Diversification of *Escherichia coli* genomes: are bacteriophages the major contributors?

Makoto Ohnishi, Ken Kurokawa and Tetsuya Hayashi

Determination of the genome sequence of enterohemorrhagic *Escherichia coli* O157 Sakai and genomic comparison with the laboratory strain K-12 has revealed that the two strains share a highly conserved 4.1-Mb sequence and that each also contains a large amount of strain-specific sequence. The analysis also revealed the presence of a surprisingly large number of prophages in O157, most of which are lambda-like phages that resemble each other. Based on these results, we discuss how the *E. coli* strains have diverged from a common ancestral strain, and how bacteriophages contributed to this process. We also describe possible mechanisms by which O157 acquired many closely related phages, and raise the possibility that such bacteria might function as 'phage factories', releasing a variety of chimeric or mosaic phages into the environment.

Makoto Ohnishi
Tetsuya Hayashi*
Dept of Microbiology, Miyazaki Medical College, 5200 Kiyotake, Miyazaki 889-1692, Japan.
*e-mail: thayash@post.miyazaki-med.ac.jp

Ken Kurokawa
Genome Information Research Center, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565-0871, Japan.

Escherichia coli is genotypically and phenotypically a highly diverse species. For example, the chromosomes of *E. coli* natural isolates vary in size from 4.5 to 5.5 Mb although their rRNA operon (*rrn*) skeletons are well conserved[1,2]. From a medical point of view, most *E. coli* strains are commensals of higher vertebrates, but some are pathogenic. Pathogenic strains cause a wide variety of diseases ranging from local infections, such as diarrhea, to systemic infections. The diarrheagenic strains are classified into at least five different subtypes, each of which causes distinct clinical symptoms[3,4]. In addition, the *Shigella* species are now known to be multiple distinct lineages of *E. coli* [5]. One of the first clues to understanding the genetic basis underlying the divergent virulence types of pathogenic *E. coli* strains came from the identification of a DNA segment specific to uropathogenic *E. coli* strains that encodes a set of virulence genes[6,7]. Such subtype-specific segments termed 'pathogenicity islands' have also been identified in other types of pathogenic *E. coli* and in many other pathogenic bacterial species[8–10]. In 1997, the complete genome sequence of the benign *E. coli* laboratory strain K-12 was determined[11]. This made it possible to perform detailed genome-wide comparative analyses of pathogenic *E. coli* strains and K-12; many strain-specific DNA sequences have been identified, and their lengths range from several hundred kb to >1 Mb (Refs 12–14).

Genomic comparison of O157 and K-12

Recently, the complete genome sequence of an enterohemorrhagic *E. coli* O157 strain, isolated from a large outbreak that occurred in 1996 in Sakai City, Japan (referred to as O157 Sakai), and the nearly complete sequence of another O157 strain (EDL933) were determined[15–17]. This provides the first opportunity to perform a direct comparison of the genomes of different *E. coli* strains at the DNA sequence level.

The O157 Sakai chromosome is 5498 kb in length, which is 859 kb larger than K-12 MG1655 (Refs 15,16). By comparing the chromosome sequences, 4.1 Mb of sequence was found to be highly conserved in both strains. This 4.1 Mb of non-contiguous sequence encodes 3729 protein-coding regions (open reading frames, ORFs), seven *rrn* operons and 82 tRNA genes, most of which are regarded as chromosomal components inherited from the common ancestor of the two strains. The locations of these conserved sequences in the two strains indicate there have been no large rearrangements of this 'backbone' sequence between the two lineages. Strain-specific sequences were also identified in both strains. These sequences ('strain-specific loops'; S- and K-loops in O157 Sakai and K-12, respectively) are inserted as DNA fragments of various sizes into a common backbone; there are 296 such insertions larger than 19 bp in O157 Sakai and 325 in K-12. The total length of S-loops in O157 Sakai is 1393 kb (25.3% of the chromosome), encoding 1632 ORFs and 20 tRNA genes that are absent in K-12. As O157 Sakai contains two plasmids of 92.7 kb and 3.3 kb (Ref. 18), the total O157 Sakai-specific sequence amounts to nearly 1.5 Mb. The K-12 genome also contains around 537 kb of strain-specific sequence.

The chromosome size of EDL933 was estimated to be about 5.5 Mb (Ref. 17), and this strain contains a 92.7-kb plasmid almost identical to that of O157 Sakai[19]. Although the EDL933 genome sequence is not yet complete, Perna *et al.* identified, also by genomic comparison between EDL933 and K-12, 4.1 Mb of common chromosome backbone sequence, 1.34 Mb of EDL933-specific sequence (177 segments being larger than 50 bp) and 0.53 Mb of K-12-specific sequence (234 segments), as well as 1576 EDL933-specific genes[17]. The chromosome of the two O157 strains is very similar, but the EDL933 chromosome contains a ~420-kb segment inversion spanning the replication terminus and a duplication of an 86-kb phage-like element (corresponding to SpLE1 of O157 Sakai, discussed later), and lacks a Mu-like phage (Sp18 of O157 Sakai). In addition, at least 16 substantial regions of the published EDL933 sequence differ from O157 Sakai, including two

http://tim.trends.com 0966-842X/01/$ – see front matter PII: S0966-842X(01)02173-4

Table 1. Prophages and prophage-like elements in *E. coli* O157 Sakai[a]

Name[b]	Length (bp)	Feature	Comments (encoded homologs)
Sp1*	10 586	Lambda-like	
Sp2*	12 887	P4-like	
Sp3	38 586	Lambda-like	Lom
Sp4*	49 650	Lambda-like	tRNAs (*ileZ*, *argN* and *argO*), SOD, Lom, EspF
Sp5	62 708	Lambda-like	tRNAs (*ileZ*, *argN* and *argO*), Stx2, Lom, Bor
Sp6	48 423	Lambda-like	SopA, YopM
Sp7	15 463		
Sp8	46 897	Lambda-like	EmrE, Bor, Lom, catalase
Sp9	58 175	Lambda-like	tRNAs (*ileZ*, pseud and *argO*), Anm, Lom, TrcA
Sp10	51 112	Lambda-like	tRNAs (*ileZ*, pseud and *argO*), SOD, Lom
Sp11	45 778	Lambda-like	tRNAs (*ileZ*, pseud and *argO*), Lom
Sp12	46 142	Lambda-like	tRNAs (*ileZ*, *argN* and *argO*), Lom
Sp13*	21 120	P2-like	
Sp14*	44 029	Lambda-like	tRNAs (*ileZ*, *argN* and *argO*), EspF, Lom
Sp15	47 879	Lambda-like	Stx1, Lom
Sp16*	8551	P22-like (?)	
Sp17*	24 199	Lambda-like	TrcA
Sp18	38 759	Mu-like	
SpLE1*	86 249	CP4-like	TraT, Urease, Ter^R, Iha, AidA-I
SpLE2	13 459		
SpLE3*	23 454		PagC, ShET2, Efa1
SpLE4*	43 450	CP4-like	LEE locus
SpLE5*	10 235		
SpLE6*	34 148		

[a]Abbreviations: LEE, locus of enterocyte effacement; SOD, superoxide dismutase; Sp, Sakai prophage; SpLE, Sakai prophage-like elements; Stx, Shiga toxin.
[b]Prophages and prophage-like elements integrated into tRNA (or tmRNA) genes are indicated by asterisks.

IS*629* loci and the internal regions of ten lambda-like prophage genomes.

The mode of genomic diversification in *E. coli*

The amount of DNA specific to O157 Sakai is comparable to the size of the whole genome of the Lyme disease pathogen, *Borrelia burgdorferi* (1.44 Mb)[20,21]. How could O157 acquire such large amounts of strain-specific DNA? Was it by duplication of pre-existing DNA or by horizontal transfer of foreign DNA? Although some DNA might represent segments that were strain- or lineage-specifically deleted from an ancestor with a larger genome, atypical base composition and codon usage in the S-loop genes indicate that many of the loops are of foreign origin. This notion is further supported by analyses of the paralogous gene families in O157 Sakai; <20% of the O157-specific ORFs have paralogs in the conserved ORFs[15]. Many of the K-12-specific sequences are also thought to have foreign origins (the average G+C content of the K-12-specific sequences is 47.7%; K. Kurokawa *et al*., unpublished). Considering that O157 and K-12 belong to distantly related *E. coli* lineages[5,22], it might be possible to deduce a common mechanism for the generation of genomic diversification of *E. coli* strains. On the one hand, the 4.1 kb common backbone sequence mainly represents the DNA that *E. coli* species inherited from a common ancestor, and the encoded gene repertoire determines the principal features of the *E. coli* species. On the other, each strain has also acquired 0.4–1.5 Mb of lineage-specific sequence, most of which is foreign DNA, and the gene repertoires encoded by these sequences determine the unique characteristics of each strain. Some sequences could have been deleted specifically both from the O157 and K-12 lineages, or the 4.1 Mb of conserved sequence might also contain some sequences acquired specifically by both lineages. This prediction will soon become testable as the genome sequences of additional *E. coli* strains (including *Shigella* species) are determined. Genomic comparison of *E. coli* strains with closely related species, such as *Salmonella*, will provide further insights into the evolution of *E. coli* species as well as the whole phylum of enteric bacteria. For example, by determining the gene content that is conserved in *E. coli* and *Salmonella* species, individually, we should be able to determine the components that define the species differences.

'...bacteriophages have played a predominant role in the emergence of [*E. coli*] O157.'

Predominant role of bacteriophages in the genomic diversification of *E. coli*

Horizontal gene transfer clearly plays a central role in generating genomic diversity among *E. coli* strains. A major finding from the genome analysis of O157 Sakai is the presence of many mobile genetic elements[15,16]. O157 Sakai contains 98 copies of IS elements, many of which were identified on the large plasmid pO157 and in strain-specific regions on the chromosome. pO157 carries many virulence-related genes, such as those encoding enterohemolysin, EspP protease and a large clostridial toxin (LCT)-like protein[18,19,23,24]. More interestingly, O157 Sakai also contains 18 prophages or prophage remnants (Sakai prophages; Sp1–18), which account for half of the strain-specific sequences (Table 1). In addition, O157 contains six large chromosome segments that appear to be prophage-like genetic elements (Table 1)[15]. This clearly implies that bacteriophages have played a predominant role in the emergence of O157. As K-12 also contains a total of 11 prophages, prophage remnants and phage-related elements[15,25], bacteriophages appear to be a major contributor to the genomic diversification of *E. coli*. Many strain-specific-sequences that are apparently unrelated to phages were identified in both O157 and K-12. They also appear to have played significant roles in *E. coli*

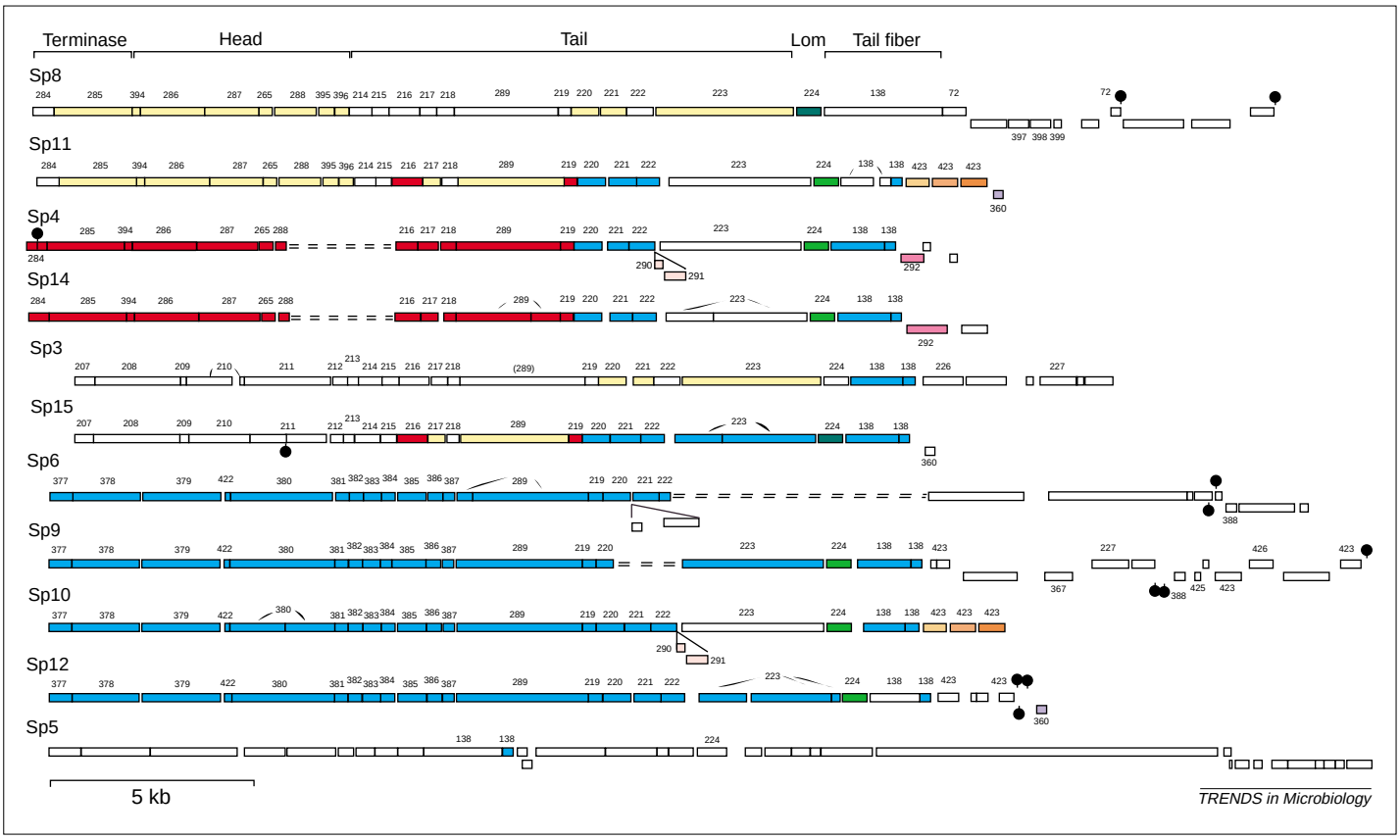

Fig. 1. Genome structures of the lambda-like phages identified in *Escherichia coli* O157 Sakai. The gene organizations of the late regions of 11 lambda-like phages are shown (two phages lack late regions). The sizes and relative locations of each gene are principally drawn to scale. Insertion sites for IS elements are indicated by closed circles, and apparently deleted segments by double broken lines. The numbers attached to each gene indicate the paralogous gene families of O157 Sakai, to which the gene belongs. The criterion used for identifying paralogous genes is that at least 60% of query amino acid sequences are aligned with more than 30% identity[15]. Open reading frames (ORFs) split by authentic frameshift mutations are connected by lines. Genes that exhibit high similarity (>90% identity) are the same color and thus genes with the same color and number are highly homologous to each other. In the regions with contiguous genes of the same numbers and colors, the level of identity is usually in the range of 95–100%, reflecting the fact that their nucleotide sequences are identical or nearly identical.

evolution, but more than half of them are smaller than 500 bp (51% of S-loops and 68% of K-loops). Among the prophages or phage-like elements of K-12, three share some sequence homology with the O157 Sakai prophages that are integrated at analogous loci, for example, Rac and Sp10. This suggests that these prophages were integrated before divergence of the two lineages, but as segments other than the shared sequences differ, these prophage genomes have undergone significant rearrangements after diversification.

The prophage-like elements of O157 Sakai (Sakai prophage-like elements; SpLE1–6, Table 1) share homology with the 'CP4 cryptic prophages' of K-12 (Refs 15,25), and also exhibit structural similarity to some pathogenicity islands[9]. They include an 86-kb element carrying a variety of genes such as those encoding proteins involved in tellurite resistance, urease and the Iha adhesin (SpLE1), and a 43-kb element (SpLE4) containing the locus of enterocyte effacement (LEE), one of the major virulence determinants of O157. SpLE3 also contains multiple potentially virulence-related genes. At present, we do not know whether all the elements are actually mobile, but SpLE1 has been duplicated in EDL933 and a second copy is integrated into another serine tRNA gene[17]. Determination of whether such elements are transferable and, if so, by what mechanism(s) will provide clues to the nature of horizontal transfer of many pathogenicity islands that also exhibit prophage-like features[8–10,26]. In this context, the presence of a P4-like phage (Sp2) and a P2-like phage remnant (Sp13) in O157 is noteworthy. P4 itself is defective, but it can package its genome into a fully infective particle by using the gene products of co-infecting P2 phage[27,28]. The P2/P4-like system could be involved in the transfer of some SpLE elements of O157.

How could O157 acquire multiple lambda-like phages?

Among the 18 prophages on the O157 Sakai chromosome, 13 are lambda-like phages (Table 1). They all contain various types of deletions and/or insertions of IS elements in the phage-essential regions, and thus are likely to be defective[16,17,29,30]. However, they encode a variety of virulence-related proteins such as Shiga toxins (Stx), zinc/copper-type superoxide dismutases (SOD), TrcA-homologs, Bor proteins and many Lom homologs[15] (Bor proteins and Lom homologs have been implicated in host serum resistance and cell adhesion, respectively[31,32]) (Table 1). tRNA genes (*ileZ*, *argN* and *argO*), which might contribute to the efficient expression of many

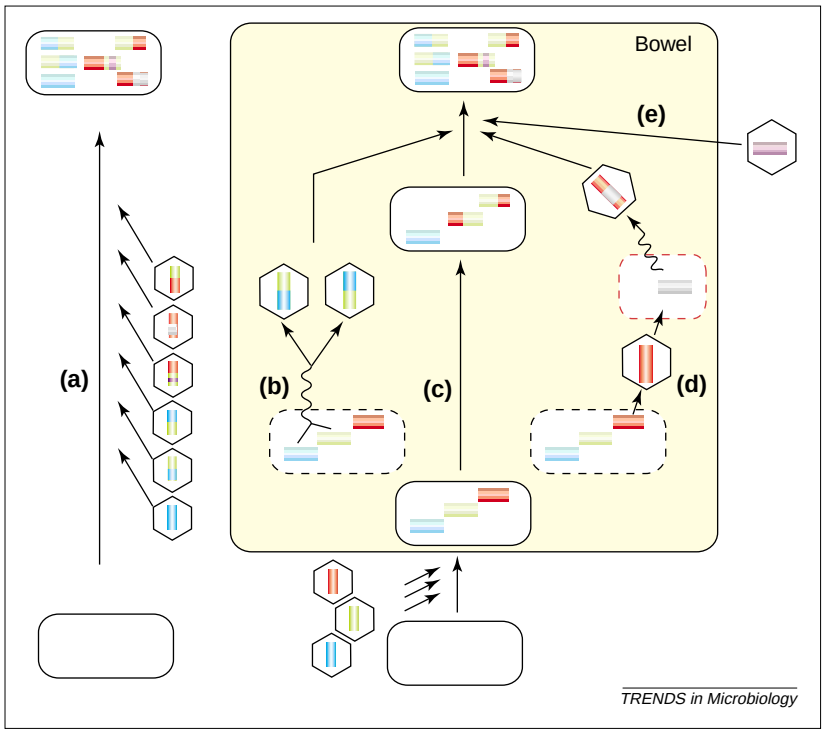

Fig. 2. Possible mechanisms for the acquisition of multiple prophages resembling each other. (a) Sequential and independent infections of closely related bacteriophages; (b–d), a more complicated model involving duplication, recombination and sequence diversification with fewer initially infecting bacteriophages. (b) Re-infection of recombinant phages generated during the co-induction of prophages. (c) Generation of mosaic phages by recombination in a single cell. (d) Re-infection of a recombinant phage generated in other strains or other species. Re-infections of recombinant phages (b and d) might not be rare events in a closed environment like the bowel. (e) Non-lytic infection by, and recombination with, a related but different phage from elsewhere. Sequence diversifications of each phage genome, which can occur both in the lytic and lysogenic cycles, are not illustrated.

Questions for future research

- Is the 4.1-Mb 'backbone' sequence also conserved in other *E. coli* strains?
- Is phage-accretion a common feature in natural *E. coli* isolates?
- Are the prophage-like elements actually mobile or transferable and if so by what mechanisms?
- Is it possible to trace the pathway that generated the lambda-like prophages of O157 by molecular phylogenetic approaches?
- How are multiple lambda-like phages with highly homologous sequences maintained in O157?
- Does O157 release new recombinant phages *in vitro* or *in vivo*?

O157-specific genes, reside on seven of the prophages[15,33]. Intriguingly, these lambda-like phages are very similar. Several phages share identical or nearly identical DNA segments of >20 kb (Fig. 1; see also Ref. 15). This raises the question of how these phages arrived in the bacterial lineage. The simplest explanation could just be independent infections by 13 different lambda-like phages (Fig. 2a). As some lambda-like phages isolated from the environment, such as HK97 and HK022, exhibit a high level of sequence similarity[34], this possibility cannot be ruled out completely. However, the presence of multiple phages exhibiting the same level or even a higher level of similarity to each other than that observed between HK97 and HK022 and the presence of an identical deletion in different phage genomes (Sp4 and Sp14 in Fig. 1) suggest that some of the phages have arisen by duplication within the lineage. Furthermore, the genomes of these phages display highly mosaic features, as shown in Fig. 1. Early and lysis regions of these lambda-like phages are also mosaic, exhibiting much more complexity (data available on our web site: http//genome.gen-info.osaka-u.ac.jp/bacteria/o157/). Therefore, we regard these phages as having undergone extensive recombination and sequence diversification after and/or during the initial infection and duplication. Recombination between the similar phages might occur rather frequently in a single cell (Fig. 2c). However, to increase the phage repertoire, re-infection of the recombinant phages generated by co-induction could take place (Fig. 2b). If the induced phages infect other strains or other species in a closed environment such as the bowel, different types of recombinant phages could be generated and re-infected (Fig. 2d). In these pathways, replacement or significant sequence diversification of phage immunity-related genes (or sequences) might have to occur to allow the re-infection. In fact, the immunity-related genes of the phages have significantly diverged. This might also be the case for the integration-related genes. All the lambda-like phages are integrated into different loci in O157. Finally, in non-lytic infections, more distantly related phages could recombine with resident prophages (Fig. 2e). Recombination can take place even between distantly related phages[35,36], and might also be possible between phages and certain pathogenicity islands. Thus, pathways (d) and (e) in Fig. 2 could introduce new genes into the O157 lineage. Once this cycle starts, the strain can increase its phage repertoire, and in turn, the cell might function as a 'phage factory' that releases a wide variety of recombinant phages into the environment[36–38]. Almost-perfect conservation of the putative tail-fiber genes (Fig. 1), which are required for re-infection of O157, would support this re-infection model.

It is clear that the genetic history of these *E. coli* prophage elements is very complex, and it will be interesting to see which of the mechanisms outlined in Fig. 2 turn out to be important in their diversification. The genome structures of most lambda-like phages in EDL933 differ significantly from the corresponding phages in O157 Sakai. This might also be additional evidence for the extensive and possibly recent recombination among the lambda-like phages in O157, but confirmation awaits the completion of the EDL933 genome sequence.

Acknowledgements
We would like to express our special thanks to Sherwood Casjens for his valuable suggestions. This work was supported by the Japan Society for the Promotion of Science 'Research for the Future Programs' (97L00101 and JSPS-RETF 00L01411), Grant-in-Aid for Scientific Research (B) (13470061 to T.H.) and Grant-in-Aid for Scientific Research on Priority Areas (C) (13206068 to T.H.) from the Ministry of Education, Culture, Sports, Science and Technology.

References

1 Bergthorsson, U. and Ochman, H. (1995) Heterogeneity of genome sizes among natural isolates of *Escherichia coli*. *J. Bacteriol.* 177, 5784–5789
2 Bergthorsson, U. and Ochman, H. (1998) Distribution of chromosome length variation in natural isolates of *Escherichia coli*. *Mol. Biol. Evol.* 15, 6–16
3 Levine, M.M. (1987) *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *J. Infect. Dis.* 155, 377–389
4 Nataro, J.R. and Kaper, J.B. (1998) Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* 11, 1–60
5 Pupo, G.M. *et al*. (1997) Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. *Infect. Immun.* 65, 2685–2692
6 Blum, G. *et al*. (1994) Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. *Infect Immun.* 62, 606–614
7 Morschhauser, J. *et al*. (1994) Adhesin regulatory genes within large, unstable DNA regions of pathogenic *Escherichia coli*: cross-talk between different adhesin gene clusters. *Mol. Microbiol.* 11, 555–566
8 Groisman, A.G. and Ochman, H. (1996) Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* 87, 791–794
9 Hacker, J. *et al*. (1997) Pathogenicity islands of virulent bacteria: structure, function, and impact on microbial evolution. *Mol. Microbiol.* 23, 1089–1097
10 Hacker, J. and Kaper, J.B. (2000) Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* 54, 641–679
11 Blattner, F.R. *et al*. (1997) The complete genome sequence of *Escherichia coli*. *Science* 277, 1453–1462
12 Rode, C.K. *et al.* (1999) Type-specific contribution to chromosome size differences in *Escherichia coli*. *Infect. Immun.* 67, 230–236
13 Ohnishi, M. *et al*. (1999) Chromosome of the enterohemorrhagic *Escherichia coli* O157:H7; comparative analysis with K-12 MG1655 revealed the acquisition of a large amount of foreign DNAs. *DNA Res.* 6, 361–368
14 Ochman, H. and Jones, I.B. (2000) Evolutionary dynamics of full genome content in *Escherichia coli*. *EMBO J.* 19, 6637–6643
15 Hayashi, T. *et al*. (2001) Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* 8, 11–22
16 Hayashi, T. *et al*. (2001) Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* (Suppl.) 8, 47–52
17 Perna, N.T. *et al*. (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 409, 529–533
18 Makino, K. *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
19 Burland, V. *et al*. (1998) The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucleic Acids Res.* 26, 4196–4204
20 Fraser, C.M. *et al*. (1997) Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* 390, 580–586
21 Casjens, S. *et al*. (2000) A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. *Mol. Microbiol.* 35, 490–516
22 Reid, S.D. *et al*. (2000) Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* 406, 64–67
23 Schmidt, H. *et al*. (1995) Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL933. *Infect. Immun.* 63, 1055–1061
24 Brunder, W. *et al*. (1997) EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves coagulation factor V. *Mol. Microbiol.* 24, 767–778
25 Campbell, A.M. (1996) Cryptic prophages. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F.C. *et al.*, eds), pp. 2041–2046, ASM Press
26 Cheetham, B.F. and Katz, M.E. (1995) A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. *Mol. Microbiol.* 18, 201–208
27 Christie, G.E. and Calendar, R. (1990) Interactions between satellite bacteriophage P4 and its helpers. *Annu. Rev. Genet.* 24, 465–490
28 Lindqvist, B.H. *et al*. (1993) Mechanisms of genome propagation and helper exploitation by satellite phage P4. *Microbiol. Rev.* 57, 683–702
29 Makino, K. *et al*. (1999) Complete nucleotide sequence of the defective prophage VT2-Sakai carrying the verotoxin2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak. *Genes Genet. Syst.* 74, 227–239
30 Yokoyama, K. *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
31 Barondess, J.J. and Beckwith, J. (1995) *bor* gene of phage lambda, involved in serum resistance, encodes a widely conserved outer membrane lipoprotein. *J. Bacteriol.* 177, 1247–1253
32 Vica Pacheco, S. *et al*. (1997) The *Iom* gene of bacteriophage lambda is involved in *Escherichia coli* K12 adhesion to human buccal epithelial cells. *FEMS Microbiol. Lett.* 156, 129–132
33 Plunkett, G., III *et al.* (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
34 Juhala, R.J. *et al.* (2000) Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *J. Mol. Biol.* 299, 27–51
35 Haggard-Ljungquist, E. *et al.* (1992) DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J. Bacteriol.* 174, 1462–1477
36 Hendrix, R.W. *et al*. (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2192–2197
37 Hendrix, R.W. *et al*. (2000) The origins and ongoing evolution of viruses. *Trends Microbiol.* 8, 504–508
38 Campbell, A. and Botsein, D. (1983) Evolution of the lambdoid phages. In *Lambda II* (Hendrix, R. *et al*., eds), pp. 365–380, Cold Spring Harbor Laboratory Press