JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 2011, p. 1495–1500 Vol. 49, No. 4
0095-1137/11/$12.00 doi:10.1128/JCM.02640-10

Multivariate Analyses Revealed Distinctive Features Differentiating Human and Cattle Isolates of Shiga Toxin-Producing *Escherichia coli* O157 in Japan[▽]

Ken-ichi Lee,[1] Nigel P. French,[2] Yukiko Hara-Kudo,[1,3] Sunao Iyoda,[4] Hideki Kobayashi,[5] Yoshiko Sugita-Konishi,[1,3] Hirokazu Tsubone,[1] and Susumu Kumagai[1]*

Graduate School of Agricultural and Life Sciences, the University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan[1]; Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand[2]; Division of Microbiology, National Institute of Health Science, 1-18-1 Kami-Yoga, Setagaya-ku, Tokyo 158-8501, Japan[3]; Department of Bacteriology, National Institute of Infectious Diseases, Toyama 1-23-1 Shinjuku-ku, Tokyo 162-8640, Japan[4]; and National Agriculture and Food Research Organization, National Institute of Animal Health, 3-1-5 Kannondai, Tsukuba, Ibaraki 305-0856, Japan[5]

Received 30 December 2010/Returned for modification 26 January 2011/Accepted 14 February 2011

Genotypes of Shiga toxin-producing *Escherichia coli* (STEC) O157 isolated from humans and cattle were analyzed by uni- and multivariable logistic regression, and population structure methods, to gain insight into transmission and the nature of human infection. Eleven genotyping assays, including PCR typing of five virulence factors (*stx*$_1$, *stx*$_2$, *stx*$_{2c}$, *eae*, and *ehxA*) and a lineage-specific polymorphism assay using six markers (LSPA6), were considered in the analyses. The prevalence of the *stx*$_1$, *stx*$_2$, and *stx*$_{2c}$ virulence factors was significantly different between human and cattle isolates. However, multivariable regression revealed that the presence of only the *stx*$_2$ gene was significantly associated with human isolates after controlling for confounding effects. LSPA6 typing demonstrated an apparent difference in the distribution of LSPA6 lineages between human and cattle isolates and a strong association between *stx* genotypes and LSPA6 genotypes. Population genetics tools identified three genetically distinct clusters of STEC O157. Each cluster was characterized by *stx* genotypes and LSPA6 genotypes. The human isolates typically comprised LSPA6 lineage I with *stx*$_1$ *stx*$_2$ strains and LSPA6 lineage I/II with *stx*$_2$ or *stx*$_2$ *stx*$_{2c}$ strains. In contrast, the cattle isolates comprised LSPA6 lineage II strains with *stx*$_{2c}$ or *stx*$_2$ *stx*$_{2c}$ strains in addition to the clusters identified for the human isolates. Our analyses provide new evidence that the *stx*$_2$ gene is the most distinctive feature in human isolates compared to cattle isolates in Japan, and only a subset of the genetically diverse population isolated from cattle is involved in human illnesses. Our results may contribute to international comparisons and risk assessments of STEC O157.

Shiga toxin-producing *Escherichia coli* (STEC) is one of the most common casual agents for food-borne illnesses worldwide. Of the STEC serotypes, O157 is the most abundant serogroup isolated from patients with food-borne illness. It often causes severe symptoms, such as hemolytic uremic syndrome (HUS) and thrombotic thrombocytopenic purpura (20, 31). The major reservoir of STEC O157 is thought to be cattle, although prevalence estimates for STEC O157 in beef and dairy cattle vary widely (0.2 to 48.8%) (9, 10). It is commonly believed that the pathogen is transmitted to humans through food or direct contact with cattle (7, 20). To assess the risk attributable to the STEC O157 population, it is important to understand the phenotypic and genotypic differences between human and cattle isolates. However, there are few reports that compare the population genetics of human and cattle isolates, and this is due partially to the lack of appropriate typing methods and analytical tools.

There are a number of genotyping methods used for epidemiological studies of STEC O157, such as PCR typing of several virulence factors, lineage-specific polymorphism assay using six markers (LSPA6), restriction fragment length polymorphism, pulsed-field gel electrophoresis, and variable-number tandem repeat analysis (11, 32, 34). Despite their resolution, many of these techniques do not reveal distinctive patterns or associations with human and cattle sources. Among them, *stx* genotyping and LSPA6 typing have been used to identify genotypic associations with bacterial origins.

The most important virulence factor in STEC O157 is a set of Shiga toxins (Stxs), comprising Stx1 and Stx2. Several variants of Stx2 have been found in STEC; in STEC O157, many of the strains carry Stx2, Stx2c, or both (27). Many previous studies have detected only the generic *stx*$_2$ gene (called general-*stx*$_2$ in this paper) and do not distinguish between variants of this gene. However, recent interest has focused on the importance of identifying variants of the *stx*$_2$ gene, such as the *stx*$_{2c}$ gene. In this paper, we refer to the *stx*$_2$ gene that does not include all other *stx*$_2$ variants as the "*stx*$_2$" gene. Some studies reported a much higher prevalence of the *stx*$_{2c}$ gene in cattle isolates than in human isolates (21). Other studies suggested that the *stx*$_2$ gene is more likely to be linked to human enteritis

* Corresponding author. Mailing address: Graduate School of Agricultural and Life Sciences, the University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. Phone and fax: 81 3 5841 5469. E-mail: askuma@mail.ecc.u-tokyo.ac.jp.

▽ Published ahead of print on 23 February 2011.

and cause severe symptoms than are other variants (13, 22). The other PCR-based method, LSPA6 typing, was developed from octamer-based genome scanning and can distinguish the human-biased lineage (LSPA6 lineage I [LI]) from the bovine-biased lineage (LSPA6 lineage II [LII]) (14, 34). Polymorphisms of six genetic loci used for LSPA6 typing would be selectively neutral and are thus more appropriate for describing bacterial divergence.

In this study, our aim was to further our understanding of the attribution of cattle STEC O157 isolates to human infection by using population genetics-based analyses. First, odds ratios of genotypes for human STEC O157 infection were estimated by multivariable logistic regression analysis comparing the distribution of virulence factors in human and cattle isolates while controlling for confounding effects. Five important virulence factors, including the stx_1, stx_2, stx_{2c}, *eae*, and *ehxA* genes, and LSPA6 genotypes were used as genotypic traits throughout our study. Second, the bacterial population structure and any association with isolate origin were evaluated. A number of analyses of population structure have been implemented, using techniques that provide insights that would not be apparent by conventional analyses (16, 23). Our results revealed the most distinctive features that characterize human STEC O157 isolates and showed significant differences between the populations of human and cattle isolates.

MATERIALS AND METHODS

Bacterial strains. A total of 144 non-sorbitol-fermenting STEC O157 strains were used in this study. They comprised 78 human isolates and 66 cattle isolates. The human isolates comprised 73 isolates from enteritis patients and enteritis-linked isolates and 5 reference strains from the American Type Culture Collection. All of these strains, except the five reference strains, were isolated in various regions of Japan from 1995 to 2009. Until use, all strains were stored at −80°C in Trypticase soya broth (Oxoid Ltd., Hampshire, United Kingdom) with 10% dimethyl sulfoxide (DMSO; Sigma Aldrich, MO) added.

Virulence factor profiling and LSPA6 typing. For DNA extraction, the strains were grown in 10 ml of Luria-Bertani broth (Becton, Dickinson and Company, NJ) overnight at 37°C. Genomic DNA was then extracted as previously described (1). Five virulence factors, including the stx_1, stx_2, stx_{2c}, *eae*, and *ehxA* genes, were detected by PCR as previously described (28, 29, 33). LSPA6 typing was conducted by the method described by Ziebell et al. (38). The LSPA6 alleles were placed in the following order: *folD-sfmA*, Z5935 gene, *yhcG*, *rtcB*, *rbsB*, and *arp-iclR*. LSPA6 genotypes 222222, 221222, 222212, and 221212 were regarded as LII.

Uni- and multivariable logistic regression analyses. To elucidate the associations between genotypes and origins of STEC O157, uni- and multivariable logistic regression models were constructed using the glm function in R version 2.11.1 (25). The origin of the strains, human or cattle, was considered a binary-outcome variable that was coded as 1 if the isolate was isolated from humans and 0 if it was isolated from cattle. The presence (coded as 1) and absence (coded as 0) of five virulence factors and LSPA6 alleles were used as explanatory variables. For the LSPA6 alleles, the LI-specific allele type was coded as 1, and all other allele types were coded as 0. In multivariable analysis, explanatory variables were selected through stepwise regression based on the Akaike information criterion using the stepAIC function of R.

Population structure analysis. To investigate the population structure of STEC O157, the 11 locus-specific test results (i.e., the presence or absence of five virulence factors and the six loci from the LSPA6 typing) were analyzed. First, rarefaction analysis was performed using the Analytic Rarefaction program, version 1.3 (http://www.uga.edu/strata/software/). This analysis is used to compare genetic diversity of two populations of different sample sizes by calculating the expected numbers of variants in a range of sample sizes, with confidence limits (8). A steeper slope of the rarefaction curve indicates a higher degree of diversity. Second, the pairwise fixation index (*F*st) was estimated using the ARLEQUIN program, version 3.5.1.2 (6), to evaluate statistical evidence for differentiation between the human and cattle populations of STEC O157. Third, to determine whether our samples could be grouped into genetic clusters and to infer the number of clusters that best fit the data, the Bayesian clustering method implemented in the STRUCTURE program (24) was used. This analysis uses departures from the Hardy-Weinberg equilibrium to detect population structure and begins by classifying individuals of unknown origin into a predefined number of populations (K_s). A Bayesian approach is used to infer the value of K that provides the best fit to the data, as measured by the estimated log probability of the data [ln $P(D)$]. Markov chain Monte Carlo searches consisted of 10,000 "burn-in" steps followed by 100,000 iterations. K values of 1 to 10 were evaluated with 10 replicate runs each, under the admixture model with correlated allele frequencies. A graph of the results was produced using the programs CLUMPP and DISTRUCT (12, 26).

TABLE 1. Distribution of five virulence factors and *stx* types among STEC O157 isolates

Genotype	No. (%) of positive strains from:[b]		Total no. (%) of positive strains
	Humans	Cattle	
Virulence factors			
stx_1	53 (68.0)[c]	27 (40.9)[c]	80 (55.6)
stx_2	69 (88.5)[c]	18 (27.3)[c]	87 (60.4)
stx_{2c}	18 (23.1)[c]	49 (74.2)[c]	67 (46.5)
eae	78 (100)	66 (100)	144 (100)
ehxA	76 (97.4)	64 (97.0)	140 (97.2)
stx genotypes[a]			
stx_1	2 (2.6)	3 (4.5)	5 (3.5)
stx_2	8 (10.3)	3 (4.5)	11 (7.6)
stx_{2c}	6 (7.7)[c]	32 (48.5)[c]	38 (26.4)
stx_1 stx_2	50 (64.1)[c]	11 (16.7)[c]	61 (42.4)
stx_1 stx_{2c}	1 (1.3)[c]	13 (19.7)[c]	14 (9.7)
stx_2 stx_{2c}	11 (14.1)	4 (6.1)	15 (10.4)

[a] Combinations of *stx* genes.
[b] Number (%) of positive strains among the isolates from the same origin.
[c] Significant difference in prevalence between human and cattle isolates ($P < 0.001$).

RESULTS

Presence patterns of virulence factors and relation to LSPA6 lineages. Prevalence estimates for the five virulence factors, including the stx_1, stx_2, stx_{2c}, *eae*, and *ehxA* genes, and the combination of *stx* virulence factors (*stx* genotypes) are shown in Table 1. Most of the isolates from humans (97.4%) and cattle (95.5%) were positive for general-stx_2, and this difference was not significant ($P = 0.66$). However, when the distinction was made between variants of the stx_2 gene, the prevalence of the stx_2 and stx_{2c} genes was markedly different between human and cattle isolates. The prevalence in human isolates was significantly higher for the stx_2 gene ($P < 0.00001$) and lower for the stx_{2c} gene ($P < 0.00001$) than in cattle isolates. All of the isolates carried the *eae* gene. Most of the isolates carried the *ehxA* gene, and there was no significant difference in prevalence between human and cattle isolates. For the combination of *stx* genes, there was also an apparent difference in the distribution between human and cattle isolates. The prevalence of the stx_1 stx_2 genotype (positive for both stx_1 and stx_2 genes) in human isolates was higher than that in cattle isolates ($P < 0.00001$), whereas the prevalence of the stx_{2c} and stx_1 stx_{2c} genotypes in human isolates was lower than that in cattle isolates ($P < 0.00001$).

LSPA6 genotypes and lineages are shown in Table 2. In human isolates, the predominant LSPA6 lineage was LI, and

TABLE 2. Distribution of LSPA6 genotypes among STEC O157 isolates

LSPA6 lineage	LSPA6 genotype	No. (%) of isolates							
		Origin		*stx* genotype					
		Human	Cattle	stx_1	stx_2	stx_{2c}	stx_1 stx_2	stx_1 stx_{2c}	stx_2 stx_{2c}
I	111111	54 (69.2)[a]	13 (19.7)[a]	5 (100)	3 (27.3)	0	59 (96.7)	0	0
I/II	211111	17 (21.8)	22 (33.3)	0	5 (45.5)	16 (42.1)	1 (1.6)	2 (14.3)	15 (100)
II	221212	1 (1.3)	6 (9.1)	0	0	6 (15.8)	0	1 (7.1)	0
	221222	0	2 (3.0)	0	0	0	0	2 (14.3)	0
	222212	0	5 (7.6)	0	0	7 (18.4)	0	0	0
	222222	0	11 (16.7)	0	0	2 (5.3)	0	9 (64.3)	0
Subtotal[b]		1 (1.3)[a]	24 (36.4)[a]	0	0	14 (51.9)	0	12 (44.4)	0
Other	111211	0	1 (1.5)	0	0	0	1 (1.6)	0	0
	212111	1 (1.3)	0	0	0	1 (2.6)	0	0	0
	212211	0	1 (1.5)	0	0	1 (2.6)	0	0	0
	221111	0	1 (1.5)	0	0	1 (2.6)	0	0	0
	221211	0	1 (1.5)	0	0	1 (2.6)	0	0	0
	221213	0	1 (1.5)	0	0	1 (2.6)	0	0	0
	231111	3 (3.9)	0	0	3 (27.3)	0	0	0	0
	241222	0	1 (1.5)	0	0	1 (2.6)	0	0	0
	252211	0	1 (1.5)	0	0	1 (2.6)	0	0	0
Total		77 (100)	90 (100)	5 (100)	11 (100)	38 (100)	2 (100)	14 (100)	15 (100)

[a] Significant difference in prevalence between human and cattle isolates ($P < 0.001$).
[b] Subtotal values represent LSPA6 lineage II results.

most of the other isolates belonged to LSPA6 lineage I/II (LI/II). Other isolates, except one isolate of LII, belonged to undefined lineages. These strains differ by only one allele from LI or LI/II strains. In contrast, LSPA6 genotypes were more evenly distributed among cattle isolates. The most predominant lineage was LII (36.7%), but the prevalence of LI/II (33.3%) was similar. A higher prevalence ($P < 0.00001$) of LI and lower prevalence ($P < 0.00001$) of LII was observed for human isolates compared to prevalences of cattle isolates.

When the associations of *stx* genotypes and LSPA6 lineage were evaluated, they were found to be strongly correlated (Table 2). LI strains carried only the stx_1, stx_2, or stx_1 and stx_2 genes, whereas LII strains did not have the stx_2 gene and comprised only stx_{2c} and stx_1 stx_{2c} strains. On the other hand, LI/II strains carried all the *stx* genotypes studied for this work, with the exception being the stx_1 genotype.

Uni- and multivariable regression analysis for five virulence factors and LSPA6 alleles. To elucidate the distinctive features differentiating human and cattle isolates, uni- and multivariable logistic regression analyses were performed using a data set of five virulence factors and six loci of LSPA6 alleles. Results of univariable logistic regression suggested strong associations between genotypes and origins of STEC O157 (Table 3). However, when these relationships were adjusted for confounding effects, only the stx_2 variable was significantly related to human isolates ($P < 0.00001$), and the *folD-sfmA* variable was weakly associated with human isolates ($P = 0.09$).

Population structure analysis. Rarefaction analysis demonstrated that the population of cattle isolates exhibited higher diversity than that of human isolates, as indicated by the steeper slope of the rarefaction curve (Fig. 1). This difference in genetic diversity was further tested by computing the pairwise *F*st. The pairwise *F*st of the population for human and cattle isolates showed significant (0.35; $P < 0.001$) differentiation, providing further evidence that these two populations are genetically distinct. For the clustering method implemented in STRUCTURE, a *K* estimate of 3 populations provided the best fit to our data. A confluent stacked bar plot, when *K* was equal to 3, is shown in Fig. 2. Typical genotypes for cluster 1 (Fig. 2, shown in red), 2 (green), and 3 (blue) were as follows: LI with the stx_1 stx_2 genotype (cluster 1), LI/II with the stx_{2c} or stx_2 stx_{2c} genotype (cluster 2), and LSPA6 genotype 222222 and its close relatives (LII) with the stx_{2c} or stx_1 stx_{2c} genotype (cluster 3). Almost all the human isolates were assigned to cluster 1 or 2, whereas the population of the cattle isolates was distributed among all three clusters. Furthermore, when the pairwise *F*st was estimated among the same clusters, cluster 1 and 2 showed significant differences between the human and cattle populations (*F*st of 0.19 and *P* value of 0.009 in cluster 1; *F*st of 0.08 and *P* value of 0.045 in cluster 2). When these analyses were conducted with only the isolates from Japan, the statistical significance was the same. An exception was the multivariable logistic regression, in which the *yhcG* and *rbsB* variables were selected in the final model, as was the stx_2 gene, but the results were not significant ($P > 0.05$).

DISCUSSION

In this study, we revealed differences in the populations of STEC O157 genotypes isolated from humans and cattle in Japan using PCR typing of virulence factors and LSPA6 geno-

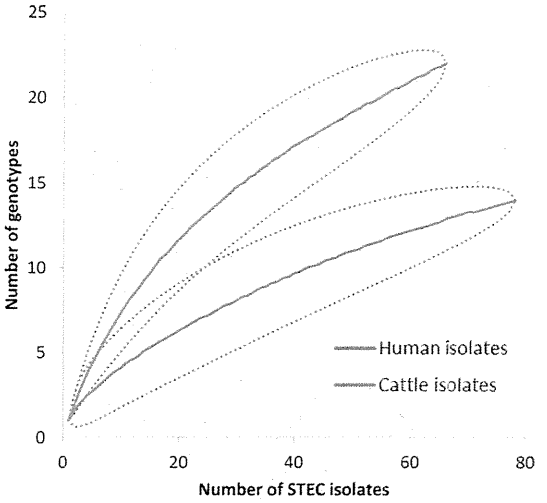

FIG. 1. Rarefaction curves of Shiga toxin-producing *Escherichia coli* populations by isolate origin. The *y* axis represents the number of genotypes from 11 locus-specific tests, including virulence factors and LSPA6 alleles. Broken lines show upper and lower 95% confidence limits.

FIG. 2. Population structure of Shiga toxin-producing *Escherichia coli* O157 from humans and cattle. The graph was produced using CLUMPP and DISTRUCT programs. The number of clusters (*K*) was predefined as three. Each haplotype is represented by a thin vertical line. Confluent stacked bar plots show the probabilities (*y* axis) that each of the individual 144 STEC O157 isolates (*x* axis) belongs to each of three Markov chain Monte Carlo (MCMC) model-derived clusters. Each color represents a population, and the color of an individual haplotype represents its proportional membership in the different clusters. Populations are color coded as follows: red, cluster 1; green, cluster 2; blue, cluster 3.

typing. Our results demonstrated that the most distinct feature of human isolates was the predominance of the stx_2 gene, whereas other *stx* virulence factors were less important. In addition, a strong correlation between isolate origin, *stx* genotype, and LSPA6 genotype was observed. These genotypic patterns were distributed differently between human and bovine isolates. Population structure analyses supported the hypothesis that the population of cattle isolates was more diverse, with only a subset of the population being linked to human disease (2, 7, 17). The importance of the stx_2 gene and the bacterial population structure adds to our understanding of the molecular epidemiology of STEC O157 and may advance global understanding of bacterial population genetics.

Our results clearly showed that the most abundant *stx* genotype among human isolates was the stx_1 stx_2 genotype, followed by the stx_2 and stx_2 stx_{2c} genotypes. STEC O157 strains carrying the stx_1 and stx_{2c} genes or the stx_{2c} gene were less likely to be involved in human illness, although they would be regarded as general-stx_2 positive. Therefore, future investigations would overestimate the risk associated with cattle isolates if stx_2 and stx_{2c} genes were not discriminated. However, this distribution of the *stx* genotypes in Japanese STEC O157 strains differs from those in Germany and Finland (3, 5), in which STEC O157 strains carrying the stx_2 and the stx_2 and stx_{2c} genes are predominant in clinical isolates.

The distribution of LSPA6 genotypes and their association with *stx* genotypes indicated a more apparent difference in the distribution of LSPA6 genotypes in Japan than that observed in North America (30, 38). Ziebell et al. (38) also showed a higher prevalence of the stx_{2c} gene in LII and LI/II strains than in LI strains. Our results are consistent with that study and provide further evidence of strong relationships between LSPA6 genotypes and *stx* genotypes. LI and LII strains did not carry the stx_{2c} and stx_2 genes, respectively. On the contrary, LI/II strains carried a wide range of *stx* virulence factors. This result is concordant with a previous report suggesting that LI/II was an intermediate genotype of LI and LII (36).

The association between these genotypic traits and the source of the isolates was examined by multivariable logistic

TABLE 3. Results of univariate and multivariate logistic regression analysis differentiating human and bovine isolates of STEC O157

Variable	Univariable analysis		Multivariable analysis	
	OR (95% CI)[a]	*P* value	OR (95% CI)	*P* value
stx_1	3.06 (1.56–6.14)	0.001	ND[c]	
stx_2	20.44 (8.83–52.03)	<0.00001	12.70 (4.82–36.76)	<0.00001
stx_{2c}	0.10 (0.05–0.22)	<0.00001	ND	
eae	NA[b]		NA	
ehxA	1.19 (0.14–10.12)	0.87	ND	
folD-sfmA	8.36 (3.99–18.42)	<0.00001	2.33 (0.86–6.06)	0.09
Z5935	9.41 (3.81–26.94)	<0.00001	ND	
yhcG	9.38 (2.98–41.56)	5.80×10^{-4}	ND	
rtcB	20.83 (6.85–90.99)	<0.00001	ND	
rbsB	NA		NA	
arp-iclR	16.25 (5.31–71.13)	1.34×10^{-5}	ND	

[a] OR, odds ratio; 95% CI, 95% confidence interval. Odds ratios above 1 indicate that the presence of virulence factors or the LSPA6 lineage I-specific allele is positively associated with human isolates.

[b] NA, not available. Odds ratio could not be calculated because all of the strains used in this study carry the *eae* gene, and none of the bovine isolates showed the LI-specific allele type in the *rbsB* loci. In the *rbsB* locus, the prevalence between human and bovine isolates was significantly different by Fisher's exact test ($P < 0.00001$).

[c] ND, not determined. These variables were eliminated from the model through backward elimination.

regression in order to control for confounding effects, a feature that has seldom been considered in previous studies. In the univariable regression, significant differences were observed for all *stx* virulence factors and LSPA6 alleles between the human and cattle isolates. Surprisingly, however, when these variables were incorporated into a multivariable regression analysis, only the stx_2 variable was significantly associated with the source, due to confounding effects between variables. For example, the apparent significant association between human isolates and the stx_1 gene revealed by univariable analysis was due to the strong correlation between the stx_1 and stx_2 genes, and simple stratification by the stx_2 status reveals no independent association between the stx_1 gene and human isolates. These results provide an interesting insight, because the stx_2 gene is the most dominant *stx* gene internationally, regardless of the variation in *stx* genotypes described above (3, 5). Other epidemiological studies showed that STEC O157 strains with the stx_2 gene are more likely to be associated with severe symptoms in humans (5, 13). It is plausible that the stx_2 gene, identified as the most distinct feature in human isolates in this study, plays the most important role in disease development as well as transmission and human infection.

Population genetics approaches clarified some characteristics of the STEC O157 population. Rarefaction analysis and pairwise *F*st showed that the bacterial population of cattle isolates was more diverse and different in structure than that of the human isolates. The Bayesian clustering method identified three distinct clusters (Fig. 2). Each cluster was characterized by certain LSPA6 genotypes and *stx* genotypes, providing further evidence for the strong correlation between LSPA6 and *stx* genotypes. In addition to LSPA6 typing, more robust typing methods, such as comparative genomic fingerprinting and microarray comparative genomic hybridization, have previously supported three genetically distinct populations in STEC O157 (14, 15, 38). Furthermore, pairwise *F*st tests indicate the possibility that clusters 1 and 2 may be further differentiated by isolate origin, although the small sample size made it difficult to draw firm conclusions. Therefore, significant differences between the human and cattle populations might become more apparent if other genetic markers are implemented in the model.

From a public health perspective, it is of interest that the human and cattle populations shared only clusters 1 and 2 and that cluster 3 was a rare genotype in human isolates. Phenotypic differences between these clusters may be a target for better control of STEC O157. The difference may be explained by recent reports describing the importance of genotypic and phenotypic heterogeneity of STEC O157 for virulence and adaptation to the host environment. Lowe et al. (18) and Zhang et al. (37) showed higher levels of adherence to intestinal cells and Shiga toxin production in LI strains than in LII strains. For *stx* genotypes, many researchers have found a correlation between the stx_2 gene and disease severity in humans (3, 5, 19). As our results suggested, if both *stx* genotypes and LSPA6 genotypes are taken into consideration, more informative observations can be made. The other phenotypic traits that may explain the genotypic difference include stress resistance; a microarray analysis of LI and LII strains revealed that there were differential expressions of several genes related to stress resistance, including heat or cold shock protein genes (4). In addition to investigations of various phenotypes, information on isolates from food and asymptomatic human carriers could explain how and where the selective pressure works. As Nakamura et al. (21) and Yokoyama et al. (35) reported, isolates from asymptomatic human carriers seemed to exhibit *stx* genotypes that were intermediate between human and cattle isolates. Further studies are required to confirm the validity of this observation.

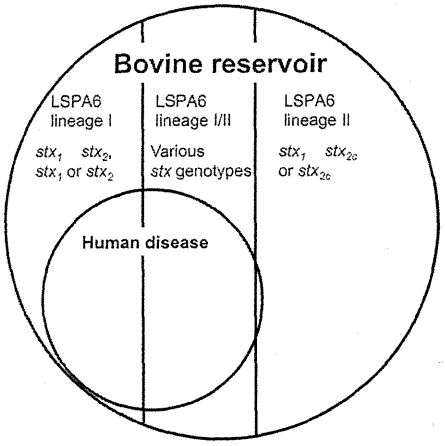

FIG. 3. Hypothetical model of association between genotypes and origins of Shiga toxin-producing *Escherichia coli* O157.

In conclusion, our analyses of genotypes in Japanese STEC O157 strains demonstrated that (i) the stx_2 gene is the most distinctive feature of human isolates compared to cattle isolates, (ii) the *stx* genotype and LSPA6 lineage are strongly correlated, and (iii) STEC O157 consists of three major genetically distinct populations, and their distribution differed significantly between human and cattle isolates. Based on those results, we suggest a hypothetical model of association between genotypes and origins of STEC O157 (Fig. 3). This model shows three genetically distinct subpopulations of STEC O157. Each subpopulation is characterized mainly by certain *stx* genotypes and LSPA6 genotypes. The population in the cattle reservoir contains all the various genotypes. However, only a subset of the population is involved in human illness, which explains the different composition of genotypes in humans and cattle. STEC isolates from a broader range of host and geographic regions would be needed to confirm and complement the hypothetical model.

ACKNOWLEDGMENTS

This research was supported by the International Training Program of the Japan Society for the Promotion of Science and a Health Sciences Research Grant from the Ministry of Health, Labor, and Welfare, Japan.

REFERENCES

1. **Beige, J., et al.** 1995. Clinical evaluation of a *Mycobacterium tuberculosis* PCR assay. J. Clin. Microbiol. **33**:90–95.
2. **Besser, T. E., et al.** 2007. Greater diversity of Shiga toxin-encoding bacteriophage insertion sites among *Escherichia coli* O157:H7 isolates from cattle than in those from humans. Appl. Environ. Microbiol. **73**:671–679.
3. **Beutin, L., et al.** 2002. Characteristics and association with disease of two major subclones of Shiga toxin (Verocytotoxin)-producing strains of *Escherichia coli* (STEC) O157 that are present among isolates from patients in Germany. Diagn. Microbiol. Infect. Dis. **44**:337–346.

4. **Dowd, S. E., and H. Ishizaki.** 2006. Microarray based comparison of two *Escherichia coli* O157:H7 lineages. BMC Microbiol. **6:**30.
5. **Eklund, M., K. Leino, and A. Siitonen.** 2002. Clinical *Escherichia coli* strains carrying *stx* genes: *stx* variants and *stx*-positive virulence profiles. J. Clin. Microbiol. **40:**4585–4593.
6. **Excoffier, L., and H. E. L. Lischer.** 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. **10:**564–567.
7. **Gyles, C. L.** 2007. Shiga toxin-producing *Escherichia coli*: an overview. J. Anim. Sci. **85:**E45–E62.
8. **Heck, K. L., G. Vanbelle, and D. Simberloff.** 1975. Explicit calculation of rarefaction diversity measurement and determination of sufficient sample size. Ecology **56:**1459–1461.
9. **Hussein, H. S., and L. M. Bollinger.** 2005. Prevalence of Shiga toxin-producing *Escherichia coli* in beef cattle. J. Food Prot. **68:**2224–2241.
10. **Hussein, H. S., and T. Sakuma.** 2005. Prevalence of Shiga toxin-producing *Escherichia coli* in dairy cattle and their products. J. Dairy Sci. **88:**450–465.
11. **Hyytia-Trees, E., S. C. Smole, P. A. Fields, B. Swaminathan, and E. M. Ribot.** 2006. Second generation subtyping: a proposed PulseNet protocol for multiple-locus variable-number tandem repeat analysis of Shiga toxin-producing *Escherichia coli* O157 (STEC O157). Foodborne Pathog. Dis. **3:**118–131.
12. **Jakobsson, M., and N. A. Rosenberg.** 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics **23:**1801–1806.
13. **Kawano, K., M. Okada, T. Haga, K. Maeda, and Y. Goto.** 2008. Relationship between pathogenicity for humans and *stx* genotype in Shiga toxin-producing *Escherichia coli* serotype O157. Eur. J. Clin. Microbiol. Infect. Dis. **27:**227–232.
14. **Kim, J., J. Nietfeldt, and A. K. Benson.** 1999. Octamer-based genome scanning distinguishes a unique subpopulation of *Escherichia coli* O157:H7 strains in cattle. Proc. Natl. Acad. Sci. U. S. A. **96:**13288–13293.
15. **Laing, C. R., et al.** 2009. *In silico* genomic analyses reveal three distinct lineages of *Escherichia coli* O157:H7, one of which is associated with hypervirulence. BMC Genomics **10:**287.
16. **Lang, P., et al.** 2010. Expanded multilocus sequence typing and comparative genomic hybridization of *Campylobacter coli* isolates from multiple hosts. Appl. Environ. Microbiol. **76:**1913–1925.
17. **Leopold, S. R., et al.** 2009. A precise reconstruction of the emergence and constrained radiations of *Escherichia coli* O157 portrayed by backbone concatenomic analysis. Proc. Natl. Acad. Sci. U. S. A. **106:**8713–8718.
18. **Lowe, R. M. S., et al.** 2009. *Escherichia coli* O157:H7 strain origin, lineage, and Shiga toxin 2 expression affect colonization of cattle. Appl. Environ. Microbiol. **75:**5074–5081.
19. **Manning, S. D., et al.** 2008. Variation in virulence among clades of *Escherichia coli* O157:H7 associated with disease outbreaks. Proc. Natl. Acad. Sci. U. S. A. **105:**4868–4873.
20. **Meng, J., M. P. Doyle, T. Zhao, and S. Zhao.** 2007. Enterohemorrhagic *Escherichia coli*, p. 249–269. *In* M. P. Doyle, L. R. Beuchat, and T. J. Montville (ed.), Food microbiology: fundamentals and frontiers, 3rd ed. ASM Press, Washington, DC.
21. **Nakamura, H., J. Ogasawara, T. Kita, A. Hase, and Y. Nishikawa.** 2008. Typing of Stx2 genes of *Escherichia coli* O157 isolates from cattle. Jpn. J. Infect. Dis. **61:**251–252.
22. **Nishikawa, Y., et al.** 2000. Relationship of genetic type of Shiga toxin to manifestation of bloody diarrhea due to enterohemorrhagic *Escherichia coli* serogroup O157 isolates in Osaka City, Japan. J. Clin. Microbiol. **38:**2440–2442.
23. **Pleydell, E., L. Rogers, E. Kwan, and N. French.** 2010. Evidence for the clustering of antibacterial resistance phenotypes of enterococci within integrated poultry companies. Microb. Ecol. **59:**678–688.
24. **Pritchard, J. K., M. Stephens, and P. Donnelly.** 2000. Inference of population structure using multilocus genotype data. Genetics **155:**945–959.
25. **R Development Core Team.** 2010. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.
26. **Rosenberg, N. A.** 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes **4:**137–138.
27. **Russmann, H., H. Schmidt, A. Caprioli, and H. Karch.** 1994. Highly conserved B-subunit genes of Shiga-like toxin II variants found in *Escherichia coli* O157 strains. FEMS Microbiol. Lett. **118:**335–340.
28. **Schmidt, H., L. Beutin, and H. Karch.** 1995. Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL 933. Infect. Immun. **63:**1055–1061.
29. **Schmidt, H., et al.** 1994. Differentiation in virulence patterns of *Escherichia coli* possessing *eae* genes. Med. Microbiol. Immunol. **183:**23–31.
30. **Sharma, R., et al.** 2009. *Escherichia coli* O157:H7 lineages in healthy beef and dairy cattle and clinical human cases in Alberta, Canada. J. Food Prot. **72:**601–607.
31. **Su, C. Y., and L. J. Brandt.** 1995. *Escherichia coli* O157:H7 infection in humans. Ann. Intern. Med. **123:**698–714.
32. **Thomson-Carter, F.** 2001. General recovery, characterisation and typing protocols for VTEC. *In* G. Duffy, P. Garvey, and D. A. McDowell (ed.), Verocytotoxigenic *E. coli.* Food & Nutrition Press, Inc., Trumbull, CT.
33. **Wang, G., C. G. Clark, and F. G. Rodgers.** 2002. Detection in *Escherichia coli* of the genes encoding the major virulence factors, the genes defining the O157:H7 serotype, and components of the type 2 Shiga toxin family by multiplex PCR. J. Clin. Microbiol. **40:**3613–3619.
34. **Yang, Z., et al.** 2004. Identification of common subpopulations of non-sorbitol-fermenting, beta-glucuronidase-negative *Escherichia coli* O157:H7 from bovine production environments and human clinical samples. Appl. Environ. Microbiol. **70:**6846–6854.
35. **Yokoyama, E., et al.** 2011. Biased distribution of IS629 among strains in different lineages of enterohemorrhagic *Escherichia coli* serovar O157. Infect. Genet. Evol. **11:**78–82.
36. **Zhang, Y., et al.** 2007. Genome evolution in major *Escherichia coli* O157:H7 lineages. BMC Genomics **8:**121.
37. **Zhang, Y. X., et al.** 2010. Lineage and host source are both correlated with levels of Shiga toxin 2 production by *Escherichia coli* O157:H7 strains. Appl. Environ. Microbiol. **76:**474–482.
38. **Ziebell, K., et al.** 2008. Genotypic characterization and prevalence of virulence factors among Canadian *Escherichia coli* O157:H7 strains. Appl. Environ. Microbiol. **74:**4314–4323.

Applied and Environmental Microbiology

Variation in Stress Resistance Patterns among *stx* Genotypes and Genetic Lineages of Shiga Toxin-Producing Escherichia coli O157

Ken-ichi Lee, Nigel P. French, Geoff Jones, Yukiko Hara-Kudo, Sunao Iyoda, Hideki Kobayashi, Yoshiko Sugita-Konishi, Hirokazu Tsubone and Susumu Kumagai

Appl. Environ. Microbiol. **2012, 78(9):3361. DOI: 10.1128/AEM.06646-11.**
Published Ahead of Print 24 February 2012.

AEM
Journals.ASM.org

Variation in Stress Resistance Patterns among *stx* Genotypes and Genetic Lineages of Shiga Toxin-Producing *Escherichia coli* O157

Ken-ichi Lee,[a] Nigel P. French,[b] Geoff Jones,[c] Yukiko Hara-Kudo,[a,d] Sunao Iyoda,[e] Hideki Kobayashi,[f] Yoshiko Sugita-Konishi,[a,d] Hirokazu Tsubone,[a] and Susumu Kumagai[a]

Graduate School of Agricultural and Life Sciences, the University of Tokyo, Yayoi, Bunkyo-ku, Tokyo, Japan[a]; Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand[b]; Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand[c]; Division of Microbiology, National Institutes of Health Sciences, Kami-Yoga, Setagaya-ku, Tokyo, Japan[d]; Department of Bacteriology, National Institute of Infectious Diseases, Toyama, Shinjuku-ku, Tokyo, Japan[e]; and National Agriculture and Food Research Organization, National Institute of Animal Health, Kannondai, Tsukuba, Ibaraki, Japan[f]

To evaluate the relationship between bacterial genotypes and stress resistance patterns, we exposed 57 strains of Shiga toxin-producing *Escherichia coli* (STEC) O157 to acid, freeze-thaw, heat, osmotic, oxidative, and starvation stresses. Inactivation rates were calculated in each assay and subjected to univariate and multivariate analyses, including principal component analysis (PCA) and cluster analysis. The *stx* genotype was determined for each strain as was the lineage-specific polymorphism assay (LSPA6) genotype. In univariate analyses, strains of the stx_1 stx_2 genotype showed greater resistance to heat than strains of the stx_1 stx_{2c} genotype; moreover, strains of the stx_1 stx_2 genotype showed greater resistance to starvation than strains of the stx_2 or stx_{2c} genotypes. LSPA6 lineage I (LI) strains showed greater resistance to heat and starvation than LSPA6 lineage II (LII) strains. PCA revealed a general trend that a strain with greater resistance to one type of stress tended to have greater resistance to other types of stresses. In cluster analysis, STEC O157 strains were grouped into stress-resistant, stress-sensitive, and intermediate clusters. In *stx* genotypes, all strains of the stx_1 stx_2 genotype were grouped with the stress-resistant cluster, whereas 72.7% (8/11) of strains of the stx_1 stx_{2c} genotype grouped with the stress-sensitive cluster. In LI strains, 77.8% (14/18) of the strains were grouped with the stress-resistant cluster, whereas 64.7% (11/17) of LII strains were grouped with the stress-sensitive cluster. These results indicate that the genotypes of STEC O157 that are frequently associated with human illness, i.e., LI or the stx_1 stx_2 genotype, have greater multiple stress resistance than do strains of other genotypes.

Shiga toxin-producing *Escherichia coli* (STEC) strains are some of the most common food-borne pathogens worldwide. Among various STEC serotypes, O157 is the most prevalent serogroup in food-borne infections, and STEC O157 infections have often been associated with severe conditions, such as hemolytic uremic syndrome (HUS) (26, 38). The low infectious dose and serious nature of STEC O157 has stimulated interest in the determinants of the survival of the organism in food and the environment. STEC O157 may encounter various environmental stresses, including nutrient depletion in the environment and changes in temperature and osmotic pressure. In the human stomach, STEC O157 must survive the low pH of gastric fluid. In order to understand the ecology of STEC O157 and establish effective control measures against the bacterium, extensive research has been done on bacterial stress responses (5, 24). Although some of these studies have shown the importance of cross-protection in stress resistance (5, 20, 33), the association between more than three types of stress has seldom been investigated. In practice, multiple stress resistance of a pathogen is important, because control measures in the food supply industry are based on the concept of employing multiple hurdles to decrease pathogen survival (14).

Recently, several studies have suggested that the bacterial populations of STEC O157 isolated from humans and cattle are genetically distinct (4, 19, 21, 45). Zhang et al. (44) and Abu-Ali et al. (1) reported that Shiga toxin production and the ability to adhere to intestinal cells are related to the genetic lineages of STEC O157. Stress resistance is an additional phenotype that may be related to genotype. Microarray and quantitative reverse transcriptase PCR analyses of STEC O157 strains have shown that several genes involved in the stress response are differentially expressed between human- and bovine-biased genotypes (7, 39). However, the difference in stress resistance among genotypes has been shown with respect to acid stress (25, 28, 34, 39). In other bacterial species, multivariate analyses, including principal component analysis (PCA) and cluster analysis, have revealed correlations between stress resistance, metabolic patterns, and their environmental niches (16, 29, 35). These methods would be useful to elucidate associations between genotype and several types of stress resistance phenotypes in STEC O157.

In this study, patterns of stress response in STEC O157 isolates were investigated in six different assays: acid, heat, freeze-thaw, high osmotic pressure, oxidative stress, and starvation. These stresses were selected because STEC O157 often encounters these stresses in food and host environments (5, 36). The relationships between stress resistance patterns and genotypes were further analyzed by univariate and multivariate methods. Genotypic traits were characterized by *stx* typing and a lineage-specific polymorphism assay with 6 markers (LSPA6), because these genotypic traits are known to reflect the divergence of STEC O157 (18, 19, 42). In this study, our objective was to examine the differences in stress resistance patterns among these genotypes of STEC O157.

Received 3 September 2011 Accepted 14 February 2012
Published ahead of print 24 February 2012
Address correspondence to Susumu Kumagai, askuma@mail.ecc.u-tokyo.ac.jp.
Supplemental material for this article may be found at http://aem.asm.org/.

doi:10.1128/AEM.06646-11

MATERIALS AND METHODS

Bacterial strains. A total of 57 STEC O157 strains, including 27 human isolates and 30 cattle isolates, were used in this study (Table 1). The human isolates comprised 23 isolates from enteritis patients and enteritis-linked isolates and four reference strains obtained from the American Type Culture Collection. All of these strains, except the four reference strains, were isolated in Japan between 1995 and 2009. The presence of stx_1, stx_2, stx_{2c}, and LSPA6 lineage has been determined previously (19). The LSPA genotype was determined using octamer-based genome scanning, which uses short sequences from the *E. coli* genome for genotyping (17). Therefore, polymorphisms in these six loci for LSPA6 are regarded as selectively neutral (42). LSPA6 lineages are associated with *stx* genotypes and the genetic clades defined by single nucleotide polymorphisms (SNP clades). In previous studies, LSPA6 lineage I (LI) strains were more likely to carry stx_2 than LSPA lineage II (LII) strains (19, 43, 45). Only limited information is available about the association between LSPA6 lineage and SNP clade; however, the hypervirulent SNP clade, clade 8, belongs to LSPA6 lineage I/II (LI/II) (12, 18).

Stress resistance assays. Strains were grown in tryptic soy broth (TSB; Becton, Dickinson and Company, NJ) at 37°C (Incubator MIR-262; Sanyo, Osaka, Japan) for 20 h. For acid stress, the culture of each strain was diluted 100-fold into 10 ml of minimal E glucose medium (EG medium) (41) acidified with hydrogen chloride (Kanto Chemical Co., Inc., Tokyo, Japan) to pH 2.5. The broth was incubated at 37°C, and viable cell counts were performed at 0, 2, 6, 8, and 12 h. To enumerate the viable cells, the inoculated broth was serially diluted with phosphate-buffered saline (PBS; Nissui Pharmaceutical Co., Ltd., Tokyo, Japan) and pour-plated onto tryptone soya agar (TSA; Oxoid Ltd., Hampshire, United Kingdom). All plates were incubated at 37°C, and colonies were counted after 48 h. For heat stress, the culture of each strain was diluted 100-fold in 1.5 ml of TSB, and the broth was submerged in a water bath (Thermominder EX; Taitec Cooperation, Saitama, Japan) at 52°C. Viable cell counts were performed at 0, 2, 4, and 6 h as described above. For freeze-thaw stress, the culture of each strain was diluted 100-fold in 10 ml of TSB, and the broth was subsequently frozen at −20°C (Medicool MPR-411 FRS; Sanyo). One freeze-thaw cycle consisted of freezing at −20°C for 22.5 h followed by thawing at 37°C for 1.5 h. Freeze-thaw cycles were repeated seven times for each sample. Viable cell counts were performed at days 0, 1, 2, 4, and 7 as described above. For high osmotic pressure stress, the culture of each strain was diluted 100-fold in 10 ml of a solution of 20% (wt/vol) NaCl (Wako Pure Chemical Industries Ltd., Tokyo, Japan). The solution was incubated at 37°C, and viable cell counts were performed at days 0, 1, 2, 4, and 7 as described above. For oxidative stress, the culture of each strain was diluted 100-fold in 10 ml of PBS containing 1 mM H_2O_2 (Wako). The solution was incubated at 37°C, and viable cell counts were performed at 0, 2, 4, 6, 8, and 12 h as described above. For starvation stress, cells were harvested by centrifugation at 4,000 × *g* for 20 min and washed twice with 10 ml of distilled water (DW) (Nihon Millipore Ltd., Tokyo, Japan). Thereafter, the cells were resuspended in 10 ml of DW at 10^3 CFU/ml. The suspension was then incubated at 25°C. Viable cell counts were performed at days 0, 7, 14, 21, and 28 as described above. DW was used, because the population of STEC decreases very slowly in other nutrient-depleted media, such as phosphate-buffered saline (11), and it is not appropriate to measure an inactivation rate. Therefore, this assay evaluated the bacterial response against starvation in the environment as well as low osmotic pressure.

Calculation and validation of inactivation rate index. We used the raw count data to construct an inactivation rate index (IRI) representing, for each strain, the rate of reduction in the bacterial count caused by each stress assay. Given an initial concentration (C_0) of bacteria in an aliquot, we assumed the subsequent concentration undergoes a constant rate of log reduction, so that at sampling time t_i the log concentration is $\log C_i = \log C_0 - \alpha t_i$, where α represents the rate of inactivation or hazard rate and i is 1, 2, 3, etc. Even if the inactivation curve is not exponential, α is still a meaningful phenotypic measure, because it represents the average rate of inactivation. We are not attempting here to produce an accurate inactivation model so have avoided multiparameter models such as the Weibull (30). The observed count at sampling time t_i is obtained from a plated volume D_i (the volume of dilution), and thus the expected count is $\mu_i = D_i C_i$. We assume then that the observed count has a Poisson distribution with mean μ_i, for which $\log \mu_i = \log D_i C_i = \log D_i + \log C_i = \log D_i + \log C_0 - \alpha t_i$.

Thus, α may be estimated as the negative slope in a Poisson regression model with $\log D_i$ as an offset. This was fitted using the glm function of R 2.11.2 (32) to obtain α values for each stress-strain combination, with smaller IRI values corresponding to greater stress resistance. Graphical inspection of α values showed that they were highly skewed, and thus they were log transformed before they were submitted for further analysis to give a reasonably symmetrical, approximately normal, distribution. This transformation can be further justified, in a wider context, by noting that modeling of the log hazard is a common approach in survival analysis (6), corresponding to an assumption of proportional hazards. Thus, the IRI used in this study was $\log \alpha$, the log of the average inactivation rate.

Univariate analyses of different STEC O157 genotypes and their resistance to stress. The IRI values of *stx* genotypes or LSPA6 lineages were compared using multiple Student's *t* tests with Tukey's adjusted *P* value.

Multivariate analyses of IRI values in all stress resistance assays. To gain insight into patterns of stress resistance, IRI values from all of the six stress resistance assays were analyzed by PCA and cluster analysis. PCA is useful for identifying a trend in a multivariate data set or a correlation between variables (40). The transformation of PCA consolidates the information of a data set into a few new variables, principal components (PCs). PCAs were computed from the following equation: $Z_i = a_{iacid}x_{acid} + a_{ifreeze}x_{freeze} + a_{iheat}x_{heat} + a_{iosmotic}x_{osmotic} + a_{ioxidative}x_{oxidative} + a_{istarvation}x_{starvation}$, where Z_i refers to a PC score in one strain in the ith PC, a refers to a factor loading, and x refers to the IRI value of the strain (Table 1) in a stress resistance assay. The a values are calculated from the correlation matrix of the observed data. First, a_1 values are computed to maximize the variance of Z_1, constrained so that the sum of the squares of factor loadings is one. This is termed the first PC, and the maximized variance the first eigenvalue. Then, a_2 is calculated to maximize the variance of Z_2 subject to orthogonality to Z_1. The computation can be continued until the sixth (the number of the variables) PC is calculated. However, PCA is usually stopped when the cumulative variance reaches 70% or when the next eigenvalue becomes less than 1.0 (13, 27). Factor loadings for each variable and PC scores for each PC have a specific interpretation. In our data set, PCs described stress resistance patterns of STEC O157. For each PC, factor loadings of each variable represent the contribution of the variable to that particular pattern. On the other hand, PC scores represent relative values of each strain among all of the strains. Therefore, the interpretations for PC scores are made according to the loadings in each PC. PCA is also useful for graphic representation. In a PCA biplot, each strain is plotted according to the first and second PC scores to show the largest variation in the data set. In the biplot, strains that have similar stress resistance patterns will be closer to each other. PCA was performed using the princomp function of R with default settings.

To see a general trend of multiple stress resistance, cluster analysis was performed using the partitioning-around-medoids (PAM) clustering method (15) in the library "cluster" in R with default settings. This method was used to inform grouping of STEC O157 strains on the basis of similarities in their IRI values for the six stress resistance assays. The number of clusters was evaluated from 2 to 10 by the silhouette width (15) and visual inspection of PCA biplots. The stress resistance pattern of each cluster was characterized by mean IRI values, and mean IRI values were compared using multiple Student's *t* tests with adjusted *P* values. From these results, the number of clusters for subsequent analysis was decided, where the majority of strains can be defined by the mean IRI values (e.g., stress resistant, susceptible, etc.). Associations between resistance patterns

TABLE 1 Genotypic characteristic and IRI values of the strains used in this study

Strain	Source	*stx* genotype	LSPA6 genotype[a]	LSPA6 lineage[a]	IRI value[b] (stress)					
					Acid	Freeze-thaw	Heat	Osmotic	Oxidative	Starvation
ESC226	Cattle	stx_1	111111	I	−0.98	−2.29	0.06	−2.54	0.75	−4.70
ESC343	Cattle	stx_1	111111	I	−0.91	−3.07	0.37	−3.94	0.09	−3.12
ATCC35150	Human	stx_1	111111	I	−1.12	−3.19	−0.20	−3.35	−0.27	−4.80
ATCC43890	Human	stx_1	111111	I	0.56	−3.15	0.58	−2.48	−0.06	−3.09
EC55	Human	stx_2	111111	I	0.52	−3.03	0.49	−3.22	−0.33	−3.13
ESC361	Human	stx_2	111111	I	−0.96	−3.23	0.15	−5.30	−0.25	−3.10
EC10	Human	stx_1 stx_2	111111	I	−0.57	−3.30	−0.03	−3.65	0.11	−4.68
EC18	Human	stx_1 stx_2	111111	I	−0.93	−3.11	0.34	−2.31	0.60	−5.95
EC28	Human	stx_1 stx_2	111111	I	−0.97	−3.05	0.80	−2.87	0.98	−4.13
EC32	Human	stx_1 stx_2	111111	I	−0.82	−2.84	0.87	−2.87	0.53	−4.03
EC33	Human	stx_1 stx_2	111111	I	−0.83	−2.42	0.20	−2.97	0.92	−4.91
EC43	Human	stx_1 stx_2	111111	I	−1.04	−3.05	−0.71	−3.79	−0.01	−5.49
EC52	Human	stx_1 stx_2	111111	I	−0.89	−3.11	−0.33	−3.35	0.13	−4.92
EC70	Human	stx_1 stx_2	111111	I	−1.84	−3.17	0.30	−3.65	−0.09	−4.40
EC170	Cattle	stx_1 stx_2	111111	I	−1.08	−3.24	−0.43	−3.46	0.12	−4.57
ESC228	Cattle	stx_1 stx_2	111111	I	−0.82	−3.17	−0.12	−2.40	0.88	−4.22
ESC342	Cattle	stx_1 stx_2	111111	I	−1.11	−3.06	0.29	−3.51	−0.05	−4.61
ATCC43895	Human	stx_1 stx_2	111111	I	−1.22	−3.41	−1.20	−3.71	−0.21	−5.01
EC160	Cattle	stx_2	211111	I/II	−0.51	−2.83	0.13	−3.64	−0.40	−3.11
ESC225	Cattle	stx_2	211111	I/II	−0.92	−2.49	−0.44	−2.04	1.00	−3.77
ESC344	Cattle	stx_2	211111	I/II	−1.01	−3.18	−0.36	−3.89	0.32	−5.34
ESC349	Human	stx_2	211111	I/II	−1.21	−3.10	−0.26	−3.53	0.18	−4.53
ESC360	Human	stx_2	211111	I/II	−0.83	−2.47	0.89	−2.51	0.74	−3.03
EC59	Human	stx_{2c}	211111	I/II	−0.93	−3.00	−0.01	−2.32	0.80	−4.13
EC157	Cattle	stx_{2c}	211111	I/II	−1.09	−3.17	−0.89	−3.61	−0.23	−5.09
ESC206	Cattle	stx_{2c}	211111	I/II	−0.38	−2.75	−0.44	−2.87	0.80	−3.96
ESC219	Cattle	stx_{2c}	211111	I/II	−0.10	−3.28	0.88	−1.96	0.79	−3.84
EC66	Human	stx_1 stx_{2c}	211111	I/II	−0.93	−2.11	0.16	−2.19	0.93	−4.14
ESC138	Cattle	stx_1 stx_{2c}	211111	I/II	−1.26	−2.89	0.92	−2.39	0.75	−4.14
EC1	Human	stx_1 stx_{2c}	211111	I/II	−1.01	−2.95	0.49	−2.32	0.71	−4.18
EC42	Human	stx_1 stx_{2c}	211111	I/II	−1.20	−3.00	0.39	−2.61	0.57	−4.16
ESC362	Human	stx_1 stx_{2c}	211111	I/II	−1.12	−2.66	0.00	−2.33	0.55	−3.94
ATCC43889	Human	stx_1 stx_{2c}	211111	I/II	−0.77	−2.85	0.94	−2.84	1.21	−3.70
ESC356	Human	stx_{2c}	221212	II	−1.05	−2.92	0.11	−3.03	0.31	−4.35
EC37	Human	stx_{2c}	222212	II	−0.23	−2.97	0.68	−1.94	0.38	−3.53
EC38	Human	stx_{2c}	222212	II	−0.60	−2.70	0.99	−2.63	1.00	−3.16
EC169	Cattle	stx_{2c}	222212	II	−0.83	−2.80	0.37	−3.87	−0.49	−3.06
EC183	Cattle	stx_{2c}	222212	II	−0.91	−2.99	0.61	−4.24	−0.43	−3.03
ESC209	Cattle	stx_{2c}	222212	II	−0.78	−3.08	0.32	−2.31	1.35	−4.07
ESC340	Cattle	stx_{2c}	222212	II	−0.84	−2.87	0.70	−2.83	0.81	−3.04
ESC211	Cattle	stx_{2c}	222222	II	−0.92	−2.88	0.48	−3.26	0.38	−3.09
EC181	Cattle	stx_1 stx_{2c}	221222	II	−1.11	−3.09	−0.33	−3.71	−0.35	−4.72
EC164	Cattle	stx_1 stx_{2c}	222222	II	−1.28	−3.26	−0.41	−3.47	−0.33	−5.04
ESC213	Cattle	stx_1 stx_{2c}	222222	II	−0.34	−2.56	1.45	−1.77	1.75	−4.02
ESC214	Cattle	stx_1 stx_{2c}	222222	II	−0.86	−2.65	1.34	−2.38	1.06	−3.02
ESC215	Cattle	stx_1 stx_{2c}	222222	II	−0.93	−2.94	1.04	−2.45	0.99	−3.07
ESC216	Cattle	stx_1 stx_{2c}	222222	II	−0.51	−2.69	1.83	−1.79	0.93	−4.32
ESC220	Cattle	stx_1 stx_{2c}	222222	II	−0.65	−2.83	0.89	−1.91	0.55	−3.37
ESC222	Cattle	stx_1 stx_{2c}	222222	II	−0.72	−2.78	0.99	−2.25	1.13	−4.17
ESC229	Cattle	stx_1 stx_{2c}	222222	II	−0.67	−2.97	1.30	−2.55	1.18	−3.05
EC44	Human	stx_2	231111	Other	−0.68	−2.65	1.14	−2.49	0.85	−3.07
EC45	Human	stx_2	231111	Other	−0.70	−2.06	0.09	−2.35	1.07	−4.37
EC175	Human	stx_{2c}	212111	Other	−1.06	−3.08	0.41	−4.36	−0.57	−3.35
ESC223	Cattle	stx_{2c}	212211	Other	−0.91	−1.38	0.86	−3.07	0.58	−3.07
ESC231	Cattle	stx_{2c}	221111	Other	−0.97	−2.90	0.08	−2.58	0.94	−4.83
ESC367	Cattle	stx_{2c}	221211	Other	−0.91	−3.23	0.35	−2.57	0.77	−4.91
ESC339	Cattle	stx_{2c}	252211	Other	−1.09	−3.12	1.09	−2.85	1.09	−4.36

[a] The LSPA6 alleles were placed in the following order: *folD sfmA*, Z5935 gene, *yhcG*, *rtcB*, *rbsB*, and *arp-iclR*. LSPA6 genotypes of 222222, 221222, 222212, and 221212 were regarded as LII.

[b] These values are calculated from inactivation rate. Because of log transformation, some values are positive, whereas the others are negative. Smaller IRI values correspond to greater stress resistance.

TABLE 2 Principal component analysis of inactivation rate index values in six stress resistance assays

Variable	Factor loading[a] in: 1st PC	2nd PC
Acid	−0.30	−0.43
Freeze-thaw	−0.36	0.12
Heat	−0.48	−0.21
Osmotic	−0.48	0.37
Oxidative	−0.47	0.44
Starvation	−0.31	−0.65
Proportion of variance	0.46	0.22
Cumulative proportion	0.46	0.68

[a] Factor loading indicates the relative weight of the individual stress resistance assay on each principal component (PC).

TABLE 3 Differences in PC scores between *stx* genotypes and LSPA6 lineage

Genotype	PC score (mean ± SD) in: 1st PC[a]	2nd PC[b]
stx genotype		
stx_1	0.33 ± 1.51	−0.67 ± 1.87
stx_2	0.07 ± 1.78	−0.59 ± 1.67
stx_{2c}	−0.14 ± 1.38	−0.25 ± 1.16
stx_1 stx_2	1.21 ± 1.37[c]	0.59 ± 0.66
stx_1 stx_{2c}	−1.11 ± 1.98[c]	0.26 ± 0.62
stx_2 stx_{2c}	−0.49 ± 0.64	0.59 ± 0.34
LSPA6 lineage		
I	0.99 ± 1.41[d]	−0.07 ± 1.49
I/II	−0.06 ± 1.44	0.34 ± 0.87
II	−0.75 ± 1.79[d]	−0.32 ± 0.92
Other	−0.60 ± 1.46	0.23 ± 1.28

[a] Higher PC scores indicate greater resistance to the six stresses.
[b] Higher PC scores indicate greater resistance to acid, heat, and starvation stresses and lesser resistance to freeze-thaw, high osmotic pressure, and oxidataive stresses.
[c] Significant ($P = 0.0099$) difference in mean PC scores in a pairwise *t* test with adjusted *P* values.
[d] Significant ($P = 0.0088$) difference in mean PC scores in a pairwise *t* test with adjusted *P* values.

and *stx* genotypes or LSPA6 lineages were explored by the distribution of the IRI-based clusters among genotypes.

RESULTS

Calculation and validation of IRI values. IRI values were obtained from the results of the six stress resistance assays (Table 1). Although IRI values were calculated from the inactivation rate, positive and negative values resulted from subsequent log transformation. These transformed IRI values gave an approximately normal distribution and thus were appropriate for multivariate analyses. The range of IRI values varied among the assays; therefore, to exclude the effect of the ranges of IRI values among the assays, IRI values were standardized before multivariate analyses. For standardization, within each assay, the mean was subtracted from each IRI value and then divided by the standard deviation (31).

Univariate analyses of STEC O157 genotypes and their stress resistance. Associations between the source or genotype of STEC O157 and the IRI values from each stress resistance assay were evaluated. No significant differences in IRI values between human and cattle isolates of STEC O157 were observed. However, in comparisons between IRI values and genotypes, significant associations were observed in the heat and starvation assays. Strains of the stx_1 stx_2 genotype (strains with both stx_1 and stx_2 genes) showed significantly greater resistance to heat than strains of the stx_1 stx_{2c} genotype ($P = 0.011$). Strains of the stx_1 stx_2 genotype also showed greater resistance to the starvation assay than strains of the stx_2 or stx_{2c} genotypes (stx_1 stx_2 strains versus stx_2 strains, $P = 0.019$; stx_1 stx_2 strains versus stx_{2c} strains, $P = 0.011$). LI and LI/II strains showed significantly greater resistance to heat than LII strains (LI strains versus LII strains, $P = 0.005$; LI/II strains versus LII strains, $P = 0.015$). In the starvation assay, LI strains showed greater resistance than LII strains ($P = 0.024$).

Multivariate analyses of IRI values in all stress resistance assays. All IRI values in the six stress resistance assays were included in the multivariate analyses. First, in PCA, the first and second PCs accounted for 45.8% and 22.1%, respectively, of the variance within the multivariate data set. Because most of the variation (67.9%) was contained in the first and second PC and no other PCs possessed eigenvalues of >1.0, we focused on these PCs. All of the factor loadings on the first PC were negative values, indicating that the first PC described multiple stress resistance, with higher scores associated with greater resistance to the six stresses, and lower scores associated with susceptibility to them (Table 2). On the other hand, higher scores in the second PC were associated with resistance to the acid, heat, and starvation stresses and susceptibility to the osmotic, oxidative, and freeze-thaw stresses. Pairwise comparisons of PC scores in the first PC and *stx* genotypes or LSPA6 lineages showed higher scores in the strains of the stx_1 stx_2 genotype ($P = 0.0099$) and LI strains ($P = 0.0088$) than in the strains of the stx_1 stx_{2c} genotype and LII strains, respectively (Table 3). This result indicates the general trend of greater multiple stress resistance of the strains of the stx_1 stx_2 genotype and LI strains than those of the stx_1 stx_{2c} genotype and LII strains, respectively. In PC scores in the second PC, there was no significant difference among genotypes.

In cluster analysis, the average silhouette width was the largest when the number of clusters was six (see Fig. S1 in the supplemental material); however, the number was not decisive, because its standard deviation was large. According to the comparison of IRI values among clusters, most of the strains (>85%) were grouped to stress-resistant or -sensitive clusters when the numbers of clusters were two and three (see Table S1 in the supplemental material). When the number of clusters was two, each strain was grouped essentially according to its PC score in the first PC (see Fig. S2 in the supplemental material). Therefore, in order to describe the diversity of stress resistance pattern, three clusters were considered appropriate for the analysis. When the number of clusters was predefined as three, 31, 7, and 19 strains were grouped into clusters 1, 2, and 3, respectively (Fig. 1). To characterize stress resistance patterns, mean IRI values were compared among these clusters (Table 4). The mean IRI value of cluster 1 was significantly lower than that of cluster 3 in all of the six assays. It suggested that cluster 1 is a group relatively resistant to the stresses in this study and that cluster 3 is a stress-susceptible group. Cluster 2 can be characterized as a group of intermediate or variable stress resistance. The mean IRI values for cluster 2 in the acid, freeze-thaw, and heat stress assays were between those for clusters 1 and 3. In

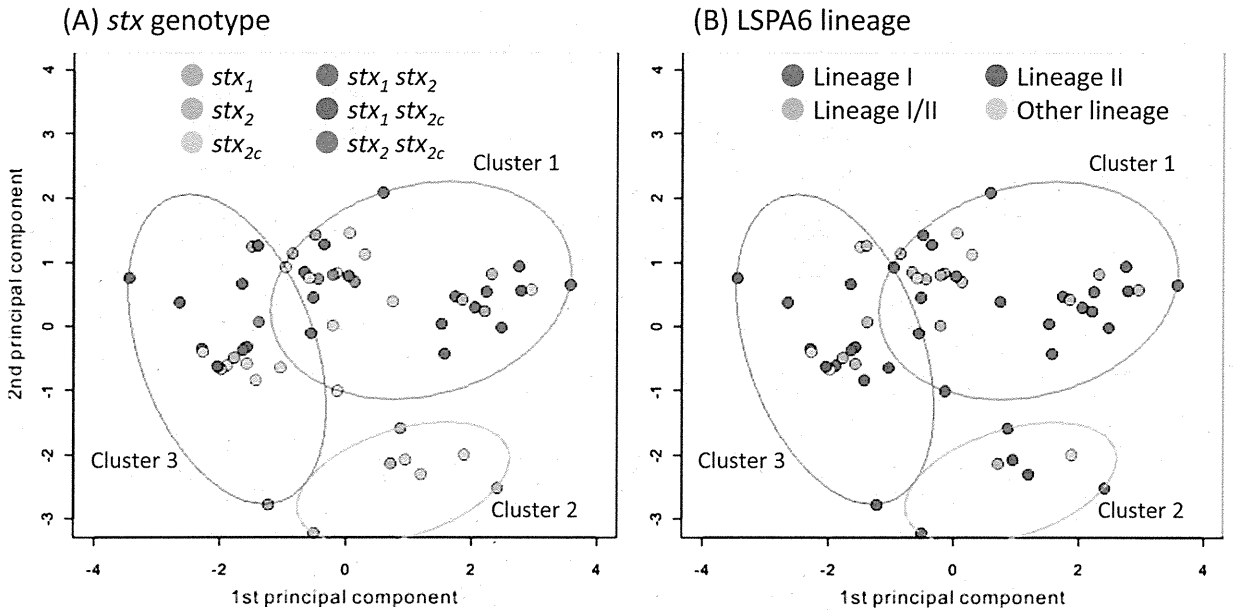

FIG 1 The results of cluster analysis in biplots by principal component analysis (PCA). Distances between plots represent the similarities between strains. Each *stx* genotype (A) and LSPA6 lineage (B) of STEC O157 corresponds to a color, as shown. Plots located inside the circle belong to each cluster: red, cluster 1; green, cluster 2; blue, cluster 3. The distance between plots shows the similarity of their stress resistance patterns. PCA shows that higher values on the *x* axis indicate greater multiple stress resistance to six stresses and higher values on the *y* axis indicate greater resistance to acid, heat, and starvation stresses and lower resistance to freeze-thaw, osmotic pressure, and oxidative stresses.

the osmotic and oxidative stress assays, strains in cluster 2 showed the greatest stress resistance; however, they showed the most susceptibility to starvation stress (Table 4). Among these clusters, genotypes of STEC O157 were not distributed evenly (Fig. 2). All strains of the stx_1 stx_2 genotype were grouped to cluster 1, whereas most of the strains (72.7%) of the stx_1 stx_{2c} genotype were grouped to cluster 3. Three of four strains (75.0%) of the stx_2 stx_{2c} genotype were grouped to cluster 1. The difference in distribution of clusters was more apparent for the LSPA6 lineage. The most common cluster for LI strains was cluster 1; on the other hand, LII was the most common in cluster 3. LI/II strains showed an intermediate pattern between LI and LII strains. In the strains carrying only one type of *stx* and strains of atypical LSPA6 lineage, none of the clusters held the majority.

DISCUSSION

In this study, we showed an association between genotypes of STEC O157 and stress resistance in six different assays by using univariate and multivariate analyses. Because only a subset of the genotypes is likely to be involved in human infection (9, 19), our results provide an insight into how selective pressures may affect the transmission of STEC O157 from cattle to humans via food and the environment.

Significant associations were found between genotypic traits of STEC O157 and their stress responses, although no significant associations were observed between the strain sources and their stress responses. Because human and cattle isolates of STEC O157 have been reported to share a subpopulation (9, 18, 19), it is potentially more revealing to examine the distribution of phenotypes of the bacterium among the genotypes rather than the sources (19). When IRI values were compared among *stx* genotypes or LSPA6 lineages, significant differences were observed in the heat and starvation stresses. Interestingly, the stx_1 stx_2 genotype and LI and LI/II strains, which showed greater resistance to these stresses, have been shown to be more frequently associated with human infection (19, 43, 45). This suggests that responses of STEC O157 to these two stresses play more important roles in their ecology than other stresses do. In *E. coli*, heat and starvation stresses induce general stress responses and affect various components of the bacterial cell via heat shock protein and *rpoS* regulons, respectively (5, 8, 37). The broad spectrum of effects on bacterial metabolism and cell components might be reflected by these results; however,

TABLE 4 Difference in IRI values between clusters in each stress resistance assay

Cluster	Mean ± SD of IRI values (stress)[a]					
	Acid	Freeze-thaw	Heat	Osmotic	Oxidative	Starvation
1	−1.02 ± 0.24 AB	−3.00 ± 0.26 B	0.02 ± 0.53 B	−3.03 ± 0.56 B	0.41 ± 0.45 B	−4.55 ± 0.57 BC
2	−0.67 ± 0.55 A	−3.00 ± 0.15	0.36 ± 0.17 A	−4.08 ± 0.66 B	−0.34 ± 0.22 B	−3.13 ± 0.10 B
3	−0.61 ± 0.37 B	−2.68 ± 0.44 B	0.90 ± 0.43 AB	−2.35 ± 0.36 B	0.91 ± 0.39 B	−3.53 ± 0.51 C

[a] A, values with the same letter designation in each assay have a statistically significant difference ($P < 0.05$) in mean IRI values in a pairwise *t* test with adjusted *P* values; B and C, values with the same letter designation in each assay have a statistically significant difference ($P < 0.01$) in mean IRI values in a pairwise *t* test with adjusted *P* values.

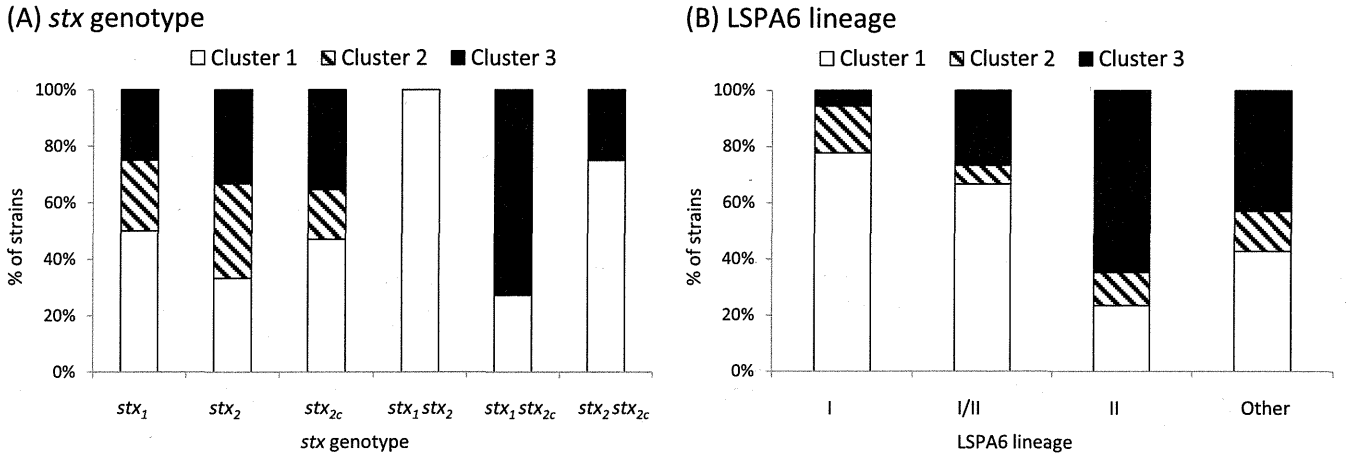

FIG 2 Associations between clusters and *stx* genotypes (A) and LSPA6 lineages (B) of STEC O157. Cluster 1 and cluster 3 were characterized as stress-resistant and stress-susceptible groups, respectively. Stress resistances of strains in cluster 2 varied depending on the stress.

rpoS sequences of the strains used in this study were highly homogeneous, and there was no apparent correlation when this was investigated by uni- and multivariate analyses (see Fig. S3 in the supplemental material). It is likely that variations in other stress-related genes are responsible for the diversity in the stress response among strains.

To characterize stress resistance patterns of the strains, all IRI values were subjected to multivariate analyses, including PCA and cluster analysis. These analyses revealed the influence of all six stresses, including four stresses that did not yield significant results in the univariate analyses. In PCA, the factor loading values for all of the six assays in the first PC were in the same direction (negative values). This finding implies that the largest contribution of variation to the stress resistance pattern is in multiple stress resistance, that is, a general trend that an STEC O157 strain that is resistant to one stress is also resistant to other stresses. This finding is in agreement with the results of Benito et al. (2), who showed a similar trend in resistance to hydrostatic pressure, mild-heat, acid, oxidative, and osmotic stresses for *E. coli* O157 strains.

In cluster analysis, STEC O157 strains were grouped into three clusters according to their stress resistance patterns. These clusters were characterized by mean IRI values (Table 4). Multiple comparisons elucidated that strains in cluster 1 showed greater resistance in all of the six assays than strains in cluster 3 did. These results corroborate the first PC in PCA, which showed that the most discriminative factor is multiple stress resistance. Interestingly, the clusters were not distributed evenly among genotypes of STEC O157. This suggests that the stress resistance patterns are associated with the phylogeny of STEC O157 to some extent. In contrast to the predominance of cluster 1 in LI strains, LII strains were predominantly comprised of cluster 3. LI/II strains showed an intermediate distribution among the clusters (Fig. 2), consistent with the intermediate genetic properties in this lineage (18, 45). Similar trends were observed in strains carrying two different types of *stx*, i.e., stx_1 stx_2, stx_1 stx_{2c}, and stx_2 stx_{2c} genotypes. This could be attributable to the strong association between the *stx* genotype and the LSPA6 lineage (12, 19, 43, 45). In a previous study, most of the strains (>85%) of the stx_1 stx_2, stx_2 stx_{2c}, and stx_1 stx_{2c} genotypes were shown to belong to LI, LI/II, and LII, respectively (19). On the other hand, strains carrying only one type of *stx* showed a more diverse stress resistance pattern. This result is in concordance with the result of a study that showed that these strains are genetically more diverse than strains carrying two types of *stx* (19). The second PC showed that the stress responses to acid, heat, and starvation stress tend to be similar to each other and likewise osmotic, oxidative, and freeze-thaw stresses. Although we cannot infer whether or not this is the result of overlapping genetic responses, these stress responses may have a similar role in the ecology of STEC O157.

An important implication of the clustering result is that human-biased genotypes, such as LI strains and strains of the stx_1 stx_2 or stx_2 stx_{2c} genotype were more likely to be stress resistant than the other genotypes. In food and the environment, pathogens encounter various types of stress. Those stressors could explain a pressure for a subset of genotypes to survive and consequently cause human infections. In *Listeria monocytogenes* and lactic acid bacteria, relationships between stress resistance patterns and their environmental niches have been shown (3, 29, 46). It is plausible that a similar selective pressure exists for STEC O157. Previously, differences in acid resistance among STEC O157 strains from different sources and of different genotypes have been reported (25, 28, 39). In our study, however, no significant difference of IRI values in acid resistance assays was observed among the genotypes. Further studies are required to draw any conclusion regarding the acid resistance in STEC O157. In other studies on stress resistance of STEC O157, the results of Benito et al. (2) are concordant with our results, while the results of Malone et al. (22) are not. In order to compare the results across studies, additional work is needed to confirm the genotypes under investigation. In further studies on the stress resistance of STEC O157, the *stx* genotype and LSPA6 lineage would be useful, because differences in distribution of these genotypes among sources have been identified in several population genetics analyses (18, 19).

The other important explanation for genetic divergence is the presence of virulence factors. Higher levels of *stx* expression and adherence to colonic cells from cattle were observed in LI strains (44). Interestingly, simultaneous upregulation of genes relating to a DNA damage repair system, SOS response, and *stx* in LI strains was observed in a microarray study (7). These associations between regulation of stress response and virulence properties may

enhance bacterial survival in the environment, as well as pathogenesis. On the other hand, Dowd and Ishizaki (7) showed that genes encoding heat and cold shock proteins were differentially expressed between LI and LII strains of STEC O157. Heat and cold shock proteins influence various cell components and regulatory genes (8, 10) and may therefore be candidates for mediators of multiple stress resistance and, possibly, virulence properties. In this study, the association between stress resistance and a highly virulent genotype, clade 8, defined by single nucleotide polymorphisms (23), remains unclear. Several studies showed that strains of clade 8 belong to LI/II (12, 18). In this study, LI/II strains showed intermediate stress resistance in cluster analysis (Fig. 2).

In conclusion, our study revealed that (i) in our collection of STEC O157 isolates, heat and starvation stresses appeared to be more important for characterizing the bacterial population than were other stresses; (ii) there was a positive correlation among the six types of stress resistance; and (iii) some genotypes showed an association with multiple stress resistance. It is of interest that human-biased genotypes, such as LI strains or strains of the stx_1 stx_2 or stx_2 stx_{2c} genotype, showed greater stress resistance than strains of other genotypes did. However, the practical importance and mechanisms of stress resistance need further investigation. Our results also showed the importance of using STEC O157 strains from different genotypes in survival studies.

ACKNOWLEDGMENTS

This research was partially supported by the International Training Program of the Japan Society for the Promotion of Science and a Health Sciences Research Grant from the Ministry of Health, Labor, and Welfare, Japan.

REFERENCES

1. **Abu-Ali GS, et al.** 2010. Increased adherence and expression of virulence genes in a lineage of *Escherichia coli* O157:H7 commonly associated with human infections. PLoS One 5:e10167.
2. **Benito A, Ventoura G, Casadei M, Robinson T, Mackey B.** 1999. Variation in resistance of natural isolates of *Escherichia coli* O157 to high hydrostatic pressure, mild heat, and other stresses. Appl. Environ. Microbiol. 65:1564–1569.
3. **Bergholz TM, den Bakker HC, Fortes ED, Boor KJ, Wiedmann M.** 2010. Salt stress phenotypes in *Listeria monocytogenes* vary by genetic lineage and temperature. Foodborne Pathog. Dis. 7:1537–1549.
4. **Besser TE, et al.** 2007. Greater diversity of Shiga toxin-encoding bacteriophage insertion sites among *Escherichia coli* O157:H7 isolates from cattle than in those from humans. Appl. Environ. Microbiol. 73:671–679.
5. **Chung HJ, Bang W, Drake MA.** 2006. Stress response of *Escherichia coli.* Compr. Rev. Food. Sci. Food Saf. 5:52–64.
6. **Collet D.** 2003. Modelling survival data in medical research. CRC Press, Boca Raton, FL.
7. **Dowd SE, Ishizaki H.** 2006. Microarray based comparison of two *Escherichia coli* O157:H7 lineages. BMC Microbiol. 6:30.
8. **Guisbert E, Yura T, Rhodius VA, Gross CA.** 2008. Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. Microbiol. Mol. Biol. Rev. 72:545–554.
9. **Gyles CL.** 2007. Shiga toxin-producing *Escherichia coli*: an overview. J. Anim. Sci. 85:E45–E62.
10. **Han MJ, Lee SY.** 2006. The *Escherichia coli* proteome: Past, present, and future prospects. Microbiol. Mol. Biol. Rev. 70:362–439.
11. **Hara-Kudo Y, Miyahara M, Kumagai S.** 2000. Loss of O157 O antigenicity of verotoxin-producing *Escherichia coli* O157:H7 surviving under starvation conditions. Appl. Environ. Microbiol. 66:5540–5543.
12. **Hartzell A, et al.** 2011. *Escherichia coli* O157:H7 of genotype lineage-specific polymorphism assay 211111 and clade 8 are common clinical isolates within Pennsylvania. Foodborne Pathog. Dis. 8:763–768.
13. **Jacobs J, Rhodes M, Sturgis B, Wood B.** 2009. Influence of environmental gradients on the abundance and distribution of *Mycobacterium* spp. in a coastal lagoon estuary. Appl. Environ. Microbiol. 75:7378–7384.
14. **Jay JM, Loessner MJ, Golden DA.** 2005. Modern food microbiology, 7th ed. Springer, New York, NY.
15. **Kaufman L, Rousseeuw PJ.** 1990. Finding groups in data: an introduction to cluster analysis. Wiley, New York, NY.
16. **Keymer DP, Miller MC, Schoolnik GK, Boehm AB.** 2007. Genomic and phenotypic diversity of coastal *Vibrio cholerae* strains is linked to environmental factors. Appl. Environ. Microbiol. 73:3705–3714.
17. **Kim J, Nietfeldt J, Benson AK.** 1999. Octamer-based genome scanning distinguishes a unique subpopulation of *Escherichia coli* O157:H7 strains in cattle. Proc. Natl. Acad. Sci. U. S. A. 96:13288–13293.
18. **Laing CR, et al.** 2009. *In silico* genomic analyses reveal three distinct lineages of *Escherichia coli* O157:H7, one of which is associated with hypervirulence. BMC Genomics 10:287.
19. **Lee K, et al.** 2011. Multivariate analyses revealed distinctive features differentiating human and cattle isolates of Shiga toxin-producing *Escherichia coli* O157 in Japan. J. Clin. Microbiol. 49:1495–1500.
20. **Leenanon B, Drake MA.** 2001. Acid stress, starvation, and cold stress affect poststress behavior of *Escherichia coli* O157:H7 and nonpathogenic *Escherichia coli.* J. Food Prot. 64:970–974.
21. **Leopold SR, et al.** 2009. A precise reconstruction of the emergence and constrained radiations of *Escherichia coli* O157 portrayed by backbone concatenomic analysis. Proc. Natl. Acad. Sci. U. S. A. 106:8713–8718.
22. **Malone AS, Yousef AE, LeJeune JT.** 2007. Association of prophage antiterminator Q alleles and susceptibility to food-processing treatments applied to *Escherichia coli* O157 in laboratory media. J. Food Prot. 70:2617–2619.
23. **Manning SD, et al.** 2008. Variation in virulence among clades of *Escherichia coli* O157:H7 associated with disease outbreaks. Proc. Natl. Acad. Sci. U. S. A. 105:4868–4873.
24. **McClure PJ, Hall S.** 2000. Survival of *Escherichia coli* in foods. Symp. Ser. Soc. Appl. Microbiol. 2000:61S–70S.
25. **McKellar RC, Knight KP.** 1999. Growth and survival of various strains of enterohemorrhagic *Escherichia coli* in hydrochloric and acetic acid. J. Food Prot. 62:1466–1469.
26. **Meng J, Doyle MP, Zhao T, Zhao S.** 2007. Enterohemorrhagic *Escherichia coli*, p 249–269. *In* Doyle MP, Beuchat LR (ed), Food microbiology: fundamentals and frontiers, 3rd ed. ASM Press, Washington DC.
27. **Oda Y, Ouchi K.** 1989. Principal component analysis of the characteristics desirable in bakers yeasts. Appl. Environ. Microbiol. 55:1495–1499.
28. **Oh DH, et al.** 2009. *Escherichia coli* O157:H7 strains isolated from environmental sources differ significantly in acetic acid resistance compared with human outbreak strains. J. Food Prot. 72:503–509.
29. **Parente E, et al.** 2010. Diversity of stress tolerance in *Lactobacillus plantarum, Lactobacillus pentosus* and *Lactobacillus paraplantarum*: a multivariate screening study. Int. J. Food Microbiol. 144:270–279.
30. **Peleg M, Cole MB.** 1998. Reinterpretation of microbial survival curves. Crit. Rev. Food Sci. Nutr. 38:353–380.
31. **Petrie A, Watson P.** 2006. Statistics for veterinary and animal science, 2nd ed. Blackwell Publishing Ltd., Oxford, United Kingdom.
32. **R Development Core Team.** 2010. R: a language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org.
33. **Rowe MT, Kirk RB.** 2000. Effect of nutrient starvation on the resistance of *Escherichia coli* O157:H7 to subsequent heat stress. J. Food Prot. 63:1745–1748.
34. **Saridakis CE, Johnson RP, Benson A, Ziebell K, Gyles CL.** 2004. Influence of animal origin and lineage on survival of *Escherichia coli* O157:H7 strains in strong and weak acid challenges. J. Food Prot. 67:1591–1596.
35. **Siezen RJ, et al.** 2010. Phenotypic and genomic diversity of *Lactobacillus plantarum* strains isolated from various environmental niches. Environ. Microbiol. 12:758–773.
36. **Stevens MP, van Diemen PM, Dziva F, Jones PW, Wallis TS.** 2002. Options for the control of enterohaemorrhagic *Escherichia coli* in ruminants. Microbiology 148:3767–3778.
37. **Stortz G, Hengge-Aronis R.** 2000. Bacterial stress responses. ASM Press, Washington, DC.
38. **Su CY, Brandt LJ.** 1995. *Escherichia coli* O157:H7 infection in humans. Ann. Intern. Med. 123:698–714.
39. **Vanaja SK, Springman AC, Besser TE, Whittam TS, Manning SD.** 2010. Differential expression of virulence and stress fitness genes between *Esch-*

erichia coli O157:H7 strains with clinical or bovine-biased genotypes. Appl. Environ. Microbiol. **76**:60–68.
40. **Venables WN, Ripley BD.** 2002. Modern applied statistics with S, 4th ed. Springer-Verlag, Berlin, Germany.
41. **Vogel HJ, Bonner DM.** 1956. Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. **218**:97–106.
42. **Yang Z, et al.** 2004. Identification of common subpopulations of non-sorbitol-fermenting, β-glucuronidase-negative *Escherichia coli* O157:H7 from bovine production environments and human clinical samples. Appl. Environ. Microbiol. **70**:6846–6854.
43. **Yokoyama E, et al.** 2011. Biased distribution of IS629 among strains in different lineages of enterohemorrhagic *Escherichia coli* serovar O157. Infect. Genet. Evol. **11**:78–82.
44. **Zhang YX, et al.** 2010. Lineage and host source are both correlated with levels of Shiga toxin 2 production by *Escherichia coli* O157:H7 strains. Appl. Environ. Microbiol. **76**:474–482.
45. **Ziebell K, et al.** 2008. Genotypic characterization and prevalence of virulence factors among Canadian *Escherichia coli* O157:H7 strains. Appl. Environ. Microbiol. **74**:4314–4323.
46. **Zotta T, Ricciardi A, Ciocia F, Rossano R, Parente E.** 2008. Diversity of stress responses in dairy thermophilic streptococci. Int. J. Food Microbiol. **124**:34–42.

International Journal of Food Microbiology 157 (2012) 95–101

Contents lists available at SciVerse ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Characteristics of a sharp decrease in *Vibrio parahaemolyticus* infections and seafood contamination in Japan

Yukiko Hara-Kudo [a,*], Shihoko Saito [b], Kayoko Ohtsuka [c], Shogo Yamasaki [d], Shunsuke Yahiro [e], Tomohiro Nishio [f], Yoshito Iwade [g], Yoshimitsu Otomo [h], Hirotaka Konuma [i], Hiroyuki Tanaka [j], Hiroshi Nakagawa [k], Kanji Sugiyama [f], Yoshiko Sugita-Konishi [a], Susumu Kumagai [l]

[a] *Division of Microbiology, National Institute of Health Sciences, Tokyo, Japan*
[b] *Akita Prefectural Institute of Public Health, Akita, Japan*
[c] *Saitama Institute of Public Health, Saitama, Japan*
[d] *Nagasaki Prefectural Institute of Environmental Sciences and Public Health, Omura, Japan*
[e] *Kumamoto Prefectural Institute of Public Health and Environmental Science, Udo, Japan*
[f] *Shizuoka Institute of Environment and Hygiene, Shizuoka, Japan*
[g] *Mie Prefectural Institute of Public Health and Environment, Yokkaichi, Japan*
[h] *Hirosaki University, Hirosaki, Japan*
[i] *Tokai University, Shizuoka, Japan*
[j] *Japan Food Research Laboratories, Tokyo, Japan*
[k] *BML Food Science Solutions, Tokyo, Japan*
[l] *Research Centre for Food Safety, University of Tokyo, Tokyo, Japan*

ARTICLE INFO

Article history:
Received 10 August 2011
Received in revised form 30 March 2012
Accepted 22 April 2012
Available online 1 May 2012

Keywords:
Vibrio parahaemolyticus
Infection
Epidemic
Seafood
Contamination
Sharp decrease

ABSTRACT

Vibrio parahaemolyticus has been one of the most important foodborne pathogens in Japan since the 1960s, and a large epidemic was caused by the pandemic serotype O3:K6 from 1997 to 2001. *V. parahaemolyticus* infections, however, have sharply declined since that time.

Data on serotypes isolated from 977 outbreaks were collected and analysed. Total and pathogenic, thermostable direct hemolysin (TDH) gene-positive *V. parahaemolyticus* were qualitatively and quantitatively detected in 842 seafood samples from wholesale markets in 2007–2009. Strains isolated from patients and seafood were analysed by serotyping, *tdh*-PCR, group-specific PCR for pandemic strains, and pulsed-field gel electrophoresis (PFGE). The sharp decrease in the infections from 1999 onwards was noted not only for O3:K6 infections but also for other serotypes. The change in the seafood contamination situation from 2001 to 2007–2009 was characterised by a decrease to three-fourths in the frequency of *tdh*-positive samples, although that decrease was small compared to the 18-fold decrease in the cases of *V. parahaemolyticus* outbreaks. PFGE detected the pandemic O3:K6 serotype in the same profile in seafood and patients from 1998 to the present. Because of no large decrease in seafood contamination by *V. parahaemolyticus* from the production to distribution stages and the presence of pandemic O3:K6 serotype in seafood to the present, it was suggested that the change of seafood contamination was unrelated to the sharp decrease in *V. parahaemolyticus* infections. *V. parahaemolyticus* infections might be prevented at the stages after the distribution stage.

1. Introduction

Vibrio parahaemolyticus is a halophilic bacterium and a major foodborne pathogen. The first *V. parahaemolyticus* outbreak in Japan, in 1950, was associated with semidried sardines. Strains carrying the *tdh* gene, which encodes the thermostable direct hemolysin (TDH); the *trh* gene, which encodes the TDH-related hemolysin (TRH); or both genes are considered virulent strains (Nishibuchi and Kaper, 1995). A new clone, the O3:K6 strain, which carries the *tdh* but not the *trh* gene, was responsible for a pandemic that was widespread over regions of Asia (Bag et al., 1999; Chiou et al., 2000; National Institute of Infectious Diseases and Tuberculosis and Infectious Diseases Control Division, Ministry of Health, Labour and Welfare, 1999; Vuddhakul et al., 2000; Wong et al., 2000), North America (Gendel et al., 2001), South America (González-Escalona et al., 2005), and Europe (Martinez-Urtaza et al., 2005; Ottaviani et al., 2008) beginning in 1996. The pandemic clone complex consisted of a number of serotypes such as O4:K68, O1:K25 and O1:KUT (Nair et al., 2007). *V. parahaemolyticus* has been a dominant cause of foodborne infections in Japan since the 1960s and the infections

* Corresponding author at: Division of Microbiology, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. Tel./fax: +81 3 3700 9497.
E-mail address: ykudo@nihs.go.jp (Y. Hara-Kudo).

0168-1605/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijfoodmicro.2012.04.019

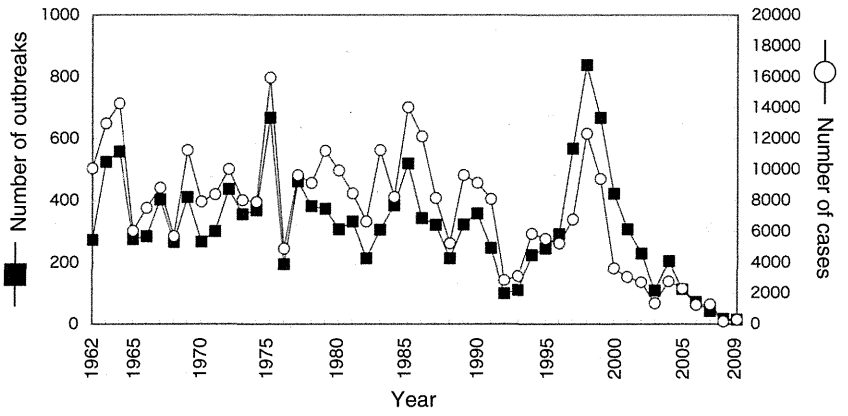

Fig. 1. Annual incidence of *V. parahaemolyticus* infections from 1962 to 2009 reported by national surveillance in Japan.This figure was prepared from data published in the annual Handbook of Health and Welfare Statistics from the Ministry of Health, Labour and Welfare in Japan.

occurred constantly (Fig. 1, this figure was prepared from data published in the annual Handbook of Health and Welfare Statistics from the Ministry of Health, Labour and Welfare in Japan). Surveillance in a prefecture started in 1963 reported that the major serotypes were O2:K3, O3:K7, O4:K8, O4:K12, O5:K15 and O4:K63, and other serotypes accounted for more than half of the infections for 22 years (Akahane et al., 1984). However, proportion of serotype O3:K6 rapidly increased from 1996, and serotype O3:K6 was the most major serotype from 1997 to 1999 (Fig. 2, this figure was prepared from the reports of Infectious Disease Surveillance Center, National Institute of Infectious Diseases in Japan). The increase was coincident a large wave of infections involving *V. parahaemolyticus* from 1997 in Japan (Fig. 1). The number of cases per year doubled (12,318 cases) and that of outbreaks per year tripled (839 outbreaks)

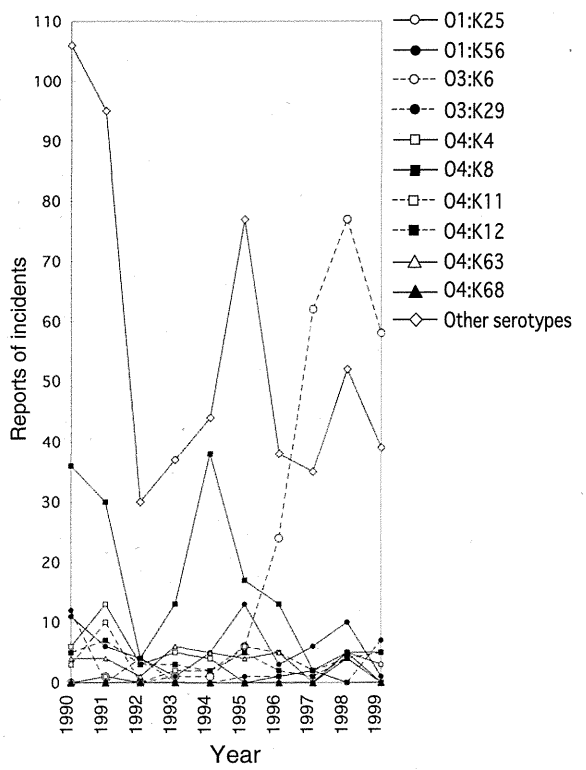

Fig. 2. Serotypes of *V. parahaemolyticus* isolated from patients in Japan from 1990 to 1999.The public health institutes of local governments (Tokyo metropolis, 46 prefectures, and 20 major cities) in Japan submit reports on isolations from patients to the Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Japan. This figure was prepared from those reports.

to a peak in 1998 compared with outbreaks in 1996 (Fig. 1). However, the numbers of cases and outbreaks of *V. parahaemolyticus* infections decreased 40-fold and 60-fold, respectively, from 1998 to 2009 (Fig. 1). Such an enormous decrease has never been previously recorded for *V. parahaemolyticus* infections.

In 1999 and 2000, the Japanese Ministry of Health, Labour, and Welfare (MHLW) advised seafood handlers to use disinfected or artificial seawater or potable water to process shellfish and finfish. The maximum temperature for seafood during distribution and storage was set at ≤10 °C and recommended to be ≤4 °C. The microbiological standards of *V. parahaemolyticus* levels in seafood were set at <100 most probable number per gram (MPN/g) and non-detectable levels per 25 g for raw consumption and ready-to-eat boiled seafood, respectively. It was also advised for consumers to consume seafood within 2 h after taking seafood from the refrigerator. Although these hygienic control measures were the first regulations to control the cooling down of boiled seafood with hygienic water and temperature for storage, and limit *V. parahaemolyticus* contamination level, the contributions of hygienic control measures on seafood are unknown.

To delineate the characteristics of the *V. parahaemolyticus* infection epidemic and the changes in the pathogen's incidence in seafood in Japan during the drastic decrease of infections, we analysed serotypes and genotypes of *V. parahaemolyticus* isolated from outbreaks and seafood. We also investigated the frequencies and levels of total *V. parahaemolyticus* and *tdh*-positive *V. parahaemolyticus* in seafood collected from wholesale markets in 2007–2009 in Japan and compared these values with those reported in 2001 (Hara-Kudo et al., 2003).

2. Materials and methods

2.1. Information on V. parahaemolyticus outbreaks

Questionnaire on *V. parahaemolyticus* outbreaks spanning 1998 to 2007; date of occurrence, place of occurrence, number of cases, implicated food, and information of *V. parahaemolyticus* isolates (serotype and *tdh* and/or *trh* genes harboring), was sent from the National Institute of Health Sciences to 30 prefectures/cities in Japan. The answers were sent to the National Institute of Health Sciences. Information on *V. parahaemolyticus* serotypes from 977 outbreaks was collected. When different serotypes were isolated from the same outbreak, each serotype was counted as an outbreak.

2.2. Analysis of total and tdh-positive V. parahaemolyticus in seafood

A total of 842 seafood samples belonging to domestic and imported (from China, Korea and Russia) bloody clams (*Scapharca broughtonii*), domestic hen clams (*Mactra sulcatari*), domestic rock oysters (*Crassostrea nippona*), domestic scallops (*Patinopecten*

yessoensis), domestic and imported (from China and Korea) short-neck clams (*Tapes japonica*), and 12 other species of domestic molluscan shellfish and domestic horse mackerel (*Trachurus japonicus*) were purchased from fresh seafood markets in various regions in Japan from May to November in 2007–2009. Meats and mantle fluid were aseptically removed from bivalves. The entire horse mackerel, including viscera, was analysed. The same qualitative and quantitative analytical methods used in our previous study were used in the current study (Hara-Kudo et al., 2003) to facilitate the comparison of results. In the quantitative analysis with an MPN method, a 25-g portion of test sample with 225 ml of alkaline peptone water (APW; Nissui Co., Tokyo, Japan) in a stomacher bag was then gently homogenised with hands. Either 10 ml, 1 ml, or 0.1 ml of the homogenate and 1 ml or 0.1 ml of a 10^{-2} dilution of the homogenate were inoculated into 10 ml of APW in triplicate and incubated at 35 to 37 °C for 18 h. One-millilitre portions of each APW culture were inoculated into 10 ml of salt polymyxin broth (SPB; Nissui Co.) and incubated for 18 h at 35 to 37 °C, and then 1 ml of the SPB culture was inoculated into 10 ml of fresh SPB and incubated for 6 h at 35 to 37 °C. A portion of the third enrichment culture in SPB was plated onto CHROMagar Vibrio agar (CHROMagar, Paris, France) for detection of total *V. parahaemolyticus* by isolation and biochemical identification. Another 1-ml portion of the third enrichment culture served as the material for PCR analysis for detection of *tdh*-positive *V. parahaemolyticus* (Tada et al., 1992). Immunomagnetic separation of K6 strains was also performed to effectively isolate TDH-producing *V. parahaemolyticus* O3:K6 (Tomoyasu, 1992). Colonies presumptively identified as *V. parahaemolyticus* were analysed for biochemical and halophilic characteristics, and *V. parahaemolyticus* species-specific sequences of the *toxR* gene (Kim et al., 1999). *Tdh*-positive isolates were tested for TDH production with a reversed passive latex agglutination assay kit (KAP-RPLA; Denka Seiken, Japan). Isolates were analysed with a group-specific PCR (GS-PCR) assay for pandemic strains (Matsumoto et al., 2000) and serotyping kits (K Serum Set and Group O Serum Set, Denka Seiken).

2.3. PFGE analysis

Twenty-four strains of serotype O3:K6 isolated from patient specimens during *V. parahaemolyticus* outbreaks in Japan from 1997 to 2009 were collected. Using PFGE with *Not*I and *Sfi*I restriction enzymes (Parsons et al., 2007), these strains were analysed together with 9 strains isolated from seafood in our previous study (Hara-Kudo et al., 2003) in 2001, and 11 strains isolated from seafood collected for this study in 2008 and 2009. PFGE was performed with a gradient of 6 V/cm at 14 °C using a CHEF-DR II apparatus (Bio-Rad Laboratories, CA). Plugs digested by *Not*I and *Sfi*I were run on two-block programme (Block 1: 4–8 s for 11 h, Block 2: 8–50 s for 9 h) and a single block programme (10–35 s for 18 h), respectively. A dendrogram was constructed from the PFGE profiles according to a Jaccard similarity coefficient with UPGMA clustering analysis using Fingerprinting II (Bio-Rad). Hundred percent concordances were considered to indicate identical patterns. *Sfi*I profiles were analysed for strains producing the same pattern from *Not*I restriction profiles.

2.4. Statistical analysis

The densities determined using the MPN of the total *V. parahaemolyticus* and *tdh*-positive *V. parahaemolyticus* as well as the frequencies of *tdh*-positive samples among domestic hen clams, domestic and imported bloody clams, and domestic and imported short-neck clams were statistically analysed with Student's *t*-test or the χ^2 test. The number of total *V. parahaemolyticus* organisms was compared between *tdh*-positive and *tdh*-negative samples in all samples, using the same methods.

3. Results

3.1. V. parahaemolyticus serotype in outbreaks

The data that we collected since 1998 on *V. parahaemolyticus* serotypes involved in outbreaks identified serotype O3:K6 in almost 50% of the total outbreaks during each year (Fig. 3). Serotypes O1:K25, O1:K56, O4:K8, and O4:K68 were relatively dominant presences among more than 60 types of serotypes. However, various serotypes were associated with the outbreaks, and there was a decrease in the number of outbreaks independent of the serotype.

3.2. Analysis of total and tdh-positive V. parahaemolyticus in seafood

V. parahaemolyticus was detected in 717 of 842 samples (85.2%; Table 1) in 2007–2009. In short-neck clams, more than 95% of both of domestic and imported samples was positive for *V. parahaemolyticus.* In scallop, *V. parahaemolyticus* was detected in less than 25% of samples. The frequencies of total *V. parahaemolyticus* in 2007–2009 in each kind of seafood and total samples were similar to those in 2001.

In total, 65 of 842 samples (7.7%) were positive for *tdh*. The highest frequency of *tdh*-positive samples was detected in imported bloody clams (18.1%; Table 1), and this frequency was significantly greater than that in domestic hen clams, domestic bloody clams, and domestic short-neck clams ($p<0.05$) and in horse mackerel ($p<0.0001$). In imported bloody clams, contamination was more prevalent in samples from Korea (26.3%) than in those from China (13.7%) and Russia (0%). In short-neck clams, four of 24 samples (16.7%) from Korea were positive for *tdh*. The prevalence in horse mackerel was significantly lower than that in hen clams ($p<0.05$), domestic and imported short-neck clams ($p<0.05$), and imported bloody clams ($p<0.05$).

In hen clams and horse mackerel, there were no significant differences between the frequencies of *tdh*-positive samples from 2001 and those from 2007 to 2009 (Table 1). However, a significant change was observed in domestic short-neck clams, with the original frequency of 22.2% significantly decreasing to 9.4% ($p<0.05$). Among the total samples, *tdh*-positive *V. parahaemolyticus* was detected in 65 of 842 samples (7.7%) in 2007–2009, and the frequency was lower than in 2001 (10%).

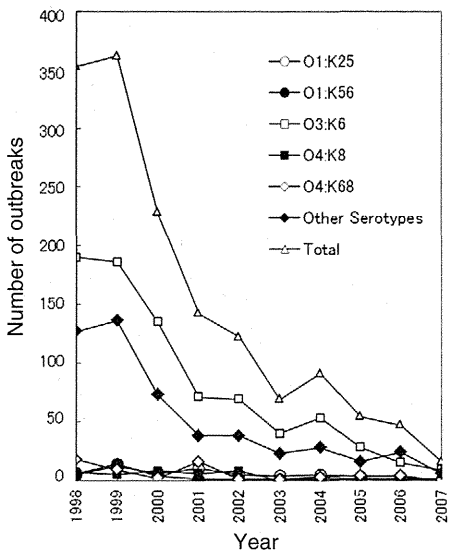

Fig. 3. Serotypes of *V. parahaemolyticus* isolates from outbreaks in Japan from 1998 to 2007.Data on serotypes isolated from outbreaks in 30 local governments in Japan were collected in this study. When different serotypes were isolated from the same outbreak, each serotype was counted as an outbreak.

Table 1
Comparison of detection of *tdh*-positive *V. parahaemolyticus* in seafood samples in Japan in 2001 and 2007–2009.

Seafood	2001[a]					2007–2009						
	tdh-positive *V. parahaemolyticus*				Total *V. parahaemolyticus*				*tdh*-positive *V. parahaemolyticus*			Total *V. parahaemolyticus*
	No. of samples tested	No. of samples *tdh*-positive[b]	No. of samples *tdh*-positive *V. parahaemolyticus* isolated	Serotype of *tdh*-positive isolate (no. of samples isolated)	No. of samples *V. parahaemolyticus* isolated/no. of samples tested	Domestic or imported	Year	No. of samples tested	No. of samples *tdh*-positive [b]	No. of samples *tdh*-positive *V. parahaemolyticus* isolated [c]	Serotype of *tdh*-positive isolate (no. of samples isolated)	No. of samples *V. parahaemolyticus* isolated/no. of samples tested
Hen-clam	128	16 (12.5%)	3 (2.3%)	O3:K6 (3)	100/104 (96.2%)	Domestic	2007, 2009	101	9 (8.9%)[A, E]	4 (4.0%)	O4:K37 (4), OUK:KUT (1)	82/101 (81.2%)
Short-neck clam	36	8 (22.2%)[F]	2 (5.6%)	O3:K6 (2)	34/34 (100%)	Domestic	2007–2009	256	24 (9.4%)[A, E, G]	4 (1.6%)	O3:K6 (2), O4:K9 (1), O5:K17 (1), O10:K52 (1)	244/256 (95.3%)
						Imported	2009	29	4 (13.8%)[E]	1 (3.4%)	O10:KUT (1)	28/29 (96.6%)
							Subtotal	285	28 (9.8%)	5 (1.8%)		272/285 (95.4%)
Bloody clam	NT	–	–	–	–	Domestic	2009	37	2 (5.4%)[A]	0	–	32/37 (86.5%)
						Imported	2009	94	17 (18.1%)[B, E]	6 (6.4%)	O1:KUT (1), O3:K6 (4), O3:K7 (1), O5:K17 (1), O8:K21(1)	78/94 (83.0%)
							Subtotal	131	19 (14.5%)	6 (4.5%)		110/131 (84.0%)
Rock oyster	78	7 (9.0%)	6 (7.7%)	O3:K6 (6)	1/1 (100%)	Domestic	2007	40	0	–	–	32/40 (80.0%)
Scallop	29	0 (0%)	–	–	–	Domestic	2007	37	0	–	–	9/37 (24.3%)
Other molluscan shellfish	41	2 (4.9%)	0	–	17/19 (89.5%)	Domestic	2007	42	3 (7.1%)	0	–	34/42 (81.0%)
Horse mackerel	17	0 (0%)	–	–	13/15 (86.7%)	Domestic	2008	206	6 (2.9%)[C, D]	3 (1.5%)	O4:KUT (1), O10:K52 (2)	178/206 (86.4%)
Total	329	33 (10.0%)	11 (3.3%)		165/173 (95.4%)			842	65 (7.7%)	18 (2.5%)		717/842 (85.2%)

ND: not detected.
[a] Samples were domestic (reference no. 11).
[b] Significant differences: A vs B, D vs E and F vs G ($p<0.05$), B vs C ($p<0.0001$).
[c] Different serotypes were isolated in three samples. O3:K6 was positive by GS-PCR while the other serotypes were negative. O1:KUT, O5:K17, O10:KUT and OUT:KUT were *trh*-positive.

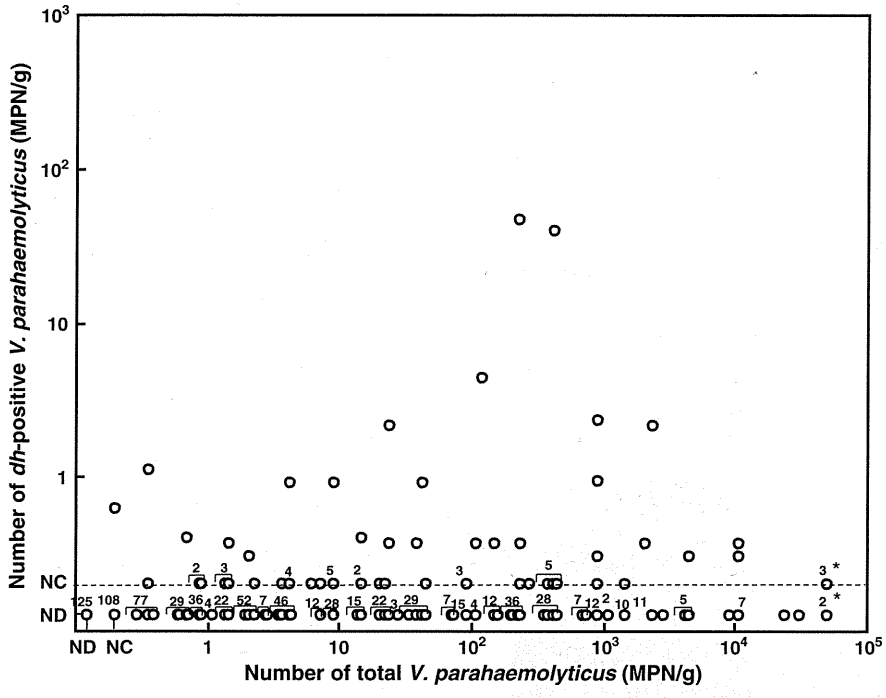

Fig. 4. Plots of the total number of *V. parahaemolyticus* organisms and the number of *tdh*-positive *V. parahaemolyticus* organisms in seafood.Each circle indicates the number in each sample. ND, *tdh* gene not detected by qualitative analysis (<1 CFU/25 g), NC, *tdh* gene not detected by quantitative analysis (<0.3 MPN/g) but detected by qualitative analysis (>1 CFU/25 g). *Total number of *V. parahaemolyticus* is >14,000 MPN/g. The frequencies of *tdh*-positive findings in samples exceeding the level of total *V. parahaemolyticus* at 100 MPN/g: 26/166 (15.7%), the frequencies of *tdh*-positive samples below the level: 39/676 (5.8%).

The total level of *V. parahaemolyticus* ranged from <0.3 to >14,000 MPN/g (Fig. 4). The density of *tdh*-positive *V. parahaemolyticus* did not exactly correspond to that of total *V. parahaemolyticus*. Focusing on 100 MPN/g as a regulation level of total *V. parahaemolyticus* for seafood consumed raw in Japan, the frequencies of *tdh*-positive findings in samples exceeding this level (26/166; 15.7%) were approximately 3 times higher than those in *tdh*-positive samples below this level (39/676; 5.8%).

Tdh-positive *V. parahaemolyticus* was detected in 65 samples, and the density ranged from <0.3 to 2.1 MPN/g. The densities of *tdh*-positive *V. parahaemolyticus* in each type of seafood were below the detection limit of 0.04 CFU/g (1 CFU/25 g) in the quantitative analysis. Among all samples, the level of total *V. parahaemolyticus* in *tdh*-positive samples (average: 13,369.3 MPN/g) was significantly greater than that in *tdh*-negative samples (average: 4017.9 MPN/g) ($p<0.01$).

3.3. TDH-producing *V. parahaemolyticus* isolates from seafood

TDH-producing *V. parahaemolyticus* was isolated from 18 of 65 *tdh*-positive samples (Table 1); serotype O3:K6 was isolated from two samples of domestic short-neck clam, a sample of bloody clam from Korea, and three samples of bloody clam from China. The frequency of serotype O3:K6 contamination was 0.7% (6 of 842 samples). Isolates of serotypes O3:K7 and O8:K21 were detected in a sample of bloody clam from Korea, from which the O3:K6 strain was isolated. Both *tdh* and *trh*-PCR assays recorded positive results for O1:KUT, O5:K17, O10:K52 and O10:KUT. The GS-PCR assay was positive for all O3:K6 serotype isolates and negative for the other isolates. Serotypes O1:K25, O1:K56, O4:K8 and O4:K68 were not isolated from seafood though the serotypes were isolated from patients (Fig. 3).

3.4. PFGE analysis

The profiles of serotype O3:K6 were classified into 18 groups, numbered 1 to 18, and 15 groups, designated "a" to "o," based on the *Not*I and *Sfi*I restriction profiles, respectively (Fig. 5). Four combinations of *Not*I and *Sfi*I restriction profiles were identical among strains isolated from seafood and patients. Strains isolated from domestic rock oysters in 2001 and bloody clams from Korea in 2009 corresponded to the strains isolated from patients in 1998, 1999, 2005 and 2009 (*Not*I: 4; *Sfi*I: b, Fig. 5). The pattern of *Not*I: 5; *Sfi*I: b (Fig. 5) was characteristic of strains isolated from domestic rock oysters in 2001, domestic short-neck clams in 2008, and patients in 2007. Strains isolated from hen clams in 2001 were in conformity with strains isolated from patients in 2006 (*Not*I: 5; *Sfi*I: g, Fig. 5). A strain isolated from domestic short-neck clams in 2008 (*Not*I: 13; *Sfi*I: h, Fig. 5) corresponded to a strain isolated from patients in 2007.

4. Discussion

V. parahaemolyticus infections in Japan rapidly increased from 1997 to 1998 and then decreased through to 2009. The surge was linked to increased incidences of the pandemic O3:K6 strain, which has spread worldwide since 1996. From the results of GS-PCR and PFGE analyses, it was demonstrated that serotype O3:K6 pandemic strains isolated from bloody clams imported from Korea in 2009 and patients in 2005, 2007 and 2009 are of the same type as O3:K6 strains isolated from seafood and patients in 1998–2001 (Fig. 5). Moreover, O3:K6 strains isolated from domestic short-neck clams in 2008 and from patients in 2007 were of the same type as strains isolated from seafood in 2001 (Fig. 5). These findings demonstrate that the pandemic O3:K6 strains continued to inhabit areas in or around Japan through 2009.

While there were few infections of the pathogen before in Southern Chile, *V. parahaemolyticus* infections greatly increased with the arrival of the pandemic strain O3:K6 in 1997 (García et al., 2009; Harth et al., 2009). The number of cases reached a peak in 2005, with 3725 cases, and then decreased to one case in 2010 (Cachicas et al., 2011). However, the background seems different from the situation in Japan. Although the epidemic decline of pandemic *V. parahaemolyticus* likely contributed heavily to the