

Published in final edited form as:

*Genom Discov.* 2014 ; 2(1): 1–15. doi:10.7243/2052-7993-2-1.

## Comparative genome analysis of non-toxicogenic non-O1 versus toxicogenic O1 *Vibrio cholerae*

Munmun Mukherjee<sup>1</sup>, Prathusha Kakarla<sup>1</sup>, Sanath Kumar<sup>2</sup>, Esmeralda Gonzalez<sup>1</sup>, Jared T. Floyd<sup>1</sup>, Madhuri Inupakutika<sup>1</sup>, Amith Reddy Devireddy<sup>1</sup>, Selena R. Tirrell<sup>1</sup>, Merissa Bruns<sup>1</sup>, Guixin He<sup>3</sup>, Ingrid E. Lindquist<sup>4</sup>, Anitha Sundararajan<sup>4</sup>, Faye D. Schilkey<sup>4</sup>, Joann Mudge<sup>4</sup>, and Manuel F. Varela<sup>1,\*</sup>

<sup>1</sup>Eastern New Mexico University, Department of Biology, Portales, New Mexico, 88130, USA.

<sup>2</sup>QC Laboratory, Harvest and Post Harvest Technology Division, Central Institute of Fisheries Education (CIFE), Seven Bungalows, Versova, Andheri (W), Mumbai 400061, India.

<sup>3</sup>University of Massachusetts Lowell, Department of Clinical Laboratory and Nutritional Sciences, Lowell, MA 01854, USA.

<sup>4</sup>National Center for Genome Resources, Santa Fe, New Mexico, 87505, USA.

### Abstract

© 2014 Varela et al; licensee Herbert Publications Ltd.

This is an Open Access article distributed under the terms of Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>). This permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

\*Correspondence: Manuel.Varela@enmu.edu.

#### Competing interests

The authors declare that they have no competing interests.

#### Publication history

Editor: Yungui Yang, Beijing Institute of Genomics, China. EIC: Viroj Wiwanitkit, Hainan Medical University, China. Received: 26-May-2014 Final Revised: 21-Jun-2014 Accepted: 24-Jun-2014 Published: 23-Jul-2014

**Citation:** Mukherjee M, Kakarla P, Kumar S, Gonzalez E, Floyd JT, Inupakutika M, Devireddy AR, Tirrell SR, Bruns M, He G, Lindquist IE, Sundararajan A, Schilkey FD, Mudge J and Varela MF. Comparative genome analysis of non-toxicogenic non-O1 versus toxicogenic O1 *Vibrio cholerae*. *Genomics Discov.* 2014; 2:1. <http://dx.doi.org/10.7243/2052-7993-2-1>

#### Authors' contributions

Authors' contributions	MM	PK	SK	EG	JTF	MI	ARD	SRT	MB	GH	IEL	AS	FDS	JM	MFV
Research concept and design	--	--	✓	--	--	--	--	--	--	✓	--	--	--	--	✓
Collection and/or assembly of data	✓	✓	--	✓	✓	✓	✓	✓	✓	--	✓	✓	✓	✓	--
Data analysis and interpretation	✓	--	--	--	--	--	--	--	--	--	--	--	--	✓	✓
Writing the article	✓	--	--	--	--	--	--	--	--	--	--	--	--	--	✓
Critical revision of the article	--	--	✓	--	--	--	--	--	--	✓	✓	✓	✓	✓	✓
Final approval of article	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Statistical analysis	✓	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Pathogenic strains of *Vibrio cholerae* are responsible for endemic and pandemic outbreaks of the disease cholera. The complete toxigenic mechanisms underlying virulence in *Vibrio* strains are poorly understood. The hypothesis of this work was that virulent versus non-virulent strains of *V. cholerae* harbor distinctive genomic elements that encode virulence. The purpose of this study was to elucidate genomic differences between the O1 serotypes and non-O1 *V. cholerae* PS15, a non-toxigenic strain, in order to identify novel genes potentially responsible for virulence. In this study, we compared the whole genome of the non-O1 PS15 strain to the whole genomes of toxigenic serotypes at the phylogenetic level, and found that the PS15 genome was distantly related to those of toxigenic *V. cholerae*. Thus we focused on a detailed gene comparison between PS15 and the distantly related O1 *V. cholerae* N16961. Based on sequence alignment we tentatively assigned chromosome numbers 1 and 2 to elements within the genome of non-O1 *V. cholerae* PS15. Further, we found that PS15 and O1 *V. cholerae* N16961 shared 98% identity and 766 genes, but of the genes present in N16961 that were missing in the non-O1 *V. cholerae* PS15 genome, 56 were predicted to encode not only for virulence-related genes (colonization, antimicrobial resistance, and regulation of persister cells) but also genes involved in the metabolic biosynthesis of lipids, nucleosides and sulfur compounds. Additionally, we found 113 genes unique to PS15 that were predicted to encode other properties related to virulence, disease, defense, membrane transport, and DNA metabolism. Here, we identified distinctive and novel genomic elements between O1 and non-O1 *V. cholerae* genomes as potential virulence factors and, thus, targets for future therapeutics. Modulation of such novel targets may eventually enhance eradication efforts of endemic and pandemic disease cholera in afflicted nations.

## Keywords

*Vibrio cholerae*; O1; non-O1; serogroup; cholera; cholera toxin; virulence; genome comparison

## Introduction

Cholera is an infectious disease characterized by profuse watery diarrhea and vomiting in humans, and the causative agent is *Vibrio cholerae*, a Gram-negative, comma-shaped, facultative anaerobic bacterium [1]. *V. cholerae* includes both pathogenic and non-pathogenic strains, and the bacteria responsible for pandemic outbreaks secrete the cholera toxin [2]. Since 1817, seven pandemics of cholera have been recorded. Cholera is a major public health concern because the disease can exhibit significant mortality if left untreated [3,4]. In the past 200 years, cholera has resulted in millions of deaths due to its ability to spread rapidly within populations, and has been capable of contaminating rivers and estuaries [5]. The most recent outbreak of *V. cholerae* was recorded in Southeast Asia, which quickly spread across the globe as the seventh pandemic [6]. In 2010 alone, 604,634 cases of cholera were reported in Haiti, raising the death toll count to 7,436 in the first two years [7].

The genomes of several pathogenic *V. cholerae* strains encode proteins that are directly or indirectly responsible for virulence. In many parts of the world, the O serogroups of *V. cholerae* are associated with diarrhea [8]. The most common mode of transmission for this bacterium is through the consumption of feces-contaminated water, fishes or crustaceans [9].

In addition to rehydration therapy, the first line of antimicrobial agent used against cholera is doxycycline, prescribed for a period of 1-3 days in order to reduce the severity of the symptoms [10,11]. Other antimicrobials which have been demonstrated to be effective in humans include cotrimoxazole, erythromycin, tetracycline, chloramphenicol, furazolidone and norfloxacin [11,12].

Unfortunately, wide spread use and misuse of these and other antimicrobials have resulted in selection of multidrug-resistant bacterial variants [13] which potentially compromise chemotherapeutic efficacy towards cholera [14]. The different mechanisms by which bacteria show resistance to antimicrobial agents include (a) biofilm production (b) drug inactivation (c) ribosome protection (d) reduced permeability (e) target alteration [15] and (f) active efflux [16]. One of the active efflux pumps of *V. cholerae* is EmrD-3, which belongs to the major facilitator superfamily (MFS) and is a drug/H<sup>+</sup> antiporter with 12 transmembrane domains [17]. Another efflux pump encoded in the genome of *V. cholerae* is VceB [18]. Drug efflux pumps are integral membrane transporters that actively efflux the toxic compounds and antibiotics out of the bacterial cell and confer resistance against multiple antibacterial agents [19-21].

The presence of the cholera toxin (CT), the *Vibrio* pathogenicity island (VPI), and the toxin co-regulated pilus (TCP) within the O1 serogroups of *V. cholerae* make these strains more virulent and pandemic than their non-O1 counterparts [22]. A significant basis for their pathogenicity is attributed to cholera toxin encoding genes. Other genes important for enhancing virulence in these organisms are *ace*, *psh*, *PIICTX*, *zot* and *cep*, which are implicated in phage morphogenesis [5,23,24]. The *Vibrio* pathogenicity island-1 (VPI-1) confers toxin release, biofilm formation, attachment to disease vectors for transmission to humans, and are receptors of CTX. The *Vibrio* pathogenicity island-2 (VPI-2) helps the cholera toxin to gain entry into the intestinal epithelium by unmasking GM1 gangliosides in the lining of the human intestine. The absence of VPI-1 and VPI-2 in non-O1 serogroups of *V. cholerae* makes them less pathogenic than the O1 serogroups [25].

Even though non-O1 *V. cholerae* strains carry certain virulence genes, the severity of disease is less compared to O1/ O139 *V. cholerae* [8]. The non-O1 serogroups of *V. cholerae* are known as the non-agglutinating *Vibrios* (NAGs) because they lack the genes coding for CT and TCP [26,27]. The presence of multidrug resistance (MDR) transporters confers resistance to ampicillin, chloramphenicol and tetracycline in non-O1 and non-O139 serogroups of *V. cholerae* species [14]. The ABC transporters present in PS15 *V. cholerae* predictably transport phosphate molecules across the periplasm and may be essential for protein synthesis, amino acid exchange, and transport of fatty acids [28].

We previously determined the genome nucleotide sequence of the non-O1 non-toxigenic *V. cholerae* PS15 (GenBank Accession No. AIJR00000000) [28]. Here, we compared non-O1 PS15 with the genetic information of virulent strains. The genome of *V. cholerae* PS15 is composed of 3,910,387 base pairs (bp) organized into 3,512 open reading frames with a G +C content of 47.55% [28]. We chose to focus our comparative analysis with *V. cholerae* PS15 [29] using *V. cholerae* El Tor N16961 because this latter genome was completely sequenced [30]. N16961 is made up of 4,033,460 base pairs (bp) organized and distributed

into two chromosomes, with a G+C content of 46.9% in chromosome 1 and 47.7% in chromosome 2 [30]. Even though the non-O1 *V. cholerae* bacterium possesses some virulence genes responsible for causing gastrointestinal infections, wound infections, septicemia and cellulitis in humans, little is known about the mechanisms that confer virulence in this microorganism. The aim of this work is to identify differences in the genetic elements between the genomes of virulent N16961 and non-virulent PS15 strains of *V. cholerae* in order to identify novel virulence mechanisms that may eventually serve as potential therapeutic targets for the ultimate purpose of fostering conditions that reduce dissemination of disease-causing virulent serotypes of *V. cholerae* through populations.

## Methods

### Comparison of non-O1 PS15 and O1 N16961 *Vibrio cholerae* genomes using RAST and UniProt

A function based genome comparison was performed between a non-toxicogenic, non-O1 *V. cholerae* PS15 environmental isolate (courtesy of Dr. Charles Kaysner) from sediment sampled in Puget Sound, WA [28,31] and O1 *V. cholerae* N16961 [30], using the RAST (Rapid Annotation using Subsystem Technology) database and Seed Viewer to predict protein function [32] focusing on comparison of categories and subsystem groupings pertaining to virulence, disease, defense, membrane transport, DNA metabolism, regulons, dormancy, sporulation, phages, prophages, transposable elements, and plasmids for both genomes of O1 and non-O1 *V. cholerae* microorganisms. The open reading frames (genes) encoding functional roles associated with a subsystem are referred as functioning parts, and a subsystem is referred as a set of predicted abstract functional roles [32]. The screening of predicted proteins encoded from elements of both genomes was performed with BLAST analysis of the amino acid sequences using UniProt [33].

### Phylogenetic analysis

The non-O1 *V. cholerae* PS15 genome sequence [28] (GenBank Accession no. AIJR000000000) was analyzed using BLAST [34] in order to generate phylogenetic trees harboring genomes of closely related organisms and virulence factors of the O1 serotypes. The BLAST pair wise alignment using Tree Neighbor Joining method [35] was used to compare the genome of PS15 to other complete *Vibrio* genome sequences in the database and is represented in **Figure 1**.

### CGView

The CGView server was used for comparative genome analysis [36]. A graphical circular genome map was constructed using CGView by BLAST analysis of the DNA sequence of *V. cholerae* non- O1 PS15 (3,910,387 base pairs) with the complete DNA sequence of *V. cholerae* El Tor N16961 (4,033,460 base pairs) [28,30].

## Results

### The genome of non-O1 *V. cholerae* PS15 is distantly related to O1 *V. cholerae* genomes

We previously determined the whole genome sequence of a non-toxicogenic, non-O1 *V. cholerae* isolate from Puget Sound, strain PS15 [28]. It had been shown that genomes of toxicogenic O1 *V. cholerae* bacteria were highly related [30], possibly implying that non-O1 genomes would be more distantly related. We tested this prediction by comparing non-O1 *V. cholerae* PS15 with other microorganisms by constructing a phylogenetic tree using BLAST pair-wise alignment in order to represent genomes that are most closely related to *V. cholerae* non-O1 PS15 and to establish relatedness of PS15 to these microorganisms (**Figure 1**). Although the non-O1 *V. cholerae* PS15 genome sequence is most closely related to those of *V. cholerae* LMA 3984-4, O395, O1 strains 2010EL-1786, MJ-1236, O1 biovar El Tor strain N16961, IEC224, and M66-2, the non-O1 *V. cholerae* PS15 strain is, nonetheless, the most distantly related member within this cluster.

### Tentative chromosome assignment in non-toxicogenic, non-O1 *V. cholerae* PS15

Since the two chromosomes of the toxicogenic O1 *V. cholerae* strain N16961 were elucidated [30], we predicted that genomic sequence alignment with the non-toxicogenic, non-O1 *V. cholerae* strain PS15 would implicate chromosome assignment in this bacterium as well. A circular genome representation was generated using the CGView server to plot the structural genome arrangement with BLAST analysis of the non-O1 *V. cholerae* PS15 genome with that of the O1 *V. cholerae* N16961 using their respective genomic nucleotide sequences in a FASTA format (**Figure 2**). Using the genome sequence data from *V. cholerae* N16961 to compare with the genome of *V. cholerae* PS15, chromosomes 1 and 2 were implicated for the non-toxicogenic PS15 strain and are shown in **Figure 2**.

### The majority of genes in the O1 N16961 and non-O1 PS15 *V. cholerae* genomes are shared

We have shown above that although the non-O1 *V. cholerae* PS15 genome is distantly related to the genomes of toxicogenic O1 *V. cholerae*, the PS15 genome is still closely related to genomes of the *Vibrio* genus. This implies a striking similarity between the non-O1 and O1 genomes, specifically regarding the commonalities within the gene space. To test this, we used RAST Seed Viewer and UniProt to compare the genome sequences of O1 *V. cholerae* N16961 and non-O1 *V. cholerae* PS15, the general features of which are shown in **Table 1**. The O1 and non-O1 *V. cholerae* genomes shared 766 genes (open reading frames) that are predicted to code for proteins within functional categories pertaining to virulence, disease, defense, membrane transport, phages, prophages, transposable elements, plasmids, DNA metabolism, dormancy, sporulation and regulons. Interestingly, when compared to the N16961 genome, the *V. cholerae* PS15 genome appears to be truncated sporadically throughout by approximately 120 kbp (**Table 1** and **Figure 2**). In **Table 2** we listed 58 of 766 genes that share 98% identity between both genomes. The remaining genes are listed in **Supplement Table S1**. Even though non-O1 *V. cholerae* PS15 is believed to be non-pathogenic compared to the known virulent O1 *V. cholerae* N16961 strain, their genomes shared 90 genes in common that code for functions pertaining to virulence, disease and defense. Some of these genes included accessory colonization factor (*acfD*), TCP pilus

virulence regulatory protein (*tcpN*), toxin coregulated pilus biosynthesis protein E (*tcpE*), TCP pilus virulence regulatory protein (*toxT*) and accessory colonization factor (*acfC*). In addition to these virulence-associated genes, both genomes shared 287 genes encoding functional properties in the DNA metabolism category, 8 genes encoding proteins for dormancy and sporulation, 366 genes encoding membrane transporters, 12 genes in the categories of phages, prophages, transposable elements and plasmids, and 3 genes pertaining to regulons. Among these shared genomic elements encoding membrane transporters are genes known to express multidrug resistance efflux pumps, including AcrA of the RND superfamily [37], SugE of the SMR superfamily [38], and NorM of the MATE superfamily [39].

### Genes present in O1 *V. cholerae* N16961 genome and absent in the non-O1 PS15 genome

The pathogenicity of the O1 *V. cholerae* serotypes suggests that they harbor genomic elements that confer virulence. For instance, the cholera toxin of toxigenic *V. cholerae* strains is the primary virulence factor in endemic and pandemic cholera cases [40]. Thus, in order to establish the association between presence of virulence-encoding genomic elements and pathogenicity, we compared the functional determinants between both PS15 and N16961 genomes. Our analysis revealed that of the 619 genes absent in the non-O1 *V. cholerae* PS15 genome [29], 56 of these genes, when compared to O1 *V. cholerae* N16961, are in the categories including virulence, disease and defense, membrane transport, DNA metabolism, dormancy and sporulation (**Table 3**). The virulence genes which were present in O1 serotypes but largely absent in the non-O1 strains, including the PS15 strain, include the accessory cholera enterotoxin (*ace*), the cholera enterotoxin subunit B (*ctxB*), the cholera enterotoxin subunit A (*ctxA*), and the zona occludens toxin (*zot*). Comparison of the predicted proteins encoded of both PS15 and N16961 genomes using UniProt revealed the absence of other virulence genes in PS15, which include genes predicted to encode accessory colonization factors A and B (*acfA* and *acfB*), and the genes encoding VceA and VceB proteins shown to confer resistance to antimicrobial agents (**Table 3**) [41]. Notably, the gene demonstrated to confer multidrug resistance and encoding a drug efflux pump, EmrD-3, of the MFS is present in N16961 but absent from the non-O1 *V. cholerae* PS15 genome [17,21].

A phylogenetic tree, which was generated by BLAST for bacterial genomes that share the cholera toxin, indicated the absence of the cholera toxin gene in the non-O1 *V. cholerae* PS15 bacterium (**Figure 3**). The most closely-related microorganisms that shared the DNA encoding the cholera toxin include *V. cholerae* IEC224, O1 biovar El Tor strain N16961, O395, MJ-1236 and the O1 strain 2010EL-1786.

Other genes that were absent in non-O1 *V. cholerae* genome but present in O1, include genes that encode glycerolipid and glycerophospholipid metabolism, and genes that code for VPI [25] (**Table 3**). Additional genes that are absent in non-O1 *V. cholerae* PS15 include those coding for the Rst operon essential for the synthesis of phage related replication protein (RstA), phage related integrase (RstB), phage related antirepressor (RstC), phage related transcriptional repressor (RstR) [24], and sulfur metabolism. Other genes that are



found in O1 *V. cholerae* but absent in non-O1 include those coding for TsaE, a protein required for the synthesis of threonylcarbamoyladenine in the presence of tRNA [42].

### Genes present in the non-O1 *V. cholerae* PS15 genome and absent in the O1 N16961 genome

Because the non-O1 *V. cholerae* PS15 environmental isolate is considered to be non-toxicogenic [31,43], this implies that genes unique to this microorganism, compared to the toxicogenic N16961 bacterium, possibly encode non-virulent functions. To test this hypothesis, we performed a function based genome comparison using RAST and UniProt for PS15 and N16961. This comparative analysis revealed that 113 genes were excluded in N16961 but present within the PS15 genome (**Table 4**). The three known genes (characterized) that are present in PS15, but absent in N16961, include the oligopeptide ABC transporter called periplasmic oligopeptide-binding protein (OppA) [44], a protein-export membrane protein (SecF) [45], and the UvrABC system protein A (*uvrA*) [46], all of which belong to the membrane transport category. Remaining genes annotated as uncharacterized hypothetical proteins as per UniProt are surprisingly predicted to code for proteins involved in functions related to virulence, pathogenesis, defense, solute transport, and DNA metabolism (**Table 4**).

### Conclusions

Upon comparison of the non-O1 *V. cholerae* PS15 genome, a non-toxicogenic strain, to that of an O1 *V. cholerae* N16961, a toxicogenic strain, we found that of the 619 missing genes, 56 of these missing genomic elements encode dormancy, sporulation, ribosome modulation in persister cells, lipid metabolism, phage infection, nucleoside metabolism, and sulfur metabolism which in turn is essential for biosynthesis of amino acids, vitamins and prosthetic groups [43]. As non-O1 *V. cholerae* lacks genes coding for metabolism of sulfur, the non-O1 serotype is predicted to be unable to convert naturally available sulfur to sulfide, which could then be incorporated into various sulfur containing metabolites. Sulfur is critical for the biosynthesis of many important compounds like amino acids (cysteine and methionine), vitamins (biotin, thiamin), and prosthetic groups (Fe-S clusters) [43]. These genetic elements and their putative gene products represent novel and promising targets for modulation of gene expression or activity and therapeutic efforts [47], in order to effectively reduce conditions that foster virulence and dissemination of *V. cholerae* pathogens through populations. These determinants, therefore, clearly also warrant further studies in order to elucidate the complete molecular mechanisms of pathogenesis in cholera infections.

Not surprisingly, also among the 56 missing genes in the non-O1 PS15 genome are those that are known to confer virulence, such as the cholera toxin [40], colonization factors [48], and antimicrobial resistance mechanisms [16]. We thus confirm that the genes encoding the cholera toxin are absent from the genome of the non-toxicogenic *V. cholerae* PS15. We confirm, however, the presence of other genes predicted to encode distinct toxins and colonization factors, as previously shown for the non-O1 *V. cholerae* strain NRT36S [49]. This latter study and our findings here are consistent with previous work demonstrating that aquatic environments are reservoirs for O1 and non-O1 *V. cholerae* [50], predicting that

such environments allow genetic exchange between unrelated strains. In order to gain valuable insights into enhancing chemotherapeutic efficacy against cholera, it is imperative to study and gain understanding into the modes of action of the toxicity-inducing factors combined with other antibacterial resistance factors in toxigenic *V. cholerae* [51].

Interestingly, we found that the genome of the nontoxigenic *V. cholerae* PS15 strain harbors genes absent from the genome of its toxigenic counterpart, N16961. Such determinants mainly include still uncharacterized genetic elements that are predicted to encode proteins that confer virulence, disease, defense, membrane solute transport and DNA metabolism, suggesting that PS15 may be pathogenic to organisms excluding humans, perhaps in environments such as estuary waters [52,53]. Among the genetic determinants unique to PS15 that have been experimentally characterized include OppA, an oligopeptide primary active transporter [44], and SecF, a protein exporter [12]. We propose that these unique genetic elements represent good targets for future development of new therapies against *V. cholerae* infections in animals other than humans.

The genome of non-O1 *V. cholerae* PS15 shares >97% identity with El Tor O1 biovar *V. cholerae* strain N16961, as per BLAST analysis at the nucleotide level. Based on the alignment of the non-O1 PS15 genome with that of O1 N16961, chromosomes 1 and 2 were assigned to the PS15 genome (**Figure 2**). This tentative chromosome assignment will require confirmation with additional experimental work. Even though the genomes of both strains are highly similar to each other, the non-O1 PS15 microorganism is considered to be non-pathogenic, compared to the O1 N16961 strain, possibly due to the absence of the cholera toxin in PS15, which is responsible for endemic and pandemic diseases [54]. More recent genomic analysis, however, has demonstrated that other genetic elements are also critical for conferring pathogenesis such as genes coding for housekeeping, homeostasis, metabolism, energy generation, and antimicrobial resistance-type functions [55]. Our phylogenetic and genome comparison analyses between the toxigenic and non-toxigenic *V. cholerae* microorganisms support both of these contentions. Further work with additional variants, such as atypical El Tor [56], NRT36S [49], and CT-producing non-O1 strains [57], will be necessary to definitively gain a complete picture of the relationships between pathogenic versus non-pathogenic *V. cholerae*.

Remarkably, we found that both of the toxigenic and non-toxigenic *V. cholerae* strains harbor a variety of genes that have previously been demonstrated to confer multidrug resistance via active drug efflux pump systems, such as AcrAB, NorM / VcaM, SugE, and VcaM [58]. All six RND transporters in *V. cholerae* N16961 have been studied physiologically [59], and our data showed that *V. cholerae* PS15 was missing only one of these pumps, called VexA. Additionally, we found a shared but uncharacterized genetic element, VC\_A0083 in the toxigenic strain and OSU\_1537 in the non-toxigenic strain, tentatively called multidrug resistance protein D and predicted to encode an MFS drug efflux pump. These multidrug resistance mechanisms may be important because of their potential selection and maintenance in environments containing antimicrobial agents, their genetic mobility to other microorganisms, and dissemination within populations [60-64].



We conclude that the study and comparison of the genomic sequences between pathogens and their non-virulent counterparts will help discover genes encoding both the classical virulence factors and those encoding novel virulence factors. Future work will focus on the study of solute transport and antibacterial resistance mechanisms of *V. cholerae* pathogenic strains and on the identification of novel housekeeping genes which may be equally significant in contributing towards the microorganisms' pathogenicity [17,65,66].

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgement and funding

This project was supported by a grant from the National Institute of General Medical Sciences (P20GM103451) of the National Institutes of Health and an Internal Research Grant from Eastern New Mexico University.

## References

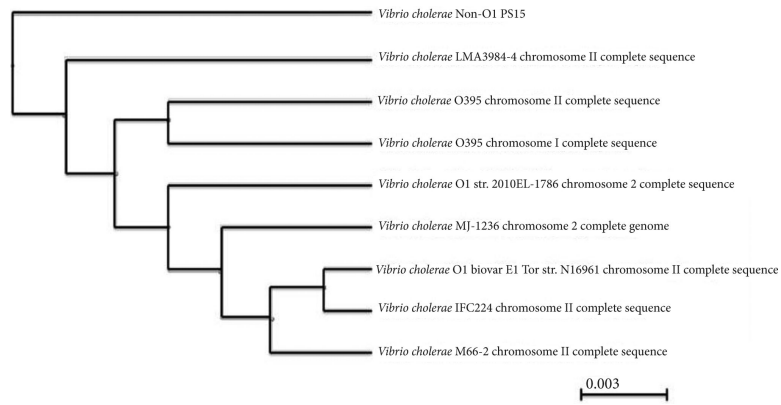
1. Mandal S, Mandal MD, Pal NK. Cholera: a great global concern. *Asian Pac J Trop Med.* 2011; 4:573–80. | Article | PubMed. [PubMed: 21803312]
2. Morris JG Jr. Cholera and other types of vibriosis: a story of human pandemics and oysters on the half shell. *Clin Infect Dis.* 2003; 37:272–80. | Article | PubMed. [PubMed: 12856219]
3. Sack DA, Sack RB, Nair GB, Siddique AK. Cholera. *Lancet.* 2004; 363:223–33. | Article | PubMed. [PubMed: 14738797]
4. Guerrant RL, Carneiro-Filho BA, Dillingham RA. Cholera, diarrhea, and oral rehydration therapy: triumph and indictment. *Clin Infect Dis.* 2003; 37:398–405. | Article | PubMed. [PubMed: 12884165]
5. Hasan NA, Choi SY, Eppinger M, Clark PW, Chen A, Alam M, Haley BJ, Taviani E, Hine E, Su Q, Tallon LJ, Prosper JB, Furth K, Hoq MM, Li H, Fraser-Liggett CM, Cravioto A, Huq A, Ravel J, Cebula TA, Colwell RR. Genomic diversity of 2010 Haitian cholera outbreak strains. *Proc Natl Acad Sci U S A.* 2012; 109:E2010–7. Article | PubMed Abstract | PubMed Full Text. [PubMed: 22711841]
6. Mutreja A, Kim DW, Thomson NR, Connor TR, Lee JH, Kariuki S, Croucher NJ, Choi SY, Harris SR, Lebens M, Niyogi SK, Kim EJ, Ramamurthy T, Chun J, Wood JL, Clemens JD, Czerkinsky C, Nair GB, Holmgren J, Parkhill J, Dougan G. Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature.* 2011; 477:462–5. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21866102]
7. Centers for Disease C. Prevention. Update: outbreak of cholera ---Haiti, 2010. *MMWR Morb Mortal Wkly Rep.* 2010; 59:1586–90. | Article | PubMed. [PubMed: 21150867]
8. Octavia S, Salim A, Kurniawan J, Lam C, Leung Q, Ahsan S, Reeves PR, Nair GB, Lan R. Population structure and evolution of non-O1/non-O139 *Vibrio cholerae* by multilocus sequence typing. *PLoS One.* 2013; 8:e65342. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 23776471]
9. Rabbani GH, Greenough WB 3rd. Food as a vehicle of transmission of cholera. *J Diarrhoeal Dis Res.* 1999; 17:1–9. | PubMed. [PubMed: 10892490]
10. Rahaman MM, Majid MA, Alam A, Islam MR. Effects of doxycycline in actively purging cholera patients: a double-blind clinical trial. *Antimicrob Agents Chemother.* 1976; 10:610–2. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 791107]
11. Okada K, Roobthaisong A, Swaddiwudhipong W, Hamada S, Chantaroj S. *Vibrio cholerae* O1 isolate with novel genetic background, Thailand-Myanmar. *Emerg Infect Dis.* 2013; 19:1015–7. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 23735934]

12. Krishna BV, Patil AB, Chandrasekhar MR. Fluoroquinolone-resistant *Vibrio cholerae* isolated during a cholera outbreak in India. *Trans R Soc Trop Med Hyg.* 2006; 100:224–6. | Article | PubMed. [PubMed: 16246383]
13. Levy SB. Antibiotic resistance: consequences of inaction. *Clin Infect Dis.* 2001; 33(Suppl 3):S124–9. | Article | PubMed. [PubMed: 11524708]
14. Kitaoka M, Miyata ST, Unterweger D, Pukatzki S. Antibiotic resistance mechanisms of *Vibrio cholerae*. *J Med Microbiol.* 2011; 60:397–407. | Article | PubMed. [PubMed: 21252269]
15. Kumar S, Parvathi A, Hernandez RL, Cadle KM, Varela MF. Identification of a novel UDP-N-acetylglucosamine enolpyruvyl transferase (MurA) from *Vibrio fischeri* that confers high fosfomycin resistance in *Escherichia coli*. *Arch Microbiol.* 2009; 191:425–9. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 19277606]
16. Kumar, S.; Varela, MF. Molecular mechanisms of bacterial resistance to antimicrobial agents.. In: Méndez-Vilas, A., editor. *Microbial pathogens and strategies for combating them: science, technology and education.* Formatex Research Center. 2013. p. 522-534. | Pdf
17. Smith KP, Kumar S, Varela MF. Identification, cloning, and functional characterization of EmrD-3, a putative multidrug efflux pump of the major facilitator superfamily from *Vibrio cholerae* O395. *Arch Microbiol.* 2009; 191:903–11. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 19876617]
18. Colmer JA, Fralick JA, Hamood AN. Isolation and characterization of a putative multidrug resistance pump from *Vibrio cholerae*. *Mol Microbiol.* 1998; 27:63–72. | Article | PubMed. [PubMed: 9466256]
19. Kumar S, Varela MF. Biochemistry of bacterial multidrug efflux pumps. *Int J Mol Sci.* 2012; 13:4484–95. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 22605991]
20. Kumar, S.; Floyd, JT.; He, G.; Varela, MF. Recent Research Developments in Antimicrobial Agents & Chemotherapy. Research Signpost, Inc.; Kerala, India: 2013. Bacterial antimicrobial efflux pumps of the MFS and MATE transporter families: A review.; p. 1-21. | Article
21. Floyd, JT.; Kumar, S.; Mukherjee, MM.; He, G.; Varela, MF. A review of the molecular mechanisms of drug efflux in pathogenic bacteria: A structure-function perspective.. In: Kerala, Shankar P., editor. *Recent Research Developments in Membrane Biology.* Vol. 3. Research Signpost, Inc.; India: 2013. p. 15-66. | Article
22. Boyd EF, Heilpern AJ, Waldor MK. Molecular analyses of a putative CTXphi precursor and evidence for independent acquisition of distinct CTX(phi)s by toxigenic *Vibrio cholerae*. *J Bacteriol.* 2000; 182:5530–8. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 10986258]
23. Baudry B, Fasano A, Ketley J, Kaper JB. Cloning of a gene (zot) encoding a new toxin produced by *Vibrio cholerae*. *Infect Immun.* 1992; 60:428–34. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 1730472]
24. Waldor MK, Rubin EJ, Pearson GD, Kimsey H, Mekalanos JJ. Regulation, replication, and integration functions of the *Vibrio cholerae* CTXphi are encoded by region RS2. *Mol Microbiol.* 1997; 24:917–26. | Article | PubMed. [PubMed: 9220000]
25. Dobrindt U, Reidl J. Pathogenicity islands and phage conversion: evolutionary aspects of bacterial pathogenesis. *Int J Med Microbiol.* 2000; 290:519–27. | Article | PubMed. [PubMed: 11100826]
26. Harris JB, LaRocque RC, Qadri F, Ryan ET, Calderwood SB. Cholera. *Lancet.* 2012; 379:2466–76. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 22748592]
27. Walker E, Carpenter J, Plemmons R, Fader R. Freshwater non-O1 *Vibrio cholerae* infection. *South Med J.* 2010; 103:1061–2. | Article | PubMed. [PubMed: 20818301]
28. Kumar S, Lindquist IE, Sundararajan A, Rajanna C, Floyd JT, Smith KP, Andersen JL, He G, Ayers RM, Johnson JA, Werdann JJ, Sandoval AA, Mojica NM, Schilkey FD, Mudge J, Varela MF. Genome Sequence of Non-O1 *Vibrio cholerae* PS15. *Genome Announc.* 2013; 1:e00227–00212. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 23409261]
29. Kumar S, Lindquist IE, Sundararajan A, Rajanna C, Floyd JT, Smith KP, Andersen JL, He G, Ayers RM, Johnson JA, et al. Genome Sequence of Non-O1 *Vibrio cholerae* PS15. *Genome Announc.* 2013:1. | Article.

30. Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, et al. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*. 2000; 406:477–83. | Article | PubMed. [PubMed: 10952301]
31. Kaysner CA, Abeyta C Jr, Wekell MM, DePaola A Jr, Stott RF, Leitch JM. Incidence of *Vibrio cholerae* from estuaries of the United States West Coast. *Appl Environ Microbiol*. 1987; 53:1344–8. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 3606111]
32. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. 2008; 9:75. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 18261238]
33. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res*. 2004; 32:D115–9. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 14681372]
34. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997; 25:3389–402. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 9254694]
35. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987; 4:406–25. | Article | PubMed. [PubMed: 3447015]
36. Grant JR, Stothard P. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res*. 2008; 36:W181–4. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 18411202]
37. Nikaido H. Structure and mechanism of RND-type multidrug efflux pumps. *Adv Enzymol Relat Areas Mol Biol*. 2011; 77:1–60. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21692366]
38. He GX, Zhang C, Crow RR, Thorpe C, Chen H, Kumar S, Tsuchiya T, Varela MF. SugE, a new member of the SMR family of transporters, contributes to antimicrobial resistance in *Enterobacter cloacae*. *Antimicrob Agents Chemother*. 2011; 55:3954–7. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21576447]
39. Kuroda T, Tsuchiya T. Multidrug efflux transporters in the MATE family. *Biochim Biophys Acta*. 2009; 1794:763–8. | Article | PubMed. [PubMed: 19100867]
40. Sanchez J, Holmgren J. Cholera toxin - a foe & a friend. *Indian J Med Res*. 2011; 133:153–63. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21415489]
41. Woolley RC, VEDIYAPPAN G, Anderson M, Lackey M, Ramasubramanian B, Jiangping B, Borisova T, Colmer JA, Hamood AN, McVay CS, Fralick JA. Characterization of the *Vibrio cholerae* vceCAB multiple-drug resistance efflux operon in *Escherichia coli*. *J Bacteriol*. 2005; 187:5500–3. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 16030246]
42. Lauhon CT. Mechanism of N6-threonylcarbamoyladenonsine (t(6) A) biosynthesis: isolation and characterization of the intermediate threonylcarbamoyl-AMP. *Biochemistry*. 2012; 51:8950–63. | Article | PubMed. [PubMed: 23072323]
43. Hall RH, Khambaty FM, Kothary M, Keasler SP. Non-O1 *Vibrio cholerae*. *Lancet*. 1993; 342:430. | PubMed. [PubMed: 8101920]
44. Lee EM, Ahn SH, Park JH, Lee JH, Ahn SC, Kong IS. Identification of oligopeptide permease (opp) gene cluster in *Vibrio fluvialis* and characterization of biofilm production by oppA knockout mutation. *FEMS Microbiol Lett*. 2004; 240:21–30. Article | PubMed. [PubMed: 15500975]
45. Fandl J, Tai PC. Protein translocation in vitro: biochemical characterization of genetically defined translocation components. *J Bioenerg Biomembr*. 1990; 22:369–87. | Article | PubMed. [PubMed: 2167893]
46. Linton KJ, Higgins CF. The *Escherichia coli* ATP-binding cassette (ABC) proteins. *Mol Microbiol*. 1998; 28:5–13. | Article | PubMed. [PubMed: 9593292]
47. Kumar S, Mukherjee MM, Varela MF. Modulation of Bacterial Multidrug Resistance Efflux Pumps of the Major Facilitator Superfamily. *International Journal of Bacteriology*. 2013; 2013:15. | Article.

48. Ghose AC. Adherence & colonization properties of *Vibrio cholerae* & diarrhoeagenic *Escherichia coli*. *Indian J Med Res.* 1996; 104:38–51. | Article | PubMed. [PubMed: 8783506]
49. Chen Y, Johnson JA, Pusch GD, Morris JG Jr, Stine OC. The genome of non-O1 *Vibrio cholerae* NRT36S demonstrates the presence of pathogenic mechanisms that are distinct from those of O1 *Vibrio cholerae*. *Infect Immun.* 2007; 75:2645–7. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 17283087]
50. Singh DV, Matte MH, Matte GR, Jiang S, Sabeena F, Shukla BN, Sanyal SC, Huq A, Colwell RR. Molecular analysis of *Vibrio cholerae* O1, O139, non-O1, and non-O139 strains: clonal relationships between clinical and environmental isolates. *Appl Environ Microbiol.* 2001; 67:910–21. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 11157262]
51. Varela MF, Kumar S, He G. Potential for inhibition of bacterial efflux pumps in multidrug-resistant *Vibrio cholerae*. *Indian J Med Res.* 2013; 138:285–7. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 24135170]
52. Colwell RR. Global climate and infectious disease: the cholera paradigm. *Science.* 1996; 274:2025–31. | Article | PubMed. [PubMed: 8953025]
53. Ramamurthy T, Bag PK, Pal A, Bhattacharya SK, Bhattacharya MK, Shimada T, Takeda T, Karasawa T, Kurazono H, Takeda Y, et al. Virulence patterns of *Vibrio cholerae* non-O1 strains isolated from hospitalised patients with acute diarrhoea in Calcutta, India. *J Med Microbiol.* 1993; 39:310–7. | Article | PubMed. [PubMed: 8411093]
54. Vanden Broeck D, Horvath C, De Wolf MJ. *Vibrio cholerae*: cholera toxin. *Int J Biochem Cell Biol.* 2007; 39:1771–5. | Article | PubMed. [PubMed: 17716938]
55. Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ. Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. *Proc Natl Acad Sci U S A.* 2002; 99:1556–61. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 11818571]
56. Ceccarelli D, Spagnoletti M, Bacciu D, Cappuccinelli P, Colombo MM. New *V. cholerae* atypical El Tor variant emerged during the 2006 epidemic outbreak in Angola. *BMC Microbiol.* 2011; 11:130. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21668969]
57. Madhusudana RB, Surendran PK. Detection of ctx gene positive non-O1/non-O139 *V. cholerae* in shrimp aquaculture environments. *J Food Sci Technol.* 2013; 50:496–504. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 24425944]
58. Levy SB. Active efflux, a common mechanism for biocide and antibiotic resistance. *J Appl Microbiol.* 2002; 92(Suppl):65S–71S. | Article | PubMed. [PubMed: 12000614]
59. Rahman MM, Matsuo T, Ogawa W, Koterasawa M, Kuroda T, Tsuchiya T. Molecular cloning and characterization of all RND-type efflux transporters in *Vibrio cholerae* non-O1. *Microbiol Immunol.* 2007; 51:1061–70. | Article | PubMed. [PubMed: 18037783]
60. Ghosh A, Ramamurthy T. Antimicrobials & cholera: are we stranded? *Indian J Med Res.* 2011; 133:225–31. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21415499]
61. Marshall BM, Levy SB. Food animals and antimicrobials: impacts on human health. *Clin Microbiol Rev.* 2011; 24:718–33. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21976606]
62. Fazil MH, Singh DV. *Vibrio cholerae* infection, novel drug targets and phage therapy. *Future Microbiol.* 2011; 6:1199–208. | Article | PubMed. [PubMed: 22004038]
63. Ploy MC, Lambert T, Couty JP, Denis F. Integrons: an antibiotic resistance gene capture and expression system. *Clin Chem Lab Med.* 2000; 38:483–7. | Article | PubMed. [PubMed: 10987194]
64. Silbergeld EK, Graham J, Price LB. Industrial food animal production, antimicrobial resistance, and human health. *Annu Rev Public Health.* 2008; 29:151–69. | Article | PubMed. [PubMed: 18348709]
65. Rambhatla P, Kumar S, Floyd JT, Varela MF. Molecular cloning and characterization of mannitol-1-phosphate dehydrogenase from *Vibrio cholerae*. *J Microbiol Biotechnol.* 2011; 21:914–20. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21952367]
66. Kumar S, Smith KP, Floyd JL, Varela MF. Cloning and molecular analysis of a mannitol operon of phosphoenolpyruvate-dependent phosphotransferase (PTS) type from *Vibrio cholerae* O395. *Arch*

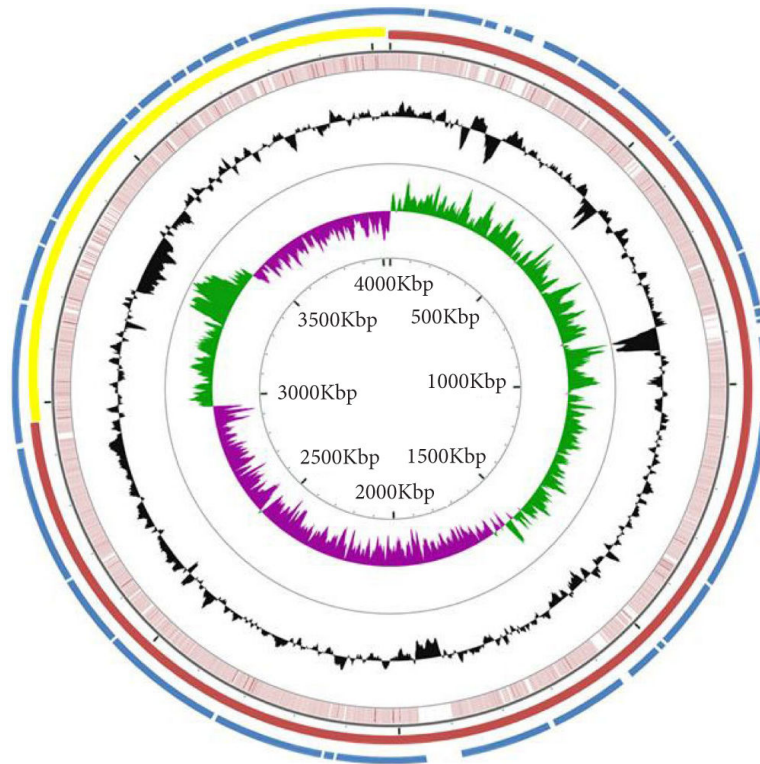
Microbiol. 2011; 193:201–8. | Article | PubMed Abstract | PubMed Full Text. [PubMed: 21184218]



**Figure 1. Phylogenetic tree showing comparison between the non-O1 *Vibrio cholerae* PS15 genome and its most closely related microorganisms**

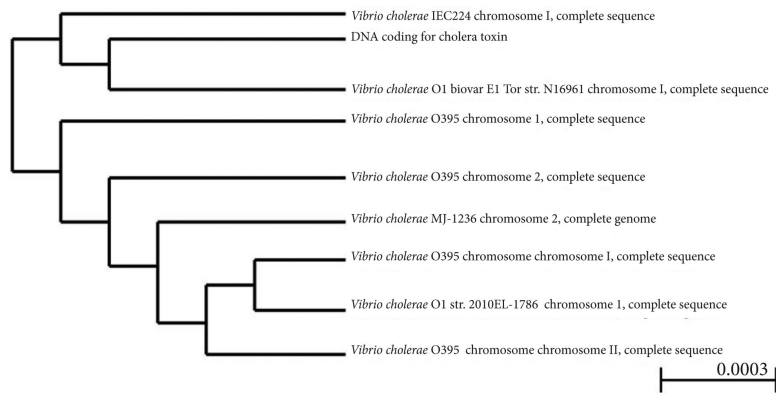
A phylogenetic tree was generated using BLAST pair wise alignment by the Tree Neighbor Joining method for the DNA sequence of PS15. The distance between the internal node for a subtree or alignment is 0.003. The accession numbers for the sequences selected are as follows: *V. cholerae* non-O1 PS15 (AIJR000000000), *V. cholerae* LMA3984-4 chromosome II complete sequence (NC\_017269.1), *V. cholerae* O395 chromosome II complete sequence (NC\_012583.1), *V. cholerae* O395 chromosome I complete sequence (NC\_009456.1), *V. cholerae* O1 str. 2010EL-1786 chromosome 2 complete sequence (NC\_016446.1), *V. cholerae* MJ-1236 chromosome 2 complete genome (NC\_012667.1), *V. cholerae* O1 biovar El Tor str. N16961 chromosome II complete sequence (NC\_002506.1), *V. cholerae* IEC224 chromosome II complete sequence (NC\_016945.1), and *V. cholerae* M66-2 chromosome II complete sequence (NC\_012580.1).





**Figure 2. Circular genome map comparing *Vibrio cholerae* non-O1 PS15 with O1 *V. cholerae* N16961**

BLAST was performed on the DNA sequence of the *V. cholerae* non-O1 PS15 (3,910,387 base pairs) with *V. cholerae* El Tor N16961 (4,033,460 base pairs) using CGView server to yield a structural representation of both genomes. The FASTA sequences of the genomes were used to generate the graphical circular genome map, and the Global Blast Settings or parameters selected were as follows: query split size=50,000, and overlap split size=0; Blast 1: non-O1 *V. cholerae* PS15, and blastn expect=0.1. The outermost circle in blue represents the entire genome of non-O1 *V. cholerae* PS15, and PS15 gaps in the genome alignment are indicated in white. In the next inner circle, the red and yellow color coded regions represent chromosomes 1 and 2, respectively, of *V. cholerae* N16961. The third inner circle in pink shows matching nucleotide base pairs representing the BLAST analysis for *V. cholerae* N16961 and *V. cholerae* PS15. The fourth inner circle represents the total G+C content color coded in black. In the fifth inner circle, the region in green represents the G+C content of the forward strand, and the region in purple represents the G+C content of the reverse strand. The innermost circle represents the base pair numbers, where kbp stands for kilo base pairs.



**Figure 3. Phylogenetic tree for genomes sharing the cholera toxin**

A phylogenetic tree representing virulence of *V. cholerae* O1 due to the presence of cholera toxin was produced by the Neighbor Joining Tree method, using BLAST pair wise alignment. Distance between internal node for a subtree or alignment is 0.0003. The distance tree of the result was generated by BLAST for sequences that share DNA encoding the cholera toxin. The accession numbers for the sequences selected are as follows: DNA coding the cholera toxin: GenBank: E00132.1, *V. cholerae* IEC224 chromosome I, complete sequence (NC\_016944.1), *V. cholerae* O1 biovar El Tor str. N16961 chromosome I, complete sequence (NC\_002505.1), *V. cholerae* O395 chromosome 1, complete sequence (NC\_009456.1), *V. cholerae* O395 chromosome 2, complete sequence (NC\_009457.1), *V. cholerae* MJ-1236 chromosome 2, complete genome (NC\_012667.1), *V. cholerae* O395 chromosome I, complete sequence (NC\_012582.1), *V. cholerae* O1 str. 2010EL-1786 chromosome 1, complete sequence (NC\_016445.1), and *V. cholerae* O395 chromosome II, complete sequence (NC\_012583.1).

**Table 1**

Bacterial strains and protein encoding genes.

Strain	<i>V. cholerae</i> N16961	<i>V. cholerae</i> PS15
Length (bp)	4,033,460	3,910,387
G+C content (%)	47.7 (chromosome 1)	47.55
	46.9 (chromosome 2)	--
Number of protein-coding genes with assigned function	424	272
Number of hypothetical	398	607
Total number of genes	822	879

Comparison of the general features for O1 *V. cholerae* N16961 and non-O1 *V. cholerae* PS15 contrasting the base pair lengths (bp), GC content, and genes associated with the biological or physiological categories such as virulence, disease, defense, membrane transport, DNA metabolism, regulons, dormancy, sporulation, phages, prophages, transposable elements and plasmids.

Table 2

Genes shared between genomes of non-O1 *Vibrio cholerae* PS15 and O1 *V. cholerae* N16961.

Description	Abbreviation N16961	Accession N16961	Abbreviation PS15	Accession PS15
DNA gyrase inhibitor YacG <sup>*</sup>	<i>yacG</i>	Q9KPE1	<i>yacG</i>	L1QXA7
Transcription elongation factor GreB <sup>*</sup>	<i>greB</i>	Q9KNL7	<i>greA</i>	L1R2F8
Autoinducer 2 sensor kinase/phosphatase LuxQ <sup>*</sup>	<i>luxQ</i>	Q9KLN7	<i>OSU_2901</i>	L1QTW3
Putative uncharacterized protein VC0929 <sup>*</sup>	<i>VC0929</i>	Q9KTH3	<i>OSU_1349</i>	L1QYD5
Putative uncharacterized protein <sup>*</sup>	<i>VC_A0118</i>	Q9KN47	<i>OSU_1575</i>	L1QXZ3
Release factor glutamine methyltransferase <sup>†</sup>	<i>prmC</i>	Q9KQ26	<i>prmC</i>	L1QU18
Probable potassium transport system protein kup <sup>†</sup>	<i>kup</i>	Q9KM59	<i>kup</i>	L1QTS9
Electron transport complex protein RnfC <sup>†</sup>	<i>rnfC</i>	Q9KT88	<i>rnfC</i>	L1QUE4
Probable oxaloacetate decarboxylase gamma chain 2 <sup>†</sup>	<i>oadG2</i>	Q9KTU4	<i>oadG</i>	L1R348
Thiamine import ATP-binding protein ThiQ <sup>†</sup>	<i>thiQ</i>	Q9KP42	<i>OSU_1092</i>	L1QYX4
Vitamin B12 import ATP-binding protein BtuD <sup>†</sup>	<i>btuD</i>	Q9KSL1	<i>btuD</i>	L1R0X6
Copper-exporting P-type ATPase A <sup>†</sup>	<i>copA</i>	Q9KPZ7	<i>OSU_0952</i>	L1QZQ5
Putative fluoride ion transporter CrcB <sup>†</sup>	<i>crcB</i>	Q9KVS9	<i>crcB</i>	L1QWW5
MSHA biogenesis protein MshG <sup>†</sup>	<i>VC_0406</i>	Q9KUV6	<i>OSU_1460</i>	L1QXZ4
MSHA biogenesis protein MshH <sup>†</sup>	<i>VC_0398</i>	Q9KUW1	<i>OSU_1452</i>	L1QYV1
Sigma-54 dependent transcriptional regulator <sup>†</sup>	<i>VC_1817</i>	Q9KR30	<i>OSU_1624</i>	L1QXJ1
Transport ATP-binding protein CydD <sup>†</sup>	<i>VC_1181</i>	Q9KSS5	<i>OSU_2736</i>	L1QUK9
Transport ATP-binding protein CydC <sup>†</sup>	<i>VC_1180</i>	Q9KSS6	<i>OSU_2737</i>	L1QVL8
Amino acid ABC transporter, permease proteint	<i>VC_A1040</i>	Q9KKR2	<i>OSU_1389</i>	L1QY57
Peptide ABC transporter, permease protein, putative <sup>†</sup>	<i>VC_A0590</i>	Q9KLN9	<i>OSU_0554</i>	L1R1Z6
Multidrug transporter, putative <sup>†</sup>	<i>VC_1391</i>	Q9KS68	<i>OSU_2315</i>	L1QVM0
Thiamin ABC transporter, periplasmic thiamin-binding protein <sup>†</sup>	<i>VC_2539</i>	Q9KP40	<i>OSU_1094</i>	L1QYM0
Cbb3-type cytochrome c oxidase subunit <sup>†</sup>	<i>VC_1439</i>	Q9KS22	<i>OSU_0068</i>	L1R226
Benzoate transport protein <sup>†</sup>	<i>VC_1970</i>	Q9KQM8	<i>OSU_3112</i>	L1QUT1
ABC transporter, permease protein <sup>†</sup>	<i>VC_A1099</i>	Q9KKK5	<i>OSU_2952</i>	L1QTR7
Transporter, LysE family <sup>†</sup>	<i>VC_1712</i>	Q9KRD0	<i>OSU_1426</i>	L1QZ42
ABC-type multidrug transport system, permease component <sup>†</sup>	<i>VC_0590</i>	Q9KUD2	<i>OSU_0545</i>	L1R0T8
Proton/glutamate symport protein / Sodium/ glutamate symport protein <sup>†</sup>	<i>VC_A0088</i>	Q9KN77	<i>OSU_1542</i>	L1QY02

Description	Abbreviation N16961	Accession N16961	Abbreviation PS15	Accession PS15
NADH dehydrogenase, putative <sup>†</sup>	VC_1581	Q9KRQ5	OSU_1036	L1R051
Na <sup>+</sup> /H <sup>+</sup> antiporter, putative <sup>†</sup>	VC_0389	Q9KUX0	OSU_1443	L1QY90
Cytochrome b561, putative <sup>†</sup>	VC_A0249	Q9KMS1	OSU_0494	L1R0V0
Osmosensitive K <sup>+</sup> channel histidine kinase KdpD/ Sensor histidine kinase <sup>†</sup>	VC_A0531	Q9KM57	OSU_3424	L1QSI3
Xanthine/uracil permease family proteint	VC_2712	Q9KNM0	OSU_0044	L1R323
DNA polymerase I/DNA polymerase II/ DNA polymerase IV <sup>§</sup>	<i>dinB</i>	Q9KPS5	<i>dinB</i>	L1R1B3
Nuclease SbcCD subunit C <sup>§</sup>	<i>sbcC</i>	Q9KM67	OSU_3415	L1QTC3
Deoxyribodipyrimidine photo-lyase <sup>§</sup>	<i>phrA</i>	Q9KNA8	OSU_1512	L1QXV4
Tyrosine--tRNA ligase 1 <sup>§</sup>	<i>tyrS1</i>	Q9KUQ0	<i>tyrS</i>	L1QY24
Formate--tetrahydrofolate ligase <sup>§</sup>	<i>fhs</i>	Q9KLX7	<i>fhs</i>	L1R108
Valine--tRNA ligase <sup>§</sup>	<i>valS</i>	Q9KP73	<i>valS</i>	L1QXG6
ADP-L-glycero-D-manno-heptose-6-epimerase <sup>§</sup>	<i>hldD</i>	Q06963	<i>hldD</i>	L1QT14
Ferrochelata <sup>§</sup>	<i>hemH</i>	Q9KTB6	<i>hemH</i>	L1QUK6
Tetraacyldisaccharide 4'-kinase <sup>§</sup>	<i>lpxK</i>	Q9KQX0	<i>lpxK</i>	L1QYU9
7-carboxy-7-deazaguanine synthase <sup>§</sup>	<i>queE</i>	Q9KS94	<i>queE</i>	L1QVP2
A/G-specific adenine glycosylase <sup>§</sup>	VC_0452	Q9KUR3	OSU_1912	L1QX56
Exodeoxyribonuclease V alpha chain <sup>§</sup>	VC_2319	Q9KPP7	OSU_0379	L1R140
Exodeoxyribonuclease V beta chain/ Exodeoxyribonuclease III <sup>§</sup>	VC_2320	Q9KPP6	OSU_0378	L1R0Z3
Exodeoxyribonuclease V gamma chain <sup>§</sup>	VC_2322	Q9KPP4	OSU_0376	L1R185
DNA helicase IV <sup>§</sup>	VC_A0204	Q9KMW4	OSU_3456	L1QTP1
Putative phosphatase YqaB <sup>§</sup>	VC_A0662	Q9KLS9	OSU_0250	L1R2A2
Non-canonical purine NTP phosphatase <sup>§</sup>	VC_0702	Q9KU27	OSU_3053	L1QU16
Non-canonical purine NTP pyrophosphatase <sup>§</sup>	VC_0456	Q9KUQ9	OSU_1907	L1QX52
Putative quercetin 2,3-dioxygenase VC_A0969 <sup>§</sup>	VC_A0969	Q9KKY1	OSU_1671	L1QYQ8
Aldose 1-epimerase <sup>§</sup>	VC_1594	Q9KRP2	OSU_1049	L1QZA8
Dihydrofolate reductase <sup>§</sup>	VC_0440	Q9KUS5	OSU_1494	L1QZC9
Cbb3-type cytochrome c oxidase subunit <sup>§</sup>	VC_1439	Q9KS22	OSU_0068	L1R226
Molybdopterin-guanine dinucleotide biosynthesis protein MobA <sup>*</sup>	<i>mobA</i>	Q9KRV8	<i>mobA</i>	L1QZNI
Molybdopterin-guanine dinucleotide biosynthesis protein MobB <sup>*</sup>	VC_1527	Q9KRV7	OSU_0936	L1QZT6
DamX-related protein <sup>//</sup>	VC_2627	Q9KNV3	OSU_1180	L1QZ10

The table represents a list of genes present in both of the genomes in which the genes share 98 % identity. Included in this table are genes predicted to code for categories including virulence, disease and defense, membrane transport, DNA metabolism, dormancy and sporulation and regulons. In the table, the first column includes protein designations and descriptions. The second and fourth columns represent abbreviated gene identifications; the third and fifth represent accession numbers for the listed genes.

\* Proteins with the symbol have putative functions in virulence, disease and defense.

† Proteins with and symbols represent proteins that have functions in membrane transport and DNA metabolism categories, respectively.

§ Proteins with and symbols represent proteins that have functions in membrane transport and DNA metabolism categories, respectively.

// Proteins with the symbol have functions within dormancy and sporulation categories, and include proteins that are part of regulons.

\* Proteins with the symbol have functions within dormancy and sporulation categories, and include proteins that are part of regulons.



**Table 3**Genes absent in non-O1 *Vibrio cholerae* genome and present in O1 *V. cholerae* genome.

Description	Abbreviation N16961	Accession N16961
Accessory cholera enterotoxin *	<i>ace</i>	P38441
Cholera enterotoxin subunit B *	<i>ctxB</i>	P01556
Cholera enterotoxin subunit A *	<i>ctxA</i>	P01555
Zona occludens toxin *	<i>zot</i>	P38442
Toxin coregulated pilus biosynthesis protein F *	<i>tcpF</i>	P0C6Q5
Toxin-coregulated pilus biosynthesis protein P †	<i>tcpP</i>	Q7BGC9
Toxin coregulated pilus biosynthesis protein I †	<i>tcpI</i>	P0C6D8
Toxin coregulated pilus biosynthesis protein H †	<i>tcpH</i>	P29489
Toxin coregulated pilus biosynthesis protein B †	<i>tcpB</i>	P23476
Toxin coregulated pilus biosynthesis protein E †	<i>tcpE</i>	P0C6C9
Transcriptional activator protein NhaR †	<i>nhaR</i>	P52692
Outer membrane lipoprotein blc †	<i>blc</i>	Q08790
ATP synthase protein I †	<i>atpI</i>	Q9KNG8
Type 4 prepilin-like proteins leader peptide-processing enzyme §	<i>tcpJ</i>	P0C6D9
N5-carboxyaminoimidazole ribonucleotide synthase §	<i>purK</i>	Q9KVT8
Coproporphyrinogen-III oxidase §	<i>hemF</i>	Q9KVT4
Aldehyde dehydrogenase §	<i>aldA</i>	P0C6D7
Putative N-acetylmannosamine-6-phosphate 2-epimerase §	<i>nanE</i>	Q9KR62
N-acetylmannosamine kinase §	<i>nanK</i>	Q9KR61
N-acetylneuraminate epimerase §	<i>nanM</i>	Q9KR69
N5-carboxyaminoimidazole ribonucleotide mutase §	<i>purE</i>	Q9KVT7
Ribosome modulation factor //	<i>Rmf</i>	Q9KRZ9
Accessory colonization factor AcfA *	<i>VC_0844</i>	H9L4S5
Accessory colonization factor AcfB *	<i>VC_0840</i>	Q9KTQ7
TagE protein *	<i>VC_A1043</i>	Q9KKQ9
TagE protein *	<i>VC_0843</i>	H9L4P5
Uncharacterized protein VC_1460 *	<i>VC_1460</i>	P38443
Fimbrial biogenesis and twitching motility protein, putative †	<i>VC_1612</i>	Q9KRM4
Type IV pilin, putative †	<i>VC_0858</i>	Q9KTP3
Fimbrial protein †	<i>VC_2423</i>	Q9KPE5

Description	Abbreviation N16961	Accession N16961
Fimbrial assembly protein <sup>†</sup>	VC_2630	Q9KNV0
Putative uncharacterized protein <sup>†</sup>	VC_1703	Q9KRD9
RTX toxin transporter <sup>†</sup>	VC_1448	Q9KS14
Uncharacterized protein similar to VCA0109 <sup>†</sup>	VC_A0109	Q9KN56
C4-dicarboxylate transport protein DctQ, putative <sup>†</sup>	VC_1928	Q9KQS0
Trk system potassium uptake protein <sup>†</sup>	VC_0042	Q9KVV7
PTS system, cellobiose-specific IIC component <sup>†</sup>	VC_1282	Q9KSH4
Multidrug resistance protein VceB <sup>†</sup>	VC_1411	Q9KS49
Iron(III) compound receptor <sup>†</sup>	VC_0200	Q9KVE6
Sugar transporter family protein <sup>†</sup>	VC_A0669	Q9KLS2
Potassium uptake protein TrkA <sup>†</sup>	VC_0043	Q9KVV6
Multidrug resistance protein, putative <sup>†</sup>	VC_1409	Q9KS51
Sodium/solute symporter <sup>†</sup>	VC_A0667	Q9KLS4
C4-dicarboxylate-binding periplasmic protein <sup>†</sup>	VC_1779	Q9KR64
Multidrug resistance protein D <sup>†</sup>	VC_A0214	Q9KMV4
Multidrug resistance protein D <sup>†</sup>	VC_A0267	Q9KMQ3
PTS system, N-acetylglucosamine-specific IIABC component <sup>†</sup>	VC_0995	Q9KTA8
Lipopolysaccharide/O-antigen transport protein <sup>†</sup>	VC_0246	Q9KVA3
Iron(III) ABC transporter, permease protein <sup>†</sup>	VC_0203	Q9KVE3
Putative uncharacterized protein <sup>†</sup>	VC_A0716	Q9KLM7
Multidrug resistance protein VceA <sup>†</sup>	VC_1410	Q9KS50
Putative uncharacterized protein <sup>†</sup>	VC_A0355	Q9KMJ3
Helicase, putative <sup>§</sup>	VC_1760	Q9KR83
DNA-damage-inducible protein J <sup>§</sup>	VC_A0324	Q9KML3
N-acetylglucosamine-6-phosphate deacetylase <sup>§</sup>	VC_1783	Q9KR60
Sigma-54 modulation protein, putative <sup>//</sup>	VC_2530	H9L4N9

Included in this table are genetic elements that are absent in the non-O1 genome but present in the O1 genome, which have putative functions in virulence, disease and defense, membrane transport, DNA metabolism and dormancy and sporulation. In the table, the first column includes gene descriptions as per UniProt. Second and fourth columns represent abbreviated gene identification; the third and fifth columns represent accession numbers for the listed genes.

\*The symbol denotes proteins that have functions in virulence, disease and defense.

<sup>†</sup>The symbol includes proteins that are putative membrane transporters.

<sup>§</sup>Symbols and include proteins that have putative functions in DNA metabolism and dormancy/sporulation categories, respectively.

<sup>//</sup>Symbols and include proteins that have putative functions in DNA metabolism and dormancy/sporulation categories, respectively.

**Table 4**Genes present in non-O1 *Vibrio cholerae* genome but absent in O1 *V. cholerae*.

Description	Abbreviation PS15	Accession PS15
Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA <sup>†</sup>	<i>oppA</i>	L1QVD3
Protein-export membrane protein SecF <sup>†</sup>	<i>secF</i>	L1QTX8
UvrABC system protein A <sup>†</sup>	<i>uvrA</i>	L1QY95
CopG protein *	<i>OSU_0951</i>	L1R0A8
Cytochrome c heme lyase subunit CcmF *	<i>OSU_1000</i>	L1QZI8
Cytochrome c heme lyase subunit CcmH *	<i>OSU_1003</i>	L1R0Q0
Cytochrome c heme lyase subunit CcmL *	<i>OSU_1002</i>	L1QZP2
Multi antimicrobial extrusion protein (Na <sup>+</sup> /drug antiporter) VcrM *	<i>OSU_3002</i>	L1QUH8
Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM *	<i>OSU_0958</i>	L1R0S0
Multidrug efflux pump component MtrF *	<i>OSU_0277</i>	L1R1D8
Putative queD like protein *	<i>OSU_1874</i>	L1QXS7
Type IIA topoisomerase, B subunit *	<i>OSU_0552</i>	L1R2I2
Arsenical resistance operon repressor *	<i>OSU_0350</i>	L1R2G7
Arsenical-resistance protein ACR3 *	<i>OSU_0349</i>	L1R1D6
Copper-sensing two-component system response regulator CusR *	<i>OSU_3536</i>	L1QTI3
DNA-binding heavy metal response regulator *	<i>OSU_2602</i>	L1QVH5
Multidrug resistance transporter, Bcr/CflA family *	<i>OSU_2210</i>	L1QWA2
MFS family multidrug transport protein, bicyclomycin resistance protein *	<i>OSU_0873</i>	L1QZR8
Macrolide export ATP-binding/permease protein MacB *	<i>OSU_3185</i>	L1QT51
P pilus assembly/Cpx signaling pathway, periplasmic inhibitor/zinc-resistance protein *	<i>OSU_1241</i>	L1QZY1
Cobalt-zinc-cadmium resistance protein *	<i>OSU_1240</i>	L1QZ70
Cobalt-zinc-cadmium resistance protein *	<i>OSU_2129</i>	L1QWD3
Cobalt-zinc-cadmium resistance protein CzcD *	<i>OSU_1105</i>	L1QYT4
Cytolysin and hemolysin, HlyA, Pore-forming toxin *	<i>OSU_0766</i>	L1R001
Metalloprotease, containing putative zinc-binding domain *	<i>OSU_0770</i>	L1R1C4
Translation initiation factor SUII-related protein *	<i>OSU_0889</i>	L1R0U9
Transcription initiation factor TFIIIB, Brfl subunit/Transcription initiation factor TFIIIB *	<i>OSU_0614</i>	L1R0L1
ABC-type tungstate transport system, ATP-binding protein <sup>†</sup>	<i>OSU_0934</i>	L1R0Z5
ABC-type tungstate transport system, periplasmic binding protein <sup>†</sup>	<i>OSU_0932</i>	L1R0H9
Phosphonate ABC transporter phosphate-binding periplasmic component <sup>†</sup>	<i>OSU_2433</i>	L1QWK1

<b>Description</b>	<b>Abbreviation PS15</b>	<b>Accession PS15</b>
AttE component of AttEFGH ABC transport system <sup>†</sup>	OSU_0877	L1R138
AttF component of AttEFGH ABC transport system / AttG component of AttEFGH ABC transport system <sup>†</sup>	OSU_0878	L1QZS3
Peptide transport periplasmic protein sapA <sup>†</sup>	OSU_1956	L1QWN6
Magnesium and cobalt transport protein CorA <sup>†</sup>	OSU_0364	L1R1E9
Mg/Co/Ni transporter MgtE / CBS domain containing protein <sup>†</sup>	OSU_1331	L1QYJ2
MSHA biogenesis protein MshO <sup>†</sup>	OSU_1466	L1QY52
MSHA biogenesis protein MshP <sup>†</sup>	OSU_1467	L1QYW4
MSHA biogenesis protein MshQ <sup>†</sup>	OSU_1468	L1QYB3
Multimodular transpeptidase-transglycosylase <sup>†</sup>	OSU_1188	L1R077
Multimodular transpeptidase-transglycosylase <sup>†</sup>	OSU_0534	L1R0P0
Type IV fimbrial biogenesis protein FimT <sup>†</sup>	OSU_0787	L1QZX9
Type IV fimbrial biogenesis protein PilV <sup>†</sup>	OSU_0784	L1R0R8
Type IV fimbrial biogenesis protein PilW <sup>†</sup>	OSU_0786	L1R195
Type IV pilin PilA <sup>†</sup>	OSU_2009	L1QWH3
Type IV pilus biogenesis protein PilE <sup>†</sup>	OSU_0788	L1R035
Type IV pilus biogenesis protein PilM <sup>†</sup>	OSU_1187	L1QZ62
Type IV pilus biogenesis protein PilN <sup>†</sup>	OSU_1186	L1QZS4
Type IV pilus biogenesis protein PilO <sup>†</sup>	OSU_1185	L1QZ15
Type IV pilus biogenesis protein PilQ <sup>†</sup>	OSU_1183	L1R071
Conjugative signal peptidase TrhF <sup>†</sup>	OSU_2230	L1QWL7
Conjugative transfer protein s043 <sup>†</sup>	OSU_2245	L1QWN3
IncF plasmid conjugative transfer pilus assembly protein TraB <sup>†</sup>	OSU_2239	L1QVW4
IncF plasmid conjugative transfer pilus assembly protein TraC <sup>†</sup>	OSU_2232	L1QX28
IncF plasmid conjugative transfer pilus assembly protein TraE <sup>†</sup>	OSU_2241	L1QW42
IncF plasmid conjugative transfer pilus assembly protein TraF <sup>†</sup>	OSU_0332	L1R1B9
IncF plasmid conjugative transfer pilus assembly protein TraH <sup>†</sup>	OSU_0331	L1R173
IncF plasmid conjugative transfer pilus assembly protein TraK <sup>†</sup>	OSU_2240	L1QWM9
IncF plasmid conjugative transfer pilus assembly protein TraL <sup>†</sup>	OSU_2242	L1QX38
IncF plasmid conjugative transfer pilus assembly protein TraU <sup>†</sup>	OSU_2228	L1QVR6
IncF plasmid conjugative transfer pilus assembly protein TraW <sup>†</sup>	OSU_2229	L1QVV2
IncF plasmid conjugative transfer protein TraD <sup>†</sup>	OSU_2247	L1QX42
IncF plasmid conjugative transfer protein TraG <sup>†</sup>	OSU_0330	L1R2J2

Description	Abbreviation PS15	Accession PS15
IncF plasmid conjugative transfer protein TraN <sup>†</sup>	OSU_2227	L1QX23
Ync <sup>†</sup>	OSU_2235	L1QWM2
Ynd <sup>†</sup>	OSU_2236	L1QW36
Toxin co-regulated pilus biosynthesis protein E, anchors TcpT to membrane <sup>†</sup>	OSU_2156	L1QXB5
T1SS associated transglutaminase-like cysteine proteinase (LapP) <sup>†</sup>	OSU_2969	L1QUR2
Membrane-fusion protein <sup>†</sup>	OSU_2970	L1QU65
ABC-type bacteriocin/lantibiotic exporter, containing an N-terminal double-glycine peptidase domain <sup>†</sup>	OSU_2967	L1QTT3
Outer membrane protein ImpK/VasF, OmpA/MotB domain containing <sup>†</sup>	OSU_1572	L1QY45
Protein ImpG/VasA <sup>†</sup>	OSU_1567	L1QY39
Type VI secretion lipoprotein/VasD <sup>†</sup>	OSU_1570	L1QXY9
Type VI secretion protein VasI <sup>†</sup>	OSU_1575	L1QXZ3
Type VI secretion-related protein VasL <sup>†</sup>	OSU_1578	L1QZ51
Uncharacterized protein ImpB <sup>†</sup>	OSU_1564	L1QXT4
Uncharacterized protein ImpC <sup>†</sup>	OSU_1565	L1QXY4
Uncharacterized protein ImpH/VasB <sup>†</sup>	OSU_1568	L1QZ45
Uncharacterized protein ImpI/VasC <sup>†</sup>	OSU_1569	L1QXT8
Uncharacterized protein ImpJ/VasE <sup>†</sup>	OSU_1571	L1QYQ3
VgrG-3 protein <sup>†</sup>	OSU_1579	L1QXU5
TRAP transporter solute receptor, TAXI family <sup>†</sup>	OSU_1483	L1QYC6
TRAP transporter solute receptor, TAXI family <sup>†</sup>	OSU_1598	L1QZ65
TRAP-type uncharacterized transport system, fused permease component <sup>†</sup>	OSU_1482	L1QYX7
TRAP-type C4-dicarboxylate transport system, periplasmic component <sup>†</sup>	OSU_1751	L1QXA6
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit E <sup>†</sup>	OSU_1606	L1QYT1
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit F <sup>†</sup>	OSU_1605	L1QY16
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit G <sup>†</sup>	OSU_1604	L1QXW5
Di-/tripeptide transporter <sup>†</sup>	OSU_0457	L1R0U7
Di/tripeptide permease DtpA <sup>†</sup>	OSU_2702	L1QVD6
4-hydroxybenzoyl-CoA thioesterase family active site protein <sup>†</sup>	OSU_1280	L1QZ99
MotA/TolQ/ExbB proton channel family protein <sup>†</sup>	OSU_3208	L1QUH0
TPR repeat containing exported protein <sup>†</sup>	OSU_1274	L1QYK6
TonB system biopolymer transport component/ Chromosome segregation ATPase <sup>†</sup>	OSU_3207	L1QTE4

Description	Abbreviation PS15	Accession PS15
TonB-dependent heme and hemoglobin receptor HutA/ TonB-dependent hemin, ferrichrome receptor <sup>†</sup>	OSU_0883	L1QZU0
PTS system, N-acetylglucosamine-specific IIA, IIB, IIC component <sup>†</sup>	OSU_2709	L1QVU6
Ferrichrome-iron receptor <sup>†</sup>	OSU_1805	L1QX54
Enterobactin receptor VctA <sup>†</sup>	OSU_0353	L1R1Z3
Putative divalent cation transport protein <sup>†</sup>	OSU_1717	L1QXK6
Tricarboxylate transport protein TctC <sup>†</sup>	OSU_2307	L1QVT4
Membrane fusion component of tripartite multidrug resistance system <sup>†</sup>	OSU_0595	L1R1A1
AmpG permease <sup>†</sup>	OSU_2915	L1QVA5
High-affinity choline uptake protein BetT <sup>†</sup>	OSU_2908	L1QUU0
Uncharacterized protein <sup>†</sup>	OSU_3399	L1QSI2
Tricarboxylate transport membrane protein TctA EMBL EKY32019 <sup>†</sup>	OSU_2305	L1QVL0
Uncharacterized protein <sup>†</sup>	OSU_2298	L1QWR3
Ca <sup>2+</sup> /H <sup>+</sup> antiporter <sup>†</sup>	OSU_1800	L1QX51
Putative permease <sup>†</sup>	OSU_2291	L1QWA8
Transporter, LysE family <sup>†</sup>	OSU_0512	L1R248
Error-prone repair protein UmuD <sup>§</sup>	OSU_2250	L1QWN9
Error-prone, lesion bypass DNA polymerase V (UmuC) <sup>§</sup>	OSU_3541	L1QX48
Error-prone, lesion bypass DNA polymerase V (UmuC) <sup>§</sup>	OSU_2251	L1QW51
Type I restriction-modification system, DNA-methyltransferase subunit M <sup>§</sup>	OSU_2542	L1QV27
Type I restriction-modification system, specificity subunit S <sup>§</sup>	OSU_2543	L1QVX0

Included in this table are genes coding for virulence, disease and defense, membrane transport and DNA metabolism. The first column includes gene descriptions as per UniProt.

\* Second and third columns represent abbreviated identifications and accession numbers for the described genes, respectively, includes proteins that have putative functions in the virulence, disease or defense categories

<sup>†</sup> symbols represent proteins with functions in membrane transport and DNA metabolism categories, respectively.

<sup>§</sup> symbols represent proteins with functions in membrane transport and DNA metabolism categories, respectively.