Mobile effector proteins on phage genomes

E. Fidelma Boyd,* Megan R. Carpenter and Nityananda Chowdhury

Department of Biological Sciences; University of Delaware; Newark, DE USA

Keywords: phage encoded exotoxins and effector proteins, Escherichia coli, Salmonella enterica

Introduction

Bacteriophage genomes found in a range of bacterial pathogens encode a diverse array of virulence factors ranging from superantigens or pore forming lysins to numerous exotoxins. Recent studies have uncovered an entirely new class of bacterial virulence factors, called effector proteins or effector toxins, which are encoded within phage genomes that reside among several pathovars of Escherichia coli and Salmonella enterica. These effector proteins have multiple domains resulting in proteins that can be multifunctional. The effector proteins encoded within phage genomes are translocated directly from the bacterial cytosol into their eukaryotic target cells by specialized bacterial type three secretion systems (T3SSs). In this review, we will give an overview of the different types of effector proteins encoded within phage genomes and examine their roles in bacterial pathogenesis.

Classical Phage-Encoded Exotoxins

The effects of different phage-encoded exotoxins depend on the specificity of their target cell and target site within the cell. Three phage-encoded AB type exotoxins such as diphtheria toxin (DT), cholera toxin (CT) and Shiga toxin (Stx) are encoded by C. diphtheria, V. cholerae and E. coli and Shigella species, respectively. Each of them has very different cellular target sites and consequently diverse disease outcomes (Fig. 1).

DT is an AB type toxin made up of a single A and B subunit where A is the catalytic subunit and B is the receptor binding subunit on the eukaryotic target cell. The tox gene which encodes DT, is encoded by corynebacteriophage β and is located directly adjacent to the phage attachment site, attR, suggesting that this gene was acquired by imprecise excision from a donor host genome. Expression of the tox gene is regulated by iron availability and is activated when iron levels are low.6,7 DT is released from the bacterial cell and attaches to a receptor on the cell surface of respiratory epithelium cells. The toxin is internalized by pore formation and transported to its site of activity, the ribosome, where it ADP-ribosylates Elongation factor 2, thereby halting protein synthesis followed by subsequent cell death (Fig. 1). Little is known about the diversity of phages that encode the tox gene but the phages that have been examined appear to be highly similar.

Cholera toxin is the main cause of secretory diarrhea, a characteristic of cholera.8 CT is an AB, toxin, consisting of a single A subunit and five B subunits. It is encoded within the genome of a filamentous phage named CTXphi, a 7 kb single-stranded DNA phage found in serogroup O1 and O139 isolates of V. cholerae.9 The ctxAB genes are found directly adjacent to the phage attachment site, attR.5 The location of these genes at the terminal region of the phage genome and the isolation of functional CTXphi genomes lacking these genes suggests that they were a later addition to the phage genome.10 CTXphi can integrate into either of the host’s two chromosomes and only at a single site, named dif1 and dif2.11-13 In V. cholerae, CT production is regulated by the bacterial encoded regulator ToxRS and its production increases under conditions encountered in vivo such as low pH, temperature and bile salt concentration.14-16 CT is secreted from V. cholerae cells by the host Type II secretion system17 and subsequently attaches to its target intestinal epithelium cell receptor, the GM1 ganglioside.18 The B subunits are required for attachment and the toxin is taken up by endocytosis within a vacuole. The toxin is retrograde trafficked to the Golgi via early and late endosomes (Fig. 1). In the Golgi, the A subunit separates from the B subunits and enters the endoplasmic reticulum (ER) via coat protein I (COP I) vesicles.19 The active A1 subunit is released into...
Stx toxin is released from the bacterial cell upon phage lysis of the cell. The Stx B subunits bind to glycolipids on the host eukaryotic cell, specifically the cell surface receptor globotriosyl ceramide (Gb3) (Fig. 1). Stx enters the host cell by a clathrin-dependent the cytosol where it stimulates adenylate cyclase, an important signaling molecule that increases cyclic AMP. Increased cAMP levels activate protein kinase A (PKA) permanently, thus opening the Cystic Fibrosis Transmembrane conductance Regulator (CFTR) chloride channel and ultimately causing chloride secretion from the cell (Fig. 1). CTXphi genomes are identical to each other in terms of gene arrangement, content and general sequence identity. The exception is the phage regulatory gene, rstR, which shows high sequence divergence. The bacterial host receptor for CTXphi attachment has been identified as a Type IV pilus, named the toxin co-regulated pilus (TCP).9

Shiga toxin (Stx) and the shiga-like toxin are also AB5 toxins consisting of one active A-subunit and five identical B-subunits, similar to CT. Stx and Stx-like toxins are present on diverse lambdoid phages that insert at multiple locations in the genome of Shigella species and select E. coli pathovars.21,22 Unlike the genes that encode DT and CT, the stx genes are not located adjacent to the phage attachment site but instead are present in the phage late gene region of the phage genome. The genome arrangement of the stx genes suggests that they were not picked up by imprecise prophage excision.23 Moreover, it has now been shown that Stx production is tightly linked to prophage induction and the phage regulatory circuit rather than by host regulatory systems.24,25 The Stx toxin is released from the bacterial cell upon phage lysis of the cell. The Stx B subunits bind to glycolipids on the host eukaryotic cell, specifically the cell surface receptor globotriosyl ceramidin (Gb3) (Fig. 1). Stx enters the host cell by a clathrin-dependent

**Table 1. Phage-encoded effector proteins in E. coli and S. enterica**

| Phage     | Effector protein | Function                              |
|-----------|------------------|---------------------------------------|
| lambdoid  | Cif              | Cyclomodulin                          |
| lambdoid  | EspJ             | Transmission                          |
| SplE3-like| EspL2            | Actin remodeling                      |
| lambdoid  | NleA             | Golgi localized                       |
| lambdoid  | NleG             | Ubiquitin ligase                      |
| lambdoid  | NleH             | NF-kappaB activity                    |
| lambdoid  | EspF (TecP)      | Actin remodeling                      |
| Gifsy-1   | GogB             | Anti-inflammatory effector            |
| Gifsy-1   | GipA             | Invasion                              |
| SopEphi   | SopE             | Guanine exchange factor mimic         |
| Phage-like| SopE2            | Guanine exchange factor mimic         |
| Phage-like| SspH2            | Inhibits actin polymerization         |
| Phage-like| SseJ             | Deacylase                             |
| ST64B     | SseK3            | Inhibition of actin polymerization    |
| Gifsy-2   | SspH1            | Inhibit actin polymerization          |
| Gifsy-3   | SseI (GtgB)      | E3 ubiquitin ligase                   |

**Figure 1.** Schematic of toxins function in eukaryotic cells. Dotted arrows indicate steps in bacterial lysis with phage/toxin release, toxin uptake by host cell and target activity. CTXphi which encodes cholera toxin (CT) does not lyse its bacterial host cell. Black solid arrows or t-bars represent activation or blocking of specified bacterial function by toxin. AB denotes a two subunit protein, subunit A is the catalytic subunit and B is the receptor binding or docking subunit on target cell. AB5 denotes a toxin that contains one A subunit and 5 B subunits. Three examples are given; the A/B diphtheria toxin (DT) and the AB5 toxins CT and shiga toxin (Stx). DT gains direct entry into target cell by pore formation. The second method of toxin uptake is by receptor mediated endocytosis, a mechanism used by CT and Stx. Only the A subunit of DT enters the cell cytosol and along with NAD, ADP-ribosylates Elongation factor 2 blocking protein synthesis. Both CT and Stx are retrograde trafficked to the Golgi apparatus. In CT, the A subunit is dissociated from B subunits in the endoplasmic reticulum (ER) and released into the cytosol where it activates adenylate cyclase increasing cAMP which in turn activates PKA resulting in the CFTR membrane Cl− channel activation. The A subunit of Stx ADP-ribosylates the 28S rRNA ribosomal subunit, thus blocking protein synthesis.
pathway and is retrograde trafficked directly from early/recycling endosomes to the Golgi and then to the ER via a COPI-independent pathway. The A1-subunit is released into the cytosol and inactivates the 60S ribosome subunit by N-glycosidic bond cleavage causing disruption of protein synthesis and cell death (Fig. 1). The bacterial receptor for Stx phage binding and uptake is the highly conserved YaeT protein required for insertion of proteins into the outer membrane of Gram-negative bacteria. Allison’s group has demonstrated that 70% of Stx phages recognize the YaeT surface molecule via conserved tail spike proteins associated with a short tailed morphology.

**Shiga Toxin Encoding E. coli Pathovars**

*Escherichia coli* is a commensal of the gastrointestinal tract of warm-blooded animals, the most common facultative anaerobe in the human intestines. In addition, many *E. coli* strains are pathogenic and cause enteric diseases ranging from severe watery diarrhea, dysentery and hemorrhagic colitis to extraintestinal infections such as cystitis, septicaemia and meningitis. Diversity in the number and type of toxins produced by *E. coli*, as well as diversity in the bacteriophages involved, has enabled *E. coli* to become a versatile pathogen. To date, eight different *E. coli* pathovars have been identified; Enterohemorrhagic *E. coli* (EHEC), enteropathogenic *E. coli* (EPEC), enteroinvasive *E. coli* (EIEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC) and diffusely adherent *E. coli* (DAEC) and the extra-intestinal pathovars Neonatal Meningitis *E. coli* (NMEC) and Uropathogenic *E. coli* (UPEC). EHEC, EPEC, ETEC and EAEC are all extracellular pathogens that cause secretory diarrhea, whereas EIEC is an intracellular pathogen that causes inflammatory diarrhea. EHEC, EIEC and EAEC are associated with infection of the large intestine whereas EPEC, ETEC, DAEC and EAEC are associated with infection of the small intestine. UPEC is associated with urinary tract infection.

In enterohemorrhagic *E. coli* (EHEC) and Shiga toxin *E. coli* (STEC) strains responsible for hemorrhagic colitis and hemolytic uremic syndrome (HUS), the shiga-toxin (Stx) encoded by *stx1* or *stx2*, and enterohemolysin encoded by the *hly* locus are present within lambdoid phage genomes. Since the first discovery of Stx in *E. coli* O157:H7 in 1983, which expresses the somatic (O) 157 and flagellar (H) 7 antigens, Stx has been identified in over 500 different *E. coli* serogroups. The lambdoid phage (denoted as Sp5 in Figure 2) containing the *stx* operon encoding *stxA* and *stxB* genes in *E. coli* O157:H7 Sakai and EDL strains, is integrated at the *yehV* gene whereas the Stx2 encoding bacteriophage (denoted as Sp5 in Figure 2) is integrated at the *wrbA* gene. The majority of *E. coli* O157:H7 strains encode Stx2 and over 75% also encode Stx1. Among these Stx1 negative strains, a truncated prophage is always present at the *yehV* locus suggesting that these strains all contained the Stx1 phage at one point that was subsequently lost in a subset of strains. In Stx1 negative strains, it was found that Stx2 was not at the *yehV* site, evidence that a diversity of phages are involved in Stx transfer. Similarly, it was shown that different STEC isolates have varying bacteriophage titers and toxin production. Structural analysis of bacteriophage-borne *stx1*, *stx2* and flanking sequences in *E. coli* O157, STEC and *Shigella dysenteriae* Type 1 strains demonstrated bacteriophage genomic variation indicating that unrelated bacteriophages encode shiga toxin genes. In *S. dysenteriae* seroType 1 strains, Stx2 has only 56% identity to Stx1.

Some of the first studies to compare the whole genomic sequence of pathogenic *E. coli* strains examined O157 strains, a pathogen that emerged as a worldwide public health threat in the past three decades, and the laboratory-maintained commensal
E. coli K12 strain.36–38 These ground breaking studies determined, for the first time, the source and cause of the genome size variation previously found among E. coli strains based on pulsed-field gel electrophoresis.39 These genomic studies were the first to show that E. coli strains contain a core-conserved sequence that is common to all genomes within the species as well as being interspersed with sequences unique to each strain. The presence of a core set of genes augmented with accessory genes has since been shown to be true for all free-living bacterial species. In one study, the E. coli O157 strain examined contained 1.3 million base pairs of strain-specific DNA, which included a bacteriophage containing stx and a pathogenicity island carrying the locus of enterocyte effacement (LEE).36 Genomic comparison of two E. coli O157 strains, Sakai and EDL933, also revealed large differences in their prophage content.37,38 Strain EDL933 contained 12 prophage genomes whereas strain Sakai contained 18 prophage genomes (Fig. 2).37,38 In silico genomic analysis of the two E. coli O157 strains identified numerous virulence factors encoded by these prophages.37,38 The pathogenic potential of these two strains correlates with their bacteriophage-encoded virulence factors, demonstrating that acquisition of phages played a decisive role in the emergence of O157 as a foodborne pathogen. Whittam and colleagues demonstrated that the O157:H7 EHEC pathogen evolved in a stepwise process by the acquisition of prophages and conversion to a new serotype.40–44 The stepwise evolutionary model proposed that the highly virulent O157:H7 EHEC strains emerged from an O55:H7 EPEC ancestor by sequential acquisition of virulence factors predominantly encoded on prophages. EPEC is a major cause of infant diarrhea in the developing world and is mainly distinguished from EHEC by the absence of the stx genes. EPEC contains the LEE island and the EPEC adherence factor (EAF) plasmid. Using whole genome microarray, Whittam and colleagues examined a number of O157:H7 strains and demonstrated that 85% of the genes that were variably absent/present were phage derived, in total 463 such phage-related genes were found.41 Their analysis demonstrated the dominance of phages in the diversification of the E. coli O157:H7 genome.41

The 2011 Germany Outbreak E. coli O104:H4 Clone

A Stx2-producing E. coli O104:H4 clone emerged in the summer of 2011 in northern Germany and was responsible for one of the largest hemolytic uremic syndrome (HUS) outbreaks.45,46 The Germany O104:H4 clone caused nearly 3,000 cases of diarrhea without HUS and nearly 900 cases with HUS leading to 50 combined deaths.47 Unlike O157:H7, the O104:H4 serotype was not associated with animals prior to the outbreak and was rarely associated with humans in Germany. Using next generation sequencing technology, several groups sequenced multiple O104:H4 outbreak strains and determined that these strains carried genes present in enteroaggregative E. coli (EAEC) and EHEC pathovars.48–50 The

![Figure 3. Location of effector genes and stx genes within prophages in E. coli and S. enterica. Partial prophage genomic maps are shown. (A) Terminal region of five lambdoid prophages from E. coli strain Sakai, (B) terminal region of S. enterica Gifsy-1 and Gifsy-2 prophages and (C) central region of the genome of Stx1 (Sp15) and Stx2 (Sp5) prophages encoding stx1AB and stx2AB genes, respectively, shown as blue arrows. Tail fiber genes are represented by white arrows and effector genes are shown as dark gray arrows. Putative additional virulence genes are shown as striped arrows. Black block arrows indicate the prophage site of genomic integration. Arrows in red represent transposase genes.](image-url)
Germany O104:H4 strains contained a plasmid carrying an aggregative adherence fimbria type (AAF/I) common to EAEC and a prophage carrying the Stx2 toxin inserted in chromosomal gene *wbrA* common to STEC. Phylogenetic analysis of 53 *E. coli* strains indicated that the O104:H4 strains from Germany formed a clade with O104:H4 strains from Africa, as well as with some EAEC strains. The O104:H4 strains differed from the EAEC strains due to the presence of the Stx2 prophage as well as the antibiotic resistance genes, TEM-1 and CTX-M-15. Similar to Stx containing prophage 933W from O157:H7, it was found that the *stxAB* genes in the O104:H4 prophage, were located between the antitermination Q gene and the S/R lysis genes of a lambdoid phage. In addition, there was a high degree of sequence identity between the O157:H7 and O104:H4 prophages suggesting recent horizontal gene transfer. Thus, emergence of this highly virulent strain is a direct consequence of phage acquisition.

**Phage-Encoded Effector Proteins**

Effector proteins or toxins are large proteins contained within a range of bacterial species that are translocated directly from the bacterial cytosol into the target eukaryotic cell cytosol using contact dependent bacterial secretion systems. Within their eukaryotic host cell, these effector proteins cause a variety of effects and in many cases encode overlapping functions. Effector proteins contain several distinct domains or modules that appear to have arisen through recombination reassortment events. Thus, the N-terminal region of one effector may show similar to the C-terminal region of another effector but lack any other homology. Most effector proteins also require a chaperone and thus have a specific chaperone-binding domain. Redundancy in functions encoded by these proteins is well documented. Effector proteins function by targeting a range of structures within the eukaryotic cell such as actin filaments, microtubules, mitochondria, lysosomes, the nucleus, the inner cell membrane and tight junctions, as well as disrupting many different cell signaling pathways such as MAPK, cyclic AMP and GTPases Rho, Rac and Cdc42. Bacterial effector proteins function in many cases by molecular mimicry, hijacking eukaryotic cell function and signaling pathways for their own advantage. To date, effector proteins have only been identified on phage genomes in two species *E. coli* and *S. enterica* (Table 1).

**E. coli** Phage-Encoded Effector Proteins

Nearly a hundred effector proteins have been identified among *E. coli* strains that are translocated via a common Type III secretion system (T3SS). Some of these effectors are encoded within prophages, some within pathogenicity islands and some elsewhere on the bacterial genome. As stated previously, EPEC and EHEC isolates are distinguished from each other mainly by the presence of the *stx* genes in EHEC strains. However, both EPEC and EHEC isolates along with ETEC isolates share the attaching and effacing (A/E) lesion phenotype. These isolates attach to the host cell surface resulting in cytoskeleton rearrangement and pedestal formation (Fig. 3). The factors involved in this host cell surface change are encoded within the LEE pathogenicity island. The canonical LEE island is a 35 kb region that encodes the Esc-Esp T3SS as well as effector proteins, their chaperones and regulators. T3SSs are contact dependent secretion systems found in a large number of Gram-negative species. These secretion systems are multi-component structures that span the bacterial cell wall and protrude from the cell surface that can inject into a eukaryotic host cell to translocate effector proteins directly into the eukaryotic cell cytoplasm.
Figure 5. Prophage and pathogenicity island location on the S. enterica serovar Typhimurium genome map. Stars indicate prophages and rectangular black boxes indicate the position of Salmonella pathogenicity islands (SPIs). Four letter designations indicate effector genes identified within prophages (red) and SPIs (black).

Of the effector proteins identified in E. coli, 60 have been identified within prophages in the EHEC Sakai strain, 39 of which were experimentally shown to be translocated into eukaryotic cells. The two shiga toxin-encoding phages Sp5 and Sp15 in strain Sakai do not encode any effector genes. Similar to many phage-encoded toxin genes, with the exception of stx genes, the majority of effector toxins in prophage genomes are found in the terminal region of the genome, after the tail fiber genes, and adjacent to the phage attachment site. The location of multiple effector genes at the end of phage genomes suggests acquisition by imprecise excision from a previous host genome. Similar to the stxAB genes that encode CT, the percent guanine-cytosine of the effector genes differs significantly from the rest of the phage genome again suggesting that they are not true phage genes per se. The E. coli effectors identified were named following the esp (E. coli secreted protein) and nle (non-leaf encoded effector) nomenclature.

The effectors identified in the E. coli Sakai strain fell into greater than 20 protein families, the largest of which is NleG and has 14 representatives in this strain. Savchenko and colleagues recently identified an ubiquitin E3 ligase function in the C-terminal region of NleG. Ubiquitination is an essential eukaryotic function that involves attachment of an ubiquitin polypeptide to a lysine residue of a target protein. The targeted protein is either destined for degradation by the host proteasome or its function is regulated. It has now been demonstrated that the NleG family function in vitro as E3 ubiquitin ligases by acting as a scaffold between E2 enzyme and the target protein in ubiquitination. The target protein(s) as well as the role in pathogenesis of NleG effectors have yet to be identified. The few previous studies on effectors with E3 ubiquitin ligase function suggest they are involved in suppression of the host immune response by targeting immune-related host proteins to proteasome degradation.

Of the many different effector genes that have been identified within pathovars of E. coli only a handful have known host cell target sites and to date, most have unknown functions. Both Tir and the phage encoded EspF (aka TccP) effector contain mitochondrial targeted sequences (MTS) within their N-terminal regions and a recent study suggested that EspF targets and disrupts nucleolus function by a mitochondrial based mechanism (Fig. 4). EspF as well as Map have also been shown to disrupt tight junctions resulting in induction of apoptosis. The phage-encoded NleA and its homologs were shown to target the cell tight junctions (disrupting cellular barrier function) as well as mediate trafficking to the Golgi. 

Salmonella enterica can cause acute infections, ranging from gastroenteritis to enteric fever depending on the serovar and strain examined. Salmonella enterica is a facultative intracellular food-borne pathogen that causes morbidity and mortality worldwide as well as economic losses due to food recalls. This bacterium can colonize birds, reptiles and mammals and is subdivided into seven subspecies: I, II, IIIa, IIIb, IV, VI and VII. Infections in warm-blooded animals such as humans are generally caused by subspecies I isolates, which encompass more than 2,300 different serovars. Salmonella enterica serovar Typhimurium is one of the most common causes of food-borne infections causing inflammatory diarrhea.

The key virulence genes of Salmonella enterica serovar Typhimurium are encoded on mobile and integrative genetic elements that were acquired by horizontal gene transfer. Numerous prophages and pathogenicity islands have been identified and characterized from Salmonella Typhimurium genome sequences. Five pathogenicity islands named Salmonella pathogenicity island 1 (SPI-1) to SPI-5 have been identified within the genome. Two of these pathogenicity islands, SPI-1 and SPI-2, contain T3SSs and effector proteins and are present in all Salmonella enterica strains. Eight different phages or phage-like region have also been characterized and shown to carry effector proteins and are variably present among strains.

Pathogenicity Islands (PAIs) are a group of integrative elements which encode multiple virulence factors that confer an advantage to the bacterium within its target eukaryotic host and were first described in an E. coli uropathogenic strain. PAIs integrate into host genome usually at tRNA sites by site-specific recombination mediated by the PAI-encoded integrase. In E. coli, PAI-encoded integrases have been shown to be unrelated.
to integrases present in phages indicating that phages and PAIs are diverse elements not closely related to one another.\textsuperscript{86,87} SPI-1 and SPI-2 encode multiple virulence genes involved in invasion and enteropathogenesis including T3SSs and several effector proteins (Fig. 5).\textsuperscript{72,88,89} SPI-3 and SPI-4 are much smaller regions and appear to have a less significant role in pathogenesis. The SPI-5 region carries a number of effector proteins that are secreted by both T3SSs encoded on SPI-1 and SPI-2 (Fig. 6).\textsuperscript{78}

Among \textit{S. enterica} strains, approximately 30 different effector proteins have been identified, most of which are encoded within the same region that contains their cognate T3SS (SPI-1 and SPI-2).\textsuperscript{72} As was the case for \textit{E. coli}, the \textit{S. enterica} phage-encoded effector genes are found at the terminal region of the phage genome but generally only occur as a single gene within the phage genome (Fig. 3B).\textsuperscript{3} Serovar Typhimurium effectors secreted by T3SS-1 trigger membrane ruffling and internalization within a vacuole and are required for host cell invasion (Fig. 6). These effectors activate cytoskeleton rearrangement by causing actin nucleation, which disrupts the microvilli on the apical cell surface and aid in bacterial uptake, creating a Salmonella containing vacuole (SCV) within the host cell cytosol (Fig. 6). Salmonella then secretes effector proteins into the cytosol to modulate intracellular survival, first by preventing lysosome fusion with the SCV (Fig. 6).\textsuperscript{91} Both T3SSs are involved in translocating proteins from the SCV into the host cell cytosol, many of which are required for modulation of the host immune response.\textsuperscript{90,91} Among \textit{S. enterica} strains examined nine effector proteins are found on prophage genomes, SseI within Gifsy-2\(\phi\), GogB and GipA within Gifsy-1\(\phi\), a homology of SspH1 on Gifsy-3\(\phi\), SseK3 on phage ST64B, SopE within SpoE\(\beta\), and SopE2, SspH2 and SseI on a phage-like region.\textsuperscript{3} T3SS-1 translocated effectors SopE and SopE2 and T3SS-2 translocated effectors SseI and SspH1 are required for modulation of the host immune response.\textsuperscript{90} SopE and SopE2 are also involved in activating Rho GTPases by mimicking guanine exchange factors.\textsuperscript{90}

\textbf{Conclusion}

Key to effector protein function is their delivery into their target cell. Bacteria encode three contact-dependent secretion systems named T3SS, T4SS and T6SS that translocate effector proteins directly from the bacterial cytosol into the eukaryotic cytosol. Interestingly the evolutionary origin of these three systems is very different as are the effector proteins translocated by each. The T3SS shows homology to the bacterial flagellar system, the T4SS shows homology to the bacterial conjugation system and the T6SS shows homology to a contractile phage tail-like structure. It was shown that this structure contains an exterior sheath wrapped around an interior tube, similar to that in phage T4. However, the structural similarities do not translate into significant primary sequence homology and it remains to be determined whether functional similarity exists.\textsuperscript{92} The T6SSs are a recent
discovery and only a handful of effector proteins have been identified and one could speculate that some of their effector proteins may be phage encoded. Although effector proteins identified to date number nearly 150, the function(s) of the vast majority of these proteins remains unknown. In addition as more species and strains are sequenced the number of phage-encoded effector proteins is sure to increase and with it our knowledge of how phages encoding cargo genes interact with their host genomes and the host regulatory systems.

References

1. Boyd EF, Briisou H. Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. Trends Microbiol 2005; 13:519-21; PMID:15943306; http://dx.doi.org/10.1016/j.tim.2005.07.002.
2. Briisou H, Chanchaya C, Hardt WD. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysozymic conversion. Microbiol Mol Biol Rev 2006; 68:560-602; PMID:15553570; http://dx.doi.org/10.1128/MMBR.68.3.560-602.2004.
3. Elharb K, Hardt WD. Bacteriophage-encoded type III effectors in Salmonella enterica subspecies 1 serovar Typhimurium. Infect Genet Evol 2005; 5:1-9; PMID:15761133.
4. Boyd EF. Bacteriophage-encoded bacterial virulence factors and phage-pathogenicity island interactions. [In Press]. Adv Virus Res 2012; 82:91-118; PMID:22420852; http://dx.doi.org/10.1016/B978-0-12-394621-0.00014-5.
5. Holmes RK, Barkdale L. Genetic analysis of tox- and tox-like bacteriophages of Corynebacterium diphtheriae. J Virol 1969; 5:58-96; PMID:4978942.
6. Boyd J, Oza MN, Murphy JR. Molecular cloning and STX-2 secretion of the diphtheria tox- bacteriophages of Corynebacterium diphtheriae. J Bacteriol 2000; 182:5530-5; PMID:10764646; http://dx.doi.org/10.1128/JB.182.17.5530-5.2000.
7. Tao X, Boyd J, Murphy JR. Specific binding of the diphtheria tox regulatory element (dtxR) from Corynebacterium diphtheriae by the TetR family repressor DtxR to the tox operon. Proc Natl Acad Sci U S A 1990; 87:5968-72; PMID:2116013; http://dx.doi.org/10.1073/pnas.87.15.5968.
8. Hallberg DR, Daniel JH, Tarr PI. Enterotoxin gene fusions to identify a pilus colonization factor of Escherichia coli. Proc Natl Acad Sci U S A 1984; 81:11687-91; PMID:2292187; http://dx.doi.org/10.1073/pnas.81.24.11687.
9. Davis BM, Lawon EH, Sandkivist M, Ali A, Sozhamannan S, Waldor MK. Convergence of the secretory pathways for cholera toxin and the ToxR regulon are essential for Vibrio cholerae pathogenesis in humans. J Exp Med 1998; 146:1847-92; PMID:9084476; http://dx.doi.org/10.1084/jem.146.1.1847.
10. Hofmeyr DT, Jansen SA, Ursonian Y, Saunders JR, McCarthy AJ, et al. Short-tailed stx phages exploit the conserved YaeT protein to disseminate Shiga toxin genes among enterobacteria. J Bacteriol 2001; 183:5109-18; PMID:11450839; http://dx.doi.org/10.1128/JB.183.13.5109-18.2001.
11. Val ME, Bouvier M, Campos J, Sherratt D, Cornet D, Menges, Jean Bernard Lubin, Serge Ongagna-Yhombi and W. Brian Whittaker. Work in the Boyd group is supported by a National Science Foundation CAREER grant DEB-0844409. Literature citations have been limited in many cases to review articles due to space consideration and we apologize to those authors whose primary research is not included or cited.

Acknowledgments

We thank members of our group for their enthusiasm and hard work: Laura Powell, Sai S. Kalburge, Luke Raber, Brandy Haines Menges, Jean Bernard Lubin, Serge Ongagna-Yhombi and W. Brian Whittaker. Work in the Boyd group is supported by a National Science Foundation CAREER grant DEB-0844409. Literature citations have been limited in many cases to review articles due to space consideration and we apologize to those authors whose primary research is not included or cited.
40. Feng PC, Monday SR, Lacher DW, Allison L, Siitonen A; Genomic profiling for detection of Escherichia coli O104:H4 outbreak strains. PLoS One 2011; 6:e22751; http://dx.doi.org/10.1371/journal.pone.0022751.

41. Wick LM, Qi W, Lacher DW, Whittam TS. Evolutionary model of Shiga-toxin-producing Escherichia coli. J Bacteriol 2011; 193:3652-61; PMID:21669121; http://dx.doi.org/10.1128/JB.00868-11.

42. Zhang W, Qi W, Albert TJ, Motiwala AS, Alland D, Huytia-Trees EK, et al. Prospective genomic characterization of Escherichia coli O104:H4 strains from nine genomes. J Clin Microbiol 2011; 49:3714-20; PMID:21880963; http://dx.doi.org/10.1128/JCM.02808-10.

43. Feng P, Abel KA, Karch H, Whittam TS. Genotypic and phenotypic changes in the emergence of Shiga toxin-producing Escherichia coli O157:H7. Emerg Infect Dis 1998; 4:615-7; PMID:9866737; http://dx.doi.org/10.1044/0715-4779.04.11.

44. Frank C, Werber D, Cramer JP, Askar M, Faber M, an den Heiden M, et al.; HUS Investigation Team. Epidemic spread of Shiga toxin-producing Escherichia coli O104:H4 outbreak in Germany. N Engl J Med 2011; 365:1770-5; PMID:21607864; http://dx.doi.org/10.1056/NEJMoa1107138.

45. Whittam TS, Reid SD, Selandar RK. Mutators and long-term evolution of pathogenic Escherichia coli. J Infect Dis 1998; 177:1750-3; PMID:9607864; http://dx.doi.org/10.1086/317438.

46. Denamur E. The 2011 Shiga toxin-producing Escherichia coli O104:H4 German outbreak: a lesson in genomic plasticity. Clin Microbiol Infect 2011; 17:1124-5; PMID:21781204; http://dx.doi.org/10.1111/j.1469-8069.2011.03620.x.

47. Muniesa M, Hammer LA, Hertwig S, Appel B, Brussow H. Shiga toxin-producing Escherichia coli O104:H4: a new pathogen? Microbiology. Appl Environ Microbiol 2012; 78:4065-73; PMID:22694186; http://dx.doi.org/10.1128/AEM.01127-12.

48. Mehlman A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Riao A, et al. Prospective genomic characterization of Salmonella enterica serovar Typhimurium DT104. J Bacteriol 2011; 193:4064-73; PMID:21607863; http://dx.doi.org/10.1128/JBC.02057-11.

49. Mills DM, Bajaj V, Lee CA. A 40 kb chromosomal fragment encoding Salmonella typhimurium invA is absent from the corresponding region of the Escherichia coli K-12 chromosome. Mol Microbiol 1995; 19:754-9; PMID:7979582; http://dx.doi.org/10.1111/j.1365-2958.1995.tb01382.x.

50. Mills DM, Bajaj V, Lee CA. A 40 kb chromosomal fragment encoding Salmonella typhimurium invA is absent from the corresponding region of the Escherichia coli K-12 chromosome. Mol Microbiol 1995; 19:754-9; PMID:7979582; http://dx.doi.org/10.1111/j.1365-2958.1995.tb01382.x.

51. Mills DM, Bajaj V, Lee CA. A 40 kb chromosomal fragment encoding Salmonella typhimurium invA is absent from the corresponding region of the Escherichia coli K-12 chromosome. Mol Microbiol 1995; 19:754-9; PMID:7979582; http://dx.doi.org/10.1111/j.1365-2958.1995.tb01382.x.
85. Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. Nat Rev Microbiol 2004; 2:414-24; PMID:15100694; http://dx.doi.org/10.1038/nrmicro884.

86. Boyd EF, Almagro-Moreno S, Parentt MA. Genomic islands are dynamic, ancient integrative elements in bacterial evolution. Trends Microbiol 2009; 17:47-53; PMID:19162481; http://dx.doi.org/10.1016/j.tirm.2008.11.003.

87. Napolitano MG, Almagro-Moreno S, Boyd EF. Dichotomy in the evolution of pathogenicity island and bacteriophage encoded integrases from pathogenic Escherichia coli strains. Infect Genet Evol 2011; 11:423-36; PMID:21147268; http://dx.doi.org/10.1016/j.meegid.2010.12.003.

88. Galán JE. Molecular and cellular bases of Salmonella entry into host cells. Curr Top Microbiol Immunol 1996; 209:43-60; PMID:8742245; http://dx.doi.org/10.1007/978-3-642-85216-9_3.

89. Shea JE, Hensel M, Gleson C, Holden DW. Identification of a virulence locus encoding a second type III secretion system in Salmonella typhimurium. Proc Natl Acad Sci U S A 1996; 93:2593-7; PMID:8637919; http://dx.doi.org/10.1073/pnas.93.6.2593.

90. McGhee EJ, Brawn LC, Hume PJ, Humphreys D, Koronakis V. Salmonella takes control: effector-driven manipulation of the host. Curr Opin Microbiol 2009; 12:117-24; PMID:19157959; http://dx.doi.org/10.1016/j.mib.2008.12.001.

91. Ibarra JA, Steele-Morrimer O. Salmonella—the ultimate insider. Salmonella virulence factors that modulate intracellular survival. Cell Microbiol 2009; 11:1579-86; PMID:19775254; http://dx.doi.org/10.1111/j.1462-5822.2009.01368.x.

92. Silverman JM, Brunet YR, Cascales E, Mouzous JD. Structure and Regulation of the Type VI Secretion System. [Epub ahead of print]. Annu Rev Microbiol 2012; PMID:22746332.