiwan National Infectious Disease Statistics System (<https://nidss.cdc.gov.tw/en/>) and the National Notifiable Diseases of Surveillance System (1) of the Taiwan Centers for Disease Control and *V. cholerae* isolates from the biobank section of Taiwan Centers for Disease Control. A total of 60 isolates were obtained for the study, among which 56 were recovered from cholera patients between 2002 and 2018 and 4 from patients of the cholera outbreak that occurred in 1962 (2).

Pulsed-field gel electrophoresis and analysis

We used the PulseNet standardized pulsed-field gel electrophoresis (PFGE) protocol (3) to characterize *V. cholerae* isolates, then analyzed PFGE patterns and performed clustering analysis of PFGE patterns using tools provided by BioNumerics 7.6.3 (Applied Maths; http://www.applied-maths.com_).

Whole-genome sequencing and sequence analysis

We conducted whole-genome sequencing of *V. cholerae* isolates using Illumina MiSeq sequencing platform (Illumina Inc. USA) with MiSeq Reagent Kit v3 (2X 300 bp), assembled sequence reads using the SPAdes assembler version 3.12.0 (<http://cab.spbu.ru/software/spades/>), identified sequence types (STs) using the *Vibrio cholerae* core genome multilocus sequence typing MLST database (<https://pubmlst.org/vcholerae>) and antimicrobial resistance genes using the ResFinder tool of the Center for Genomic Epidemiology (<http://www.genomicepidemiology.org/>).

Construction of a whole genome single nucleotide polymorphism tree for *V. cholerae* strains from Taiwan

We used BioNumerics version 7.6.3 to construct a minimal spanning tree with whole genome single nucleotide polymorphism (wgSNP) profiles of *V. cholerae* strains. The sequences of raw reads were mapped to the reference genomic sequence of *V. cholerae* strain N16961 (GenBank accession no. GCA_000006745.1) and the mapped sequences of strains and the reference were aligned for SNP calling by using the option of strict SNP filtering (closed SNP set). By using this SNP calling criteria, SNPs are called by removing positions with at least one ambiguous base (non-ATGC base), one unreliable base (N), one gap and non-informative SNPs. Each retained SNP position has minimum 5x coverage, at least covered once in both forward and reverse direction. The minimum distance between retained SNP positions is 12 bp. A dendrogram was constructed with the whole genome SNP profiles using the categorical (SNPs) option for similarity coefficient and minimum spanning tree algorithm for cluster analysis.

Construction of a core genome multilocus sequence typing tree for *V. cholerae* strains from Taiwan and NCBI database

We downloaded genomic sequences from the SRA and Assembly databases of the National Center for Biotechnology Information (NCBI), assembled raw reads using the SPAdes assembler version 3.12.0, and generated core genome multilocus sequence typing (cgMLST) profiles (based on 2,951 core genes) using the in-house developed cgMLST Profiling tool in the cgMLST@Taiwan Web service platform (<https://rdvd.cdc.gov.tw/cgMLST>) (unpublished). We compared the cgMLST profiles of 60 isolates from Taiwan with those (5,048) from the NCBI database and selected the most match profiles with the Taiwanese isolates to construct a cgMLST genetic relatedness tree using the unweighted pair-group method with arithmetic means algorithm.

References

1. Jian SW, Chen CM, Lee CY, Liu DP. Real-time surveillance of infectious diseases: Taiwan's experience. *Health Secur.* 2017;15:144–53.
2. Yen CH. A recent study of cholera with reference to an outbreak in Taiwan in 1962. *Bull World Health Organ.* 1964;30:811–25.

3. Cooper KL, Luey CK, Bird M, Terajima J, Nair GB, Kam KM, et al. Development and validation of a PulseNet standardized pulsed-field gel electrophoresis protocol for subtyping of *Vibrio cholerae*. Foodborne Pathog Dis. 2006;3:51–8.

Appendix Table. Details of *Vibrio cholerae* strains from cholera cases in Taiwan, 2002–2018

Clade	Sequence type	Variants	No. of isolates	Resistance genes	Country of origin
1	ST69		18	<i>catB9</i> [§]	Taiwan, Thailand, Philippines, Malaysia
2	ST75	ST725*, ST726*, ST728*, ST727 [†]	38	<i>qnrVC4</i> [¶] (in 35 of 38 isolates)	Taiwan, Vietnam
3	ST75		2	None	Taiwan
4	ST723		2 [‡]	None	Indonesia

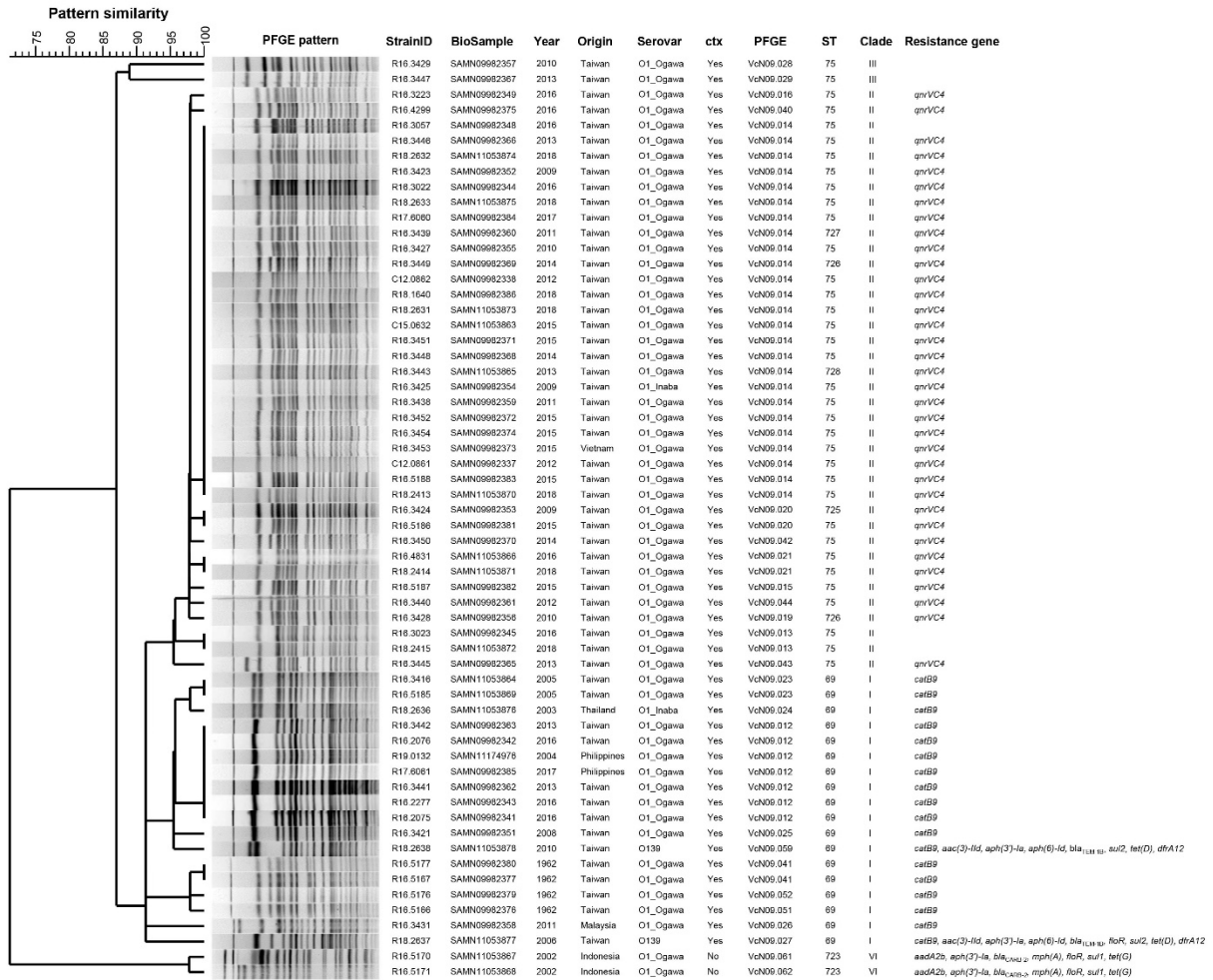
*Single-locus variant

†Double-locus variant

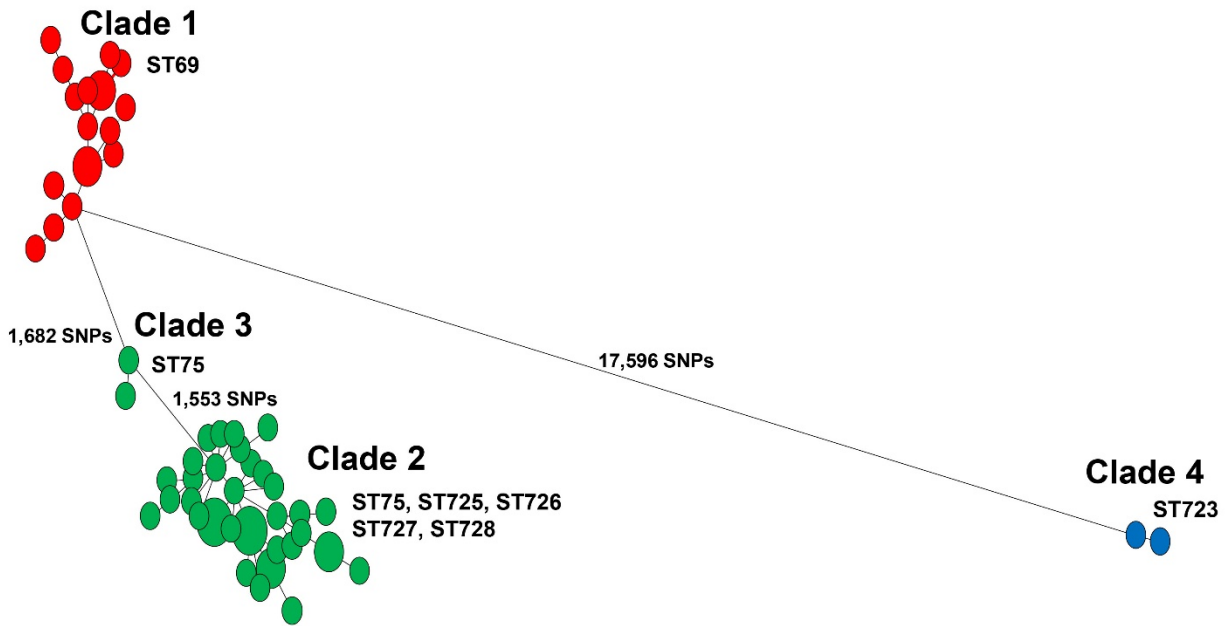
‡Nontoxicogenic

§Resistant to chloramphenicol

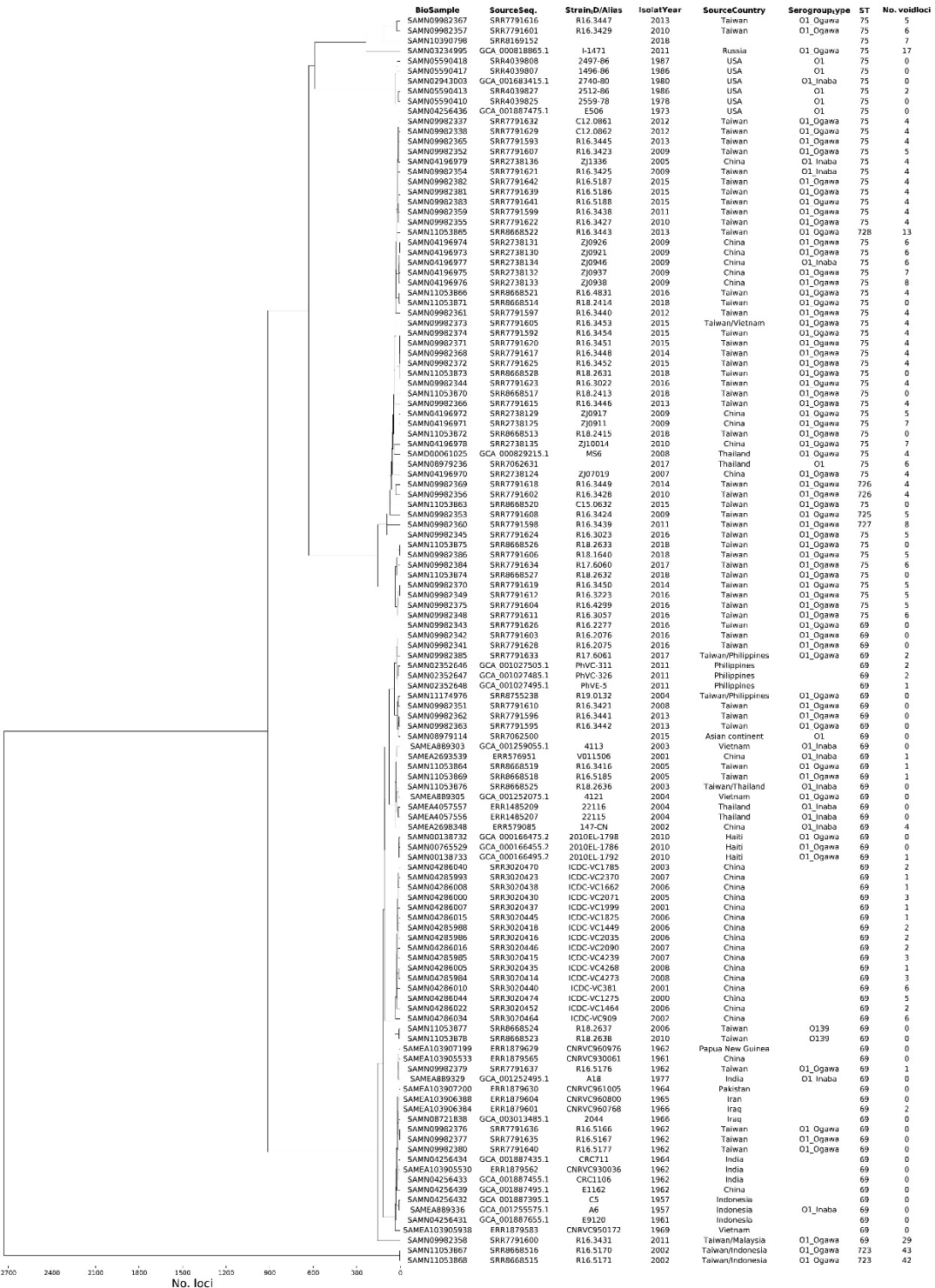
¶Resistant to quinolone



Appendix Figure 1. Genetic relationships among 60 *Vibrio cholerae* isolates, with corresponding information. The dendrogram was constructed using PFGE patterns and the single linkage algorithm provided in BioNumerics software version 7.6.3, with settings of 1.5% optimization and 0.95% tolerance.



Appendix Figure 2. A minimum spanning tree for 60 *Vibrio cholerae* isolates from Taiwan collected in 1962 and in 2002–2018. The tree was constructed with whole genome single-nucleotide polymorphism profiles comprising 20,639 SNPs.



Appendix Figure 3. A genetic dendrogram for 60 *Vibrio cholerae* isolates collected in Taiwan and their most closely related strains in the NCBI database. Three strains from the 2010 Haiti cholera outbreak were included for comparison. The tree was constructed with cgMLST profiles using the unweighted pair-group method with arithmetic means algorithm.