



## OPEN ACCESS

## EDITED BY

Marcin Łoś,  
University of Gdansk, Poland

## REVIEWED BY

Sabine Delannoy,  
Agence Nationale de Sécurité Sanitaire de  
l'Alimentation, de l'Environnement et du  
Travail (ANSES), France  
Joseph M. Bosilevac,  
Agricultural Research Service (USDA),  
United States  
Susan Leonard,  
United States Food and Drug Administration,  
United States

## \*CORRESPONDENCE

Keiji Nakamura  
✉ nakamura.keiji.046@m.kyushu-u.ac.jp

RECEIVED 24 June 2024

ACCEPTED 25 July 2024

PUBLISHED 06 August 2024

## CITATION

Nakamura K, Taniguchi I, Gotoh Y, Isobe J, Kimata K, Igawa Y, Kitahashi T, Takahashi Y, Nomoto R, Iwabuchi K, Morimoto Y, Iyoda S and Hayashi T (2024) Diversity of Shiga toxin transducing phages in *Escherichia coli* O145:H28 and the different Shiga toxin 2 production levels associated with short- or long-tailed phages.  
*Front. Microbiol.* 15:1453887.  
doi: 10.3389/fmicb.2024.1453887

## COPYRIGHT

© 2024 Nakamura, Taniguchi, Gotoh, Isobe, Kimata, Igawa, Kitahashi, Takahashi, Nomoto, Iwabuchi, Morimoto, Iyoda and Hayashi. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Diversity of Shiga toxin transducing phages in *Escherichia coli* O145:H28 and the different Shiga toxin 2 production levels associated with short- or long-tailed phages

Keiji Nakamura<sup>1\*</sup>, Itsuki Taniguchi<sup>1</sup>, Yasuhiro Gotoh<sup>1</sup>, Junko Isobe<sup>2</sup>, Keiko Kimata<sup>2</sup>, Yukiko Igawa<sup>3</sup>, Tomoko Kitahashi<sup>4</sup>, Yohei Takahashi<sup>5</sup>, Ryohei Nomoto<sup>6</sup>, Kaori Iwabuchi<sup>7</sup>, Yo Morimoto<sup>8</sup>, Sunao Iyoda<sup>9</sup> and Tetsuya Hayashi<sup>1</sup>

<sup>1</sup>Department of Bacteriology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan, <sup>2</sup>Toyama Institute of Health, Imizu, Japan, <sup>3</sup>Nagano Prefecture Suwa Public Health and Welfare Office, Suwa, Japan, <sup>4</sup>Chiba City Institute of Health and Environment, Chiba, Japan, <sup>5</sup>Aomori Prefectural Institute of Health, Aomori, Japan, <sup>6</sup>Kobe Institute of Health, Kobe, Japan, <sup>7</sup>Iwate Prefectural Research Institute for Environmental Sciences and Public Health, Morioka, Japan, <sup>8</sup>Hokkaido Institute of Public Health, Sapporo, Japan, <sup>9</sup>National Institute of Infectious Diseases, Tokyo, Japan

Shiga toxin (Stx)-producing *Escherichia coli* (STEC) causes serious gastrointestinal illness, including hemorrhagic colitis and hemolytic uremic syndrome. Two types of Stxs (Stx1 and Stx2) are known and both are encoded by bacteriophages (Stx phages), but the production of Stx2 is known to be a major risk factor for severe STEC infections. The production of Stx2, but not Stx1, is tightly coupled with the induction of Stx phages, and Stx2 production levels vary between STEC strains even within the same serotype. Here, we analyzed the genomic diversity of all Stx phages in 71 strains representing the entire O145:H28 lineage, one of the often highly pathogenic STECs, and the relationship between the variations in Stx phage genomes and the levels of Stx2 production by host strains. Our analysis reveals highly dynamic natures of Stx phages in O145:H28, including the independent acquisition of similar Stx phages by different sublineages, the recent transfer of Stx phage between different sublineages, and the frequent gain and loss of Stx phages in some sublineages. We also show the association of the Stx2 phage types with the Stx2 production levels of host strains: strains carrying short-tailed Stx2 phages exhibited significantly higher Stx2 production levels than those carrying long-tailed Stx2 phages. Detailed analyses of the Stx2 phage genomes revealed that both of short- and long-tailed phages exhibited sequence diversification and they were divided into two groups, respectively, based on the sequence similarity of the phage early region encoding genes responsible for phage induction, short-tailed phages contained early regions clearly different in genetic organization from those in long-tailed phages. Therefore, the variations in the early regions between short- and long-tailed Stx2 phages appeared to be linked to a striking difference in Stx2 production levels in their host strains. These results broaden our understanding of the diversification and dynamism of Stx phages in O145:H28 and the association of Stx2 phage types with the Stx2 production level in this STEC lineage.

## KEYWORDS

Shiga toxin-producing *Escherichia coli*, Stx phage, Stx2 production, O145:H28, comparative genomics

## Introduction

Shiga toxins (Stxs) are the key virulence factors of Stx-producing *Escherichia coli* (STEC), which causes diarrhea and hemorrhagic colitis with life-threatening complications, such as hemolytic uremic syndrome. Stxs are classified as Stx1 or Stx2, each of which include several subtypes (Stx1a, Stx1c-Stx1e; Stx2a-Stx2l) (Scheutz et al., 2012; Koutsoumanis et al., 2020). While STEC strains produce one or more Stx subtypes (Ogura et al., 2007; Steyert et al., 2012), epidemiological studies suggest that Stx2-producing strains cause more severe STEC infections than strains producing only Stx1 (Fuller et al., 2011; Bielaszewska et al., 2013).

The *stx* genes are encoded by bacteriophages (Stx phages), and STEC strains acquire these genes via the lysogenization of Stx phages. Although the integration sites and genome sequences of Stx phages are highly variable even within the same serotype (Delannoy et al., 2017; Yara et al., 2020; Miyata et al., 2023; Yano et al., 2023), Stx phages are morphologically divided into two types based on their tail structures, which are defined by late genes: lambda-like long-tailed phages (L-phages) and short-tailed phages (S-phages) similar to the Stx2a phages of O157:H7, such as Sp5 and 933W (Plunkett et al., 1999; Teel et al., 2002; Muniesa et al., 2004; Strauch et al., 2004; Asadulghani et al., 2009; García-Aljaro et al., 2009; Bonanno et al., 2016; Pinto et al., 2021). Of 279 publicly available Stx phages, 99% can be classified into either type based on genetic content (Pinto et al., 2021). Both types of Stx phages encode several lambda-like regulator genes that modulate early and late gene expression, such as the *cl* and *q* genes (Neely and Friedman, 1998; Makino et al., 1999; Plunkett et al., 1999; Strauch et al., 2008).

The *stx* genes are located downstream of the late gene promoter (Plunkett et al., 1999; Yokoyama et al., 2000; Shi and Friedman, 2001; Strauch et al., 2008). The expression of *stx1* is primarily under the control of the iron-regulated authentic promoter (Calderwood and Mekalanos, 1987), although prophage induction-dependent production of Stx1 has been described in some O157:H7, O26:H11, and O48:H21 strains (Wagner et al., 2002; Steyert et al., 2012; Berger et al., 2019; Yano et al., 2023). In contrast, the expression of *stx2* depends on the late promoter (Wagner et al., 2001; Tyler et al., 2004) and Stx2 production is tightly coupled with phage induction; thus, variations in Stx2 phage genomes can affect the amount of Stx2 production by each strain, as has been shown in several STEC lineages (Ogura et al., 2015; Nishida et al., 2021; Miyata et al., 2023; Nakamura et al., 2023). For example, the variation in Stx2 production levels in O157:H7 STEC strains was associated with the subtypes of Stx2a phages (all are S-phages) as defined by their early regions (Ogura et al., 2015). In particular, in O157:H7 clade 8, a highly pathogenic lineage of O157:H7 STEC, the  $\gamma$  subtype of the Stx2a phage confers increased Stx2 production and pathogenicity to host strains than do other clade 8 strains, which carry the Stx2a phages belonging to the  $\delta$  subtype (Miyata et al., 2023).

O145:H28 is one of the often highly pathogenic STECs (Karmali et al., 2003; Valilis et al., 2018). We previously analyzed the whole-genome sequences (WGSs) of 239 O145:H28 strains, including a systematic

analysis of the prophages in seven finished genomes, and revealed notable variations in the sequences and integration sites of Stx phages among O145:H28 strains (Nakamura et al., 2020). In that study, we found that although the distribution of *stx1a* genes was biased toward specific clades, the *stx2a* genes were widely but variably distributed throughout the entire O145:H28 lineage. However, the precise variation in Stx phages and its impact on Stx production by host strains of this lineage have not been elucidated. In the present study, we performed a systematic analysis of Stx phages of STEC O145:H28 and analyzed the diversity and dynamism of Stx phages in O145:H28 and the association of Stx2 phage types with Stx2 production by host strains.

## Materials and methods

### Bacterial strains

The initial O145:H28 strain set included 59 strains that were available in our laboratory and sequenced in our previous study (Nakamura et al., 2020) and 18 strains with complete genome sequences (the plasmid genome was not available for strain 2015C-3125). The genome sequences of the 18 previously reported strains (Cooper et al., 2014; Patel et al., 2018; Tyson et al., 2019; Carter et al., 2021) were downloaded from NCBI. Six of the 18 strains were excluded from the strain set because their recombination-free core genome sequences were identical to those of other strains. Therefore, in this study, 71 strains were analyzed (see Supplementary Table S1 for the strains included in the final set).

### Determination of complete genome sequences

The genomic DNA of strains 16003 and 12E115 was purified using Genomic-tip 100/G (Qiagen). Libraries for Illumina sequencing (average insert size: 700 bp) were prepared using the NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs) and sequenced using Illumina MiSeq to generate 300 bp paired-end reads. These genomes were additionally sequenced using MinION with R9.4.1 flow cells (Nanopore) for 68 (16003) or 96 h (12E115). MinION reads were trimmed and filtered as described previously (Nishida et al., 2021) and assembled along with the Illumina reads of each strain using Unicycler v0.4.8 (Wick et al., 2017) to obtain the finished genome sequences.

### Phylogenetic analysis

Phylogenetic analysis of the initial strain set ( $n=77$ ) was performed based on the SNPs identified on the prophage/integrative element/IS-free and recombination-free chromosome backbone that was conserved across all genomes by Gubbins (Croucher et al., 2015) and MUMmer (Kurtz et al., 2004) using the genome of strain 10942 as a reference. A maximum likelihood (ML) tree was constructed with RAXML v8.2.12

(Stamatakis, 2006) as previously described (Nakamura et al., 2021) and displayed using FigTree v1.4.4.<sup>1</sup>

## Analyses of integration sites and the sequencing of Stx phages

Stx phages integrated into the *attB* in the *ompW* prophage and *yecE* loci in 54 O145:H28 draft genomes were previously described (Nakamura et al., 2021). Stx phages integrated into *argW*, *wrbA*, and *sbcB* in these 54 genomes were identified by the same strategy as that we previously employed. Briefly, we first examined each site by BLASTN search, and when some phage (or other genetic element) was found to be integrated into the site, long PCR amplification using primers targeting the *stx* genes and sequences adjacent to the integration site was performed to determine whether the integrated phage was Stx phage or not (schematically shown in Supplementary Figure S1; see Supplementary Table S2 for the primers used). The entire genome sequence of each phage was determined by sequencing long PCR products as previously described (Nakamura et al., 2021). DFAST v1.2.18 (Tanizawa et al., 2018) and GenomeMatcher v3.0.2 (Ohtsubo et al., 2008) were used for annotating phage genomes and comparing Stx phage genomes, respectively. Based on the late gene organization, Stx phages that contained late genes similar to those of phage lambda were defined as L-phages and Stx phages that contained late genes similar to those of Sp5 and 933W were defined as S-phages.

## Clustering analyses of Stx phage genomes and sequence analysis of Q proteins

All-to-all phage genome comparison was performed for the three sets of genomes (the entire genomes of 83 Stx phages, the entire genomes of 54 Stx2 phages, and the early regions of 54 Stx2 phages) using Mash v2.0 (Ondov et al., 2016) with default parameters to generate pairwise Mash distance matrices. Based on each matrix, Stx phages and the early regions of Stx phages were clustered with a cutoff of 0.05 as previously described (Nagano et al., 2023). The amino acid sequences of Q proteins encoded by each Stx2 phage were aligned with the Q proteins from phages lambda, Sp5, and 933W by MUSCLE in MEGA v10.1.8 (Kumar et al., 2018), and a dendrogram was generated based on the alignment using the UPGMA algorithm in MEGA v10.1.8.

## Determination of Stx2 production levels

Cell lysates of all tested strains were prepared as described previously (Nishida et al., 2021), except for the final concentration (1.0 µg/mL) of mitomycin C (MMC; Wako Chemicals). The MMC concentration and sampling time were optimized based on the results of exploratory analyses using seven O145:H28 strains (see Supplementary Figure S2 for details). The Stx2 concentration in the lysate of each O145:H28 strain was determined by sandwich ELISA as previously described (Nishida et al., 2021). An unpaired *t* test was performed to compare Stx2 production levels between the O145:H28 strains carrying S-Stx2a phages and those

carrying L-Stx2 phages using Prism 9 software (GraphPad Software).  $p < 0.05$  was considered to indicate statistical significance.

## Results and discussion

### Strain set and Stx phages

For detailed analyses of Stx phages in O145:H28 strains, we selected 64 strains from the 239 strains analyzed in our previous study (Nakamura et al., 2020), 59 of which were sequenced in our laboratory. This set included eight finished and 56 draft genomes and covered seven of the eight clades in ST32 (clades A-H) and the ST137/ST6130 lineage in O145:H28 (ST32 clade D strains were not available in our laboratory). Among the 56 strains, for which only draft genomes were available, two were subjected to Nanopore long-read sequencing to obtain finished sequences by hybrid assembly. Seven genome-finished strains recently deposited in the NCBI database (Tyson et al., 2019; Carter et al., 2021) were also included in the dataset (Supplementary Table S1). Thus, the final set included 71 strains (Figure 1 and Table 1), of which 48 were isolated in Japan and the remainder were isolated in the USA, Belgium, or Canada. Most strains were human isolates, but five bovine and four environmental/food isolates were included. There were five *stx* genotypes, and four strains carried two copies of the *stx2a* gene (Table 1).

As the genome sequences of the Stx phages of 27 strains were already known ( $n = 38$ ), we determined those of the Stx phages ( $n = 45$ ) in the remaining 43 strains. The Stx1a phage in an *stx1a/stx2d*-positive strain (strain 112808, whose Stx2d phage was previously sequenced but Stx1a phage was not) was also sequenced to obtain the full set of genome sequences of Stx phages ( $n = 84$ ) and determine their integration sites (Figure 1 and Supplementary Table S1).

Of the 84 phages, 33 were S-Stx2a phages, and 21 were L-Stx2a phages. The remaining 30 phages encoded other types of Stxs, and all were L-phages (27 Stx1a, two Stx2c, and one Stx2d). For the integration sites, five loci (*argW*, *wrbA*, *attB* in the *ompW* prophage, *yecE*, and *sbcB*) were identified. The L-Stx2a phage was duplicated in strain 112648 and integrated into the *attB* site in the *ompW* prophage (referred to as *attB*<sub>in\_PPompW</sub>) and the *yecE* loci (Nakamura et al., 2021); thus, these phages were considered one L-Stx2a phage. Although three strains (12E129, RM12367-C1, and H27V05) contained two Stx2a phages, they were included as different phages because they showed considerable sequence variation (Supplementary Figure S3); thus, a total of 83 Stx phages were analyzed in subsequent analyses.

### Dynamics of Stx phages in O145:H28 isolates

To analyze the sequence similarities of the 83 Stx phages, we performed all-to-all sequence comparisons using the Mash program (Ondov et al., 2016) and constructed a dendrogram using the complete linkage method based on pairwise Mash distances. The 83 phages were clearly divided into L-phages and S-phages (Figure 2 and Supplementary Table S3), but multiple phage clusters were detected with a threshold of 0.05 in both types of phages (PC1-PC5 in L-phages and PC6 and PC7 in S-phages).

In five phage clusters (PC1, 3, 4, 6, and 7), the encoded *stx* was the same subtype. However, variations in *stx* subtype were found

<sup>1</sup> <http://tree.bio.ed.ac.uk/software/figtree/>

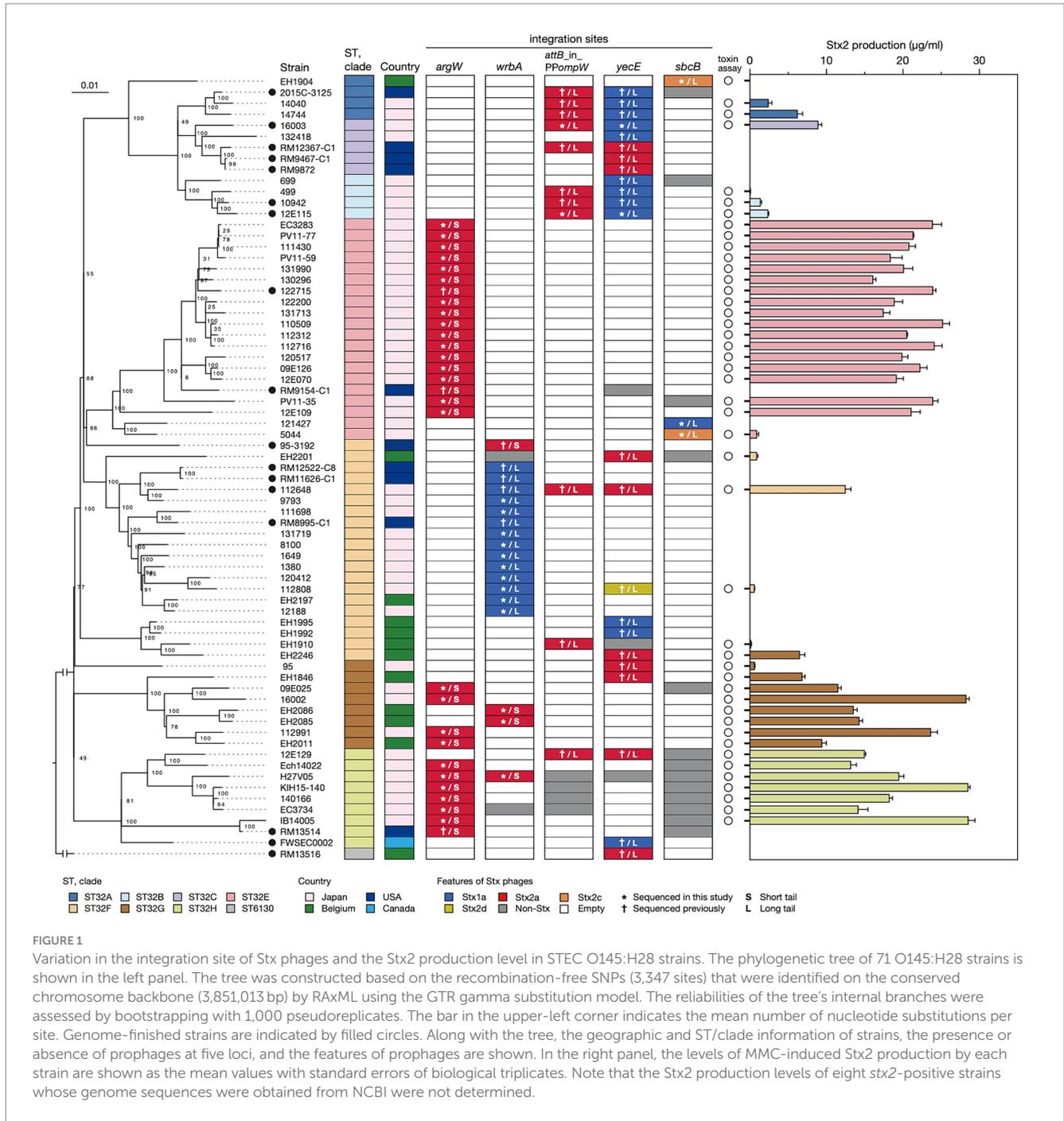
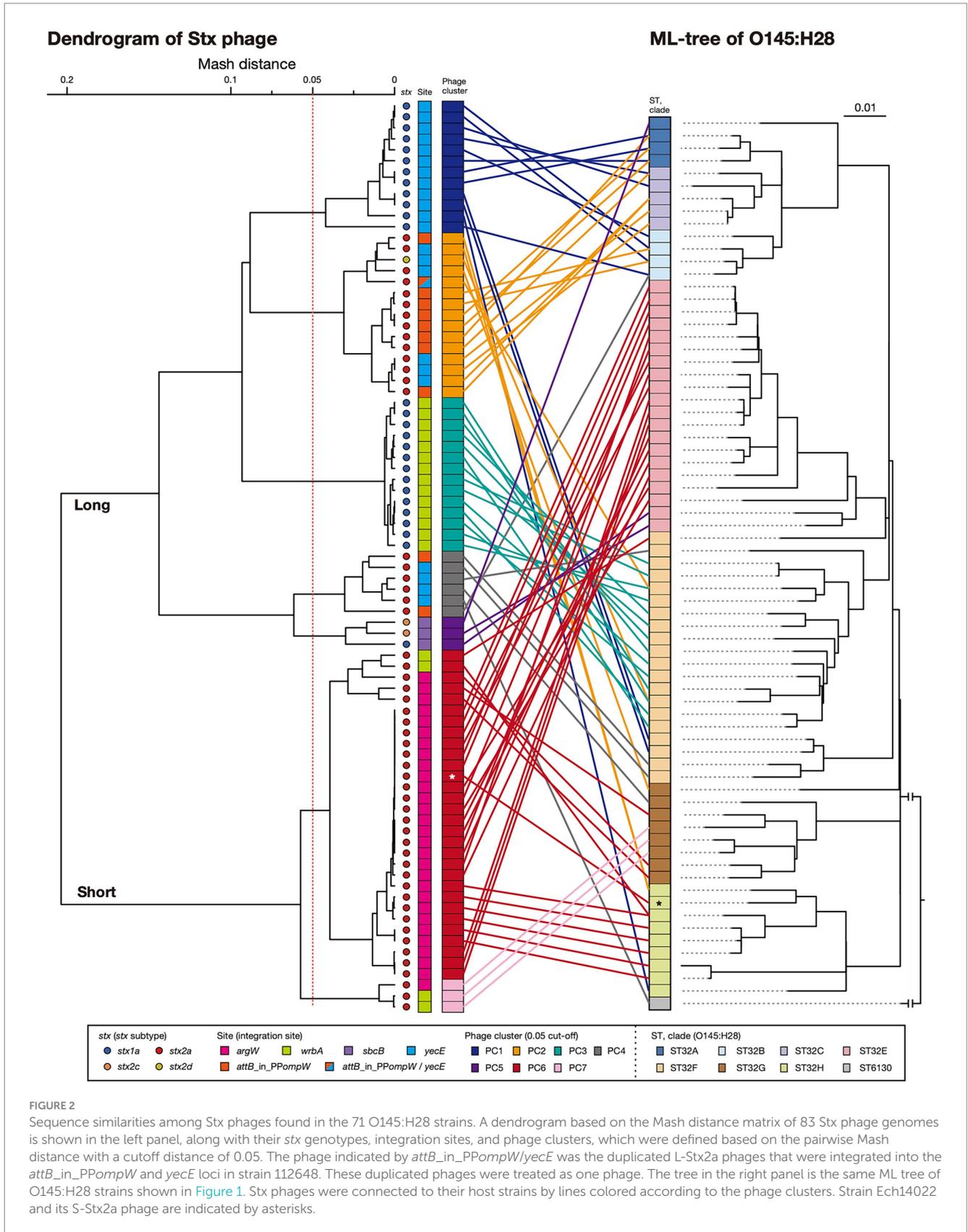


TABLE 1 The stx genotypes of the O145:H28 strain analyzed in this study.

Country	stx1a	stx1a stx2a	stx1a stx2a (x2) <sup>a</sup>	stx1a stx2d	stx2a	stx2a (x2) <sup>a</sup>	stx2c	Total
Japan	11	6	1	1	26	2	1	48
United States	3	1	0	0	5	1	0	10
Belgium	3	0	0	0	8	0	1	12
Canada	1	0	0	0	0	0	0	1
Total	18	7	1	1	39	3	2	71

<sup>a</sup>2 copies of stx2a genes.



in PC2 and PC5 (12 *stx2a* and one *stx2d* in PC2 and one *stx1a* and two *stx2c* in PC5), suggesting the replacement of *stx* in these two phage clusters. While the Stx1a phages at *yecE* and *wrbA* (PC1 and

PC3) and the Stx1a and Stx2c phages at *sbcB* (PC5) formed distinct clusters, the L-Stx2a and S-Stx2a phages were separated into two clusters, respectively (Figure 2). Interestingly, although the L-Stx2a

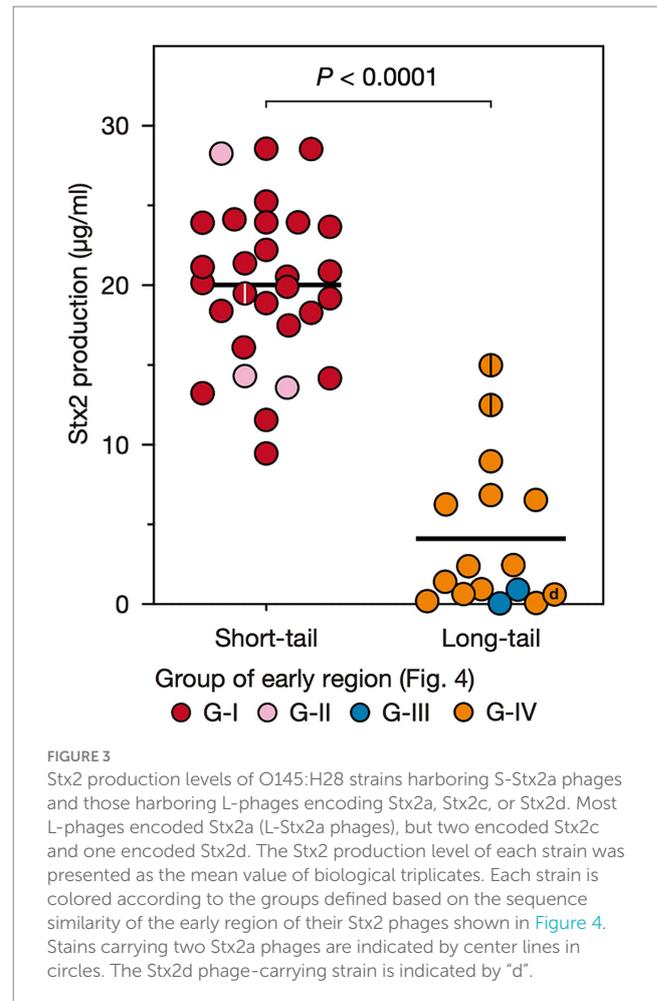
phages were separated into PC2 and PC4, both included phages at *attB\_in\_PPompW* and *yecE*. Similarly, S-Stx2a phages were separated into PC6 and PC7, but both included phages at *argW* and *wrbA*. Variations in PC2 and PC4 can be easily generated because the *attB* sequences in *attB\_in\_PPompW* and *yecE* are essentially the same (Nakamura et al., 2021). In contrast, the variation in PC2 and PC4 was apparently generated by replacement of the integrase gene.

The within-cluster heterogeneity of Stx phages was more evident when the phylogeny of their host strains was considered (Figure 2). While two clusters (PC3 and PC7) were comprised of phages found in the same host clade, respectively, the remaining five clusters included Stx phages found in multiple host clades. For example, PC1 phages were found in strains belonging to clades A, B, C, F, and H, and PC2 phages were found in clades A, B, C, F, G, and H. This finding indicates dynamic changes in Stx phages in each clade. The most striking case was the S-Stx2a phage of strain Ech14022 of clade H (indicated by an asterisk in Figure 2), which had a sequence nearly identical to those of the S-Stx2a phages of clade E strains, suggesting the recent interclade transfer of this phage.

These findings suggest that wide circulation of each phage cluster likely allowed the acquisition of similar phages by host strains with various phylogenetic backgrounds. Repeated acquisition of Stx phages may induce the loss of resident Stx phages, leading to within-host clade variation in Stx phages. Similar changes/variations were observed for Stx1a phages in the ST21 lineage of STEC O26:H11 (Yano et al., 2023) and Stx2a phages in clade 8 of STEC O157:H7 (Miyata et al., 2023). Notably, however, different distribution patterns of Stx phages were also observed in STEC O121:H19 and the STEC belonging to clonal complex 119 (CC119). In these STECs, systematic analyses of Stx phages have been conducted and revealed the stable maintenance of an S-Stx2a phage and an L-Stx2a phage in the major lineage of each STEC, respectively (Nishida et al., 2021; Nakamura et al., 2023).

## Stx2 production was greater in the strains carrying S-Stx2 phages than in the strains carrying L-Stx2 phages

To examine the variation in the level of MMC-induced Stx2 production across O145:H28 strains, we optimized the MMC concentration and sampling time using seven O145:H28 strains (Supplementary Figure S2) and measured the Stx2 concentrations using this condition in the cell lysates of 45 *stx2*-positive strains available in our laboratory (29, 13, 2, and 1 strains carried S-Stx2a, L-Stx2a, L-Stx2c, and L-Stx2d phages, respectively), which covered seven of the eight clades in ST32. As shown in Figure 1 and Supplementary Table S3, the Stx2 production levels were highly variable between the strains (0.06–28.6 µg/mL). Moreover, the comparison of strains carrying S-Stx2a phages and those carrying L-Stx2 phages (including Stx2a, Stx2c, and Stx2d phages) revealed that the former strains produced significantly more Stx2 than the latter strains (19.8 vs. 4.1 µg/mL on average,  $p < 0.0001$ ) (Figure 3). The Stx2 production level of strain H27V05, which carried two S-Stx2a phages, was average (19.5 µg/mL) among the S-Stx2a



phage-carrying strains. However, the two strains (112648 and 12E129) that carried two L-Stx2a phages produced greater amounts of Stx2 (12.6 µg/mL and 15.0 µg/mL, respectively) than the other strains carrying L-Stx2 phages. It has recently been reported that two O113:H21 strains carrying L-Stx2 phages showed a significant difference in Stx2 production level and the strain producing a greater amount of Stx2 carried two L-Stx2 phages (Allué-Guardia et al., 2022). Although this finding may be linked to our finding, it should be noted that these two O113:H21 strains belonged to different *E. coli* lineages and the two strain L-Stx2 phages of the strain showing a higher Stx2 production level encoded the *stx2a* and *stx2d* genes, respectively.

## Variations in the early genes of S- and L-Stx2 phages

Stx2 production levels were clearly different between the strains harboring S- and L-phages. However, as mentioned before, Stx2 production is tightly coupled with phage induction, which is achieved by the expression of early genes and the activation of late gene promoters (Wagner et al., 2001; Tyler et al., 2004). Therefore, to examine the relationship between the variation in Stx2 phages and the Stx2 production levels in host strains in more detail, we performed an additional clustering analysis of the Stx2 phages

( $n = 56$ ; the Stx2 production levels of their host strains were determined) based on the pairwise Mash distances of their early regions (Figure 4). These phages were classified into four groups (referred to as G-I, G-II, G-III, and G-IV) with a threshold of 0.05, the same threshold that was used for the analysis of full-length phage genomes. This grouping correlated well with that based on full-length phage genomes; G-I contained all S-Stx2a phages in PC6, G-II contained all S-Stx2a phages in PC7, and G-III contained both of the two L-Stx2c phages in PC5. However, all L-Stx2a phages and one L-Stx2d phage in PC2 and PC4 were grouped together into G-IV, indicating that they contained similar early regions, which were distinct from those in the S-Stx2a and L-Stx2c phages. The sequence of Q anti-terminator, which is involved in the coupling of phage induction and Stx2 production (Tyler et al., 2004; Pacheco and Sperandio, 2012), also differed between S- and L-Stx2 phages (Supplementary Figure S4). Although the Q proteins of all S-Stx2a phages (G-I and G-II groups) were similar to those of Sp5 and 933 W, the S-Stx2a phages of O157:H7 strains EDL933 and Sakai, the Q proteins of L-Stx2c phages (G-III) and those of the L-Stx2a and L-Stx2d phages (G-IV group) formed distinctive clusters except for a unique Q protein of an L-Stx2a phage in the G-IV group.

Consistent with the result of the Mash distance-based clustering analysis, the early regions exhibited distinct genetic organizations between four groups and overall sequence similarities within each group (Figure 4). However, in the G-I and G-IV groups, several genes including the *cl* gene exhibited notable within-group variations. Therefore, although the Stx2 phages of O145:H28 strains were classified into S- and L-phages, both were further divided into several groups based on the sequences of the early genes, which may have some impact on the Stx2 production levels in host strains. Similar findings related to the association of increased Stx2 production with the carriage of Stx2a phages with specific types of early regions have been reported for the S-Stx2a phages of O157:H7 STEC (Ogura et al., 2015; Miyata et al., 2023). Notably, of the L-Stx2a phages that were found in the two strains (112648 and 12E129) that carried two L-Stx2a phages and showed higher Stx2 production levels than other L-Stx2 phage-containing strains, the duplicated L-Stx2a phage in strain 112648 and one of the L-Stx2a phages in strain 12E129 formed a distinct subgroup in G-IV (Figure 4). Although the effect of the copy number of *stx* and the difference in host genetic background (ST32F and ST32H) should be considered, there is a possibility that the acquisition of these phages contributes to the increased Stx2 production level in these strains.

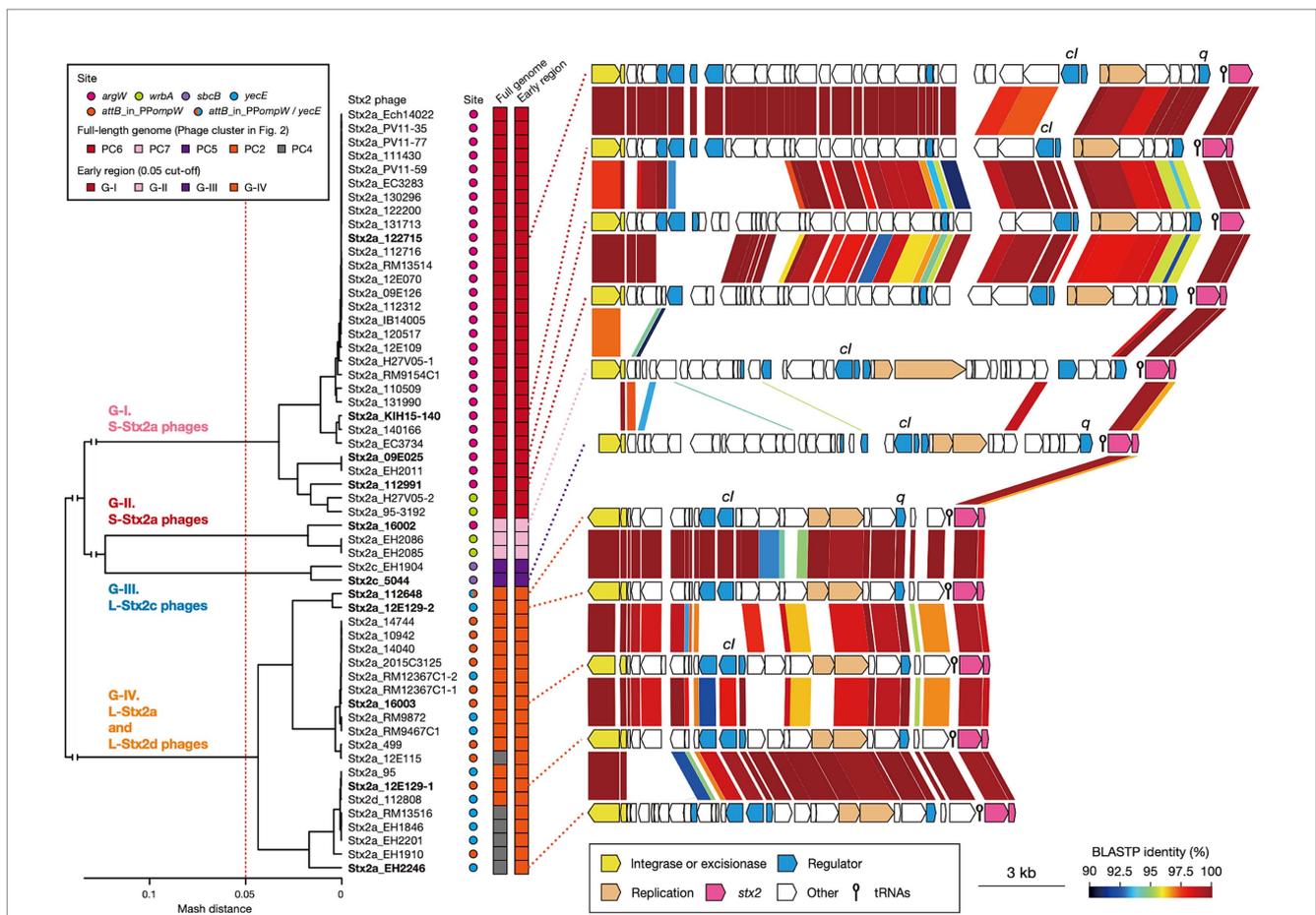


FIGURE 4 Comparisons of early regions between Stx2 phage genomes. A dendrogram constructed based on pairwise Mash distances of early regions is shown in the left panel. Stx2 phages were divided into four groups (threshold: 0.05). Phage clusters, which were defined based on the sequence similarity of full-length genomes (shown in Figure 2), and integration sites are also indicated. The genetic structures of the early regions of representative Stx2 phages (indicated in bold) are drawn to scale in the right panel. Amino acid sequence homologies are shown by shading with a heatmap.

## Conclusion

Our systematic analysis of the full set of Stx phages in 71 STEC strains covering the entire O145:H28 lineage revealed marked genomic diversity and dynamism of Stx phages. While all Stx1 phages were long-tailed phages, Stx2 phages are divided into short- and long-tailed phages (referred to as S- and L-Stx2 phages, respectively, in this article) and the strains carrying S-Stx2 phages produced significantly more Stx2 than those carrying L-Stx2 phages. However, by detailed sequence comparison in the early region, the Stx2 phages were further classified into the four groups; G-I and G-II, each comprising S-Stx2a phages, and G-III and G-IV, each comprising L-Stx2 phages. Thus, although the variations in the early regions between S-Stx2a and L-Stx2 phages appeared to be linked to a striking difference in Stx2 production levels in the host strains and distinguishing S-Stx2a and L-Stx2a phages is important for the surveillance of STEC, the impact of the variation in the early regions within each group and the carriage of two Stx2 phages on the levels of Stx2 production by host strains should be further analyzed in future.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary material](#).

## Author contributions

KN: Conceptualization, Formal analysis, Funding acquisition, Methodology, Visualization, Writing – original draft, Writing – review & editing. IT: Formal analysis, Software, Writing – review & editing. YG: Formal analysis, Writing – review & editing. JI: Resources, Writing – review & editing. KK: Resources, Writing – review & editing. YI: Resources, Writing – review & editing. TK: Resources, Writing – review & editing. YT: Resources, Writing – review & editing. RN: Resources, Writing – review & editing. KI: Resources,

## References

- Allué-Guardia, A., Koenig, S. S. K., Martinez, R. A., Rodriguez, A. L., Bosilevac, J. M., Feng, P., et al. (2022). Pathogenomes and variations in Shiga toxin production among geographically distinct clones of *Escherichia coli* O113:H21. *Microb. Genom.* 8:e000796. doi: 10.1099/mgen.0.000796
- Asadulghani, M., Ogura, Y., Ooka, T., Itoh, T., Sawaguchi, A., Iguchi, A., et al. (2009). The defective prophage pool of *Escherichia coli* O157: prophage-prophage interactions potentiate horizontal transfer of virulence determinants. *PLoS Pathog.* 5:e1000408. doi: 10.1371/journal.ppat.1000408
- Berger, M., Aijaz, I., Berger, P., Dobrindt, U., and Koudelka, G. (2019). Transcriptional and translational inhibitors block SOS response and Shiga toxin expression in enterohemorrhagic *Escherichia coli*. *Sci. Rep.* 9:18777. doi: 10.1038/s41598-019-55332-2
- Bielaszewska, M., Mellmann, A., Bletz, S., Zhang, W., Köck, R., Kossow, A., et al. (2013). Enterohemorrhagic *Escherichia coli* O26:H11/H-: a new virulent clone emerges in Europe. *Clin. Infect. Dis.* 56, 1373–1381. doi: 10.1093/cid/cit055
- Bonanno, L., Petit, M.-A., Loukiadis, E., Michel, V., and Auvray, F. (2016). Heterogeneity in induction level, infection ability, and morphology of Shiga toxin-encoding phages (Stx phages) from dairy and human Shiga toxin-producing *Escherichia coli* O26:H11 isolates. *Appl. Environ. Microbiol.* 82, 2177–2186. doi: 10.1128/AEM.03463-15
- Calderwood, S. B., and Mekalanos, J. J. (1987). Iron regulation of Shiga-like toxin expression in *Escherichia coli* is mediated by the fur locus. *J. Bacteriol.* 169, 4759–4764. doi: 10.1128/jb.169.10.4759-4764.1987
- Carter, M. Q., Pham, A., Du, W.-X., and He, X. (2021). Differential induction of Shiga toxin in environmental *Escherichia coli* O145:H28 strains carrying the same genotype as the outbreak strains. *Int. J. Food Microbiol.* 339:109029. doi: 10.1016/j.ijfoodmicro.2020.109029
- Cooper, K. K., Mandrell, R. E., Louie, J. W., Korlach, J., Clark, T. A., Parker, C. T., et al. (2014). Comparative genomics of enterohemorrhagic *Escherichia coli* O145:H28 demonstrates a common evolutionary lineage with *Escherichia coli* O157:H7. *BMC Genomics* 15:17. doi: 10.1186/1471-2164-15-17
- Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 43:e15. doi: 10.1093/nar/gku1196
- Delannoy, S., Mariani-Kurkdjian, P., Webb, H. E., Bonacorsi, S., and Fach, P. (2017). The mobilome; a major contributor to *Escherichia coli* stx2-positive O26:H11 strains intra-serotype diversity. *Front. Microbiol.* 8:1625. doi: 10.3389/fmicb.2017.01625
- Fuller, C. A., Pellino, C. A., Flagler, M. J., Strasser, J. E., and Weiss, A. A. (2011). Shiga toxin subtypes display dramatic differences in potency. *Infect. Immun.* 79, 1329–1337. doi: 10.1128/IAI.01182-10
- García-Aljaro, C., Muniesa, M., Jofre, J., and Blanch, A. R. (2009). Genotypic and phenotypic diversity among induced, stx2-carrying bacteriophages from environmental *Escherichia coli* strains. *Appl. Environ. Microbiol.* 75, 329–336. doi: 10.1128/AEM.01367-08

Writing – review & editing. YM: Resources, Writing – review & editing. SI: Resources, Writing – review & editing. TH: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was supported by AMED under Grant Nos. 21fk0108611h0501, 22fk0108611h050, and 23fk0108611h0503 to TH and 24fk0108702h0701 to KN and a KAKENHI from the Japan Society for the Promotion of Science (18K07116) to KN.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2024.1453887/full#supplementary-material>

- Karmali, M. A., Mascarenhas, M., Shen, S., Ziebell, K., Johnson, S., Reid-Smith, R., et al. (2003). Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. *J. Clin. Microbiol.* 41, 4930–4940. doi: 10.1128/JCM.41.11.4930-4940.2003
- Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bover-Cid, S., Chemaly, M., Chemaly, M., et al. (2020). Pathogenicity assessment of Shiga toxin-producing *Escherichia coli* (STEC) and the public health risk posed by contamination of food with STEC. *EFSA J.* 18:e05967. doi: 10.2903/j.efsa.2020.5967
- Kumar, S., Stecher, G., Li, M., Nkayaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12
- Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C. H., Kimura, S., Kurokawa, K., et al. (1999). Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak. *Genes Genet. Syst.* 74, 227–239. doi: 10.1266/ggs.74.227
- Miyata, T., Taniguchi, I., Nakamura, K., Gotoh, Y., Yoshimura, D., Itoh, T., et al. (2023). Alteration of a Shiga toxin-encoding phage associated with a change in toxin production level and disease severity in *Escherichia coli*. *Microb. Genom.* 9:e000935. doi: 10.1099/mgen.0.000935
- Muniesa, M., Blanco, J. E., de Simon, M., Serra-Moreno, R., Blanch, A. R., and Jofre, J. (2004). Diversity of *stx2* converting bacteriophages induced from Shiga-toxin-producing *Escherichia coli* strains isolated from cattle. *Microbiology* 150, 2959–2971. doi: 10.1099/mic.0.27188-0
- Nagano, D. S., Taniguchi, I., Ono, T., Nakamura, K., Gotoh, Y., and Hayashi, T. (2023). Systematic analysis of plasmids of the *Serratia marcescens* complex using 142 closed genomes. *Microb. Genom.* 9:e001135. doi: 10.1099/mgen.0.001135
- Nakamura, K., Murase, K., Sato, M. P., Toyoda, A., Itoh, T., Mainil, J. G., et al. (2020). Differential dynamics and impacts of prophages and plasmids on the pangenome and virulence factor repertoires of Shiga toxin-producing *Escherichia coli* O145:H28. *Microb. Genom.* 6:e000323. doi: 10.1099/mgen.0.000323
- Nakamura, K., Ogura, Y., Gotoh, Y., and Hayashi, T. (2021). Prophages integrating into prophages: a mechanism to accumulate type III secretion effector genes and duplicate Shiga toxin-encoding prophages in *Escherichia coli*. *PLoS Pathog.* 17:e1009073. doi: 10.1371/journal.ppat.1009073
- Nakamura, K., Seto, K., Lee, K., Ooka, T., Gotoh, Y., Taniguchi, I., et al. (2023). Global population structure, genomic diversity and carbohydrate fermentation characteristics of clonal complex 119 (CC119), an understudied Shiga toxin-producing *E. coli* (STEC) lineage including O165:H25 and O172:H25. *Microb. Genom.* 9:e000959. doi: 10.1099/mgen.0.000959
- Neely, M. N., and Friedman, D. I. (1998). Arrangement and functional identification of genes in the regulatory region of lambdoid phage H-19B, a carrier of a Shiga-like toxin. *Gene* 223, 105–113. doi: 10.1016/S0378-1119(98)00236-4
- Nishida, R., Nakamura, K., Taniguchi, I., Murase, K., Ooka, T., Ogura, Y., et al. (2021). The global population structure and evolutionary history of the acquisition of major virulence factor-encoding genetic elements in Shiga toxin-producing *Escherichia coli* O121:H19. *Microb. Genom.* 7:e000716. doi: 10.1099/mgen.0.000716
- Ogura, Y., Mondal, S. I., Islam, M. R., Mako, T., Arisawa, K., Katsura, K., et al. (2015). The Shiga toxin 2 production level in enterohemorrhagic *Escherichia coli* O157:H7 is correlated with the subtypes of toxin-encoding phage. *Sci. Rep.* 5:16663. doi: 10.1038/srep16663
- Ogura, Y., Ooka, T., Asadulghani, T., Terajima, J., Nougaryède, J.-P., Kurokawa, K., et al. (2007). Extensive genomic diversity and selective conservation of virulence-determinants in enterohemorrhagic *Escherichia coli* strains of O157 and non-O157 serotypes. *Genome Biol.* 8:R138. doi: 10.1186/gb-2007-8-7-r138
- Ohtsubo, Y., Ikeda-Ohtsubo, W., Nagata, Y., and Tsuda, M. (2008). GenomeMatcher: a graphical user interface for DNA sequence comparison. *BMC Bioinformatics* 9:376. doi: 10.1186/1471-2105-9-376
- Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., et al. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 17:132. doi: 10.1186/s13059-016-0997-x
- Pacheco, A. R., and Sperandio, V. (2012). Shiga toxin in enterohemorrhagic *E. coli*: regulation and novel anti-virulence strategies. *Front. Cell. Infect. Microbiol.* 2:81. doi: 10.3389/fcimb.2012.00081
- Patel, P. N., Lindsey, R. L., Garcia-Toledo, L., Rowe, L. A., Batra, D., Whitley, S. W., et al. (2018). High-quality whole-genome sequences for 77 Shiga toxin-producing *Escherichia coli* strains generated with PacBio sequencing. *Genome Announc.* 6:e00391. doi: 10.1128/genomeA.00391-18
- Pinto, G., Sampaio, M., Dias, O., Almeida, C., Azeredo, J., and Oliveira, H. (2021). Insights into the genome architecture and evolution of Shiga toxin encoding bacteriophages of *Escherichia coli*. *BMC Genomics* 22:366. doi: 10.1186/s12864-021-07685-0
- Plunkett, G. 3rd, Rose, D. J., Durfee, T. J., and Blattner, F. R. (1999). Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J. Bacteriol.* 181, 1767–1778. doi: 10.1128/JB.181.6.1767-1778.1999
- Scheutz, F., Teel, L. D., Beutin, L., Piérard, D., Buvens, G., Karch, H., et al. (2012). Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J. Clin. Microbiol.* 50, 2951–2963. doi: 10.1128/JCM.00860-12
- Shi, T., and Friedman, D. I. (2001). The operator-early promoter regions of Shiga-toxin bearing phage H-19B. *Mol. Microbiol.* 41, 585–599. doi: 10.1046/j.1365-2958.2001.02507.x
- Stamatakis, A. (2006). RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446
- Steyert, S. R., Sahl, J. W., Fraser, C. M., Teel, L. D., Scheutz, F., and Rasko, D. A. (2012). Comparative genomics and *stx* phage characterization of LEE-negative Shiga toxin-producing *Escherichia coli*. *Front. Cell. Infect. Microbiol.* 2:133. doi: 10.3389/fcimb.2012.00133
- Strauch, E., Hammerl, J. A., Konietzny, A., Schneiker-Bekel, S., Arnold, W., Goesmann, A., et al. (2008). Bacteriophage 2851 is a prototype phage for dissemination of the Shiga toxin variant gene 2c in *Escherichia coli* O157:H7. *Infect. Immun.* 76, 5466–5477. doi: 10.1128/IAI.00875-08
- Strauch, E., Schaudinn, C., and Beutin, L. (2004). First-time isolation and characterization of a bacteriophage encoding the Shiga toxin 2c variant, which is globally spread in strains of *Escherichia coli* O157. *Infect. Immun.* 72, 7030–7039. doi: 10.1128/IAI.72.12.7030-7039.2004
- Tanizawa, Y., Fujisawa, T., and Nakamura, Y. (2018). DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34, 1037–1039. doi: 10.1093/bioinformatics/btx713
- Teel, L. D., Melton-Celsa, A. R., Schmitt, C. K., and O'Brien, A. D. (2002). One of two copies of the gene for the activatable Shiga toxin type 2d in *Escherichia coli* O91:H21 strain B2F1 is associated with an inducible bacteriophage. *Infect. Immun.* 70, 4282–4291. doi: 10.1128/IAI.70.8.4282-4291.2002
- Tyler, J. S., Mills, M. J., and Friedman, D. I. (2004). The operator and early promoter region of the Shiga toxin type 2-encoding bacteriophage 933W and control of toxin expression. *J. Bacteriol.* 186, 7670–7679. doi: 10.1128/JB.186.22.7670-7679.2004
- Tyson, S., Peterson, C.-L., Olson, A., Tyler, S., Knox, N., Griffiths, E., et al. (2019). Eleven high-quality reference genome sequences and 360 draft assemblies of Shiga toxin-producing *Escherichia coli* isolates from human, food, animal, and environmental sources in Canada. *Microbiol. Resour. Announc.* 8:e00625. doi: 10.1128/MRA.00625-19
- Valilis, E., Ramsey, A., Sidiq, S., and DuPont, H. L. (2018). Non-O157 Shiga toxin-producing *Escherichia coli*-a poorly appreciated enteric pathogen: systematic review. *Int. J. Infect. Dis.* 76, 82–87. doi: 10.1016/j.ijid.2018.09.002
- Wagner, P. L., Livny, J., Neely, M. N., Acheson, D. W. K., Friedman, D. I., and Waldor, M. K. (2002). Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. *Mol. Microbiol.* 44, 957–970. doi: 10.1046/j.1365-2958.2002.02950.x
- Wagner, P. L., Neely, M. N., Zhang, X., Acheson, D. W., Waldor, M. K., and Friedman, D. I. (2001). Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain. *J. Bacteriol.* 183, 2081–2085. doi: 10.1128/JB.183.6.2081-2085.2001
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:e1005595. doi: 10.1371/journal.pcbi.1005595
- Yano, B., Taniguchi, I., Gotoh, Y., Hayashi, T., and Nakamura, K. (2023). Dynamic changes in Shiga toxin (Stx) 1 transducing phage throughout the evolution of O26:H11 Stx-producing *Escherichia coli*. *Sci. Rep.* 13:4935. doi: 10.1038/s41598-023-32111-8
- Yara, D. A., Greig, D. R., Gally, D. L., Dallman, T. J., and Jenkins, C. (2020). Comparison of Shiga toxin-encoding bacteriophages in highly pathogenic strains of Shiga toxin-producing *Escherichia coli* O157:H7 in the UK. *Microb. Genom.* 6:e000334. doi: 10.1099/mgen.0.000334
- Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S., Yutsudo, C. H., et al. (2000). Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic *Escherichia coli* O157:H7 strain derived from the Sakai outbreak. *Gene* 258, 127–139. doi: 10.1016/S0378-1119(00)00416-9