PROPHAGE ARSENAL OF *SALMONELLA ENTERICA* SEROVAR TYPHIMURIUM

Lionello Bossi and Nara Figueroa-Bossi

8

"There can be little reasonable doubt that the lysogenic bacteria or, more accurately, the symbiotic phages they harbour, are the origin and the continuing source of all [temperate] phages in nature."
J. S. K. Boyd, 1950

The genus *Salmonella* comprises a large collection of enteric bacteria that infect a wide range of animals from reptiles to mammals. Most isolates from warm-blooded animals are grouped into a single species, *Salmonella enterica*, and classified according to their antigenic formulas (serovars) (104). The existence of >2,500 serovars illustrates the high degree of diversity of this bacterial species. *S. enterica* serovars differ greatly in their host range and in the type of disease they cause. Some serovars can infect hosts as distantly related as birds and humans; others show variable degrees of adaptation to specific hosts (68, 106, 124). The broad-host-range serovars Typhimurium and Enteritidis represent a primary source of food-borne disease in humans and livestock (8, 68, 106). They cause a self-limiting gastroenteritis in most of their hosts and a systemic disease resembling typhoid fever in rodents. On the opposite side of the spectrum, serovar Typhi, the causative agent of typhoid fever, is an exclusively human pathogen (99). The wealth of knowledge that has accumulated during more than half a century of use of *S. enterica* serovar Typhimurium for genetic research has made this organism a favored model system for the study of pathogenicity. This work has shed light on the molecular mechanisms that allow bacteria to invade host cells and elude the immune response. Many of the genes that compose the bacterium's pathogenic arsenal are organized into discrete units called the *Salmonella* pathogenicity islands (SPI) (10, 23, 43, 47, 96, 115, 137, 139; for a recent review, see reference 112); other genes are scattered around the chromosome as individual loci or as small clusters (44). This archipelago of virulence-related loci is thought to result for the most part from horizontal acquisition, making the study of virulence intimately linked to that of genome evolution and differentiation (7, 43, 69, 97). The emergence in the last two decades of new epidemic *Salmonella* strains with enhanced virulence traits is indicative of the fast pace of the evolutionary process (122). The purpose of this chapter is to review evidence

Lionello Bossi and Nara Figueroa-Bossi, Centre de Génétique Moléculaire, CNRS, 91198 Gif-sur-Yvette, France.

Phages: Their Role in Bacterial Pathogenesis and Biotechnology
Edited by Matthew K. Waldor, David I. Friedman, and Sankar L. Adhya

pointing to a central role played by temperate phages in the dissemination of virulence determinants in the *Salmonella* complex.

The lysogenic condition of most *Salmonella* strains was recognized prior to an understanding of the genetic bases of lysogeny. J. S. K. Boyd described the release of different types of "symbiotic" phages from clinical isolates of *S. enterica* serovar Typhimurium (14). These phages appeared to be in a state of latency in the vast majority of cells in culture. Cells harboring symbiotic phages could occasionally give rise to active viruses, forming plaques on suitable indicator strains, while the donor strain was immune to superinfection (14). We know now that Boyd's symbiotic phages were integral parts of the host chromosome (prophages) and that their latency status reflected the repression of most of the viral functions. Early studies also indicated that some genes of certain prophages escape lysogenic repression and express functions that modify the host bacterium. The modification of somatic antigens by prophage-encoded proteins constitutes the earliest example of lysogenic conversion in *Salmonella* (48, 63, 142, 148). Modification often affects the receptor of the converting phage and thus renders the lysogenic bacterium resistant to that phage and related phages. In one study, possibly the first report linking a phage to *Salmonella* pathogenicity, a prophage-mediated increase in O antigen chain length was shown to result in enhanced serum resistance and mouse virulence (95). Overall, the study of *Salmonella* phages during the second part of the past century has largely focused on phage P22. While this work has contributed enormously to our current knowledge of phage biology and evolution, it may have suppressed the interest in exploring the diversity of the *Salmonella* phage-prophage pool suggested by Boyd's seminal studies. An appreciation for such diversity has occurred only recently due to a combination of fortuitous findings and whole-genome sequence analyses.

A conspicuous feature of phage genomes is their extensive genetic mosaicism (13, 21, 54, 55, 65). There can be no reasonable doubt that DNA shuffling constitutes a primary force driving the evolution of these elements. The profusion of broken or fragmented "tiles" in the mosaics is indicative of the inherent sloppiness of the remodeling process. Indeed, the sequence organization near the right ends of some of the prophages described here underscores the restraint in a popular quote by S. Brenner: "Anything that is produced by evolution is bound to be a bit of a mess." Mosaicism raises a problem for phage nomenclature. Prophages found at identical positions in closely related strains often contain different modules and encode different types of immunity. Conversely, prophages located at different chromosomal locations can share extensive sequence identity throughout most of their genomes. The only way to identify a phage unambiguously is to name it after the strain from which it was originally isolated. However, this solution seems unappealing and loses accuracy upon propagation of the phage. For the sake of simplicity, in this chapter we use the name given in the initial description of a phage for all phages found at the same chromosomal location as the original (except when multiple designations already existed). Nonetheless, one should bear in mind that homonymous phages are not necessarily identical if they are isolated from different strains.

Figure 1 shows a schematic representation of the *Salmonella* chromosome indicating the positions of all functional prophages identified in serovar Typhimurium strains thus far. Interestingly, all of these elements are oriented according to the polarity of bidirectional chromosomal DNA replication, proceeding always from *attL* to *attR*. Most strains carry a subset of these prophages (typically between four and five) in a variable assortment. This variability reflects differences in prophage distribution. Some prophages are found in all strains (Gifsy-1 and Gifsy-2), others have a more limited distribution (SopEΦ, St64B, and P22), and still others are specific to certain isolates (Fels-1 and Gifsy-3). Some lines of evidence indicate that the "rare" prophages may occur frequently in other *S. enterica* serovars or subspecies, suggesting that they were acquired during occasional excursions of *S. enterica* serovar Typhimurium

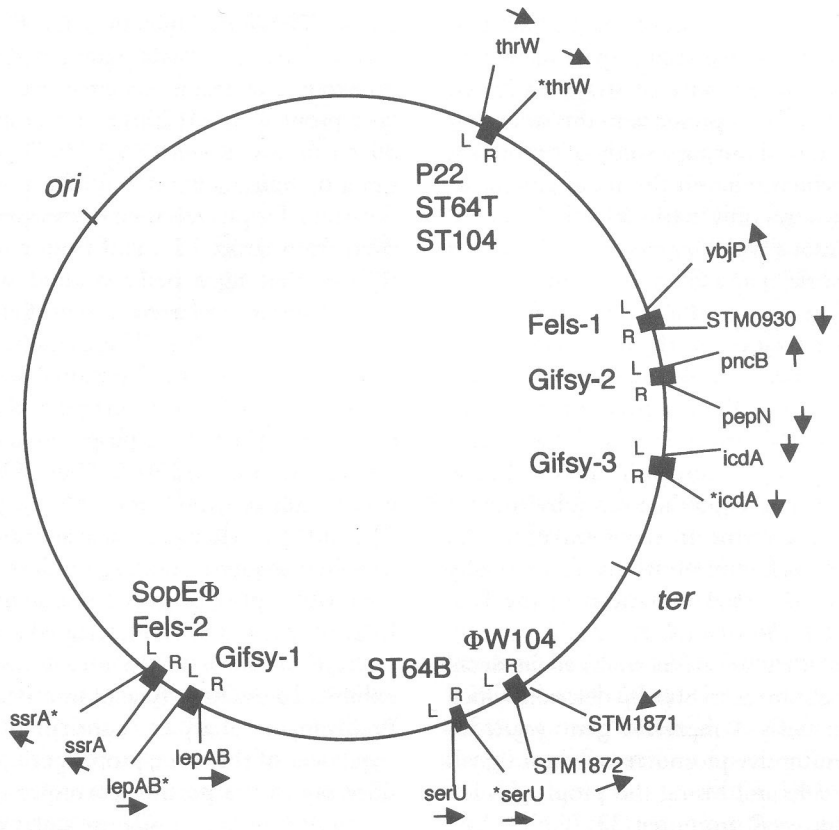

FIGURE 1 Schematic diagram showing the prophages in the *S. enterica* serovar Typhimurium chromosome. The prophage left-right orientation (L/R) is shown according to the convention used for the prophage map of bacteriophage lambda (120). Genetic symbols specify the genes flanking the insertion sites, with arrows indicating their orientations. An asterisk on the left or right side of the symbol indicates that the gene is truncated at its 5′ end or its 3′ end, respectively.

strains into the reservoirs of these lineages. Still, phage circulation among different serovars is constrained by restriction barriers. Interestingly, the effectiveness of these barriers appears to correlate with the host range of the serovar. Phages released from the broad-host-range serovars Typhimurium and Enteritidis can efficiently multiply in and lysogenize the host-adapted serovars Typhi, Gallinarum, and Abortusovis. However, when released from these newly lysogenized strains, the very same phages can no longer grow in their initial donors (39).

Since phage and chromosomal sequences near the attachment sites of most prophages are conserved, PCR can be used to assess the phage occupancy of these sites. This approach is particularly attractive because the reaction can be designed in such a way as to always give a signal, and the presence or absence of the prophage can be deduced from the size of the amplified fragment. We have used this type of analysis extensively in our studies and will refer to it as *att*-site PCR.

Gifsy PHAGES

(i) Discovery

Gifsy-1 and Gifsy-2 are lambdoid phages that have been found in the chromosomes of all *S. enterica* serovar Typhimurium strains analyzed

thus far. The DNA sequences of the two prophages were determined as part of the sequencing of the genome of strain LT2 (83) (Color Plate 1). Their presence in this strain was originally inferred during a study of suppressor mutations which relieved the requirement for RecBCD for recombinational DNA repair. These mutations were mapped to Gifsy-1 (at a locus named *sbcE*) and expressed their suppressor phenotype only in the presence of Gifsy-2 (33). Characterizations of the mutants unveiled the nature of the two elements and provided some insights into their regulation. The suppression of *recBCD* defects by *sbcE* mutations results from the activation of a Gifsy-1-borne *recE*-like gene, whose product can substitute for the RecBCD enzyme in recombination and repair. Thus, *sbcE* mutations are functionally equivalent to the *sbcA* mutations of the Rac prophage of *Escherichia coli* K-12 (72, 79, 80). One mutant that has been analyzed in detail (*sbcE21*) results from an 800-bp deletion which disrupts the Gifsy-1 repressor gene *gogR*, removing presumptive promoter-operator signals from its 3′ side and fusing the prophage's left operon to the *gogR* promoter (33) (Color Plate 1). Intriguingly, this promoter was found to depend on the presence of Gifsy-2 for its activity, thus explaining the requirement for Gifsy-2 for *sbcE*-mediated suppression. The product responsible for this activation was recently identified as the Gifsy-2-encoded repressor protein GtgR (75). The *gtgR* gene is a perfect duplicate of the *gogR* gene and is located in a region where the two prophages share complete sequence identity (Color Plate 1). Thus, these results suggest that the GogR and GtgR proteins autogenously activate their own expression, similar to the phage λ *c*I repressor (105). It is noteworthy that *sbcE* selection would not have revealed the existence of Gifsy-1 and Gifsy-2 if the two prophages had not been homoimmune.

(ii) Inducibility and Immunity

Concomitant to derepressing the *recE* gene, the *sbcE21* deletion activates the transcription of the Gifsy-1 *xis* gene, thus causing the prophage to become unstable and to excise at a high frequency. The *sbcE21* mutation destabilizes Gifsy-2 as well, but to a lesser extent, making it possible to isolate strains that are cured of either or both prophages (33). This pattern can be reproduced in any *S. enterica* serovar Typhimurium strain by transducing the *sbcE21* mutation into the strain. Prophage-cured derivatives were isolated from strain LT2 and from two virulent isolates that have been studied worldwide, ATCC 14028s (referred to hereafter as 14028) (32) and SL1344 (61). These strains were used as hosts to assess the functional status of the Gifsy-1 and Gifsy-2 prophages in their respective parents. The Gifsy-2 phage could only be recovered from strain 14028. Thus, in spite of appearing full-size, the Gifsy-2 prophage in strains LT2 and SL1344 must contain alterations that affect their ability to undergo induction and/or form visible plaques. In contrast, all three strains, LT2, 14028, and SL1344, released active Gifsy-1 phage, but surprisingly, in each case the virus exhibited a different type of immunity (Fig. 2). Preliminary analyses confirmed that the sequences of the three prophages diverge considerably in the portion corresponding to the immunity module. These modules can be exchanged by recombination between homologous flanking regions, explaining why the *sbcE* phenotype could be reproduced in strains other than LT2. A recombination-conversion event of the same type might have been involved in the transfer of Gifsy-2 immunity determinants to Gifsy-1 in strain LT2 (Fig. 2).

(iii) Attachment Sites and Occurrence

The integration site of phage Gifsy-1 in the *Salmonella* chromosome lies within the *lepA* gene. This gene encodes a membrane-associated GTPase of unknown function and is cotranscribed with the *lepB* gene for signal peptidase I (81, 130). The Gifsy-1 attachment site is defined by a 14-bp core sequence present both in the phage and in the chromosome and duplicated in direct order at the two ends of the prophage (Color Plate 2). Unlike most phage insertions within genes, which typically target the 3′ end of the transcribed region, the Gifsy-1 attachment site lies on the 5′ side of *lepA*. The

Host Strain Background ⇩	Donor Strain ⇨	LT2	14028			SL1344
	Phage ⇨	Gifsy-1	Gifsy-1	Gifsy-2	Gifsy-3	Gifsy-1
	Prophage ⇩					
LT2	Gifsy-1	●	○	●	○	○
	Gifsy-2	●	○	●	○	○
14028	Gifsy-1	○	●	○	○	○
	Gifsy-2	●	○	●	○	○
	Gifsy-3	○	○	○	●	●
SL1344	Gifsy-1	○	○	○	●	●
	Gifsy-2	●	○	●	○	○

FIGURE 2 Immunity relationships among Gifsy phages from three representative strains of *S. enterica* serovar Typhimurium. Phages isolated from the indicated strains were used to infect strains that carried or lacked their resident Gifsy prophages. Open circles, phage forms plaques on the strain carrying the specified prophage; closed circles, phage does not form plaques unless the strain is cured of the specified prophage. Phage Gifsy-2 could not be obtained from strains LT2 and SL1344.

insertion does not inactivate the gene because a sequence encoding a polypeptide identical to the N-terminal portion of the LepA protein is found adjacent to the crossover site in the phage DNA and provides the *lepA* gene with a new 5′ end upon integration (Color Plate 2). This integration strategy is analogous to that of the SXT element of *Vibrio cholerae*, which inserts at the 5′ end of the *prfC* gene for the translation release factor RF-3 (60). The RF-3 protein is also a GTPase and shows similarity to LepA, particularly in its GTP-binding domains (82). In both genes, the attachment core sequences correspond to the regions encoding the first of these domains. In both systems, the foreign element appropriates the control of the target gene upon integration, raising the possibility that lysogenization by Gifsy-1 results in different regulation of the *lep* operon. However, the levels of epitope-tagged LepA and LepB proteins were found to be closely comparable in cells carrying or lacking the prophage under different conditions (in vitro or inside epithelial cells) (127).

The chromosomal attachment site of the Gifsy-2 phage lies in the intercistronic region between the *pncB* and *pepN* genes. Integration of the phage results in the duplication of a 15-bp sequence (5′-TTATAAAAATGTAGC-3′). A PCR-based survey of the occupancy of the Gifsy-2 *att* site detected an insert in the vast majority of *S. enterica* serovars (37). However, from the few cases for which DNA sequence data are available, it appears that the identity of the inserted element within Gifsy-2 breaks down past a distance from both ends of the prophage. In a serovar Abortusovis isolate, an IS*1414* transposon encoding a heat-stable enterotoxin replaces a segment at the right end of the prophage, including two tail genes (5). More dramatically, the element occupying the Gifsy-2 *att* site in serovar Typhi is only distantly related to its *S. enterica* serovar Typhimurium counterpart. The identity between the two elements is limited to the regions specifying the integration functions (*attP*, *int*, and *xis* genes) and immunity. In place of the *recE-recT* module, the *S. enterica* serovar Typhi prophage contains a recombination module similar to the phage λ *red* operon (27, 98). The lack of recognizable head and tail genes in the *S. enterica* serovar Typhi prophage suggests its defective nature (26, 98).

Although it is seldom found in serovars other than Typhimurium, the Gifsy-1 prophage is nearly always present in strains from this serovar. An analysis of 72 *S. enterica* serovar Typhimurium

strains isolated in France from human or animal sources found that all of them contained Gifsy-1, and this prophage was also present in 19 of 21 strains from the *Salmonella* Reference Collection A (SARA) (9, 37).

(iv) Regulatory and Structural Features

The genome organization of the Gifsy-1 and Gifsy-2 phages is typical of members of the lambdoid family. Both prophages are induced in response to DNA-damaging treatments. Induction requires the product of the *recA* gene, suggesting that it results from the RecA-mediated cleavage of the respective repressors, like in most lambdoid phages (109). Yet some observations point to an additional intricacy in the relationships between Gifsy prophage induction and the global response to DNA damage. In enteric bacteria, this response involves the activation of transcription of the SOS regulon, which includes genes for the repair of damaged DNA (133). Activation results from cleavage of the repressor protein LexA mediated by the RecA* protein. A member of the SOS regulon, the *dinI* gene, encodes a small protein (81 amino acids [aa]) that negatively regulates the activation cascade by diverting the RecA* protein from LexA cleavage (132, 147). Interestingly, Gifsy-1 and Gifsy-2 both harbor a *dinI* gene homologue that is directly controlled by LexA (19) (Color Plate 1). An epitope-tagged version of the DinI protein from the Gifsy-1 prophage was shown to accumulate rapidly following exposure to the DNA-damaging agent mitomycin C (3). These findings raise the possibility that the Gifsy prophages can negatively modulate the SOS induction signal.

Gifsy-1 and Gifsy-2 have been difficult to study because of their fragility and poor plaque-forming efficiencies. The latter results in part from inefficient receptor recognition. Ho and Slauch (57) identified the major outer membrane porin protein OmpC as the receptor for both phages. They found that phage adsorption and the plaque-forming efficiency could be increased significantly when a *galE* mutation was present in the recipient strain (58). Since the *galE* gene product is required for the synthesis of the outer core and the O antigen of lipopolysaccharide, these results suggest that the O antigen partially blocks OmpC recognition. With *galE* mutants as hosts, Gifsy-1 and Gifsy-2 virions were obtained in sufficient amounts to allow for structural studies. Negative-stain electron microscopy revealed the Gifsy-2 capsid to be icosahedral and ~550 Å in diameter. The Gifsy-2 capsid is connected to a long flexible tail, much like the coliphages λ and HK97 (29). N-terminal sequencing of the Gifsy-2 major coat protein revealed that this protein corresponds to the C-terminal part of the predicted product of locus STM1033 (83) (Color Plate 1). This 703-aa polypeptide carries a Clp protease motif in its N-terminal portion, suggesting that the mature coat protein is generated by proteolytic cleavage around position 399 of the STM1033 polypeptide. The existence of similar protease-coat fusion arrangements in other prophages (including the *Salmonella* phage Fels-1 [see below] and the *E. coli* phages CP-933K and CP-933U) suggests that this represents a novel head shell construction paradigm for lambdoid phages (20).

(v) Role in *Salmonella* Pathogenicity

Initial evidence for the contribution of both the Gifsy-1 and Gifsy-2 prophages to *Salmonella* pathogenicity came from mouse infection experiments with strains that were cured of either or both elements (34). Strains lacking Gifsy-2 appeared to be significantly attenuated in mice, regardless of whether the bacteria were administered orally or inoculated through the intraperitoneal route. Strains with a deletion of Gifsy-1 were attenuated as well, but in this case the effects were much less severe and were dependent on the strain background and the route of infection. Subsequent work identified most of the prophage genes responsible for these effects as well as additional loci that can be linked to pathogenicity by in vitro assays or sequence analogies (34, 35, 59, 117). These various genes are found scattered throughout the genomes of the two prophages and tend to concentrate at their right ends, at a position corresponding to the dispensable "b" region in phage λ (24). This

region also contains truncated stretches of identity with various genes or elements, including genes involved in DNA transposition and inversion, suggesting that it has been subject to extensive DNA scrambling.

Gifsy-1

Thus far, the Gifsy-1 prophage has been the only such element whose implication in mouse pathogenesis can be linked to the intestinal phase of the infection. Using in vivo expression technology, J. Slauch and coworkers identified a Gifsy-1 locus that is transiently and specifically induced during *Salmonella* colonization of the small intestine (117). This locus, named *gipA*, is located in the phage tail operon and transcribed in the opposite direction (Color Plate 1). The *gipA* gene is needed for the proliferation or survival of bacteria in Peyer's patches, a primary site of infection in the small intestine. A *gipA* null mutant was slightly attenuated in mice when delivered by the oral route but showed no virulence defects when inoculated intraperitoneally. To date, the function of GipA remains unknown. The protein shows similarity to a family of DNA transposases; however, transposition was shown not to be required for GipA's function in Peyer's patches (117). It has been suggested that GipA is a site-specific recombinase that regulates the expression of virulence genes (117).

Some lines of evidence suggest a further contribution to intestinal invasiveness by a factor(s) encoded in the Gifsy-1 b region. Cloning of a 5.6-kb fragment covering the interval between the Gifsy-1 *gogD* and *xis* genes (Color Plate 1) into the *pir*-dependent plasmid pGP704 (87) yielded a plasmid that specifically integrates at the Gifsy-1 *attB* site (39). A strain carrying this inserted element outcompeted the Gifsy-1-cured parental strain more than threefold in orally infected mice, whereas the two strains were recovered from spleens in comparable numbers when injected intraperitoneally (39). Experiments assessing the involvement of the two major b region genes, *gogB* and *gogC*, in these effects were inconclusive. The *gogB* gene encodes a 56-kDa protein that resembles several type III-translocated effectors of the leucine-rich repeat family (23a, 35). The gene is upregulated by the SPI-2-encoded SsrB activator and strongly induced in bacteria growing inside cultured epithelial cells (125). Transfection experiments with a vector expressing a *gogB-gfp* gene fusion showed that the hybrid protein localized to the cytoplasm of eukaryotic cells (23a, 74).

Although the Gifsy-1 prophage appears dispensable for the systemic phase of murine infection in a wild-type background, it plays a small but definite role in a strain that lacks Gifsy-2 (34). This suggests the presence in the Gifsy-1 genome of one or more loci that are functionally redundant with a gene(s) that is present in Gifsy-2. Alternatively, the defect resulting from the lack of Gifsy-1 may be subtle and require an attenuated background in order to be detected.

Gifsy-2

The initial finding of the involvement of the Gifsy-2 prophage in *S. enterica* serovar Typhimurium pathogenicity was closely followed by the recognition that the prophage carries the *sodC1* gene for periplasmic [Cu, Zn] superoxide dismutase. This enzyme had been previously shown to enhance *Salmonella* virulence by protecting bacteria against products of macrophage respiratory burst (26, 31). Thus, initial efforts concentrated on determining whether *sodC1* accounted for the entirety of the Gifsy-2 contribution to virulence. In one line of experiments, the *sodC1* gene was reintroduced as a single copy into the chromosome of a Gifsy-2-cured strain. The resulting strain remained significantly attenuated in mice, although it exhibited increased persistence in infected organs, causing an acute splenomegaly (34). A separate study, comparing the effects of deleting the *sodC1* gene to those of removing the entire Gifsy-2 prophage, showed that the *sodC1* single mutant was attenuated approximately 5-fold whereas the Gifsy-2-cured strain was attenuated >100-fold (59). Overall, these results confirmed the general role of *sodC1* in mouse pathogenesis and at the same time suggested the existence of one or more additional virulence

loci in the Gifsy-2 prophage genome. The completion of the Gifsy-2 genome sequence in strain LT2, combined with the use of the one-step inactivation technique for in vivo gene replacement (25), brought rapid progress in the search for this gene(s). Testing of strains carrying nested deletions of the Gifsy-2 genome determined that the region involved in mouse virulence was a segment at the far right end of the prophage (Color Plate 3). The deletion of a single open reading frame (ORF), named *gtgE*, was found to confer a sevenfold attenuation in competition assays, and the virulence defect of the *gtgE* mutant could be abolished by the reintroduction of the locus as a single copy (59). The *gtgE* sequence does not match any entry in the database, and to date no clues have been found concerning its function. Both the *sodC1* and *gtgE* genes are arranged in the opposite orientation relative to the Gifsy-2 late operon, and their expression is uncoupled from viral regulation. Both genes are expressed efficiently in the lysogen in vitro as well as in vivo (125, 126). Although inactivation of the *sodC1* or *gtgE* gene separately attenuates virulence to a moderate extent, when combined the two mutations act synergistically, causing a defect close to that of a strain with a deletion of Gifsy-2. This has led to the conclusion that the *sodC1* and *gtgE* genes account for most, if not all, of the contribution of the Gifsy-2 prophage to mouse virulence (59).

S. enterica serovar Typhimurium expresses a second periplasmic [Cu, Zn] superoxide dismutase, SodC2, which is encoded by a chromosomal gene. Inactivation of the *sodC2* gene was reported to sensitize bacteria to killing by peritoneal macrophages, suggesting that it also participates in protection against reactive oxygen species generated by the immune response (30, 108). However, subsequent studies failed to substantiate a requirement for the SodC2 protein for mouse virulence. Strains carrying *sodC2* deletions appeared unaffected in their ability to systemically infect mice (71, 126). The *sodC2* gene is highly transcribed in stationary-phase cultures grown in laboratory medium. Unlike the *sodC1* gene, *sodC2* is poorly expressed when bacteria proliferate intracellularly. Furthermore, even when expressed from the *sodC1* promoter, the SodC2 protein does not relieve the requirement for SodC1 for pathogenicity (3, 71). Thus, both transcriptional and enzymatic differences account for the nonequivalence of the two *sodC* genes in pathogenesis. It seems possible that the acquisition of *sodC1* has allowed the two [Cu, Zn] superoxide dismutases to diversify and become specialized for the different oxidative environments encountered by *Salmonella* bacteria.

The Gifsy-2 prophage also includes an antivirulence gene. A disruption of the *grvA* locus was shown to render *S. enterica* serovar Typhimurium more virulent in mice (58). The effect required the presence of a wild-type *sodC1* gene, suggesting that the *grvA* function is also somehow connected with the response to oxidative stress. However, the function of GrvA remains elusive. The *grvA* ORF lies between the presumptive capsid precursor gene (STM1033) and a tail protein homologue (*Z*) and is oriented in the opposite direction from that of these genes (Color Plate 1). Intriguingly, the strand opposite *grvA* also has coding potential (ORFs STM1034 and STM1035 by the strain LT2 annotation), raising the formal possibility that either or both of these ORFs are involved in the antivirulence phenotype.

Although a number of additional loci in the Gifsy-2 genome can be linked to virulence based on sequence analogies or other evidence, they are not required for pathogenesis in the mouse model (Table 1). These loci might encode redundant products or products whose functions are not important for mouse infection. A gene named *sseI* or *srfH* was identified independently in two laboratories as a gene specifying a type III-secreted protein (*sseI*) (85) and as being under the control of the SPI-2-encoded SsrA/SsrB two-component system (*srfH*) (140). This gene is strongly activated in bacteria that proliferate within host cells or in mouse tissues (125, 140). The SseI (SrfH) protein was shown to specifically interact with the actin cross-linking protein filamin through its N-terminal domain and to localize to the polymerizing actin cytoskeleton of eukaryotic cells (86). A

TABLE 1 *Salmonella* prophage genes that can be linked to pathogenicity

Phage	Locus	Link to virulence	Reference(s)
Gifsy-1	*gogA*	Similar to *pipA* gene of SPI-5	35, 139
	gipA	Needed for growth in Peyer's patches	117
	gogD	Similar to *pagJ* and *pagK*	35, 46
	gogB	Similar to type III secreted proteins of LRR family	35, 125
Gifsy-2	*gtgA*	Similar to *pipA*	35, 139
	grvA	Antivirulence gene	58
	ailT	Similar to attachment-invasion locus (*ail*) and to serum resistance proteins	35, 45, 53
	sodC1	Periplasmic [Cu, Zn] superoxide dismutase; protects against macrophage oxidative burst	26, 31, 34, 59, 126
	sseI(*srfH, gtgB*)	Type III translocated protein under SPI-2 control	85, 125, 140
	gtgE	Needed for mouse virulence	59, 125
	gtgF	Similar to macrophage survival gene (*msgA*)	45, 59
Gifsy-3	*pagJ*	PhoP/PhoQ-activated locus	35, 46
	SspH1	Type III translocated protein; downregulates interleukin-8	35, 49, 84, 123
Fels-1	*sodC3*	Periplasmic [Cu, Zn] superoxide dismutase; similar to *sodC1*	35, 125
	nanH	Neuraminidinase; involved in nutrient scavenging, host cell adhesion, and toxin action	35, 41, 42, 62, 64
SopEΦ	*sopE*	Type III translocated G nucleotide exchange factor; promotes epithelial cell invasions	50–52, 88

more recent study unveiled a different type of interaction involving the host factor TRIP6, a protein that localizes to the plasma membrane and regulates cell adhesion and motility. Based on these findings and on additional evidence, it was proposed that SseI (SrfH) stimulates the phagocyte-mediated systemic spread of bacteria by modulating TRIP6 activity (141). This model of active subversion of phagocytes to promote the dissemination of bacteria within the host may constitute a new paradigm for host-pathogen interplay.

Gifsy-3

During the initial characterization of the Gifsy-1 prophage, drug resistance markers located in the prophage genome were observed to recombine with a site distinct from both Gifsy-1 and Gifsy-2 when transferred by phage P22 transduction into strain 14028. Since no such class of recombinants was observed with other recipient strains, this suggested that 14028 carried an additional prophage that was homologous to Gifsy-1. The work that followed led to the identification of the Gifsy-3 phage. Although this phage remains incompletely characterized at the sequence level, partial sequence data confirmed its relatedness to Gifsy-1. Gifsy-3 virions are morphologically indistinguishable from Gifsy-1 and Gifsy-2 virions, and the phage is subjected to the same immunity as the Gifsy-1 phage from strain SL1344 (Fig. 2). As far as attachment is concerned, Gifsy-3 appears to be the *Salmonella* equivalent of coliphage 21. It inserts in the 3′ region of the isocitrate dehydrogenase gene (*icd*) at the same position as phage 21 and the e14 element in *E. coli* (15, 56, 134) (Color Plate 4). Like these two cases, Gifsy-3 insertion does not inactivate the *icd* gene. A 162-bp segment corresponding to the terminal 54 aa of the protein is present at the phage attachment site and becomes fused to the rest of the gene in the correct frame upon integration. *att*-site PCR analysis detected a Gifsy-3-related insert in 2 of the 21 *S. enterica* serovar Typhimurium strains from the SARA collection; however, none was found in a group of 72 clinical strains isolated in France in 2002 (37). Thus, the presence of Gifsy-3 in strain 14028 can be regarded as a highly specific feature of this strain.

The Gifsy-3 genome includes at least three genes that have been linked to pathogenicity. At the right end of the prophage map, one finds the *pagJ* locus, a member of the PhoP/PhoQ regulon (46), and the *sspH1* gene (84). The latter encodes a protein containing leucine-rich repeats (LRR) that is translocated into the host cell by both SPI I and SPI II type III secretion systems (84, 85). The SspH1 protein was shown to localize to the nuclei of mammalian cells and to interfere with the activation of transcription factors needed for the production of interleukin-8 (49). It was proposed that this interference with the host inflammatory response could promote pathogenesis. However, a deletion of the *sspH1* gene did not affect *Salmonella* virulence in a series of assays except in a strain that also lacked *sspH2*, which encodes another type III-translocated protein of the LRR family (123). Unlike either single mutant, the doubly deleted strain was significantly attenuated in its capacity to elicit enterocolitis in calves. A study of the occurrence of the *sspH1* gene across the *Salmonella* genus detected the sequence in only a few isolates from *S. enterica* subspecies I; in contrast, the locus was present at a considerably higher frequency in strains from other subspecies (123). The evolutionary significance of this distribution pattern remains elusive.

The disruption of a different locus in the middle of the Gifsy-3 genome impairs the ability of a strain to multiply inside cultured macrophages and epithelial cells. This phenotype is correlated with a defect in the ability of mutant bacteria to adapt to acid stress (22). The affected gene, *irsA* (intracellular response to stress), lies downstream of the *dinI* homologue of Gifsy-3 and shows significant sequence identity on the amino acid level with loci found at the corresponding position in several enteric phages, including Gifsy-1 and Gifsy-2. The *irsA* gene was shown to be highly upregulated within epithelial cells and macrophages and in serum-supplemented medium (22). The sequence of the IrsA protein includes a helix-turn-helix DNA binding motif in its C-terminal portion, suggesting its function as a transcriptional regulator. The protein may therefore be needed for the transcription of genes involved in the stress response. However, the possibility also exists that IrsA is in fact a negative regulator of prophage genes whose products are deleterious for bacterial growth.

Fels PHAGES

In the mid-1960s, N. Yamamoto, then at the Fels Research Institute in Philadelphia, Pa., identified two phages released by the *S. enterica* serovar Typhimurium laboratory strain LT2, which he named Fels-1 and Fels-2 (145, 146). Although morphologically and serologically unrelated to each other and to phage P22, both Fels phages could recombine at a low frequency with P22, yielding hybrid phages (143–146). Fels-1 and Fels-2 attracted little attention until recently, when their DNA sequences and sites of attachment were determined by the LT2 genome sequencing project (83).

Fels-1

The Fels-1 phage is a long-tailed lambdoid virus that bears considerable similarity to phages Gifsy-1 and Gifsy-2. Like these phages and Gifsy-3, it belongs to the *Siphoviridae* family (1). The phage is inserted between the loci *ybjP* and STM0930 of strain LT2 (83). The insert is flanked by the two imperfect 17-bp repeats 5′-TCCTTTCAGTGATTGCA-3′ (*attL*) and 5′-TCCTTTCAATGATAGCG-3′ (*attR*). The latter sequence is found in strains that lack the prophage, thus defining the core region of the chromosomal attachment site (*attB*). A survey of the occupancy of the Fels-1 *attB* site in 72 *S. enterica* serovar Typhimurium clinical isolates and in 21 strains from the SARA collection showed the site to be free of inserts in all but the LT2 entry of the collection (37). Further evidence presented below confirms the rarity of Fels-1 in serovar Typhimurium and suggests that the phage originated from a separate *Salmonella* subspecies.

Fels-1 phage appear to employ the same head construction strategy as that already described for Gifsy-2. The putative major capsid protein of Fels-1 is encoded by an ORF (STM0912) that specifies a Clp protease motif

at its 5′ end. Like Gifsy-2, Fels-1 also carries a [Cu, Zn] superoxide dismutase gene, *sodC3*, which is expressed in lysogenic cells at levels comparable to those of *sodC1* (125). However, some observations suggest that the SodC3 protein cannot complement a *sodC1* deletion in mouse virulence assays (35). At the far right end of the Fels-1 genome lies the *nanH* gene, which encodes neuraminidase, an enzyme capable of removing sialic acid from glycolipids, glycoproteins, and poly- and oligosaccharides (131). Neuraminidases are found in many bacterial pathogens and have been suggested to play a role in virulence through diverse mechanisms. NanH was proposed to improve bacterial survival inside the host by making sialic acid available as a carbon source (42, 91), to increase adhesion to host cells by decreasing mucus viscosity (64), and in *Vibrio cholerae*, to act synergistically with cholera toxin to facilitate binding penetration of the toxin to enterocytes (41, 112). However, to date, there is no direct evidence linking the *nanH* gene of Fels-1 to *Salmonella* pathogenicity. A survey of >200 *S. enterica* serovar Typhimurium strains, including 22 isolates from the original LT collection (77), found that only strain LT2 contains the gene. In contrast, *nanH* was present in about 60% of isolates from *S. enterica* subspecies III (formerly known as *Salmonella arizonae*) (62). Assuming a tight association of *nanH* with the Fels-1 prophage, these results suggest that strain LT2 acquired the prophage from a subspecies III strain.

Fels-2

The sequence of the Fels-2 prophage shows significant similarity to that of coliphage 186, a member of the P2 family. Members of this family have considerably smaller genome sizes (<35 kb) than those of the lambdoid group (>45 kb). Like those of phage P2, Fels-2 virion particles have rigid, tubular tails with a contractile sheath (a characteristic of *Myoviridae*) (1). A distinctive feature of the Fels-2 genome is the presence of a putative invertase gene (*pin*) in the tail fiber region, suggesting the presence of an inversion system analogous to that found in the e14 element (103) and in phage Mu (66) (Color Plate 5). The DNA sequence in the region to the left of the *pin* gene homologue contains two 18-bp segments with imperfect dyad symmetry that are noticeably similar to sequences found at the boundaries of the phage Mu invertible G segment (102). The two sites, 5′-CACATAC-CTCGGTTTAGG-3′ and 5′-CCTAAACC-GAGGTTTATG-3′, are located 2,792 and 58 bp upstream of the *pin* gene, respectively, suggesting that the 2,734 bp within this interval constitute the invertible segment. Interestingly, the outer edges of this segment contain nearly perfect 163-bp inverted repeats, raising the possibility that homologous recombination contributes to the inversion. The left boundary of the putative invertible segment lies within an ORF with a strong similarity to phage tail fiber genes. A sequence encoding a tail fiber motif is also found at the right boundary, but on the opposite strand, so that inversion would generate a tail fiber protein with an alternative carboxyl end and a different tail assembly protein. Thus, this system bears strong analogies to the mechanism that allows phage Mu to alternate its host range, switching between hosts as different as *E. coli* and *Citrobacter freundii* (128). Whether this inversion mechanism is active in Fels-2 remains to be determined. The finding that SopEΦ, a *Salmonella* phage related to Fels-2 (see below), has an inferred tail structure with the opposite configuration of that of Fels-2 (Color Plate 5) suggests that the switching mechanism did operate at some point in time and that host specificity changes occurred within the *Salmonella* genus.

Another feature of the Fels-2 genome is the presence of a homologue of the *tum* gene, which encodes the antirepressor of coliphage 186 (116). The sequence of this ORF is preceded by two potential binding sites for the LexA repressor protein (19), suggesting that, like that of phage 186 (16, 73), Fels-2 prophage induction is directly coupled to SOS induction. Derepression of this locus, resulting from mutations that inactivated the LexA protein, was shown to cause a lethal phenotype for *S. enterica* serovar Typhimurium strain LT2 (19).

The Fels-2 prophage insert is delimited by a 47-bp repeat which defines the *att* core sequence (83). This sequence corresponds to the terminal portion of the *ssrA* gene coding for transfer-messenger RNA (tmRNA), which targets nascent polypeptides for degradation under conditions of aberrant translation (136). Since the *att* core sequence includes the 3′ end of mature tmRNA, the presence of the prophage is not expected to inactivate the *ssrA* gene. However, different transcription termination signals are used in strains that carry or lack the prophage, making it possible that differences in the processing of the primary transcript by RNase E affect tmRNA levels (78). The *ssrA* gene is a favored target for the insertion of various genetic elements in bacteria (135). Besides the phages Fels-2 and SopEΦ, the CP4-57 element of *E. coli* K-12 (70) and the VPI pathogenicity island of *V. cholerae* (67) have also been found integrated at this locus. The presence of a CP4-57-related integrase gene immediately adjacent to the Fels-2 *att* site in most, if not all, *S. enterica* serovars (STM2740 in Color Plate 5) suggests that this site has already been used during *Salmonella* evolution.

SopEΦ

The invasion of epithelial cells by *Salmonella* requires the activity of effector proteins that are translocated into the host cell cytosol, where they stimulate signaling pathways that result in bacterial uptake (40). One such effector is SopE, a 25-kDa protein secreted by the SPI I type III secretion system (50, 138). SopE is a guanine exchange factor that activates at least two members of the Rho GTPase family, Cdc42 and Rac-1, leading to actin cytoskeleton rearrangements, membrane ruffling, and nuclear responses (52). A characterization of the DNA sequence surrounding the *sopE* gene in strain SL1344 revealed that the gene was inserted into the tail fiber region of a P2-related prophage named SopEΦ (51). Recent work has shown that SopEΦ is a close relative of the Fels-2 phage (101). An alignment of the sequences of these two prophages suggested that the *sopE* gene entered the genome of a common ancestor as a result of an illegitimate double-crossover event that removed a portion of the tail fiber inversion module along with the majority of the *pin*-like gene (Color Plate 5). The exchange, which might have been caused by an aberrant invertase reaction, locked the tail fiber switch in a configuration similar to the "off" configuration of present-day Fels-2.

In spite of the importance of the SopE function, neither this protein nor the entire SopEΦ phage is absolutely required for virulence in mice, and the prophage is found in only a limited subset of serovar Typhimurium strains. A functional redundancy with other proteins, namely, SopE2 (6, 118) and SopB (90, 149), which are both present in all *Salmonella* species, likely accounts for the dispensability of SopE. However, the introduction of SopEΦ into strain 14028, which normally lacks the prophage, was reported to cause a small but significant increase in fluid accumulation in infected bovine ligated ileal loops, suggesting that SopE plays a definite role in cattle enteropathogenicity (150). The occurrence of SopEΦ in a group of cattle-associated *S. enterica* serovar Typhimurium isolates responsible for epidemic outbreaks in the United Kingdom and the former country of East Germany in the 1970s and 1980s (88) led to the proposal that SopE activity is especially relevant to bovine enterocolitis (150). However, this idea is difficult to reconcile with the presence of the *sopE* gene in *S. enterica* serovar Typhi (27, 98, 101), a strictly human pathogen, and in various poultry-associated serovars such as Gallinarum, Hadar, and Enteritidis (89). In serovar Typhi, the SopE prophage appears to have translocated from the *ssrA* locus (~59 cs) to a large pathogenicity island (SPI-7, located at 93 cs) that also includes operons involved in the synthesis of the Vi antigen and type IV pili (17, 27, 98). The prophage contains one intact copy of the 47-bp *att* core sequence at its right end but only the innermost 8-bp portion of the sequence on its left end, suggesting that it has lost the capacity to excise (27, 98, 101). In poultry-associated serovars, the *sopE* gene cassette is not part of a P2-like prophage; rather, it is inserted within a defec-

tive lambdoid prophage with similarity to the Gifsy phages (89). Data from *S. enterica* serovar Enteritidis genome sequence projects show that this prophage (which also carries the *sodC1* gene) lies adjacent to a tandem array of previously acquired elements, namely, a defective prophage which contains the *pagK* and *pagO* loci (46) and a SopE2-encoding islet (90). These two elements are found throughout the *Salmonella* genus, suggesting that they were incorporated, possibly as a result of separate events, into the genome of *S. enterica* early on (90). The repeated 23-bp sequence 5′-GGAATCGTATTCGGTCTCTTTTT-3′, which flanks the *pagK*- and *pagO*-carrying prophage, likely constitutes the common *att* core region (see Color Plate 7). Significantly, the same sequence is used as an attachment site by a phage found in the genome of serovar Typhi strain CT18 (98) and by the ΦW104 prophage in serovar Typhimurium DT104 isolates (see below).

INTERPLAY OF Fels-2 AND SopEΦ: LYSOGENIC RELAY AND DOUBLE LYSOGENY

In spite of their relatedness, the Fels-2 and SopEΦ phages specify distinct immunity determinants and escape mutual repression. In particular, SopEΦ can grow on strain LT2, which is normally a Fels-2 lysogen, provided that this strain is first cured of its resident Gifsy-1 prophage, which otherwise inhibits SopEΦ growth by an unknown mechanism (35). Strain LT2-derived bacteria exposed to SopEΦ become lysogenic for this phage, raising the question of which integration site is used when the primary attachment site is occupied. Recent work revealed that SopEΦ has the ability to dislodge the Fels-2 prophage from its chromosomal position and to insert itself in its place (38). We refer to this phenomenon as "lysogenic relay" (Color Plate 6). Both the SopEΦ and Fels-2 genomes include loci similar to the phage 186 excisionase gene (*apl*) and an *int* gene (28, 107). Like the case in phage 186, the SopEΦ *apl* homologue is the first gene in the right operon. This suggests that the Apl protein is made shortly after DNA injection and mediates Fels-2 prophage excision upon forming a complex with the SopEΦ Int or Fels-2 Int protein (the two Int proteins are 93% identical). Deleting the Fels-2 *int* gene causes a reduction in the frequency of dislodgement. However, dislodgement is still observed when the entire Fels-2 prophage is replaced by a *lacZ* gene cassette, indicating that no Fels-2 sequence (other than the *attL* and *attR* sites) is absolutely required for this process (38). It is interesting to consider that infection by a phage may lead to the removal and replacement of any genomic element inserted at that phage's attachment site on the chromosome.

In the experiment described above, a fraction of the bacteria became lysogenic for SopEΦ without losing Fels-2. These double lysogens contained the two prophages in a tandem array (38). In all isolates, the SopEΦ prophage was found on the right side of Fels-2 (*attR*), suggesting that *ssrA* gene sequences upstream of the 47-bp *att* core region contributed to the integration specificity (Color Plate 6). Double lysogens are unstable and segregate cells that have lost one or the other of the two prophages. Segregation occurs independently of the RecA protein and is also observed when the entire Fels-2 sequence is replaced by a *lacZ* gene cassette (38). Thus, the formation and resolution of the double lysogen constitute another pathway leading to lysogenic relay. Interestingly, resolution of the tandem array requires the presence of a functional *int* gene from either SopEΦ or Fels-2, but no *apl* (excisionase) gene. Presumably, the lack of a requirement for excisionase reflects the fact that, mechanistically, the resolution reaction is closer to an integration reaction than to an excision. This is because the fusion of Fels-2 *attR* and SopEΦ *attL* reconstitutes a bona fide *attP* site and because the Int protein alone is sufficient to promote synapsis and recombination between this site and a second *att* core sequence (Color Plate 6). The recovery of both possible resolution products suggests that, unlike the case for the formation of the double lysogen (with only *attR* used for attachment [see above]), both Fels-2 *attL* and SopEΦ *attR* can

participate in a synaptic complex with the chromosomal *attP* (38).

Conceivably, the phenomena described above represent events that participate in the shaping and evolution of the *Salmonella* genome. For example, the tandemly arranged islets that include the *sopE2* gene and *pagK-pagO* loci are probably remnants of ancestral prophages that inserted successively at the same attachment site (Color Plate 7). Both elements contain putative integrase genes at their left boundaries; however, both genes appear to be truncated and defective, which may explain their present-day stability. This site has been the target of further insertions, as shown by the presence of the *sopE* prophage in serovar Enteritidis (Color Plate 7) and of the ΦW104 prophage in epidemic strains of serovar Typhimurium.

ST64B

A widely accepted method for typing *S. enterica* isolates is based on scoring strain sensitivities to a reference collection of bacteriophages. For serovar Typhimurium, a more recent version of this method allows the differentiation of 207 definite types (DTs) (4). Increasing efforts are being made to correlate the phage type of a strain with its prophage complement. In one such study, a DT64 strain exposed to mitomycin C was found to release a tailless virus that was named ST64B (92). Although the viral particles were defective for infection, they were recovered in sufficient amounts to allow sequencing of the phage genome (93). The sequence data made it possible to infer the presence of an ST64B-related prophage in laboratory strains 14028 and SL1344 (36). The virions produced by these isolates were also defective for the infection of strains with a deletion of the prophage. However, when the strains carrying the defective prophage were cultivated in a mixture with a susceptible strain, phage variants with a recovered ability to form plaques accumulated. Compared to their parents, these variants appeared to result from the reversion of a +1 frameshift mutation in a presumptive tail assembly gene (36). Thus, promiscuous growth, which is likely a common feature of *Salmonella* lifestyles, creates conditions that positively select for the regeneration of active phage. Further studies have demonstrated the wide distribution of the ST64B prophage among strains of epidemiological relevance. These include multidrug-resistant DT104 strains, which represent the main epidemic clone worldwide. Significantly, a study of three independent DT104 isolates found that all of them harbor a normal copy of the presumptive tail assembly gene and release a fully functional ST64B phage (36).

The ST64B phage uses the *serU* gene for $tRNA_2^{Ser}$ as an attachment site. Since the sequence at the phage *attP* site is not identical to that of *serU*, recombination generates a new version of the gene, which is predicted to specify a tRNA molecule that differs from the wild-type species at 10 of 90 bases. However, most of these changes are mutually compensatory changes that do not affect the overall secondary structure of the mature tRNA.

Strains with a deletion of the ST64B prophage competed equally well with their wild-type parents for colonization of mouse organs after mixed infections, leading to the conclusion that the prophage does not play a relevant role in murine salmonellosis (2). Still, we notice that the ORF sb26, at the right end of the prophage map, corresponds to a locus that was previously identified as being highly upregulated during murine infection (18).

P22, ST64T, AND ST104

The bacteriophage P22 occupies a relevant place in the history of microbial genetics both as a model system (119) and as a genetic tool (151). The biology of the phage is the subject of a comprehensive monograph (119), and some additional information on its genomic structure was published recently (100, 129). P22 is a member of the family *Podoviridae*, which is characterized by the presence of a short, noncontractile tail (1). Among the phages described here, P22 is the one that forms the largest plaques by far, and to the best of our knowledge, it is the only one that is capable of generalized transduction. The isolation of P22 phage variants with increased transduction efficiencies

(113) has made the virus an irreplaceable tool in *Salmonella* genetics. The phage also provides a classical example of lysogenic conversion, as it modifies the antigenic formula of the host bacterium upon establishing itself as a prophage. Serotype conversion results from the addition of a glucosyl residue to galactose moieties in the O antigen repeats of lipopolysaccharide, which is mediated by the products of the *gtrABC* operon, located at the right end of the prophage map (129, 148). This change prevents further binding of the phage to its O antigen receptor, thus contributing to the exclusion of superinfecting phage from the same group. Serotype conversion may also influence the interaction of lysogenic bacteria with the animal host, as shown for other phages. However, to date, this possibility lacks experimental support. The P22 genome includes a sequence identical to the last 46 bp of the *thrW* gene for $tRNA_2^{Thr}$, which contains the recombination site used for integration into the host chromosome (76). Unlike the ST64B phage, P22 does not modify the target gene upon insertion.

Several lines of evidence point to the high incidence of P22-related prophages in *Salmonella* genomes (110, 111, 114). The DNA sequences of two such phages, ST64T and ST104, released from a DT64 strain and a DT104 strain, respectively, were published recently (94, 121). A study examining the configuration downstream of the *thrW* gene of *S. enterica* serovar Typhimurium isolates from the SARA collection found a P22-like insert in 7 of 21 strains analyzed. The incidence of a P22-like insert was higher in isolates with epidemiological relevance. In particular, all strains from DT104 and DT120 that were tested were found to contain a P22-related insert (37).

ΦW104

DNA sequence data from the Sanger Institute (*Salmonella* spp. comparative sequencing project) first revealed the existence of a new prophage in an *S. enterica* serovar Typhimurium DT104 strain. The phage genome lies adjacent to the *pagK-pagO* islet and is flanked by the same 23-bp repeat that delimits the islet (Color Plate 7). A recent survey of a large number of epidemic strains from different sources showed this prophage to be specifically associated with isolates of the DT104 group (37). The possibility that the prophage—tentatively named ΦW104—was implicated in phage typing was ruled out upon the isolation of derivatives with a deletion of ΦW104 and the finding that their responses to typing phages were essentially unchanged (37). The strains with a deletion of ΦW104 were used as indicator strains to assess the release of plaque-forming particles from their respective parents. Although they were very small, plaques were observed in all instances, indicating that ΦW104 is fully competent for undergoing induction and forming active virus. The phage genome includes ORFs similar to the *irsA* gene of the Gifsy-3 phage (22) and a cluster of genes involved in lipopolysaccharide synthesis. Whether these loci or others that have no homologues in the DNA databases contribute to the enhanced virulence traits of DT104 strains remains to be determined.

CONCLUDING REMARKS

Prophages represent between 4 and 6% of the *Salmonella* genome. Pathogenicity loci and other genes that are useful to the bacterium contribute only a tiny portion of this material, suggesting that other selective forces are responsible for preserving prophage integrity. A main force is likely the very one that selects for the initial prophage acquisition, namely, viral killing. A complete or partial loss of the prophage causes a loss of immunity against superinfection and exposes the bacterium to killing. Due to spontaneous prophage induction, sufficiently dense cultures of lysogenic strains inevitably contain active viral particles; such phage may selectively destroy cells that have lost the corresponding prophage. Besides this patrolling role, spontaneously released phage become especially relevant whenever strains with different prophage repertoires grow together. In this case, each strain is susceptible to killing by viruses released by the other strains, creating conditions that support viral proliferation, which, in turn, can deeply affect the evolution of the bacterial pop-

ulation. On the one hand, these conditions fuel the formation of recombinant phage variants and can also select for revertants of inactive phage (12, 36). On the other hand, they favor the expansion of lysogenic clones, generating strains whose prophage contents are enriched relative to the starting strains. Although it is based on laboratory models, the above scenario is likely to exist in nature, where strain mixing may not be uncommon. In fact, one may predict the prophage arsenal of natural isolates to be a direct reflection of their lifestyles. Strains circulating in promiscuous environments, such as the mammalian gut and its excretions, are expected to be subjected to the strongest pressure to increase and diversify their prophage repertoires. In contrast, these repertoires may be invariant or decay in strains occupying more secluded niches, such as *S. enterica* serovar Typhi strains colonizing the surfaces of gallbladder stones in chronically infected humans. Except for such special cases, there can be little doubt that phage trafficking is a primary mechanism fostering genomic evolution and diversification in the *Salmonella* complex.

ACKNOWLEDGMENTS

We are grateful to Theresa Ho for critically reading the manuscript. We also thank Fred Heffron for communicating unpublished data and S. Casjens for suggestions.

The work from our laboratory was supported by the "Centre National de la Recherche Scientifique" in France and by a grant from the French Ministry of Research and Technology.

REFERENCES

1. **Ackermann, H. W.** 1998. Tailed bacteriophages: the order *Caudovirales*. *Adv. Virus Res.* **51:**135–201.
2. **Alonso, A., M. G. Pucciarelli, N. Figueroa-Bossi and F. García del Portillo.** Increased excision of the *Salmonella* defective prophage ST64B caused by a deficiency in Dam methylase. Submitted for publication.
3. **Ammendola, S., N. Figueroa-Bossi, A. Battistoni, and L. Bossi.** Unpublished data.
4. **Anderson, E. S., L. R. Ward, M. J. de Saxe, and J. D. de Sa.** 1977. Bacteriophage-typing designations of *Salmonella typhimurium*. *J. Hyg.* **78:**297–300.
5. **Bacciu, D., G. Falchi, A. Spazziani, L. Bossi, G. Marogna, G. S. Leori, S. Rubino, and S. Uzzau.** 2004. Transposition of the heat-stable toxin *astA* gene into a gifsy-2–related prophage of *Salmonella enterica* serovar Abortusovis. *J. Bacteriol.* **186:**4568–4574.
6. **Bakshi, C. S., V. P. Singh, M. W. Wood, P. W. Jones, T. S. Wallis, and E. E. Galyov.** 2000. Identification of SopE2, a *Salmonella* secreted protein which is highly homologous to SopE and involved in bacterial invasion of epithelial cells. *J. Bacteriol.* **182:**2341–2344.
7. **Bäumler, A. J.** 1997. The record of horizontal gene transfer in *Salmonella*. *Trends Microbiol.* **5:**318–322.
8. **Bäumler, A. J., R. M. Tsolis, T. A. Ficht, and L. G. Adams.** 1998. Evolution of host adaptation in *Salmonella enterica*. *Infect. Immun.* **66:**4579–4587.
9. **Beltran, P., S. A. Plock, N. H. Smith, T. S. Whittam, D. C. Old, and R. K. Selander.** 1991. Reference collections of strains of the *Salmonella typhimurium* complex from natural sources. *J. Gen. Microbiol.* **137:**601–606.
10. **Blanc-Potard, A. B., and E. A. Groisman.** 1997. The *Salmonella selC* locus contains a pathogenicity island mediating intramacrophage survival. *EMBO J.* **16:**5376–5385.
11. **Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **277:**1453–1474.
12. **Bossi, L., J. A. Fuentes, G. Mora, and N. Figueroa-Bossi.** 2003. Prophage contribution to bacterial population dynamics. *J. Bacteriol.* **185:**6467–6471.
13. **Botstein, D.** 1980. A theory of modular evolution for bacteriophages. *Ann. N.Y. Acad. Sci.* **354:**484–491.
14. **Boyd, J. S.** 1950. The symbiotic bacteriophages of *Salmonella typhi-murium*. *J. Pathol. Bacteriol.* **62:**501–517.
15. **Brody, H., and C. W. Hill.** 1988. Attachment site of the genetic element e14. *J. Bacteriol.* **170:**2040–2044.
16. **Brumby, A. M., I. Lamont, I. B. Dodd, and J. B. Egan.** 1996. Defining the SOS operon of coliphage 186. *Virology* **219:**105–114.
17. **Bueno, S. M., C. A. Santiviago, A. A. Murillo, J. A. Fuentes, A. N. Trombert, P. I. Rodas, P. Youderian, and G. C. Mora.** 2004. Precise excision of the large pathogenicity island, SPI7, in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* **186:**3202–3213.
18. **Bumann, D.** 2002. Examination of *Salmonella* gene expression in an infected mammalian host using the green fluorescent protein and two-colour flow cytometry. *Mol. Microbiol.* **43:**1269–1283.