

Prevalence and genetic diversity of clinical *Vibrio parahaemolyticus* isolates from China, revealed by multilocus sequence typing scheme

Dongsheng Han^{1†}, Hui Tang^{2†}, Chuanli Ren¹, Guangzhou Wang¹, Lin Zhou¹ and Chongxu Han^{1*}

¹ Department of Clinical Microbiology, Clinical Medical Examination Center, Northern Jiangsu People's Hospital, Yangzhou, China, ² Department of Biobank, Northern Jiangsu People's Hospital, Yangzhou, China

OPEN ACCESS

Edited by:

Julio Alvarez,
University of Minnesota, USA

Reviewed by:

Ronald Paul Rabinowitz,
University of Maryland School of
Medicine, USA

Li Xu,
Cornell University, USA

*Correspondence:

Chongxu Han,
Department of Clinical Microbiology,
Clinical Medical Examination Center,
Northern Jiangsu People's Hospital,
Nangtongxi Road, Yangzhou 225001,
China
huisheng113905@163.com

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Infectious Diseases, a section of the
journal *Frontiers in Microbiology*

Received: 06 February 2015

Accepted: 24 March 2015

Published: 09 April 2015

Citation:

Han D, Tang H, Ren C, Wang G, Zhou
L and Han C (2015) Prevalence and
genetic diversity of clinical *Vibrio*
parahaemolyticus isolates from China,
revealed by multilocus sequence
typing scheme.
Front. Microbiol. 6:291.
doi: 10.3389/fmicb.2015.00291

The population structure of clinical *Vibrio parahaemolyticus* isolates spreading in China remains undefined. We brought 218 clinical isolates from the pubMLST database originating from different regions of China collected since the year of 1990, analyzed by multilocus sequence typing (MLST), to elucidate the prevalence and genetic diversity of *V. parahaemolyticus* circulating in Chinese population. The MLST scheme produced 137 sequence types (STs). These STs were clustered into six clonal complexes (CCs), six doublets, and 91 singletons, exhibiting a high level of genetic diversity. However, less diversity was displayed on the peptide level: only 46 different peptide sequence type (pST) were generated, with pST2 (44.0%, 96/218) and pST1 (15.1%, 33/218) the predominant. Further analysis confirmed all the pSTs belong to a single complex founded by pST1, pST2, pST3, and pST4. *recA* presented the highest degree of nucleotide diversity (0.026) and the largest number of variable sites (176) on the nucleotide level. *pyrC* was the most diverse locus on the peptide level, possessing the highest percentage of variable sites (9.2%, 15/163). Significant linkage disequilibrium with the alleles was detected when the Standardized Index of Association (I_A^S) was calculated both for the entire isolates collection (0.7169, $P < 0.01$) and for the 137 STs ($I_A^S = 0.2648$, $P < 0.01$). In conclusion, we provide an overview of prevalence and genetic diversity of clinical *V. parahaemolyticus* spreading in Chinese population using MLST analysis. The results would offer genetic evidences for uncovering the microevolution relationship of *V. parahaemolyticus* populations.

Keywords: *Vibrio parahaemolyticus*, multilocus sequence typing, phylogenetic analysis, clonal complex, peptide sequence types

Introduction

Vibrio parahaemolyticus is a leading cause of food-borne outbreaks and acute gastroenteritis throughout the world, especially in coastal countries and regions. Consumption of raw or undercooked seafood is the major route of transmission for *V. parahaemolyticus* infection (Pal and Das, 2010). *V. parahaemolyticus* infection is caused by diverse serotypes; however, since the pandemic serovar O3:K6 emerged in Asia in 1996, it was confirmed as the predominant cause of outbreaks of

V. parahaemolyticus infection on a global scale (Okuda et al., 1997; Bag et al., 1999; Chowdhury et al., 2000; Nair et al., 2007). In recent years, at least 21 pandemic serotypes were identified as being associated with the outbreaks of *V. parahaemolyticus* infection (Nair et al., 2007).

V. parahaemolyticus is the first leading cause of food-borne outbreaks and bacterial infectious diarrhea in China, especially in the southeast coastal area (Lin et al., 2011). During 2000–2009, a multicentric surveillance for *V. parahaemolyticus* diarrhea at 11 provinces (Beijing, Jiangsu, Shanghai, Zhejiang, et al.) was conducted by China CDC, they collected stool or rectal swab specimens from 79,075 diarrhea patients to detect the prevalence ratio of *V. parahaemolyticus*, and found the average ratio of *V. parahaemolyticus* was 3.11% in these diarrhea patients. Studies also confirmed clinical isolates in China typically correspond to the pandemic serovar O3:K6 and several serovariants of it (e.g., O1:KUT, O1:K56, and O4:K68) (Chao et al., 2009; Yu et al., 2011; Fan et al., 2013; Shi et al., 2013; Li et al., 2014). So it's critical to clarify the prevalence and genetic diversity of this pathogen circulating in a particular population for minimizing both the risk of infection and economical burden. For the molecular genetic studies of *V. parahaemolyticus*, a number of molecular typing techniques have been developed and applied (Marshall et al., 1999; Gonzalez-Escalona et al., 2008; Chen et al., 2012). Multilocus sequence typing (MLST) of *V. parahaemolyticus* was developed by González-Escalona et al in 2008 (Gonzalez-Escalona et al., 2008). Gonzalez-Escalona's own and a number of subsequent studies demonstrated MLST was a powerful tool with a high resolution rate in identification of clonal complexes (CCs) of *V. parahaemolyticus* population and in understanding the processes leading to the emergence and spread of pathogenic isolates (Harth et al., 2009; Yu et al., 2011; Ellis et al., 2012; Banerjee et al., 2014).

Although several studies already used MLST analysis to study the genetic diversity of *V. parahaemolyticus* in China in recent years, they were restricted to specific regions (Chao et al., 2009; Han et al., 2012; Fan et al., 2013; Shi et al., 2013), focused exclusively on pandemic pathogenic isolates (Chao et al., 2011; Yan et al., 2011) or were based on a limited isolate number (Yu et al., 2011). In the present study, we brought 218 Chinese clinical isolates from the pubMLST database (<http://pubmlst.org/vparahaemolyticus/>) into our analyses. These clinical isolates were collected mostly from the provinces of Beijing, Jiangsu, Zhejiang, Fujian, Guangdong, and Taiwan. With these isolates, we aimed to elucidate the prevalence and genetic diversity of clinical *V. parahaemolyticus* isolates circulating in Chinese population. We would investigate the sequence/peptide polymorphisms of the isolates and analyze the probable evolutionary relationships among the isolates. The differences in regard to CC and sequence type (ST)/peptide sequence type (pST) affiliation in the analyzed isolates were considered. We provide a broader overview of the genetic population structure of clinical *V. parahaemolyticus* from China, and predict that the results will provide genetic evidences for uncovering the microevolution relationship among different isolates and might be conducive to the early warning and prevention of *V. parahaemolyticus* infection.

Materials and Methods

Sampling of *V. parahaemolyticus* Isolates

A total of 218 clinical *V. parahaemolyticus* isolates from Chinese patients with acute gastroenteritis were selected as the research subject of this study, they were all available in the pubMLST database (<http://pubmlst.org/vparahaemolyticus/>). These isolates were both temporally (collected from 1990 to November 2014) and geographically (collected from the provinces of Beijing, Jiangsu, Zhejiang, Fujian, Guangdong, and Taiwan) diverse (see Table S1 in the Supplemental Material).

MLST Analysis and Phylogenetic Analysis

There is a PCR protocol of internal fragments of the seven house-keeping genes [*recA*(729bp), *dnaE*(557bp), *gyrB*(592bp), *dtdS*(458bp), *pntA*(430bp), *pyrC*(493bp), and *tnaA*(423bp)] on *V. parahaemolyticus* pubMLST website (<http://pubmlst.org/vparahaemolyticus/>), the allele designations and STs of all the 218 isolates had been determined. Based on the related STs, all the isolates were subdivided into CCs or groups by goeBURST analysis using Phyloviz software (<http://www.phyloviz.net>). We also implemented "population snapshot" analysis on the basis of STs and pSTs by using goeBURST. Isolates that shared 100% identity in six of the seven loci with at least one other member of the group, the single locus variants (SLVs), were assigned to a single CC. The primary founder of a CC, SLVs, double locus variants (DLVs), and singletons were defined as in previous study (Han et al., 2014).

When a nucleotide sequence was translated in frame, a peptide sequence could be obtained, in other words, each nucleotide sequence correspond to a unique peptide sequence, an individual isolate contains a unique ST as well as a pST. So translating the in-frame nucleotide sequences into peptide sequences allows a phylogenetic analysis based on pSTs. In this study, the assignment of pSTs of the analyzed isolates to CCs was carried out as previously (Theethakaew et al., 2013; Urnersbach et al., 2014) and predicted also by goeBURST algorithm.

Minimum-evolution (ME) trees for the in-frame concatenated sequences (*recA-dnaE-gyrB-dtdS-pntA-pyrC-tnaA*) of each (p)ST were constructed by Mega 5 software, genetic distance of the analyzed isolates was estimated by the Kimura two-parameter model, as did in the other study (Han et al., 2014).

Population Genetic Analysis

DnaSP V5 was used to calculate the following parameters: the number of alleles, the number of polymorphic sites and nucleotide diversity (π), for evaluating the varying degrees of the loci in our selected isolates (Librado and Rozas, 2009). START V2 was implemented to calculate the ratio of non-synonymous-to-synonymous substitutions (dN/dS) through the Nei and Gojori method (Jolley et al., 2001). dN/dS < 1 indicates that the relative gene was mainly affected by purifying selection during the population evolution, dN/dS = 1 indicates neutral selection and dN/dS > 1 indicates positive selection. The value of Standardized Index of Association (I_A^S) was calculated by START2, in order to access the population structure of *V. parahaemolyticus*. $I_A^S = 0$ indicates alleles are in linkage equilibrium

(alleles are independently distributed at all loci analyzed) and recombination occurred frequently (Gonzalez-Escalona et al., 2008).

Results

Diversity of Sequence Types (STs)

The data on diversities of the seven loci in the 218 *V. parahaemolyticus* isolates are showed in **Table 1**. All these analyzed isolates resulted in 137 unique STs by applying MLST analysis. Among these STs, Individual STs were mostly recovered once (118 STs), ST3 was most frequent (52 isolates), 18 STs was constituted of between two and five isolates (see Table S1 in the Supplemental Material). When the geographical subsets were considered, the number of different STs was high in Guangdong, Jiangsu, Taiwan and Shanghai (**Table 2**).

The number of alleles observed for each locus ranged from 51 (*pntA* and *tnaA*) to 79 (*gyrB*). *gyrB* possessed the most number of alleles (79 alleles) but only 62 (10.47%) variable sites. The nucleotide diversity ranged from 0.011(*pntA*) to 0.026(*recA*). The dN/dS ratio for every locus was close to zero (**Table 1**), this suggests that the housekeeping genes are mainly affected by purifying selection during the evolutionary process.

Diversity of Peptide Sequence Types (pSTs)

A total of 46 different pSTs were obtained from the analyzed isolates (see Table S1 in the supplemental material), occurred with a frequency of 0.5 to 44.0%. pST2 (44.0%, 96/218) and pST1 (15.1%, 33/218) were the two predominant pSTs. One particular pST could be comprised of numerous STs, in this study, we found that pST1 were translated by the nucleotide sequences of 27 different STs (ST62, ST91, ST120, etc.), pST2 by 38 STs (ST3, ST189, ST345, etc.), pST3 by 7 STs (ST328, ST332, ST444, etc.), pST4 by 9 STs (ST8, ST224, ST301, etc.), and the other pSTs by four or less STs.

On pST level, the proportion of allele ONE was more than 90.0% for most of the seven loci, except for *dnaE* (allele ONE accounting for 35.8%) and *pyrC* (allele ONE accounting for 73.4%) (see Table S1 in the Supplemental Material). The individual loci possessed 2 (*gyrB*) to 16 (*pyrC*) unique alleles. *pyrC* possessed the highest percentage of variable sites (9.2%, 15/163) (**Table 1**).

Clonal Relationships of the Collected Isolates

In this study, the calculated I_A^S value was 0.7169 ($P < 0.01$) for all of the 218 analyzed isolates, that is to say, the alleles in the seven housekeeping genes are in linkage disequilibrium. When we calculated the I_A^S value repeatedly, using one isolate to

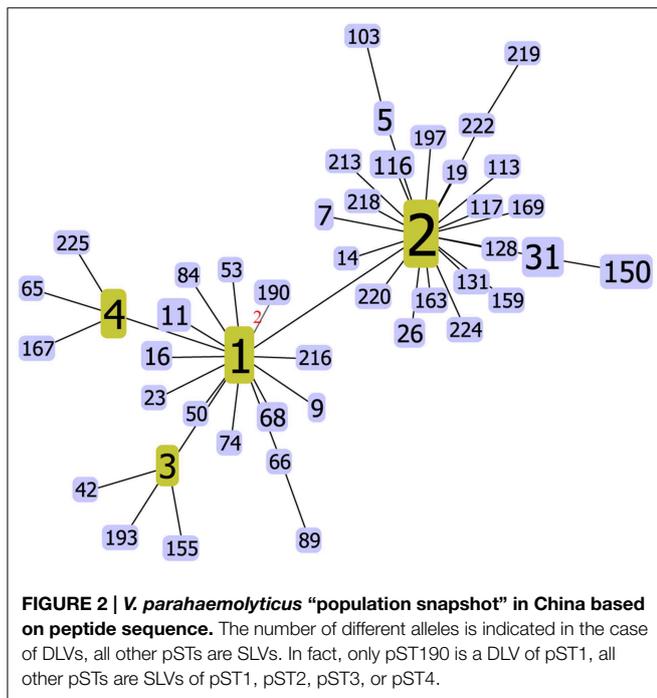
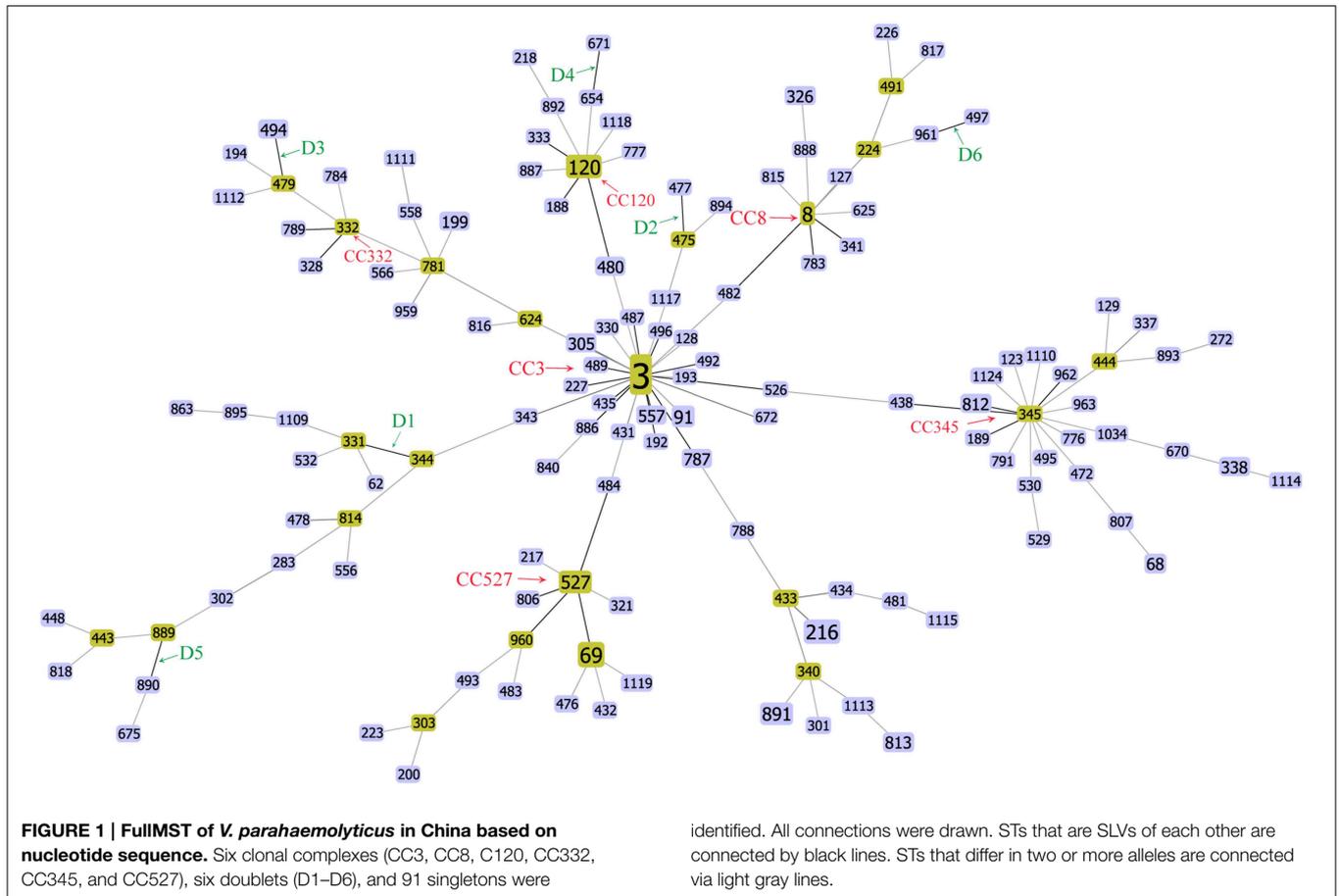
TABLE 1 | Diversities of the seven loci in the 218 *V. parahaemolyticus* isolates.

Locus	Fragment size		No of alleles		No (%) of variable sites		nucleotide diversity (π)		dN/dS	
	ST	pST	ST	pST	ST	pST	ST	ST	ST	ST
I										
<i>dnaE</i>	557bp	185aa	67	6	63 (11.31)	5 (2.70)	0.012		0.028	
<i>gyrB</i>	592bp	197aa	79	2	62 (10.47)	–(–)*	0.015		0.001	
<i>recA</i>	729bp	242aa	75	4	176 (24.14)	19 (7.85)	0.026		0.017	
II										
<i>dtbS</i>	458bp	152aa	68	3	47 (10.26)	2 (1.32)	0.023		0.002	
<i>pntA</i>	430bp	143aa	51	12	40 (9.30)	10 (6.99)	0.011		0.031	
<i>pyrC</i>	493bp	163aa	62	16	52 (10.55)	15 (9.20)	0.013		0.066	
<i>tnaA</i>	423bp	141aa	51	7	42 (9.93)	6 (4.26)	0.013		0.013	

*For the occurring alleles no variable sites could be determined.

TABLE 2 | Geographic distribution of the 218 *V. parahaemolyticus* isolates.

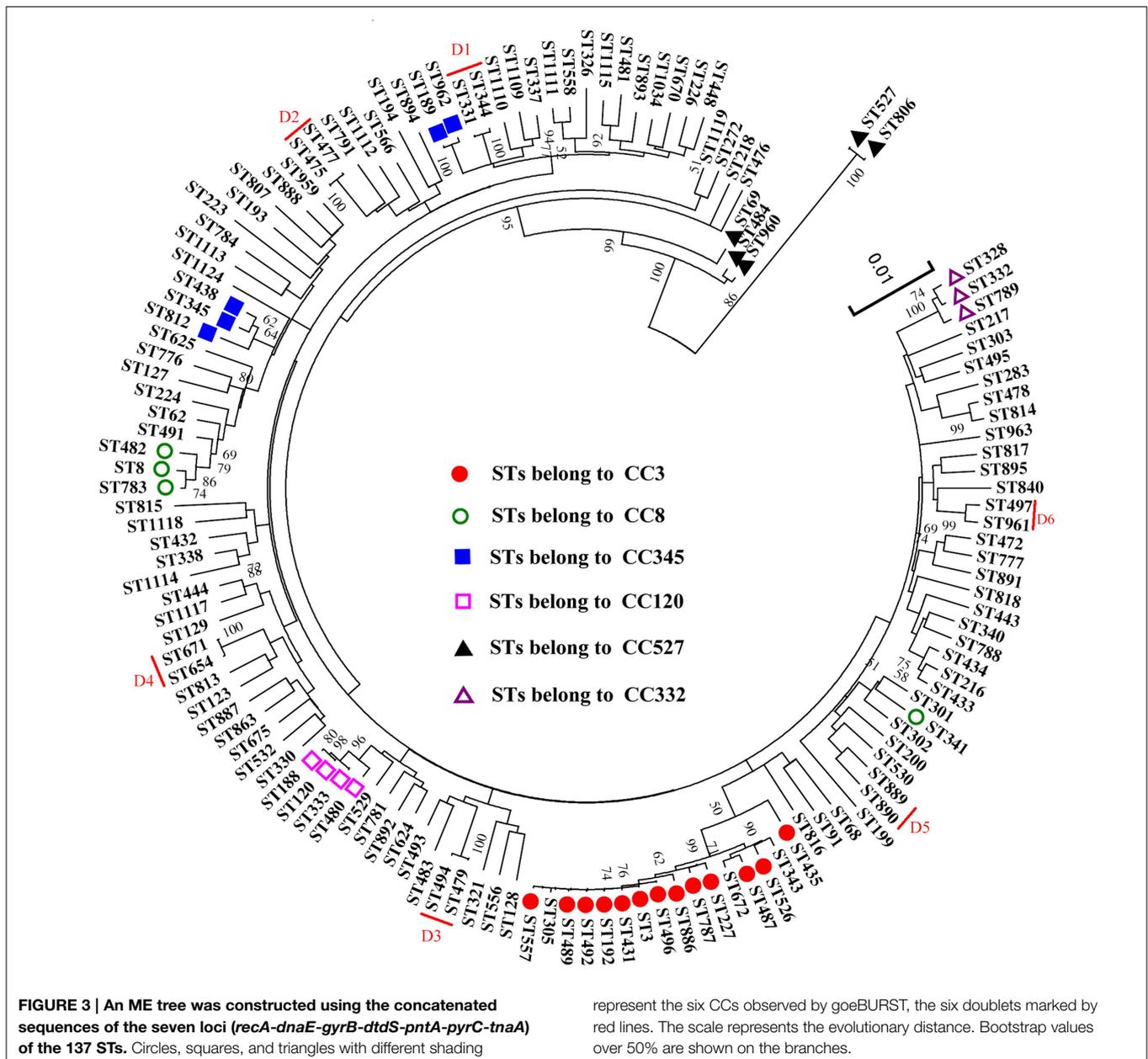
Geographical subsets	No. of isolates	No. of STs	Mainly STs	No. of pSTs	Mainly pSTs
Guangdong	51	33	ST3, ST305, ST480, etc.	16	pST1,pST2,pST11, pST116, etc.
Jiangsu	37	22	ST3,ST8,ST120, etc.	14	pST1,pST2,pST4, pST26, etc.
Taiwan	30	11	ST3,ST8, ST120, etc.	7	pST1,pST2,pST4, pST31, etc.
Shanghai	17	17	ST283,ST301,ST302, etc.	6	pST1,pST2,pST3, pST4, etc.
zhejiang	5	5	ST654,ST670,ST671, etc.	3	pST1,pST2,pST16
Beijing	4	4	ST431,ST432,ST478,ST479	3	pST1,pST2,pST163
Hong Kong	3	2	ST3,ST91	2	pST1,pST2
Fujian	2	2	ST963,ST1034	2	pST2, pST197
Unknown	69	53	ST3,ST69,ST216, etc.	21	pST1,pST2,pST3, pST4, etc.
ALL isolates	218	137	ST3,ST8, ST120, etc.	46	pST1,pST2,pST3, pST4, etc.



represent each of the 137 STs, it's 0.2648 ($P < 0.01$). Although the value represented a markedly decrease from 0.7169, these alleles are in linkage disequilibrium. This indicates a nonrandom distribution of alleles in the *V. parahaemolyticus* population in general.

Identification of Clonal Complexes

The goeBURST algorithm used in our study resolved the 137 STs into six CCs (CC3, CC8, C120, CC332, CC345, and CC527) and six doublets (D1–D6). The remaining 91 STs were singletons (Figure 1). CC3 was the most prevalent CC, including 66 isolates with 13 STs. However, neither the relationships between the 91 singletons themselves nor the relationships between the singletons and the defined CCs or doublets could be deduced here. It suggested that goeBURST had limited utility on nucleotide level for identifying related isolates. To counter this, we implemented a "population snapshot" analysis by using goeBURST on the basis of pSTs. The result showed that only pST190 differed in more than one allele to all the other 45 pSTs, leading to a single complex founded by pST1, pST2, pST3, and pST4 (Figure 2). So the relationship among the isolates appears more closely when analyzed on peptide level than on nucleotide level.

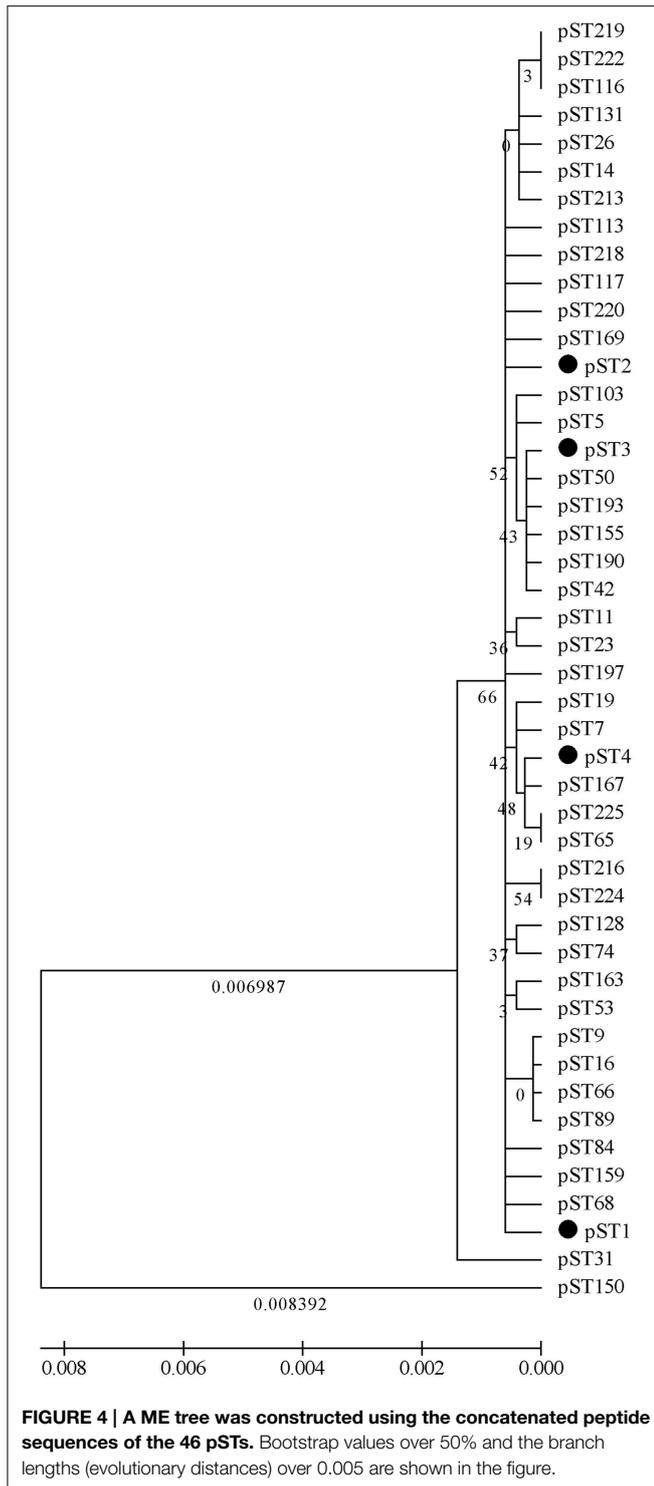


Phylogenetic Analysis

The result of a phylogenetic analysis using ME tree on nucleotide level revealed a high genetic diversity among the 218 analyzed isolates (Figure 3). The isolates belong to the same CCs and Doublets (D1–D6) in the goeBURST analysis were also clustered together in the ME tree, except for the isolates of CC345 and CC527. CC345 was divided into two different clusters, ST438 and ST812 exhibited a relatively closer evolutionary distance to ST345 (the ancestral type of CC345) than ST962 and ST189. When compared the number of single nucleotide polymorphisms (SNPs), we found ST962 and ST189 showed differences of 25 SNPs (found in *gyrB* allele), and 24 SNPs (found in *recA* allele) from ST345 respectively, whereas ST438 and ST812 only differs from ST345

by 9 SNPs (found in *recA* allele) and 15 SNPs (found in *recA* allele), respectively. Similarly, the difference between SNP numbers had an important impact on the classification of CC527, ST806 (only one SNP different from ST527) showed a closer evolutionary relationships to ST527 than ST69 (139 SNPs differs), ST484 (136 SNPs differs), and ST960 (135 SNPs differs). Thus, a phylogenetic analysis based on the nucleotide sequences provides a better resolution and elucidated some genetic relationships among isolates that were not resolved by goeBURST analysis.

An ME tree analysis was also conducted based on the concatenated sequences of peptide sequences among the 46 pSTs (Figure 4). The low bootstrap values (all < 70%) indicated that



the topology of the ME tree was poorly supported. However, the longest branch length (evolutionary distance) was only 0.008392 (found between ST150 and the cluster of other pSTs); the genetic similarity of all the clusters of the pSTs was higher than 99.99%. Thus, the 46 analyzed pSTs belong to a single CC when a phylogenetic analysis carried out on peptide level.

Discussion

In the present study, we analyzed the prevalence and extent of genetic diversity of *V. parahaemolyticus* among 218 clinical isolates collected in different regions of China. The diversity of the *V. parahaemolyticus* isolates was analyzed on nucleotide as well as peptide level. It is clear that multiple sequence/peptide types are contributed to human infection in China, and the genetic relationship among the isolates appears more closely on peptide level than on nucleotide level.

The observed alleles, variable sites, dN/dS-value, nucleotide diversity (π), I_A^S -value and (p)STs of our isolates were similar to those derived from a collection of global isolates in the pubMLST database, which were presented in a study conducted by Urmersbach et al. (2014). This reveals a high diversity of clinical isolates collection in Chinese population.

The low dN/dS ratios (all close to zero) obtained for all the seven genes indicating purifying selection for all the loci, as shown by others (Turner et al., 2013). It means synonymous substitutions were dominant in the nucleotide sequences, which do not alter amino acid sequences. This finding could explain the fact that the numbers of different alleles per locus were reduced and the loci were dominated by a single allele when analyzed on peptide level.

V. parahaemolyticus population was extremely genetic diverse in China. A total of 137 STs were identified in this study. ST3 was the most common ST, in agreement with the finding of a previous study on the basis of a global clinical isolates collection (Han et al., 2014). As we know, ST3 has been a sequence type of *V. parahaemolyticus* with an international distribution. Numerous reports revealed that it was widely distributed and played an important role in *V. parahaemolyticus* infection in China (Yu et al., 2011; Han et al., 2012; Fan et al., 2013; Shi et al., 2013). However, when carried out an analysis based on the peptide sequences, the diversity decreased. Only 46 pSTs were generated, and pST1, pST2, pST3, and pST4 were the most prevalent pSTs, which also the predominant pSTs in the pubMLST database (Urmersbach et al., 2014).

Eight CCs and 11 doublets have been identified when we analyzed the clinical isolates with a global collection in the pubMLST dataset (Han et al., 2014). In this study, six CCs and six doublets were found in the “population snapshot” of the 137 STs. CC3, a global pandemic clone of *V. parahaemolyticus* (Martinez-Urtaza et al., 2004, 2005; Ansaruzzaman et al., 2005), was also the most prevalent CC in China, being comprised of 69 isolates with 13 different STs in this study, posing a significant public health threat. However, the “population snapshot” of pSTs consists of only one unique CC. with pST1, pST2, pST3, and pST4 being the ancestral types at the same time. Other pSTs might have arisen from the four types by genetic drift associated with genetic changes (Osorio et al., 2012).

As discussed above, when we analyzed the genetic relationships among different isolates based on STs, goeBURST algorithm showed a decreased ability in identifying the related genotypic relationships due to the high degree of allelic diversity. Relationships are reliable only for identical or closely related isolates. When isolates are more distantly related, such as the 91

singletons, little information can be gained about their relationships. However, when the goeBURST analysis was implemented on the basis of pSTs, all the isolates were classified into one unique CC. This result maybe more representative of the real relationships among the isolates on the phenotypic level. Using pSTs instead of STs might be more efficient in reaching a reliable identification of related isolates.

Previous study confirmed that linkage disequilibrium could be observed when a recent, more epidemic clone arises (Smith et al., 1993). The calculated I_A^S value was 0.7169 ($P < 0.01$) for all of the analyzed isolates in this study, suggesting that all alleles in the seven housekeeping genes were in linkage disequilibrium or were non-randomly distributed. However, even the analysis was repeated using one isolate to represent each of the 137 STs, which would weaken the influence of the potential pandemic isolates in the data set, these alleles are still in linkage disequilibrium ($I_A^S = 0.2648$, $P < 0.01$). These results indicate that a non-randomly distribution of alleles in the *V. parahaemolyticus* population in general, even though recombination might be occurring in different subtypes (Gonzalez-Escalona et al., 2008). These observations are also typical for epidemic populations (Ellis et al., 2012; Theethakaew et al., 2013; Turner et al., 2013), thus to some extent, our data support the hypothesis that the population structure of *V. parahaemolyticus* follows the epidemic model of clonal expansion (Yu et al., 2011).

The results of goeBURST and ME tree shared a high similarity in the identifying of CCs. In the MLST scheme, the isolates of four CCs (CC3, CC8, C120, and CC332) and the six doublets were also clustered together in the ME tree, indicating that they were genetically exclusive complexes or groups. However, the isolates of CC345 and CC527 were resolved into different clusters. It could be explained by the different approaches used in ME tree and goeBURST algorithm. The ME tree is sequence-based, all sequences with fewer differences could be clustered together. goeBURST algorithm is allelic profile-based, only the SLVs were assigned to a single CC or groups. So the ME tree seems to be more suitable for analyzing genetic relationships of *V. parahaemolyticus* populations (Yan et al., 2011).

An ME tree was also constructed for phylogenetic analysis of *V. parahaemolyticus* on peptide level. All the 46 analyzed pSTs are grouped into a single CC, with the genetic similarity of all the clusters of the pSTs higher than 99.99%. A similar result has been observed by Osorio et al. who aimed to deduce putative ancestral relationships between different *Brachyspira hyodysenteriae* isolates (Osorio et al., 2012).

In this study, isolates without a clear source or STs were screened out. This may has some influence on the genetic characteristics of *V. parahaemolyticus* in China in general.

References

Ansaruzzaman, M., Lucas, M., Deen, J. L., Bhuiyan, N. A., Wang, X. Y., Safa, A., et al. (2005). Pandemic serovars (O3:K6 and O4:K68) of *Vibrio parahaemolyticus* associated with diarrhea in Mozambique: spread of the pandemic into the African continent. *J. Clin. Microbiol.* 43, 2559–2562. doi: 10.1128/JCM.43.6.2559-2562.2005

However, our findings would facilitate the researchers in this field to understand the population structure of *V. parahaemolyticus* in China. As the *V. parahaemolyticus* PubMLST database is not mandatory for uploading laboratory data, not all the research data of *V. parahaemolyticus* in the world (including China) have been completely uploaded into this public database. Only when a new ST is discovered, the researcher will sent the isolate information to the manager of the database to get identification for a new subtype. Therefore, the pubMLST database contains all the STs of *V. parahaemolyticus* all around the world, but does not contain all the discovered isolates. Here, we recommend that the database should be made some mandatory improvement, for example, when a new subtype is uploaded, the corresponding biological characteristics of isolates (sample source, resistance, serotype, virulence gene, etc.) should be also updated. We believe this will enable the pubMLST database to play a more important role in studies on infection and molecular epidemiology of *V. parahaemolyticus*.

In summary, we provide an overview of prevalence and genetic diversity of clinical *V. parahaemolyticus* spreading in Chinese population using MLST analysis. We implemented the identification of CCs and phylogenetic analysis both on nucleotide level and on peptide level. The results in this study will provide genetic evidences for uncovering the microevolution relationship among different pathogenic isolates of *V. parahaemolyticus*. The pubMLST database provides a platform for the comprehensive analysis of genetic relationships of *V. parahaemolyticus*. With the growing number of the uploaded isolates, more molecular biological characteristics of *V. parahaemolyticus* in china and other counties would be revealed.

Acknowledgments

The authors thank Dr. Narjol Gonzalez-Escalona, who is from FDA, USA, for development of the *V. parahaemolyticus* MLST scheme. This study was jointly supported by grant NO.81400899 from the National Natural Science Foundation of China, grant NO.YZ2014061 from the Science and Technology Foundation Program of Yangzhou city, and a research project (NO.yzucms201425) of Northern Jiangsu People's Hospital, Yangzhou, China.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2015.00291/abstract>

Bag, P. K., Nandi, S., Bhadra, R. K., Ramamurthy, T., Bhattacharya, S. K., Nishibuchi, M., et al. (1999). Clonal diversity among recently emerged isolates of *Vibrio parahaemolyticus* O3:K6 associated with pandemic spread. *J. Clin. Microbiol.* 37, 2354–2357.

Banerjee, S. K., Kearney, A. K., Nadon, C. A., Peterson, C. L., Tyler, K., Bakouche, L., et al. (2014). Phenotypic and genotypic characterization of Canadian clinical isolates of *Vibrio parahaemolyticus* collected from

- 2000 to 2009. *J. Clin. Microbiol.* 52, 1081–1088. doi: 10.1128/JCM.03047-13
- Chao, G., Jiao, X., Zhou, X., Yang, Z., Huang, J., Pan, Z., et al. (2009). Serodiversity, pandemic O3:K6 clone, molecular typing, and antibiotic susceptibility of foodborne and clinical *Vibrio parahaemolyticus* isolates in Jiangsu, China. *Foodborne Pathog. Dis.* 6, 1021–1028. doi: 10.1089/fpd.2009.0295
- Chao, G., Wang, F., Zhou, X., Jiao, X., Huang, J., Pan, Z., et al. (2011). Origin of *Vibrio parahaemolyticus* O3:K6 pandemic clone. *Int. J. Food Microbiol.* 145, 459–463. doi: 10.1016/j.ijfoodmicro.2011.01.022
- Chen, W., Xie, Y., Xu, J., Wang, Q., Gu, M., Yang, J., et al. (2012). Molecular typing of *Vibrio parahaemolyticus* isolates from the middle-east coastline of China. *Int. J. Food Microbiol.* 153, 402–412. doi: 10.1016/j.ijfoodmicro.2011.12.001
- Chowdhury, N. R., Chakraborty, S., Ramamurthy, T., Nishibuchi, M., Yamasaki, S., Takeda, Y., et al. (2000). Molecular evidence of clonal *Vibrio parahaemolyticus* pandemic isolates. *Emerg. Infect. Dis.* 6, 631–636. doi: 10.3201/eid0606.000612
- Ellis, C. N., Schuster, B. M., Striplin, M. J., Jones, S. H., Whistler, C. A., and Cooper, V. S. (2012). Influence of seasonality on the genetic diversity of *Vibrio parahaemolyticus* in New Hampshire shellfish waters as determined by multilocus sequence analysis. *Appl. Environ. Microbiol.* 78, 3778–3782. doi: 10.1128/AEM.07794-11
- Fan, Y. Y., Zhu, M., Shang, X. R., Wang, M., Huang, Y. F., Gu, H. T., et al. (2013). Virulence characteristics and multilocus sequence type of *Vibrio parahaemolyticus* isolated from clinic (in Chinese). *Chin. J. Epidemiol.* 36, 548–552. doi: 10.1371/journal.pone.0107371
- Gonzalez-Escalona, N., Martinez-Urtaza, J., Romero, J., Espejo, R. T., Jaykus, L. A., and DePaola, A. (2008). Determination of molecular phylogenetics of *Vibrio parahaemolyticus* isolates by multilocus sequence typing. *J. Bacteriol.* 190, 2831–2840. doi: 10.1128/JB.01808-07
- Han, D., Tang, H., Lu, J., Wang, G., Zhou, L., Min, L., et al. (2014). Population structure of clinical *Vibrio parahaemolyticus* from 17 coastal countries, determined through multilocus sequence analysis. *PLoS ONE* 9:e107371. doi: 10.1371/journal.pone.0107371
- Han, D. S., Chen, X., Li, Z. J., Zheng, S. F., Yu, F., and Chen, Y. (2012). Prevalence and multi-locus sequence typing of *Vibrio parahaemolyticus* associated with acute diarrhea in Zhejiang (in Chinese). *J. Environ. Health* 29, 917–919.
- Harth, E., Matsuda, L., Hernandez, C., Rioseco, M. L., Romero, J., Gonzalez-Escalona, N., et al. (2009). Epidemiology of *Vibrio parahaemolyticus* outbreaks, southern Chile. *Emerg. Infect. Dis.* 15, 163–168. doi: 10.3201/eid1502.071269
- Jolley, K. A., Feil, E. J., Chan, M. S., and Maiden, M. C. (2001). Sequence type analysis and recombinational tests (START). *Bioinformatics* 17, 1230–1231. doi: 10.1093/bioinformatics/17.12.1230
- Li, Y., Xie, X., Shi, X., Lin, Y., Qiu, Y., Mou, J., et al. (2014). *Vibrio parahaemolyticus*, Southern Coastal Region of China, 2007–2012. *Emerg. Infect. Dis.* 20, 685–688. doi: 10.3201/eid2004.130744
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Lin, X. H., Ran, L., Ma, L., Wang, Z. J., and Feng, Z. J. (2011). Analysis on the cases of infectious diarrhea (other than cholera, dysentery, typhoid, and paratyphoid) reported in China in 2010 (in Chinese). *Chin. J. Food Hyg.* 23, 385–389.
- Marshall, S., Clark, C. G., Wang, G., Mulvey, M., Kelly, M. T., and Johnson, W. M. (1999). Comparison of molecular methods for typing *Vibrio parahaemolyticus*. *J. Clin. Microbiol.* 37, 2473–2478.
- Martinez-Urtaza, J., Lozano-Leon, A., DePaola, A., Ishibashi, M., Shimada, K., Nishibuchi, M., et al. (2004). Characterization of pathogenic *Vibrio parahaemolyticus* isolates from clinical sources in Spain and comparison with Asian and North American pandemic isolates. *J. Clin. Microbiol.* 42, 4672–4678. doi: 10.1128/JCM.42.10.4672-4678.2004
- Martinez-Urtaza, J., Simental, L., Velasco, D., DePaola, A., Ishibashi, M., Nakaguchi, Y., et al. (2005). Pandemic *Vibrio parahaemolyticus* O3:K6, Europe. *Emerg. Infect. Dis.* 11, 1319–1320. doi: 10.3201/eid1108.050322
- Nair, G. B., Ramamurthy, T., Bhattacharya, S. K., Dutta, B., Takeda, Y., and Sack, D. A. (2007). Global dissemination of *Vibrio parahaemolyticus* serotype O3:K6 and its serovariants. *Clin. Microbiol. Rev.* 20, 39–48. doi: 10.1128/CMR.00025-06
- Okuda, J., Ishibashi, M., Hayakawa, E., Nishino, T., Takeda, Y., Mukhopadhyay, A. K., et al. (1997). Emergence of a unique O3:K6 clone of *Vibrio parahaemolyticus* in Calcutta, India, and isolation of isolates from the same clonal group from Southeast Asian travelers arriving in Japan. *J. Clin. Microbiol.* 35, 3150–3155.
- Osorio, J., Carvajal, A., Naharro, G., La, T., Phillips, N. D., Rubio, P., et al. (2012). Dissemination of clonal groups of *Brachyspira hyodysenteriae* amongst pig farms in Spain, and their relationships to isolates from other countries. *PLoS ONE* 7:e39082. doi: 10.1371/journal.pone.0039082
- Pal, D., and Das, N. (2010). Isolation, identification and molecular characterization of *Vibrio parahaemolyticus* from fish samples in Kolkata. *Eur. Rev. Med. Pharmacol. Sci.* 14, 545–549.
- Shi, X. L., Wang, Y., Hu, Q. H., Li, Y. H., Lin, Y. M., Qiu, Y. Q., et al. (2013). Molecular characterization of *Vibrio parahaemolyticus* collected from human infections in Shenzhen, between 2002 and 2008 (in Chinese). *Chin. J. Epidemiol.* 34, 609–613.
- Smith, J. M., Smith, N. H., O'Rourke, M., and Spratt, B. G. (1993). How clonal are bacteria? *Proc. Natl. Acad. Sci. U.S.A.* 90, 4384–4388.
- Theethakaew, C., Feil, E. J., Castillo-Ramirez, S., Aanensen, D. M., Suthienkul, O., Neil, D. M., et al. (2013). Genetic relationships of *Vibrio parahaemolyticus* isolates from clinical, human carrier, and environmental sources in Thailand, determined by multilocus sequence analysis. *Appl. Environ. Microbiol.* 79, 2358–2370. doi: 10.1128/AEM.03067-12
- Turner, J. W., Paranjpye, R. N., Landis, E. D., Biryukov, S. V., Gonzalez-Escalona, N., Nilsson, W. B., et al. (2013). Population structure of clinical and environmental *Vibrio parahaemolyticus* from the Pacific northwest coast of the United States. *PLoS ONE* 8:e55726. doi: 10.1371/journal.pone.0055726
- Urmersbach, S., Alter, T., Koralage, M. S., Sperling, L., Gerdts, G., Messelhauser, U., et al. (2014). Population analysis of *Vibrio parahaemolyticus* originating from different geographical regions demonstrates a high genetic diversity. *BMC Microbiol.* 14:59. doi: 10.1186/1471-2180-14-59
- Yan, Y., Cui, Y., Han, H., Xiao, X., Wong, H. C., Tan, Y., et al. (2011). Extended MLST-based population genetics and phylogeny of *Vibrio parahaemolyticus* with high levels of recombination. *Int. J. Food Microbiol.* 145, 106–112. doi: 10.1016/j.ijfoodmicro.2010.11.038
- Yu, Y., Hu, W., Wu, B., Zhang, P., Chen, J., Wang, S., et al. (2011). *Vibrio parahaemolyticus* isolates from southeastern Chinese coast are genetically diverse with circulation of clonal complex 3 isolates since 2002. *Foodborne Pathog. Dis.* 8, 1169–1176. doi: 10.1089/fpd.2011.0865

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Han, Tang, Ren, Wang, Zhou and Han. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.