Short Communication

Sequence Analysis of Stx2-Converting Phage VT2-Sa Shows a Great Divergence in Early Regulation and Replication Regions

Hiroshi Miyamoto,* Wataru Nakai, Naoto Yajima, Akemi Fujibayashi, Tomonori Higuchi, Koki Sato, and Aizo Matsushiro

Faculty of Biology-Oriented Science and Technology, Department of Molecular Genetics, Kinki University, Uchita, Wakayama 649-6493, Japan

(Received 6 June 1999; revised 5 July 1999)

Abstract

In enterohemorrhagic Escherichia coli, Shiga toxin is produced by lysogenic prophages. We have isolated the prophage VT2-Sa that is responsible for production of Shiga toxin type 2 protein, and determined the complete nucleotide sequence of this phage DNA. The entire DNA sequence consisted of 60,942 bp, exhibiting marked similarity to the 933W phage genome. However, several differences were observed in the immunity and replication regions, where cl, cII, cIII, N, cro, O, and P genes were present: Predicted amino acid sequences of N, cl, cro, O and P in the VT2-Sa genome did not show significant similarity to the counterparts of the 933W genome; however its cl showed higher similarity to λ. Furthermore, O and P closely resembled those of phage HK022. These observations suggest that the various degrees of homology observed in the immunity and replication regions of VT2-Sa could have resulted from frequent recombination events among the lambdoid phages, and that these regions play a key role as a functional unit for phage propagation in competition with other lambdoid phages.

Key words: enterohemorrhagic Escherichia coli; Shiga toxin 2; VT2-Sa

Infection with enterohemorrhagic Escherichia coli (EHEC) serotype O157:H7 is associated with hemor- rhagic colitis, hemolytic uremic syndrome (HUS), and thrombotic thrombocytopenic purpura (TTP). Virulence determinants involved in these symptoms include pathogenicity island LEE locus on EHEC chromosome, extrachromosomal factors such as plasmids pO157 and pColD157, and prophage-encoded Shiga toxin (Stx). Stx of EHEC and Shigella dysenteriae type 1 have similar structures and biological functions. Each consists of a single A subunit and multiple B subunits, and binding to a glycolipid receptor on the cell surface mediates N-glycosylation of the 60S ribosomal subunit. EHEC O157 contains two converting prophages encoding Shiga toxin types 1 and 2, respectively. To understand the production mechanisms of Stx encoded by those genes is the urgent prerequisite for establishing medical treatment of EHEC 0157 infection.

In Japan, a large outbreak of EHEC O157 occurred at Sakai city, Osaka in 1996; subsequently, the Ministry of Health and Welfare in Japan issued a recommendation for the use of kanamycin, fosfomycin, and a new quinolone antimicrobial agent, norfloxacin (NFLX). In our previous paper, however, we have reported that induction of the prophage and production of Stx occurred concomitantly with administration of NFLX. To understand the production mechanism further, we have investigated the genomic structure of the phage. 933W is one of the converting phages that produce Stx2, and its complete nucleotide sequence was determined by Plunkett III et al., who noted that prophage 933W has a genomic organization similar to that of the lambdoid phage.

We have identified an Stx2-converting phage, VT2-Sa, from EHEC O157 strain RIMD0509894 which had been isolated during the EHEC outbreak in Sakai city in 1996, and here we report the entire nucleotide sequence (DDBJ/EMBL/GenBank accession number: AP000363) and precise comparisons of its immunity and early regulation regions with those of phages 933W and λ.

For large-scale propagation of VT2-Sa, E. coli K12 was used. The cells containing prophage VT2-Sa were treated with UV for prophage induction. E. coli cells were infected with the induced phage and VT2-Sa phage particles were collected by the plate lysate method. The collected phage was precipitated with polyethylene glycol and subjected to ultracentrifugation in stepwise gradients of CsCl. After collection of banded phage particles, the phage DNA were extracted with phenol, and...
precipitated in ethanol. The phage DNAs obtained were then subjected to fragmentation by sonication or digestion with restriction endonucleases, followed by cloning the phage DNA into a pBluescript II plasmid vector. DNA sequences were determined by the chain termination method using ABI PRISM 373S and 377 automated DNA sequencers. Assembling sequence data, we have found that the entire nucleotide sequence of the VT2-Sa genome was 60,942 bp in size. This sequence was extremely homologous to the genomic sequence of 933W (Fig. 1). However, there were several differences found between these two. Then, we searched open reading frames (ORFs) longer than 50 amino acids, which were subjected to search for similarities against the protein database of DDBJ using FASTA and BLAST programs. Putative ORFs are listed in Table 1. As can be seen in the 933W sequence, VT2-Sa is mainly composed of five gene clusters: recombination, early regulation, replication, lysis, and possible head and tail structural gene regions. In the recombination region, we have identified the int, xis, exo, bet and gam genes. int functions in the lysogenic pathway, and is required for insertion of phage DNA into the host chromosome. In the excision process, int and xis gene products are required. Both of these proteins of VT2-Sa were identical to those of 933W. exo, bet and gam were highly conserved among VT2-Sa, 933W and λ phages. Each gam ORF of VT2-Sa and λ has an N-terminal extension of 40 amino acid residues, which is absent in 933W because of a frameshift due to the addition of one nucleotide to 117 bp from the initiation sequence of gam ORF of 933W. ORF16 encoded the Kil protein, which was identical to the Kil of 933W except for Arg and Met which were changed to Lys and Ile in VT2-Sa, respectively.

As for the early regulation region, several differences were observed between VT2-Sa and 933W. The structure and function of the early regulation region have been investigated for phage λ in detail. Phage λ has two alternative life cycles, the lytic and the lysogenic. In the lytic pathway, Cro protein is synthesized, and progeny virions are produced following the replication of phage genomic DNA and synthesis of head and tail proteins. However, in the lysogenic pathway, CI protein is synthesized, and the phage is repressed and does not enter vegetative growth. Instead, the phage DNAs are inserted into the host chromosome. For the choice between the lysis and lysogeny, the early regulation region plays important roles. The early regulation region contains five genes: cl, cII, cIII, N, and cro. CI repressor, a helix-turn-helix DNA binding protein, binds to two operators, OR, and accordingly inhibits transcription from the adjacent two early promoters, P0 and Pr. The CII protein binds to promoter PrE, inducing the cl gene transcription. The CIII protein contributes to stabilization of the CII protein. The N protein is an antiterminator which allows transcription through Rho-dependent termination signals such as tL and tR. The Cro protein, an antirepressor, binds to operators OR and OR2, and prevents synthesis of the CI repressor. Thus, the proteins encoded by the early regulation region greatly affect the synthesis of nearly all the phage transcripts, and dominate behaviors of the phage in the host.

ORF17 was almost identical to cIII of 933W and λ. ORF21 of VT2-Sa showed 32% identity to N of 933W (Fig. 2A). This low identity of N was in contrast to 98% identity between 933W and H19B phage. In 933W, a sequence similar to an eukaryotic serine/threonine protein kinase was found, but we were unable to detect the corresponding sequence in VT2-Sa. ORF23 encoded a CI repressor of 169 amino acids, of which a C-terminal part...
| ORF | start | end  | length of amino acids | gene | description | % identity |
|-----|-------|------|-----------------------|------|-------------|------------|
| 1   | 1363  | 32   | 444                   | int  | *E. coli* ORF b1579 phage Rac integrase | 52         |
| 2   | 1691  | 99   | 99                    | *xis* | *E. coli* YdaQ | 40         |
| 3   | 2073  | 103  | 299                   |      |              |            |
| 4   | 3032  | 299  | 95                    |      |              |            |
| 5   | 3323  | 328  | 72                    |      |              |            |
| 6   | 3543  | 3548 | 95                    |      |              |            |
| 7   | 3832  | 765  | 155                   |      |              |            |
| 8   | 4229  | 51   | 257                   |      |              |            |
| 9   | 4875  | 4105 | 31                    | *phage λ ea22* | *E. coli* enterohemolysin2 | 38         |
| 10  | 5473  | 5195 | 93                    | *phage λ orf 61* | *E. coli* YdaQ | 97         |
| 11  | 5675  | 5487 | 63                    | *phage λ orf 63* | *E. coli* enterohemolysin2 | 90         |
| 12  | 5836  | 651  | 62                    | *phage λ orf 60a* | *E. coli* YdaQ | 96         |
| 13  | 6507  | 98   | 226                   | *phage λ exonuclease* | *E. coli* YdaQ | 96         |
| 14  | 7289  | 261  | 183                   | *phage λ bet* | *E. coli* enterohemolysin2 | 99         |
| 15  | 7711  | 138  | 93                    | *phage λ gam* | *E. coli* enterohemolysin2 | 96         |
| 16  | 7935  | 89   | 90                    | *phage λ kil* | *E. coli* enterohemolysin2 | 100        |
| 17  | 7942  | 54   | 122                   | *phage λ cII* | *E. coli* enterohemolysin2 | 100        |
| 18  | 8383  | 8018 | 51                    | *phage λ ea10* | *E. coli* enterohemolysin2 | 99         |
| 19  | 8624  | 8472 | 83                    | *phage λ N* | *E. coli* enterohemolysin2 | 32         |
| 20  | 8817  | 8569 | 173                   | *phage λ 933W N* | *E. coli* enterohemolysin2 | 32         |
| 21  | 9220  | 8879 | 114                   | N    | *phage λ 933W N* | 32         |
| 22  | 10398 | 9880 | 173                   |      | *phage λ 933W N* | 32         |
| 23  | 11499 | 10903| 169                   | c I  | *phage λ CI repressor* | 81         |
| 24  | 11671 | 11883| 71                    | *cro* | *phage λ CI repressor* | 16         |
| 25  | 12028 | 12321| 98                    | c II | *phage λ CI II* | 29         |
| 26  | 12496 | 13392| 139                   | O    | *phage λ HK022 O* | 98         |
| 27  | 13370 | 14818| 483                   | P    | *phage λ HK022 P* | 85         |
| 28  | 14821 | 15087| 89                    |      |              |            |
| 29  | 15161 | 15436| 92                    |      |              |            |
| 30  | 15792 | 16031| 80                    |      |              |            |
| 31  | 15892 | 16434| 181                   |      |              |            |
| 32  | 16434 | 16958| 175                   |      |              |            |
| 33  | 16958 | 17137| 60                    |      |              |            |
| 34  | 17245 | 17415| 57                    |      |              |            |
| 35  | 17415 | 18146| 244                   | ant  | *phage P22 antirepressor* | 41         |
| 36  | 18218 | 18943| 242                   | roi  | *phage HK022 roi* | 79         |
| 37  | 18946 | 19548| 201                   |      |              |            |
| 38  | 19548 | 19739| 64                    |      |              |            |
| 39  | 19696 | 20166| 157                   | Q    | *phage H19B Q* | 88         |
| 40  | 20953 | 21909| 319                   |      |              |            |
| 41  | 21924 | 22190| 89                    |      |              |            |
| 42  | 22680 | 24614| 645                   |      |              |            |
| 43  | 24801 | 25217| 139                   |      |              |            |
| 44  | 25223 | 25510| 96                    | S    | *E. coli* lys gene S | 60         |
| 45  | 25518 | 26048| 177                   | R    | *E. coli* ORF b1554 | 89         |
| 46  | 26322 | 26888| 189                   | ant  | *phage P22 antirepressor* | 28         |
| 47  | 27045 | 27509| 155                   | Rz   | *phage λ Rz* | 68         |
| 48  | 27500 | 27757| 86                    |      |              |            |
| 49  | 27816 | 28619| 268                   |      |              |            |
| 50  | 28603 | 30306| 568                   |      |              |            |
| 51  | 30309 | 32450| 714                   |      |              |            |
showed marked homology to λ CI repressor (Fig. 2B). The entire amino acid sequence of VT2-Sa CI appeared to be irrelevant to those of 933W (18% identity) and H19B (10% identity). As expected from the immunity difference between these two phages VT2-Sa and 933W, the VT2-Sa phages were indeed able to form plaques on the E. coli K12 lysogenic for 933W (data not shown). In addition, VT2-Sa showed a different immunity with VT2-141 which is a Stx2-converting phage identified independently of VT2-Sa in Japan (data not shown). ORF24 could be a homolog of cro of λ and 933W judging from its position in the early regulation region. However, it did not show any similarity to cro of both λ (14% identity) and 933W (18% identity). In contrast, its N-terminal 39 amino acid residues exhibited 54% identity to C2, a repressor-like protein of P22-related Salmonella phage L.21 ORF25 encoded the ell homolog (95% identity to 933W; 94% identity to H19B; and 29% identity to λ) (Fig. 2C). Between 933W and λ, O and P are highly conserved. However, there was no significant corresponding similarity in VT2-Sa, whereas its ORF26 and ORF27 showed marked identities to O (98%) and P (85%) of phage HK022, respectively. ORF35, which was not found in 933W, was most similar to a C-terminal part of the antirepressor protein of bacteriophage P2222 (41% identity). ORF39 was completely identical to 933W Q and showed 88% identity to H19B Q. However, the amino acid sequences of Q proteins of these three phages were not so similar to λ Q (15% identity). The Q protein allows readthrough transcription by the pR\textsuperscript{3} promoter beyond the antitermination signal, t\textsubscript{R}, in λ phage, enabling expression of the lysis genes.23,24 ORF40 and ORF41 encoded A subunit and B subunit of Stx2, respectively. The two ORFs were, as reported by Plunkett III et al.14 and Kanjo and Inokuchi,25 located downstream of Q and three tRNA genes. Lysis genes (S, R, Rz) were located downstream of the stx2 gene. The fact that the stx2 genes are located within the lysis segment indicates that toxin production and lysis proceed concomitantly, leading to massive extracellular release of toxin, which may confer an evolutionary advantage in connection with EHEC O157 epidemics.

Judging the present findings comprehensively, it is speculated that stx-converting phages like VT2-Sa and 933W belong to lambdoid phages, and that diverse lambdoid genomes are elaborated through free assortment, i.e., frequent occurrence of recombination among them. This concept has been presented by A. Campbell (1988),26 who argued that the lambdoid phage genome is chimeric in origin.

Acknowledgements: We would like to thank A. Tanaka for critical reading of the manuscript. This work was supported in part by a Grant-in-Aid for Scientific Re-
Figure 2. Alignments of the predicted amino acid sequences in the immunity and early regulation regions. Comparisons of VT2-Sa with 933W, H19B and λ were carried out by the program ClustalW. Gaps introduced to improve the alignments are indicated by broken lines. In (A) and (C), the residues are boxed when conserved among three or four phages. In (B), identical residues are boxed. (A) N. (B) C. (C) CI.

search from the Ministry of Education, Science, Sports and Culture of Japan.
References

1. Griffin, P. M. and Tauxe, R. V. 1991, The epidemiology of infections caused by Escherichia coli O157:H7, other enterohemorrhagic E. coli, and the associated hemolytic uremic syndrome, Epidemiol. Rev., 13, 60–98.

2. McDaniel, T. K., Jarvis, K. G., Donnenberg, M. S., and Kaper, J. B. 1995, A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens, Proc. Natl. Acad. Sci. USA, 92, 1664–1668.

3. Makino, K., Ishii, K., Yasunaga, T. et al. 1998, Complete nucleotide sequences of 93-kb and 3.3 kb plasmids of an enterohemorrhagic Escherichia coli O157:H7 derived from Sakai Outbreak, DNA Res., 5, 1–9.

4. Hofinger, C., Karch, H., and Schmidt, H. 1998, Structure and function of plasmid pColD175 of enterohemorrhagic Escherichia coli O157 and its distribution among strains from patients with diarrhea and hemolytic-uremic syndrome, J. Clin. Microbiol., 36, 24–29.

5. O’Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W., and Formal, S. B. 1984, Shiga-like toxin-converting phages from Escherichia coli strains that cause hemolytic colitis or infantile diarrhea, Science, 226, 694–696.

6. Newland, J. W., Stockbine, N. A., Miller, S. F., O’Brien, A. D., and Holmes, R. K. 1985, Cloning of shiga-like toxin structural genes from a toxin converting phage of Escherichia coli, Science, 230, 179–181.

7. Donohue-Rolfe, A., Keusch, G. T., Edson, C., Thorley-Lawson, D., and Jacewicz, M. 1984, Pathogenesis of Shigella diarrhea. IX. Simplified high yield purification of Shigella toxin and characterization of subunit composition and function by the use of subunit-specific monoclonal and polyclonal antibodies, J. Exp. Med., 160, 1767–1781.

8. Kozlov, Y. V., Kabishev, A. A., Lukyanov, E. V., and Bayev, A. A. 1988, The primary structure of the operon coding for Shigella dysenteriae toxin and temperate phage I36 shiga-like toxin, Gene, 75, 213–221.

9. Reisbig, R., Olson, S., and Eiklid, K. 1997, The cytotoxic activity of Shigella toxin, Evidence for catalytic inactivation of the 60 S ribosomal subunit, J. Biol. Chem., 262, 8739–8744.

10. Stockbine, N. A., Marques, L. R. M., Newland, J. W., Smith, H. W., Holmes, R. K., and O’Brien, A. D. 1986, Two toxin-converting phages from Escherichia coli O157:H7 strain 933 encode antigenically distinct toxins with similar biologic activities, Infect. Immunol., 53, 135–140.

11. Jackson, M. P., Neil, R. J., O’Brien, A. D., Holmes, R. K., and Newland, J. W. 1987, Nucleotide sequence analysis and comparison of the structural genes for Shiga-like toxin I and Shiga-like toxin II encoded by bacteriophages from Escherichia coli 933, FEMS Microbiol. Lett., 44, 109–114.

12. O’Brien, A. D., Marques, L. R., Kerry, C. F., Newland, J. W., and Holmes, R. K. 1989, Shiga-like toxin converting phage of enterohemorrhagic Escherichia coli strain 933, Microb. Pathog., 6, 381–390.

13. Matsuhiro, A., Sato, K., Miyamoto, H., Yamamura, T., and Honda, T. 1999, Induction of prophages of enterohemorrhagic Escherichia coli O157:H7 with Norfloxacin, J. Bacteriol., 181, 2257–2260.

14. Plunkett III, G., 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 phase lategene product, J. Bacteriol., 181, 1767–1778.

15. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989, Molecular cloning: A laboratory manual, 2nd ed., Cold Spring Harbor Laboratory, N.Y.

16. Sanger, F., Nicklen, S., and Coulson, A. R. 1977, DNA sequencing with chain-terminating inhibitors, Proc. Natl. Acad. Sci. USA, 74, 5463–5467.

17. Hoes, R. H., Foeller, C., Bidwell, K., and Landy, A. 1980, Site-specific recombination functions of bacteriophage λ: DNA sequence of regulatory regions and overlapping structural genes for Int and Xis, Proc. Natl. Acad. Sci. USA, 77, 2482–2486.

18. Reichardt, L. F. 1975, Control of bacteriophage lambda repressor synthesis after phage infection: The role of the N, cII, cII and cro products, J. Mol. Biol., 93, 267–288.

19. Reichardt, L. F. 1975, Control of bacteriophage lambda repressor synthesis: Regulation of the maintenance pathway of the cro and cI products, J. Mol. Biol., 93, 289–309.

20. Neely, M. N. and Friedman, D. I. 1998, Functional and genetic analysis of regulatory regions of coliphage H-19B: location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release, Mol. Microbiol., 28, 1255–1267.

21. Schicklimaier, P. and Schmiedie, H. 1997, Sequence comparison of the genes for immunity, DNA replication, and cell lysis of the P22-related Salmonella phages ES18 and L, Gene, 195, 93–100.

22. Sauer, R. T., Krovatin, W., DeAnda, J., Youderian, P., and Susskind, M. M. 1983, Primary structure of the ImmI immunity region of bacteriophage P22, J. Mol. Biol., 168, 699–713.

23. Daniels, D. L. and Blattner, F. R. 1982, Nucleotide sequence of the Q gene and the Q to S intergenic region of bacteriophage lambda, Virology, 117, 81–92.

24. Grayhack, E. J., Yang, X., Lau, L. F., and Roberts, J. W. 1985, Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site, Cell, 42, 259–269.

25. Kanjo, N. and Inokuchi, H. 1999, Genes for tRNA Arg located in the upstream region of the Shiga toxin II operon in enterohemorrhagic Escherichia coli O157:H7, DNA Res., 6, 71–73.

26. Campbell, A. 1988, Phage evolution and speciation, In The Bacteriophages Vol. 1 (ed. Calendar, R.), Plenum Press, New York, 1–14.