

Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion

Harald Brüssow,^{1*} Carlos Canchaya,¹ and
Wolf-Dietrich Hardt^{2*}

*Nestlé, Research Center, Vers-chez-les-Blanc, Lausanne,¹
and Institute for Microbiology, D-BIOL,
ETH Zürich, Zürich,² Switzerland*

INTRODUCTION	561
EVOLUTION OF BACTERIAL PATHOGENS	561
Medical Microbiology: Adaptation over a Limited Timescale.....	561
From Virulence to Fitness Factors	565
Evolutionary Origin of Fitness Factors	565
MOLECULAR MECHANISMS SHAPING BACTERIAL EVOLUTION	565
Comparative Genomics: Macroscopic Effects of Phages on Genome Structure	566
Intraspecies Bacterial Genome Comparisons	566
Colinear genomes.....	566
Rearranged bacterial genomes.....	566
Interspecies Bacterial Genome Comparisons	566
Bacterial Genomes Lacking Prophages	569
EVOLUTION OF PHAGES.....	569
General Considerations	569
Phage and Prophage Genomics: Evidence for the Modular Theory of Phage Evolution	570
Mechanism of Phage Module Exchange: Illegitimate versus Homologous Recombination.....	570
Loss and Decay of Prophages	574
Role of Phages in Bacterial Evolution	574
PHAGE-MEDIATED GENE TRANSFER AND INACTIVATION	574
Transduction.....	576
Lysogenic Conversion	576
Gene Disruption.....	576
LYSOGENIC CONVERSION: ROLE OF PROPHAGE MORONS	577
Morons of Temperate Phages, Prophages, and Phage Remnants in Pathogenic Bacteria	577
Examples of Distribution and Function of Morons	577
Prophages in <i>S. enterica</i> serovar Typhimurium	577
(i) The <i>sopE</i> moron of SopEΦ	578
(ii) Other <i>S. enterica</i> serovar Typhimurium prophages: Fels-1, Fels-2, GIFSY-1, GIFSY-2, GIFSY-3, E34, and ST64B	578
(iii) Phage remnants encoding SopE2 and SspH2	581
<i>V. cholerae</i> toxin moron	581
Shiga toxin-producing <i>E. coli</i>	582
<i>C. botulinum</i>	586
<i>S. pyogenes</i>	587
<i>S. aureus</i>	590
<i>C. diphtheriae</i>	593
Fitness Cassettes in Nonpathogenic Bacteria?	594
Evidence for Increased or Modulated Virulence after Phage Conversion	595
CONCLUSION AND OUTLOOK	595
ACKNOWLEDGMENTS	596
REFERENCES	596

* Corresponding author. Mailing address for Harald Brüssow: Nestlé, Research Center, Vers-chez-les-Blanc, CH-1000 Lausanne, Switzerland. Phone: 41 21 785 86 76. Fax: 41 21 785 85 49. E-mail: harald.bruessow@rdls.nestle.com. Mailing address for Wolf-Dietrich Hardt: Institute for Microbiology, D-BIOL, ETH Zürich, Schmelzbergstr. 7, CH-8092 Zürich, Switzerland. Phone: 41 1 632 5143. Fax: 41 1 632 1129. E-mail: hardt@micro.biol.ethz.ch.

INTRODUCTION

In the early days of molecular biology, phages were studied intensively for their own sake and as simple model systems. In the following years, many molecular biologists shifted their attention from phages and bacteria to higher organisms. Currently, we are witnessing a renaissance of phage research. The stage has been set by a recent shift toward ecology-oriented phage research covering the impact of phages on subjects as diverse as the cycling of matter in the oceans and bacterial pathogenesis. In fact, phages that carry key virulence factors were discovered long ago. Two examples are phages gamma and C1, which encode key virulence factors of *Corynebacterium diphtheriae* and *Clostridium botulinum* (11, 85). However, progress in the molecular biology of higher organisms and the continued advances in bacterial research have now opened the door to studies of "host-pathogen interactions" at a new level of detail. For some time, this research had focused on a few selected standard strains or on the molecular mechanism of selected toxins. The discovery of the cholera toxin phage CTXΦ (233) was one of the landmarks of this new phage research. It was becoming increasingly clear that phages play an important role in the evolution and virulence of many pathogens (Table 1). The interest in phage research was further fueled by the plethora of prophage sequences, which were discovered as a by-product of bacterial genome sequencing. The analysis of these sequences revealed that phages affect the bacterial genome architecture. In addition, phages are important vehicles for horizontal gene exchange between different bacterial species and account for a good share of the strain-to-strain differences within the same bacterial species (Fig. 1) (67, 136). In fact, two-thirds of all gamma-proteobacteria and low-G+C gram-positive bacteria harbor prophages (46, 47); these include *Escherichia coli* O157 Sakai (171) and *Salmonella* spp. (26). An example is shown in Fig. 1: the genomes from two *Streptococcus pyogenes* strains belonging to two different M serotypes and associated with different disease types were aligned at the DNA sequence level. All major genome differences were attributed to prophage sequences.

The early studies had indicated that some prophages carry additional cargo genes (termed morons or lysogenic conversion genes), and recent genomic analyses have revealed many more examples. These morons are not required for the phage life cycle. Instead, many morons from prophages in pathogenic bacteria encode proven or suspected virulence factors. They are postulated to change the phenotype or fitness of the lysogen. As a consequence, phages have emerged as prime suspects in the adaptation of pathogens to new hosts and the emergence of new pathogens or epidemic clones. This has led to a conceptual shift from host-pathogen interactions to host-pathogen-phage interactions.

It is important to note that lysogenic conversion is only one of at least five different ways by which temperate phages affect bacterial fitness: (i) as anchor points for genome rearrangements, (ii) via gene disruption, (iii) by protection from lytic infection, (iv) by lysis of competing strains through prophage induction, and (v) via the introduction of new fitness factors (lysogenic conversion, transduction).

We begin our review with a brief introduction to concepts of bacterial evolution and the evolution of pathogens, and we

highlight the role of phages in shaping the bacterial genome architecture. Then we review the mechanisms of phage evolution and the mechanisms which allow phages to modify the host bacterium. A special focus is placed on lysogenic conversion, which is illustrated with seven portraits of human pathogens containing prophages that encode important virulence factors. Lysogenic conversion is thought to have a great impact on the evolution of pathogenic bacteria and results in a very interesting situation of bacterium-phage coevolution.

EVOLUTION OF BACTERIAL PATHOGENS

Before we start with the discussion of phages, we provide a conceptual framework by giving a brief introduction to basic concepts of the evolution of pathogenic bacteria.

Medical Microbiology: Adaptation over a Limited Timescale

Bacterial evolution started long before the emergence of animals. In fact, bacteria were the first, and for some time the only, inhabitants of Earth. Therefore, early evolution involved competition, genetic exchange, and selection only between bacteria. One can assume that phages also took part in this early phase of evolution. As discussed below, many of the virulence factors found in contemporary pathogens might actually date back to this time. Multicellular eukaryotes evolved only during the past 1 billion years, and mammals proliferated massively only during the past 65 million years (244). Human-restricted pathogens such as *S. pyogenes*, *Shigella* spp., and the human-adapted *Salmonella* strains must have become adapted to their hosts in 1 million years or less (the time frame of human evolution). This is approximately the timescale for the separation of *E. coli* strain K-12 from *Shigella flexneri*. It should be noted that the traditional bacterial nomenclature can be misleading. The *Shigella* genus attribution roots back to the historical recognition of shigellosis as a distinct medical entity. Modern taxonomy and comparative genomics would classify *S. flexneri* at best as a subspecies within *E. coli*. The emergence of *S. flexneri* as a human pathogen thus occurred at the level of the outmost twigs of the bacterial phylogenetic tree.

Based on bacterial genome analyses, the very idea of a phylogenetic tree as a description for bacterial evolution has recently come under heavy attack. It was proposed that horizontal gene transfer has occurred at such a level that the tree analogy should be replaced by a web analogy. However, also the validity of the web arguments has recently been challenged (60). Without going into this hotly debated area, we insist that most of the current evidence for the involvement of phages in shaping bacterial genomes, bacterial fitness, and host-pathogen interactions deals with events at this lowest taxonomy level. Figure 2A illustrates this for the comparison between *E. coli* K-12 and *Shigella*. The two genomes differ in DNA insertions (including many prophages), loss of genes (discussed below), genome inversions, and translocations (some are flanked on one side by prophage sequences). Most prophages found in these two genomes differ in their DNA sequence, suggesting that they were acquired (and then often degraded) after the separation of the two bacterial lineages. However, at the DNA sequence level both genomes could still be aligned over essentially the entire chromosomes (Fig. 2A). Figure 2B

TABLE 1. Common themes among phage-encoded virulence factors^a

Protein	Gene	Phage	Bacterial host
Extracellular toxins			
Diphtheria toxin	<i>tox</i>	β-Phage	<i>C. diphtheriae</i>
Neurotoxin	<i>C1</i>	Phage C1	<i>C. botulinum</i>
Shiga toxins	<i>stx1, stx2</i>	H-19B	<i>E. coli</i>
Enterohaemolysin	<i>hly2</i>	φFC3208	<i>E. coli</i>
Cytotoxin	<i>ctx</i>	φCTX	<i>P. aeruginosa</i>
Enterotoxin	<i>see, sel</i>	NA	<i>S. aureus</i>
Enterotoxin P	<i>sep</i>	φN315	<i>S. aureus</i>
Enterotoxin A	<i>entA</i>	φ13	<i>S. aureus</i>
Enterotoxin A	<i>sea</i>	φMu50A	<i>S. aureus</i>
Exfoliative toxin A	<i>eta</i>	φETA	<i>S. aureus</i>
Toxin type A	<i>speA</i>	T12	<i>S. pyogenes</i>
Toxin type C	<i>speC</i>	CS112	<i>S. pyogenes</i>
Cholera toxin	<i>ctxAB</i>	CTXφ	<i>V. cholerae</i>
Leukocidin	<i>pvl</i>	fPVL	<i>S. aureus</i>
Superantigens	<i>speA1, speA3, speC, speI, speH, speM, speL, speK, ssa</i>	8232.1	<i>S. pyogenes</i>
Cytolytic distending toxin	<i>cdt</i>	Unnamed	<i>E. coli</i>
Proteins altering antigenicity			
Membrane proteins	Mu-like	Pnm1	<i>N. meningitidis</i>
Glucosylation	<i>rfb</i>	ε ³⁴	<i>S. enterica</i>
Glucosylation	<i>gtr</i>	P22	<i>S. enterica</i>
O-antigen acetylase	<i>oac</i>	Sf6	<i>S. flexneri</i>
Glucosyl transferase	<i>gtrII</i>	SfII, SfV, SfX	<i>S. flexneri</i>
Effector proteins involved in invasion			
Type III effector	<i>sopE</i>	SopEΦ	<i>S. enterica</i>
Type III effector	<i>sseI (gtgB)</i>	GIFSY-2	<i>S. enterica</i>
Type III effector	<i>sspH1</i>	GIFSY-3	<i>S. enterica</i>
Enzymes			
Superoxide dismutase	<i>sodC</i>	Sp4, 10	<i>E. coli</i> O157
Superoxide dismutase	<i>sodC-I</i>	GIFSY-2	<i>S. enterica</i>
Superoxide dismutase	<i>sodC-III</i>	Fels-1	<i>S. enterica</i>
Neuraminidase	<i>nanH</i>	Fels-1	<i>S. enterica</i>
Hyaluronidase	<i>hylP</i>	H4489A	<i>S. pyogenes</i>
Leukocidin	<i>pvl</i>	φPVL	<i>S. aureus</i>
Staphylokinase	<i>sak</i>	φ13	<i>S. aureus</i>
Phospholipase	<i>sla</i>	315.4	<i>S. pyogenes</i>
DNase/streptodornase	<i>sdn, sda</i>	315.6, 8232.5	<i>S. pyogenes</i>
Serum resistance			
OMP ^b	<i>bor</i>	λ	<i>E. coli</i>
OMP	<i>eib</i>	λ-like	<i>E. coli</i>
Adhesions for bacterial host attachment			
Vir	<i>vir</i>	MAV1	<i>M. arthritidis</i>
Phage coat proteins	<i>pblA, pblB</i>	SM1	<i>S. mitis</i>
Others			
Mitogenic factors	<i>mf2, mf3, mf4</i>	370.1, 370.3, 315.3	<i>S. pyogenes</i>
Mitogenic factor	<i>toxA</i>	Unnamed	<i>P. multocida</i>
Mitogenic factor	Unnamed	phisc 1	<i>S. canis</i>
Virulence	<i>gtgE</i>	GIFSY-2	<i>S. enterica</i>
Antivirulence	<i>grvA</i>	GIFSY-2, Fels-1	<i>S. enterica</i>

^a Updated from reference 26 with permission.^b OMP, outer membrane protein.

illustrates this for two even more closely related bacterial strains. They belong to the same O serotype of *E. coli*, the recently emerged food-borne pathogen O157:H7. Molecular evidence suggested that these strains evolved from an enteropathogenic O55 *E. coli* strain, perhaps over the last few decades (see below), and the prominent role of prophages for strain differentiation can be directly gleaned from the dot plot alignment (Fig. 2B).

Many of the concepts derived from the studies of human-pathogen-phage interactions will most probably also apply to phage-bacterium coevolution in other ecological niches. However, we cannot exclude that the view derived from pathogenic interactions is skewed by two facts. First, humans are evolutionary newcomers. Therefore, the time for adaptation to this new niche (i.e., balancing the pathogen-host interaction) has been relatively short and will be enriched for short-term adap-

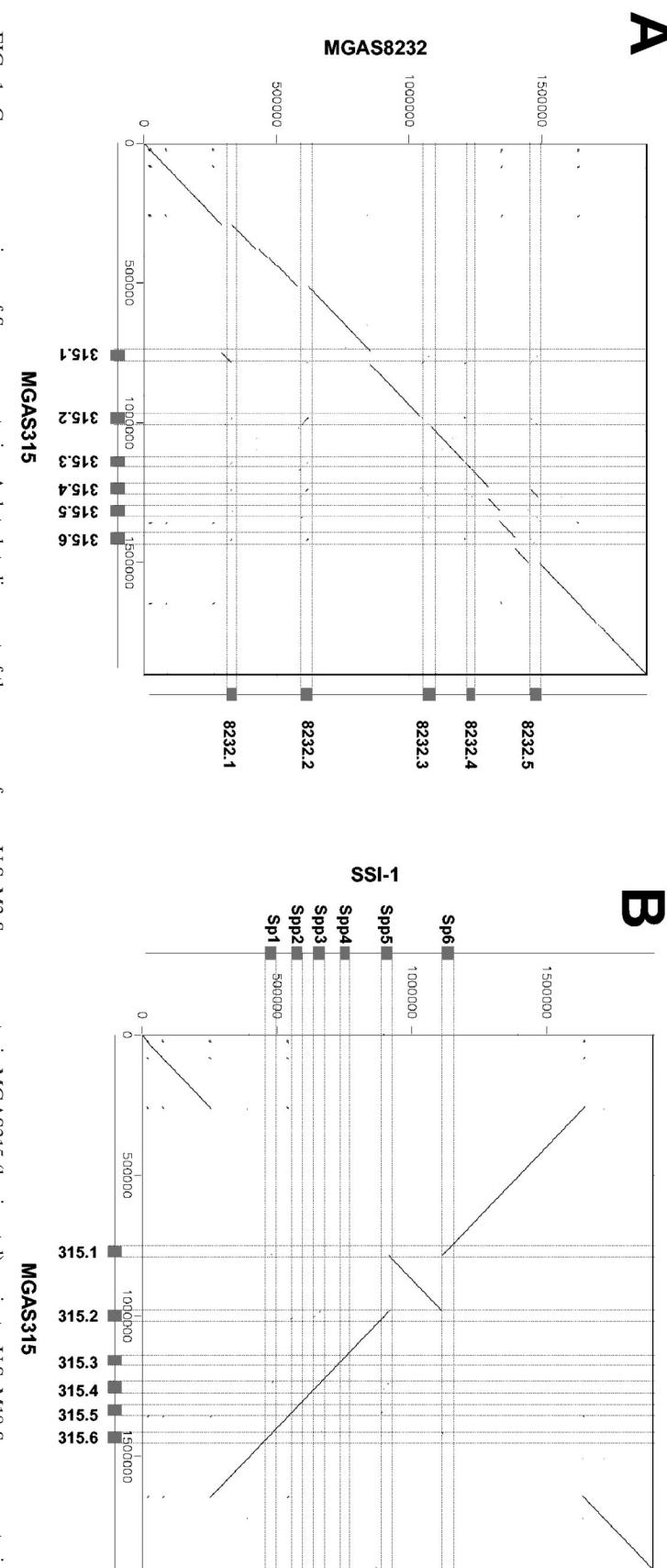


FIG. 1. Genome comparisons of *S. pyogenes* strains. A dot plot alignment of the genomes from a U.S. M3 *S. pyogenes* strain MGAS315 (horizontal) against a U.S. M18 *S. pyogenes* strain MGAS8232 (vertical axis) (A) or a Japanese M3 *S. pyogenes* strain SSI-1 (vertical axis) (B) is shown. The position and name of the prophages are indicated by boxes on the axes and by thin horizontal and vertical lines crossing the dot plot. The dot plot was done with the Dottup program (<http://www.emboss.org/>); the word size was 15, the output format was Postscript, the program was run in the direct and reverse directions, and the figures were combined. The numbers and ticks on the axes give the scale for the genomes in base pairs.

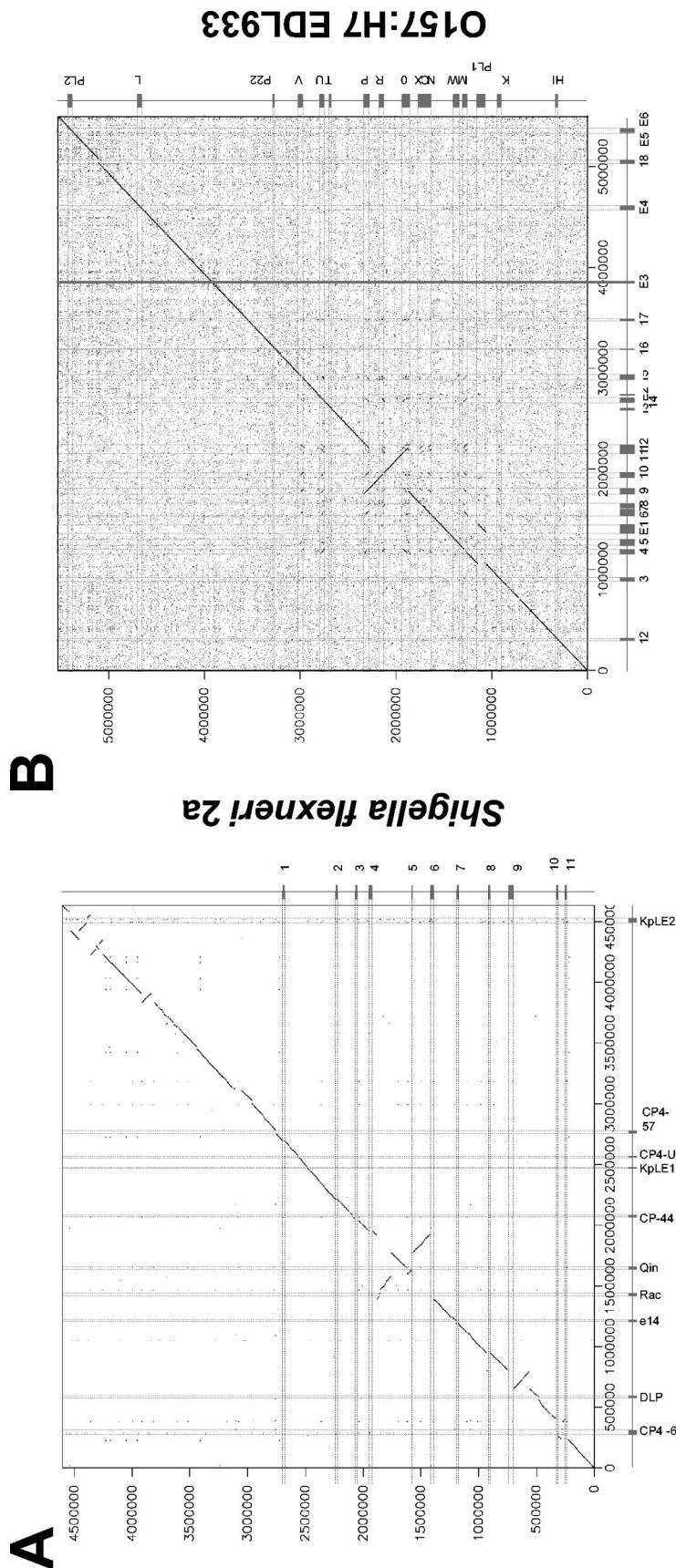


FIG. 2. Genome comparisons of *E. coli* strains. (A) Dot plot alignment of the laboratory *E. coli* strain K-12 (horizontal axis) against the *S. flexneri* strain 2a (vertical axis). *Shigella* is treated here as a subspecies of *E. coli*. (B) Dot plot alignments of a U.S. (horizontal axis) *Shiga-toxin* producing O157:H7 *E. coli* strain. The prophages and prophage-like elements are indicated on the axes by boxes and by thin horizontal and vertical lines crossing the dot plot.

tation processes. Second, pathogens attacking mammals and birds confront a formidable defense barrier generated by the adaptive immune system. Not surprisingly, numerous virulence factors from human and veterinary pathogens target this evolutionarily recent resistance system.

From Virulence to Fitness Factors

How do bacteria adapt to the life-style of a pathogen? At first glance, mammalian hosts are ecological niches like any other. However, in some respects they are difficult niches. In addition to the "normal" challenges encountered in nonliving ecological niches, mammals have defenses, which have been shaped by coevolution with microbes. These defenses include simple physical barriers like the dead cell layers of the skin and the mucus-covered epithelia, as well as the elaboration of antimicrobial peptides, iron-sequestering mechanisms, and immune responses. The factors and mechanisms that pathogens have evolved to circumvent these defenses are termed virulence factors. Virulence factors come in a great variety of forms and include factors that neutralize defenses of the host and factors that help to engage, subvert, or destruct host cells. Generally, the interaction of a pathogen with the host is a multistage process, which includes searching for an entry site, targeting a place for multiplication in the body of the host, and becoming persistent in the original host or finding ways to reach the next host. Survival or multiplication in the environment and transmission to the next host (which might include insect vectors) certainly affect the overall success of a pathogen. However, many bacterial functions enhancing fitness in these stages of the life cycle are not considered bona fide virulence factors. For this reason, we classify these factors, together with the bona fide virulence factors, as fitness factors. A fitness factor is thus broader in its scope than a virulence factor and is not limited to pathogenic bacteria. It could be defined in the following terms. (i) A gene, factor or allele is found in successful representatives of a bacterial species. Successful clones are those that outcompete their kin. (ii) Disruption of this gene reduces the chances for multiplication of the bacterium in its specific environment. (iii) Complementation (reintroduction of the fitness factor) restores the capacity for this multiplication.

A conceptual division of fitness factors into the three classes of survival, defensive, and offensive factors is illustrated in Fig. 3. Fitness factors can evolve "vertically" by gene duplication, mutation, and even gene disruption. However, quite long periods are needed to achieve results, especially when a complex combination of genes is required for a new niche. Combination of new gene constellations by a permutation principle is, in contrast, readily achieved by horizontal gene transfer (i.e., conjugation, DNA uptake, phage transduction, or lysogenic conversion). In this review, we discuss the role of phages in these processes. Several temperate phages which carry bacterial fitness factors in their genomes are discussed in detail. In these very interesting cases, the prophage-lysogen interaction is in fact a two-way process since the evolutionary success of prophages and their lysogens are intimately linked. In many cases it is not clear which partner exploits which in this relationship. Actually, much as the phylogenetic relationships between bacteria changed from linear (trees) to two-dimensional

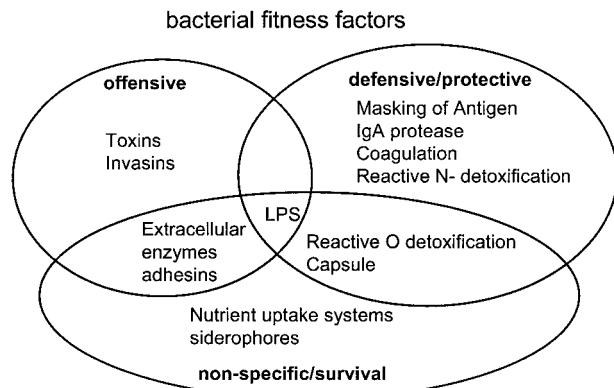


FIG. 3. Fitness factors commonly found in pathogenic bacteria. LPS, lipopolysaccharide.

displays (webs), we might also soon see web-like ecological relationships which define the interplay among phages, lysogens, and mammalian hosts.

Evolutionary Origin of Fitness Factors

Where do the virulence factors and fitness factors found in contemporary pathogens come from? Certainly, some of the virulence factors have evolved recently as a result of the ongoing "arms race" between higher eukaryotes and bacteria. However, many may stem from the coevolution of bacteria with unicellular eukaryotes during the past 1 billion years. This is supported by the notion that numerous virulence factors attack host cell proteins, molecules, and structures present in all eukaryotic cells. For example, there is a large number of toxins specific for heterotrimeric G-proteins, small G-proteins, or actin. Many membrane-disrupting toxins might also belong to this group. It is plausible that these virulence factors originally served to paralyze unicellular eukaryotic predators and were later adopted by animal-pathogenic bacteria. In fact, some virulence factors of animal pathogens are still active against unicellular eukaryotes. For example, the type III secretion system (specifically the type III effector protein ExoU) of *Pseudomonas aeruginosa* mediates killing of the slime mold *Dictyostelium discoideum* (188) but also plays a key role in the interaction with mammalian hosts (83, 102). Similarly, toxins injected into host cells via the *Legionella pneumophila* Icm/Dot type IV secretion system seem to be active in both mammalian cells and amoebae (110, 161).

MOLECULAR MECHANISMS SHAPING BACTERIAL EVOLUTION

Bacterial evolution requires the modification of "old" functions and the development of new ones. Nucleotide exchange, insertion, and deletion are the most frequent events. Mutation rates in bacteria are generally in the range of 10^{-6} to 10^{-9} per nucleotide per generation. In addition, module exchange between different genes, gene disruptions, and deletions occur at appreciable frequency. These mechanisms are common to all living organisms and allow modification of existing functions to optimize fitness in an existing niche or to adapt to a new niche.

In contrast to many higher eukaryotes, bacteria have no sexual life cycles to facilitate the exchange of alleles within a population. In bacteria, this function is fulfilled by horizontal gene transfer: in this way, entire functional units can be imported from other sources, which are not restricted by species barriers. The transferred DNA can range in size from less than 1 to more than 100 kb. It can encode entire metabolic pathways or complex surface structures. These genes can be taken up as naked DNA or transferred in the form of plasmids, conjugative transposons, or phages. Here, we focus on the role of phages in horizontal gene transfer.

Comparative Genomics: Macroscopic Effects of Phages on Genome Structure

For a number of bacterial species, more than one genome sequence is available in the National Center for Biotechnology Information database. This allows us to investigate the genetic differences between strains within a bacterial species when taking a bird's eye view of whole genome comparisons. There are also completed genome sequences from different bacterial species belonging to the same genus. These data thus permit a similar first glimpse of the evolution of bacterial species at the genomic level.

Intraspecies Bacterial Genome Comparisons

Colinear genomes. When the genomes of two bacterial strains belonging to the same species were aligned, two different outcomes were observed. One type of dot plot alignment shows essentially a straight line across the entire genome length, indicating the matching of closely related DNA sequences. The colinear alignment is illustrated for a gram-positive pathogen in Fig. 1. The genomes of two *S. pyogenes* strains yielded a straight diagonal line when aligned in the dot plot display (Fig. 1) (203). The line is interrupted by small regions of nonalignment. In the depicted cases, most of the conspicuous gaps represent prophage sequences. The relative contribution of the different mobile DNA elements (prophages, transposons, integrative plasmids, pathogenicity islands, and IS elements) varies from one bacterial system to the next. In *S. pyogenes* and *E. coli* genome comparisons, most of the gaps are caused by prophages. In other systems, the proportion of prophages among the mobile DNA elements is less prominent (e.g., *Streptococcus agalactiae*) (216). Differences were seen even between strains belonging to the same species: the comparison of *S. aureus* strain NCTC 8325 and MSSA-476 showed prophages as the main determinants of difference (National Center for Biotechnology Information unfinished genomes) while the comparison between strains N315 and MW2 revealed transposons and small pathogenicity or genomic islands as major contributors of diversity (7). This important contribution of mobile DNA to the strain-specific DNA was confirmed by microarray analysis to occur in a number of pathogenic bacteria (e.g., *S. pyogenes*, *S. agalactiae*, *S. aureus*, and *E. coli*) (84, 172, 203, 216), but also in gut commensal bacteria (*Lactobacillus johnsonii*) (226).

Rearranged bacterial genomes. The other type of intraspecies comparison also shows a high level of DNA sequence identity across the two compared genomes, but the two ge-

nomes are not longer colinear. The degree of rearrangement varies substantially. Some comparisons showed only one inversion, like the intraserovar comparison between the two *Salmonella enterica* serovar Typhi sequences (Fig. 4A), (66), or two inversions, like the two M3 serotypes of *S. pyogenes* (Fig. 1B). The interserovar comparison between *S. enterica* serovar Typhimurium and Typhi demonstrated multiple inversions and a less close DNA sequence relatedness (Fig. 4B). Even more complicated genomic rearrangements were observed between different strains of *Yersinia pestis* (65) or the plant pathogen *Xylella fastidiosa* (223). The rearrangements are in many cases probably the consequence of homologous recombination between repeat sequences in the bacterial genome. Sometimes these are duplicated bacterial genes (*com* genes in the M3 *S. pyogenes* strains) or genes which occur naturally in multiple copies (rRNA genes in *S. enterica* serovar Typhi). However, in some cases prophages with limited DNA sequence identity (*S. pyogenes* [Fig. 1B and see below]) or essentially duplicated prophages (*X. fastidiosa*) have served as anchoring points for homologous recombination reactions leading to major genomic rearrangements. In fact, in the plant pathogen *X. fastidiosa*, five of the six deduced recombination sites of three genome inversions were located in prophages (45, 223). The two sequenced *X. fastidiosa* strains represent different pathovars specialized for different plant species. It is currently unknown whether the genome rearrangement represents an adaptation of the bacteria to the different plant hosts. Recently, it was demonstrated that artificial chromosome inversions can significantly modulate the fitness of *Lactococcus lactis* (43). This supports the notion that the phage-mediated genome inversions observed in some genomes can indeed affect the fitness of the lysogen.

Interspecies Bacterial Genome Comparisons

When the genomes of two different species belonging to the same bacterial genus were aligned, different outcomes were observed. On one hand, some genomes aligned perfectly across the entire genome length (e.g., *Mycobacterium tuberculosis* and *M. bovis* [Fig. 5A]; note the absence of prophagelike elements in these two genomes) or aligned except for positions occupied by mobile DNA elements (*Listeria monocytogenes*-*L. innocua* [Fig. 5B]). Still other genomes from two different species belonging to the same bacterial genus aligned over a major part of the two genomes, but only at a lower level of DNA sequence identity (e.g., *Bacillus cereus*-*B. anthracis* and *Staphylococcus aureus*-*S. epidermidis* [Fig. 5C]).

Finally, several genomes from different bacterial species belonging to the same genus yielded only small segments of DNA sequence alignment either scattered over the two genomes or in an X-like constellation (*S. pyogenes*-*S. agalactiae* [Fig. 5D]). Within the *Streptococcus* genus, *S. pyogenes*, *S. pneumoniae*, *S. agalactiae*, and *S. mutans* did not have longer segments of DNA sequence similarity whereas *S. pyogenes* and *S. equi* are sister species which were still closely related at the DNA sequence level (J. Parkhill, unpublished data). This observation illustrates gradients of relatedness between bacteria. A similar pattern was seen in the *Lactobacillus* genus: *L. johnsonii* and *L. gasseri* showed DNA sequence similarity over most of their genomes, while *L. johnsonii* and *L. plantarum* had only seven

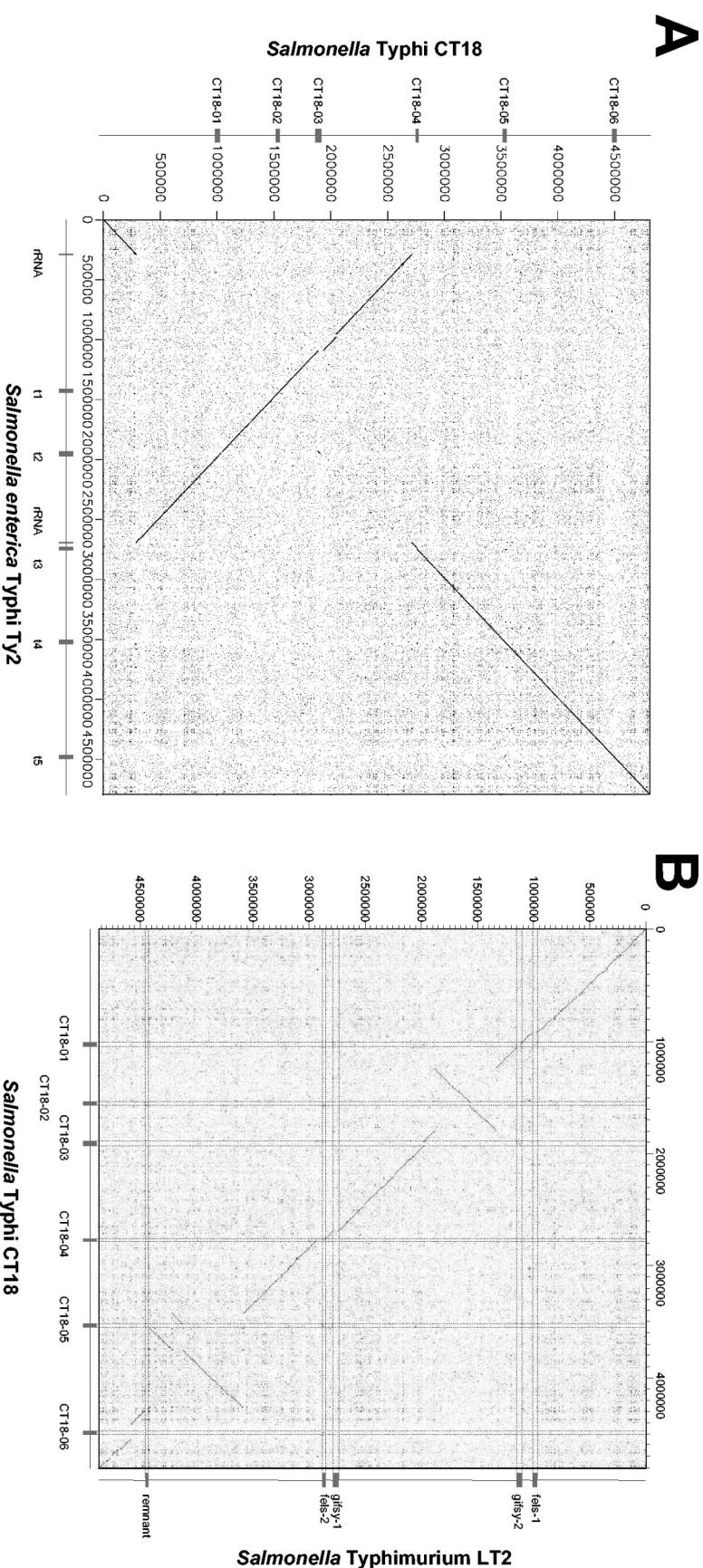


FIG. 4. Genome comparisons of *S. enterica* strains. (A) Dot plot alignment of the *S. enterica* serovar Typhi strain Ty2 (horizontal axis) versus serovar Typhi strain CT18 (vertical axis). The positions of prophages and prophage remnants are indicated by boxes. The two rRNA gene clusters which served as anchoring points for genome rearrangement are marked with a tick. (B) Dot plot alignment of the *S. enterica* serovar Typhi strain CT18 (horizontal axis) versus *S. enterica* serovar Typhimurium strain LT2 (vertical axis). The locations and names of the prophages are given.

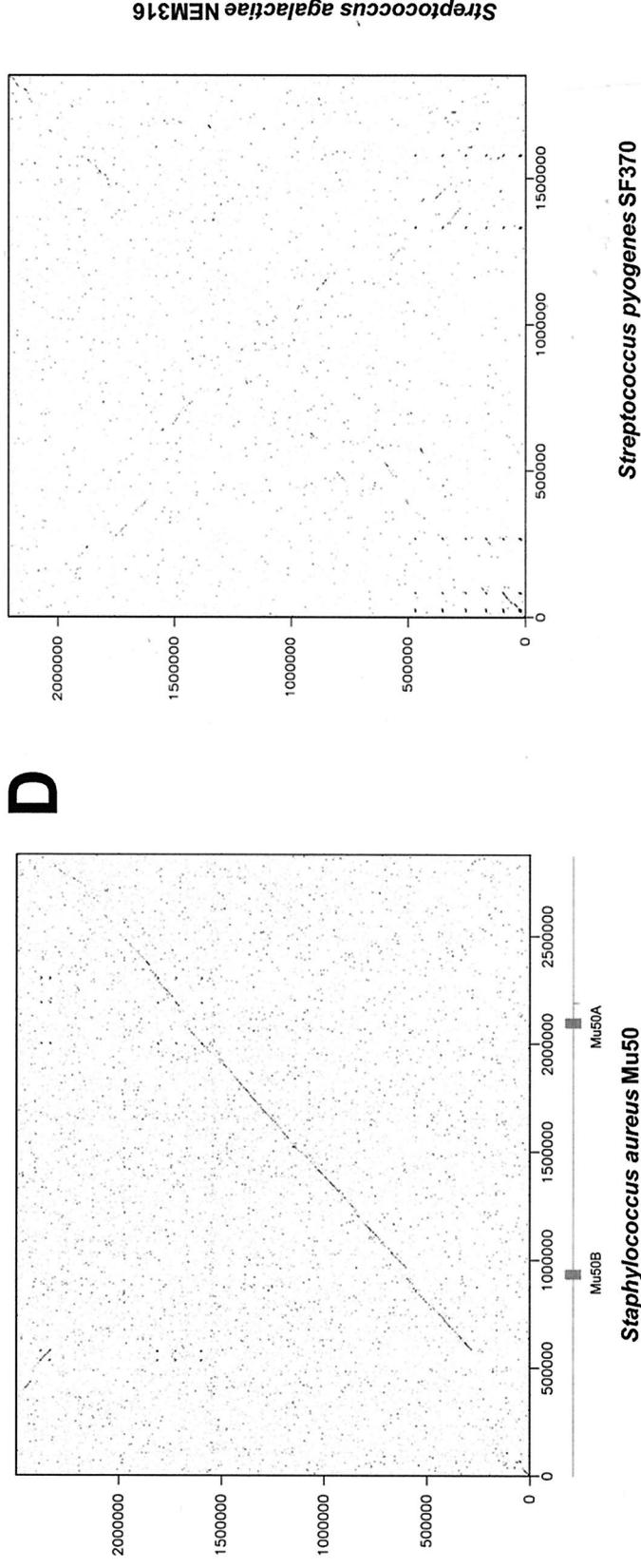
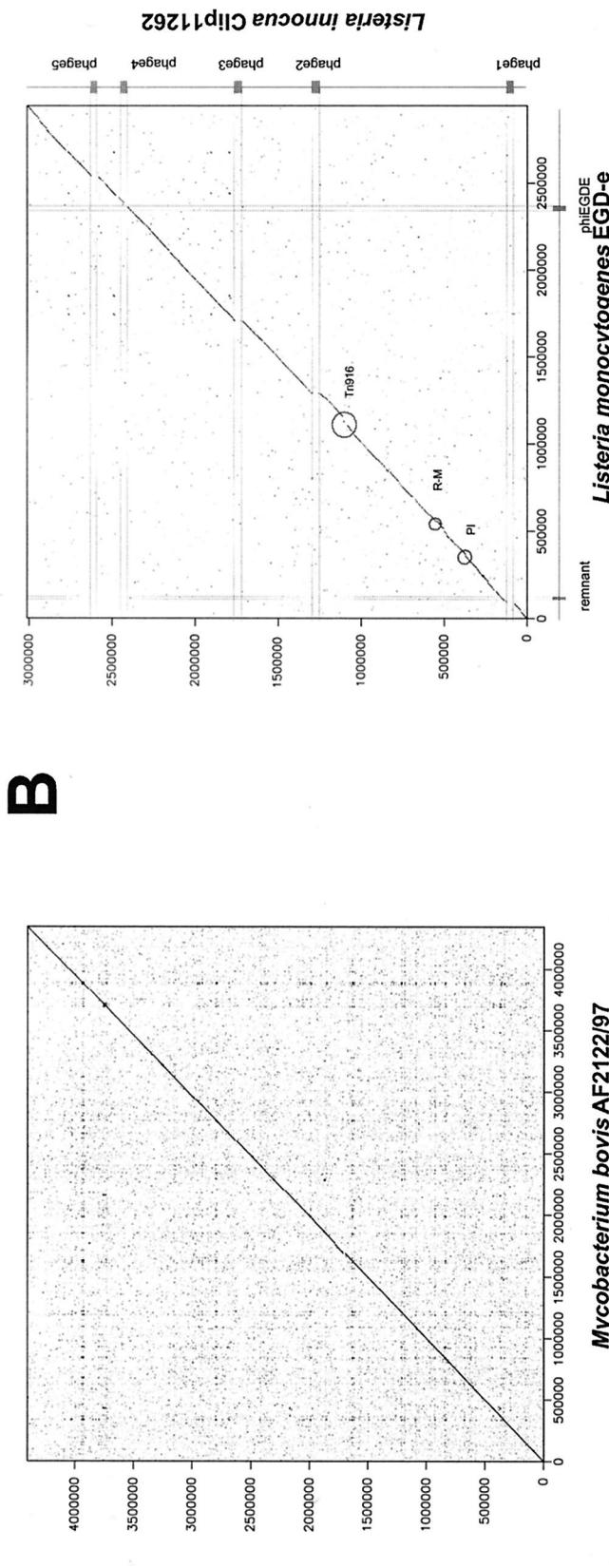


FIG. 5. Genome comparisons of gram-positive bacteria belonging to different species. The four panels compare two species from the high-G+C-content gram-positive bacterial genus *Mycobacterium* (*M. tuberculosis* versus *M. bovis* [A]) and two species from three different genera of low-G+C-content gram-positive bacteria, namely, *Listeria* (*L. innocua* versus *L. monocytogenes* [B]), *Staphylococcus* (*S. aureus* versus *S. epidermidis* [C]), and *Streptococcus* (*S. pyogenes* versus *S. agalactiae* [D]). In panels B and C, the positions of the prophages are indicated by boxes at the axes and thin lines crossing the dot plot. Further mobile elements are circled (transposon Tn916, restriction-modification genes R-M, and pathogenicity island PI).

small segments of DNA sequence similarity. At the protein level, about 20 conserved clusters comprising about 500 genes were identified and a cross-like organization of these conserved clusters was seen in a two-genome alignment (J. Boekhorst et al., unpublished data). Comparative genomics will probably soon assist bacterial taxonomy in the definition of the higher taxa at the genus and family levels. Overall, there is often less evidence for phage involvement in genome diversification in interspecies than in intraspecies alignments. Presumably this is due to the transient nature of prophage presence and the fast decay of defective prophages in bacterial genomes (see below).

Bacterial Genomes Lacking Prophages

In a number of sequenced bacterial pathogens, no prophages were detected (e.g., *Helicobacter pylori*, *Campylobacter jejuni*, *Streptococcus pneumoniae*, *Mycobacterium leprae*, and different *Mycoplasma* and *Chlamydia* spp.). Various reasons might account for this observation. Despite the lack of prophages in three sequenced *S. pneumoniae* strains, prophages are present in 76% of all clinical isolates of *S. pneumoniae* (189). This shows that the choice of bacterial strain for a sequencing project can result in nonrepresentative observations. Another case is presented by *Mycoplasma* spp. *Mycoplasma* phages are known, but the intracellular location of these bacterial pathogens might reduce the chances for lysogenic conversion, and lysogeny might therefore be a rare event in this group of pathogens. Similar arguments apply to *Chlamydia* spp. A further case is illustrated by *H. pylori*. To our knowledge, no *H. pylori*-specific phages have been described in the literature. Maybe nobody has looked for phages, but we cannot exclude the possibility that certain bacteria have eliminated prophages or that the peculiar habitat of a bacterium (e.g., the human stomach for *H. pylori*) is not favorable for phage infections. *M. leprae* represents another special case since its genome shows evidence of dramatic gene losses. Mobile DNA might have been a prime target for deletion.

In conclusion, the absence of prophages might simply reflect the chance event in the choice of a strain for sequencing. However, the fact remains that prophages may not be required for the evolution of a pathogenic life-style in every bacterial species.

EVOLUTION OF PHAGES

General Considerations

Bacterial viruses (phages) depend for a number of functions on the energy production and biosynthetic activities of their host bacteria. This obligate dependence of viruses on host activity led many biologists to deny organismal status to viruses. This distinction between living and nonliving biological material is more of philosophical than of biological interest. Biochemically, viruses are composed of the same building blocks as their host cells and their genomes consist of nucleic acids, although sometimes in configurations (e.g., double-stranded RNA) not commonly encountered in cellular genomes. Furthermore, phages and bacteria are linked by a long history of coevolution. The time dimension of this coevolution

can at present not be defined. We simply do not know when bacterial viruses evolved, i.e., whether they pre-date modern bacteria and represent remnants of former cellular life-forms that lost the competition with the modern forms of cellular life and persisted only as dependant nuisances to modern life. Some phages may have originated from assemblages of host genes that split billions of years ago from bacterial genomes, escaped from cellular control, and now lead a selfish life. Other phages might have originated recently. Most importantly, there is ample evidence for continued exchange of genetic elements between phages, bacterial genomes, and various other mobile genetic elements. This is illustrated below, and it explains the sometimes fuzzy distinction between phages, plasmids, and pathogenicity islands and the chimeric nature of some extant phages.

Phages have no fossil record and no molecular clock. DNA sequence analysis and sequence comparisons between extant phages are currently our only tools to look back into phage evolution. However, this view is blurred for two reasons. First, phage sequencing is a latecomer in the genomics revolution, and we currently have only about 200 complete phage genomes. This number may seem large, but actually it is extremely small considering that phages outnumber bacteria in many environments by a factor of 10 and represent, with approximately 10^{31} tailed phage particles, numerically the largest share of biological material on Earth (240). Up to 10^7 particles/ml were found in ocean water and sediment (30, 36, 176, 238, 240). Considering these vast numbers of phages and the number of bacteria coexisting in these niches, it has been estimated that 10^{25} phage infections are initiated every second worldwide (177); this has probably occurred for the last 3 billion years.

Second, we have no consensus model for phage evolution. Random sequencing of viral DNA in different environments ("metagenome" analysis [29, 30]) argues in favor of 2 billion undiscovered phage open reading frames, thus representing a large fraction of the total DNA sequence space (195). In addition, phage taxonomy is currently in turmoil. The official International Committee for Taxonomy of Viruses (ICTV) taxonomy based on phage morphology and genome organization was challenged by a taxonomy based on individual modules (135), a phage proteome tree (196), or elements of horizontal evolution revealed in the head structural gene cluster of phages (187). Originally, "lambdoid" phages were defined by their ability to form recombinant hybrids with phage lambda DNA (42, 212). In the meantime, this original definition has been extended on various occasions to include phages having DNA or protein sequence similarity to phage lambda or any other lambdoid phage. As in the bacterial phylogeny discussion, models of vertical and horizontal evolution and combination of the two have been proposed (106). Even if we add the about 200 prophage sequences identified in bacterial genomes to the phage database, the small number of phage sequences on which these models are based makes them tentative at best. Therefore, coverage of phage genomes is extremely low, while the number of phages is likely to be very large and phage evolution seems to be extremely fast. Therefore, it is no surprise that the phages analyzed so far show great variability.

Phage and Prophage Genomics: Evidence for the Modular Theory of Phage Evolution

The availability of more than 200 phage and 200 prophage genomes has facilitated the study of phage evolution by using the techniques developed for bacterial genome comparisons. Figure 6A shows a dot plot alignment for two prophages from two different *S. aureus* strains, N315 and Mu50A. Once again we see a familiar pattern for an intraspecies comparison: a long straight line interrupted by small regions of nonalignment. However, the type of nonalignment in prophages differs from the nonalignments in closely related bacterial genomes. In the bacterial case the gaps are frequently insertions of mobile DNA elements, while in prophage alignments the gaps are mainly DNA replacements. A DNA segment, found in one phage, is replaced in another phage by a sequence-unrelated DNA segment that frequently fulfills the same or a related function. These modular exchanges in phages had already been described by electron microscopy heteroduplex analysis of phages from enterobacteria in the pregenomics era and became the basis of one of the most popular hypotheses on phage evolution, the modular theory (24, 109, 212, 237). According to that theory, it makes no sense to speak about the evolutionary history of an entire phage genome. The genomes from lambdoid coliphages, for example, can be subdivided into 11 modules, each representing an independent genetic functional unit (head or tail genes, integration and excision; homologous recombination, and so forth). Each functional unit is represented by several alleles (modules). Only modules have an evolutionary history, which can be traced back over longer timescales. In particular, ancestral functions such as the building of a phage head or the phage tail may have a vertical evolutionary history. The order of modules on the phage genome map is fairly well conserved, while different alleles of the modules can be freely assorted (Fig. 7). This gives phage genomes a substantial genomic variability. Actually, one could describe the prophage from different *E. coli* isolates as a "swarm" of lambda genomes integrated at different chromosomal sites and sharing variable amounts of genome segments on a relatively random basis. It is currently not very clear how many different alleles exist for each module in lambdoid coliphages. Heteroduplex mapping and phage sequencing suggest that in this phage group there are perhaps about 10 different alleles for each module (48), but this might be a serious underestimation due to the incomplete sampling of lambdoid coliphages from the environment. Nevertheless, phage genomics (e.g., in lambdoid coliphages [125]) has largely confirmed the modular theory of phage evolution.

Mechanism of Phage Module Exchange: Illegitimate versus Homologous Recombination

Today, it is widely accepted that phages evolve mainly via the exchange of modules. However, the genetic mechanisms (illegitimate versus homologous or even site-specific recombination) driving this process are still a matter of debate (Fig. 8).

The Pittsburgh phage group argues for the dominance of illegitimate recombination between phages infecting a wide range of bacteria. Illegitimate recombination takes place at

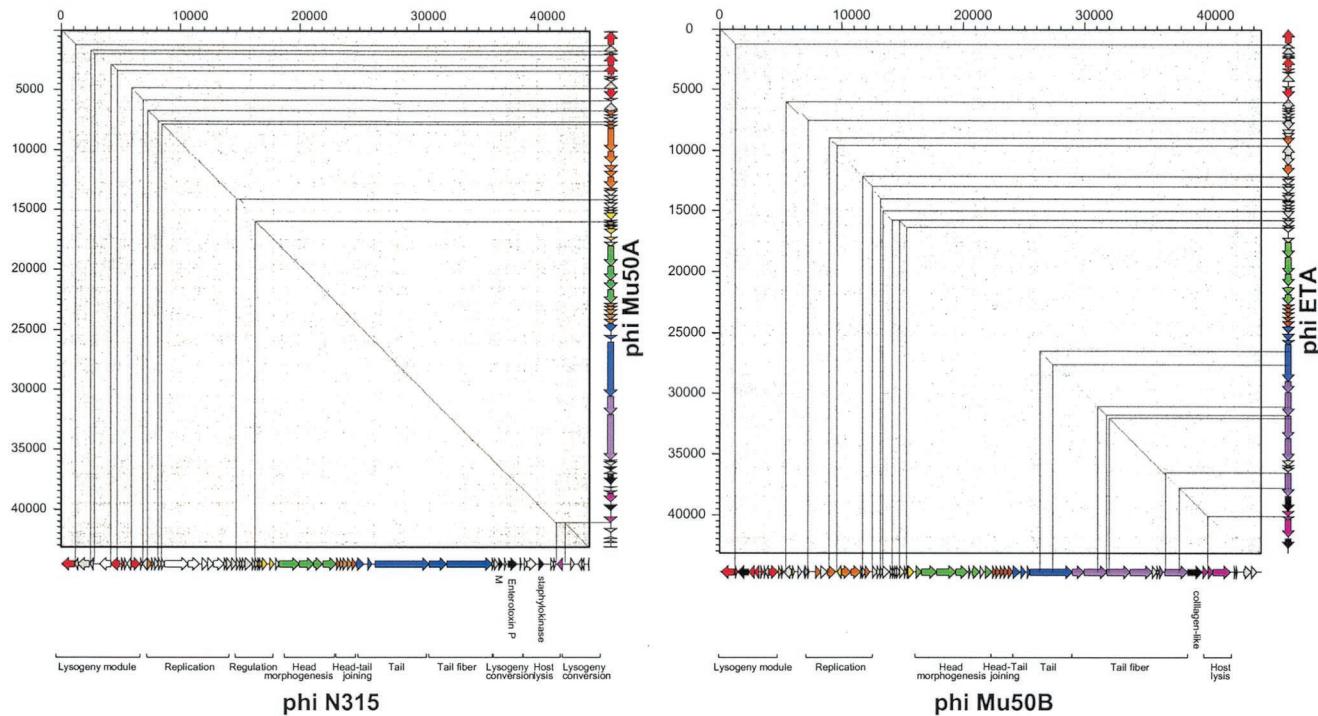


FIG. 6. Genome comparison of *S. aureus* prophages. (A) Dot plot alignment of prophages φN315 (horizontal axis) versus prophage φMu50A (vertical axis) extracted from two closely related *S. aureus* strains. The prophage maps were projected at the axes to facilitate the orientation. The modular structure of the prophages is indicated by a color code, and the putative function of the modules is indicated for prophage φN315. Candidate virulence factors are marked in black. The thin lines locate the gaps in the alignment on the prophage maps. (B) Dot plot alignment of distantly related *S. aureus* prophages φMu50B (horizontal axis) versus φETA (vertical axis). The prophage genome maps and their color-coded modular organization are shown on the axes. The thin lines locate the regions of DNA sequence identity on the prophage genome maps.

random sites between different phages. However, most of these recombination events occur within open reading frames, change the phage genome size beyond useful limits, or disrupt gene clusters, rendering the recombinant phage nonfunctional. Only a few recombination events lead to viable phage. These "productive" recombinations most probably occur in intergenic regions, which do not disrupt functional modules. In this model, order is only the consequence of selection, discarding all unsuccessful recombination events that do not lead to viable phages (105, 125, 177). Furthermore, module exchange (at low frequency) is not restricted to specific families of tailed phages or phages infecting the same bacterial species (108).

Which enzymatic functions are driving illegitimate recombination? It has been suggested that special recombination systems catalyzing illegitimate recombination may exist (107). DNA restriction enzymes encountered on injection of phage DNA into a new host bacterium can efficiently fragment the phage genome. Normally these fragments are nonfunctional and are degraded. However, at a certain low frequency, these fragments might recombine with prophages present in the bacterial chromosome. It is unclear whether this mechanism is important in nature and which enzymatic functions might be involved.

Alternatively, DNA nonhomologous end-joining mechanisms present in the bacterial cytosol might play a role. DNA fragments could originate from different phages coinfecting the same bacterium, plasmids and DNA fragments

originating from the host bacterial chromosome, or foreign DNA taken up from the environment. In this case, end-joining DNA ligase activities would be required to "assemble" these DNA fragments to yield new "mosaic" DNA assemblies. Some of the reassembled DNA fragments might turn out to be functional phages. DNA ligases of bacterial and phage origin have been known for a long time and are standard tools in the molecular biology laboratory. It is interesting to speculate whether this process might be enhanced by Ku-like proteins, which catalyze the recruitment of DNA ligase to DNA ends and ligation (68). Ku homologs have recently been identified in bacteria (68, 235, 236), and the Gam subfamily of Ku-like proteins is present in bacteria (*Campylobacter* spp., *Neisseria* spp., *Haemophilus* spp., *E. coli* O157:H7, and *S. enterica* serovar Typhi) and bacteriophage Mu (56).

Other observations argue that homologous recombination can also drive module exchange between phages (53, 155). This is supported by the presence of conserved linker sequences between modules of lambdoid coliphages, which could facilitate the exchange of DNA segments. Conserved linkers may also drive the transfer of modules between different phage families or even chromosomal loci. The *sopE* module (a moron [see below]) in *Salmonella* spp. has been studied in considerable detail. It has been found in P2-like prophages, lambdoid prophages, and the bacterial chromosome (Fig. 9). The surrounding sequences are very different, but the border regions

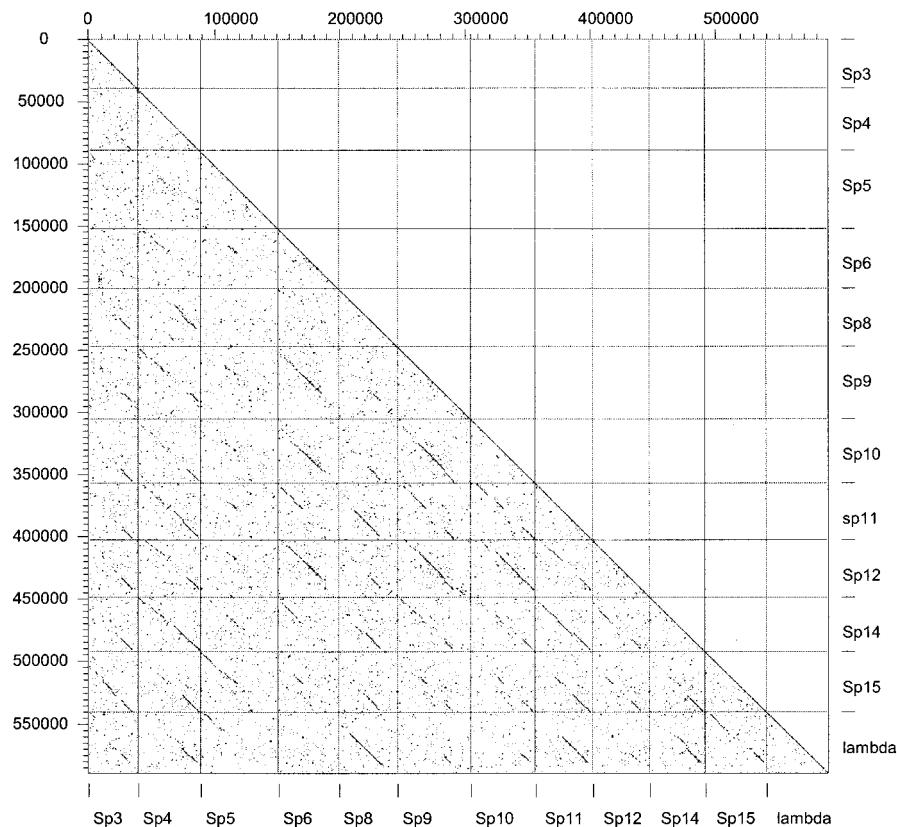


FIG. 7. Dot plot matrix for the lambda-like prophages from *E. coli* strain O157:H7 VT-2 Sakai. The genomes of the lambda-like prophage from the Sakai strain were plotted versus themselves and phage lambda. At the bottom and on the right axis, the prophage names are given as in the original publication. Some prophage genomes were inverted to correct for the distinct orientation of the prophage genomes located on the right and left sides of the terminus of replication.

flanking the *sopE* module on either side have considerable sequence similarity (157, 178) (Fig. 9). This suggests that the *sopE* cassette has been transferred between these loci by homologous recombination between flanking sequences (Fig. 8B and C). Similar observations have been made with dairy phages (35), and a recent study of moron gene cassettes in P2-like phages from *E. coli* lends further support to this notion (170). The moron gene cassettes were highly diverse in length, DNA sequence, and encoded gene products. However, the vast majority of these gene cassettes were flanked by conserved sequences (Fig. 8C). Therefore, these gene cassettes could travel between phages by homologous recombination (170).

Which recombination systems might facilitate module exchange by homologous recombination? The enzymes might be provided by the host bacterium (RecA), or they could be phage encoded (*lambda red*). Future laboratory work will have to determine which enzymatic functions of the phages and the host bacteria are involved.

In silico analyses of prophages, phage remnants, and functional phages have clearly demonstrated that all tailed phages have a common gene pool. At this stage, it is difficult to discern whether homologous or illegitimate recombination is more important. Many arguments seem to favour illegitimate recombination (Fig. 8A). Extensive alignments between prophages have identified phage pairs that share only rather limited segments of DNA identity, as demonstrated by the comparison of

the *S. aureus* prophages Mu50B and ETA from different strains (Fig. 6B). The apparent units of DNA exchange between phages are small and might cover a few adjacent genes, only a single gene, or even a gene fragment encoding a single protein domain. The great variability seen in many phage genome alignments seems to exclude the use of a few predetermined conserved linker sites for recombination. Illegitimate recombination occurs with much lower frequency than homologous recombination and would thus require longer periods or very large populations of phages to materialize. The estimated size of the global phage population (approximately 10^{31} particles) makes even unlikely recombination events observable, provided that they are conferring a sufficiently strong selective advantage to the recombinant phage. In view of the vast global number of phage particles and their fast evolution, the current database represents the phage population only poorly. Therefore, many phages with intermediate sequence similarity to the existing ones (Fig. 8B) have simply not yet been identified. If one takes these “missing links” into account, one could envision that even the transfer of one phage module into an entirely different phage could be the result of a long sequence of homologous (not illegitimate) recombinations (compare Fig. 8A and B). Much more work is needed to settle this question. However, it is very likely that all types of recombination (illegitimate, homologous, and site specific) will contribute to some extent to the module exchange between phages.

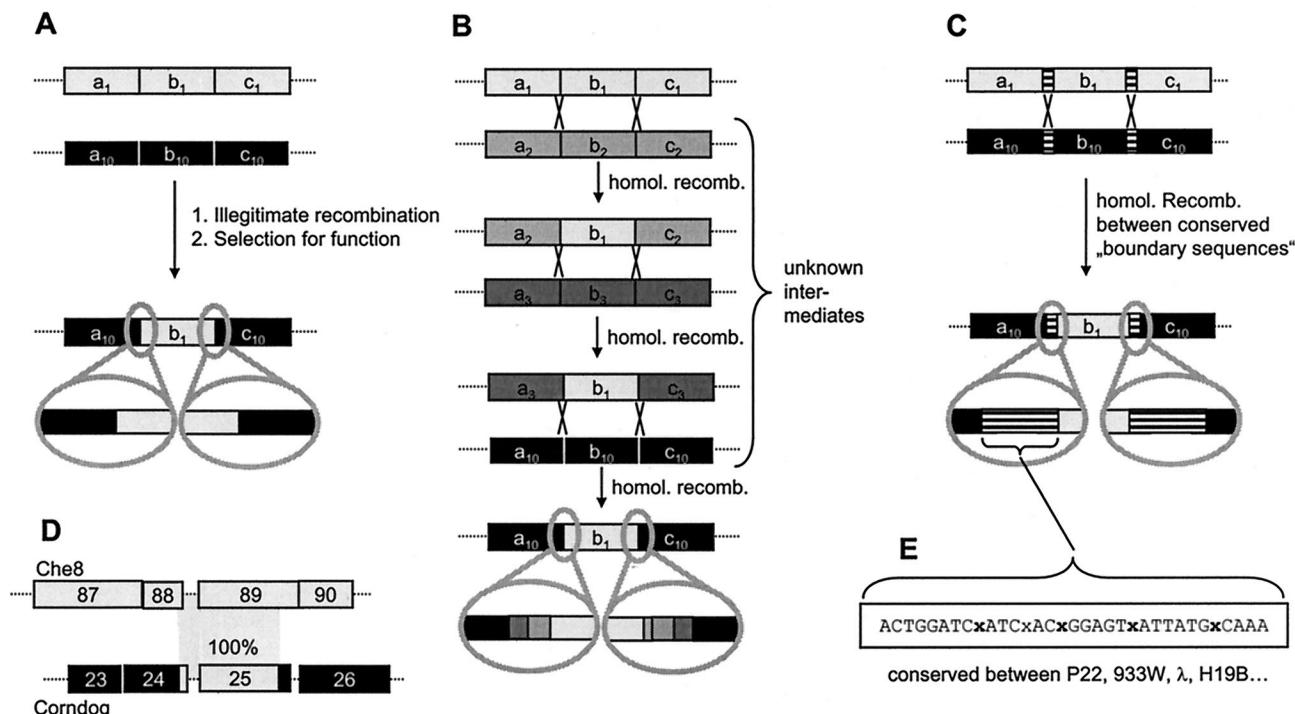


FIG. 8. Molecular models for module exchange between phages. (A) Module exchange by illegitimate recombination. (B) Module exchange by repeated homologous recombination. This results in blocks of lower sequence identity at the border between modules. The more recombinations occur, the less easily these blocks can be recognized. (C) Module exchange by recombination via conserved linker sequences. (D) Identical (100% identity) sequence elements in different regions of two unrelated phages (177). (E) Example of a conserved linker sequence in a lambdoid phage.

It can be taken for granted that modular exchange reactions diversify phage and prophage genomes at an enormous frequency; even closely related strains of the same species almost never harbor 100% identical prophages. For example, for the recently emerged food pathogen *E. coli* O157:H7, a Japanese strain and a U.S. strain were sequenced. The two strains could be perfectly aligned at the DNA sequence level (Fig. 2B), in accordance with the hypothesis that they split quite recently from a common ancestor strain resembling *E. coli* O55 strains.

Most of the differences were prophage related (Fig. 2B). Both genomes contain a prodigious number of lambda-like prophages (11 in the Sakai strain). However, despite the very short evolutionary distance separating the two strains, only one lambdoid prophage pair could be aligned over the entire length without any modular exchanges (prophage Sp3 and CP-933K). Also, the numerous lambdoid prophages within the Sakai strain differed substantially from each other, and not a single one is an exact copy of the other (Fig. 7). The acquisition of

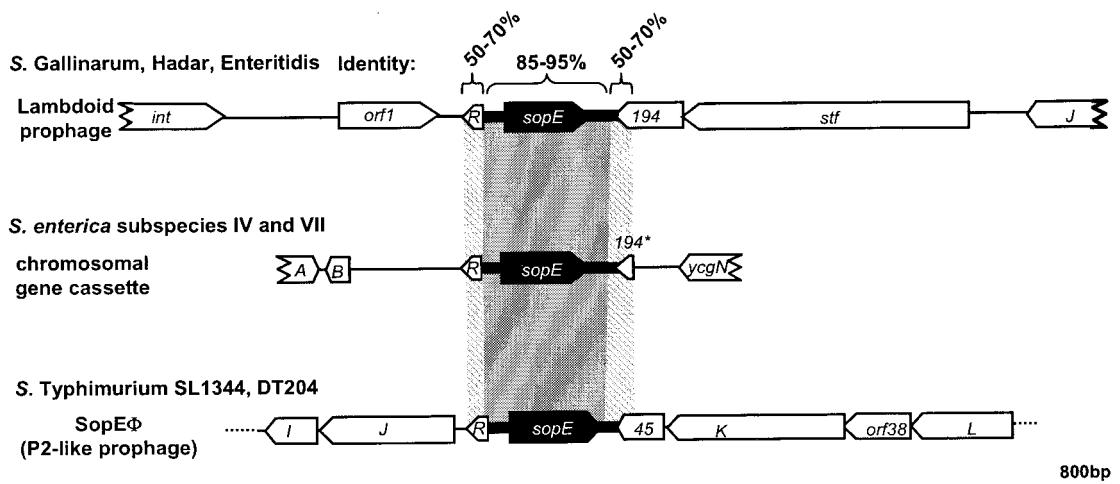


FIG. 9. Localization of the sopE moron in different *Salmonella* spp. Reprinted from reference 157 with permission.

new prophages seems to be a rapid process, as indicated by the sequential appearance of a specific set of prophages in *S. pyogenes* strains collected by hospital laboratories over the last 70 years. Historical surveys revealed the emergence of highly virulent *S. pyogenes* strains over the last half century that was correlated with the acquisition of a specific set of prophages (14).

Loss and Decay of Prophages

Prophages seem to be only transient passengers on the bacterial chromosomes, at least when seen on an evolutionary timescale. Theoretical arguments suggested a series of events leading from the accumulation of mutations to massive loss of prophage DNA and ultimate disappearance of the prophage (46, 136). This scenario is backed by a number of observations. Many prophages are no longer inducible: only one of the many lambda-like prophages in *E. coli* O157 EDL933 and none in strain Sakai could be induced. In two different *Lactobacillus* species, none of the multiple prophages could be induced (225). In some cases, inactivating point mutations were identified by bioinformatic analysis, e.g., introduction of stop codons into the replisome organizer gene and the portal protein-encoding gene in *S. pyogenes* prophages SF370.2 and SF370.3 (67) or inactivation of the N antitermination genes in the lambdoid coliphages from O157 (136). However, prophage inactivation is not a universal process, as demonstrated by an M3 serotype *S. pyogenes* strain in which all five prophages could be induced, although with different inducers and at different efficiencies (10).

Also, the frequent observation of prophage remnants (*E. coli* K-12, *S. flexneri*, *Lactococcus lactis*, and the *S. enterica* serovar Typhimurium *sopE2* and *sspH2* loci [Fig. 10]) could be easily explained by an ongoing prophage decay process. If the prophage decay process is slow with respect to bacterial speciation, one would expect the conservation of closely related prophage remnants in different bacterial isolates from a given species. A few closely related prophage remnants were detected in *S. pyogenes* (45) and *E. coli* (47). However, they are the exceptions and are not even widely distributed in the investigated species. In fact, even closely related bacteria generally do not share prophage remnants: the remnants in *E. coli* K-12 and *S. flexneri* (Fig. 2A), which diverged from K-12 perhaps just 1 million years ago, are distinct, as was the case when K-12 was compared to other sequenced *E. coli* strains. This observation suggests that the average time for acquisition and subsequent loss of prophages is shorter than the timescale of strain differentiation within a bacterial species.

Role of Phages in Bacterial Evolution

It is conceivable that bacteria evolve with two gears. The slow mode (often in 1-million-year range) is based on the usual mechanisms of vertical evolution mediating a step-by-step genetic adaptation to their approximate environment.

Horizontal gene transfer can be regarded as the fast mode of evolution (timescale of years to decades). New sets of genes are acquired by transduction, transposition, transformation, and, last but not least, lysogenization with phages. Most of these gains might be ephemeral, and the genes are as easily

gained as lost (134, 137). What counts is a momentary selective advantage over competing bacteria, especially in environments that are quickly changing. This changing environment can be the body surfaces of new host species that are rapidly proliferating in the ecosphere (e.g., humans) or settings with unusual host densities (animal farming or human urbanization). New industrial food preparation and health care techniques, international travel and transportation, and wars create enormous possibilities for exploitation by microbes that can invade these new niches. The impact of lateral gene transfer is so obvious in some of these settings (e.g., antibiotic resistance gene acquisition via plasmids and transposons in the hospital environment) that it would be surprising if it were restricted to these examples. The acquisition of mobile DNA can provide the necessary genetic material on a timescale that allows bacteria to quickly exploit these ecological "opportunities." Phage DNA fulfills a number of criteria for being an ideal vehicle for lateral gene transfer. Actually, some bacteria seem to use phages as gene transfer particles to shuttle pathogenicity islets (*S. aureus* phage ϕ 80 α [see below]) or random samples of chromosomal DNA (*Bacillus subtilis* prophage PBSX) (173).

In this model, different combinations of mobile DNA can be explored, and suitable combinations are maintained and further developed, leading to genotypes that suddenly fill old and newly created niches. In the context of pathogenic bacteria, we see these events as emerging infectious diseases (flesh-eating streptococci) or emerging food pathogens (*E. coli* O157). As discussed below, phages play a key role in these short-term adaptation processes.

The effect of this quick process on the long-term evolution of bacteria is less certain. Apparently only very small amounts of prophage DNA are fixed in the bacterial chromosome. However, it is a matter of speculation why phages do not accumulate to large numbers (or do so only in rare cases, e.g., *E. coli* O157). Are entire phages simply lost by excision reactions? Or are their genomes degraded so that they leave behind only a few remnants and possibly some fitness factors? The widespread occurrence of isolated phage-like integrase genes in bacterial chromosomes might be markers of previous prophage integration events (46, 67). However, only in some of the cases are a few phage-like genes (frequently repressor genes) found in the vicinity of these isolated integrase genes, which are sometimes still transcribed (224). At this stage, it is almost impossible to discern whether phage-encoded fitness factors might also have remained in the chromosome. Their phage origin cannot be recognized easily by simple sequence analyses. Some virulence genes or small genes clusters encoding fitness functions flanked by isolated phage genes might thus represent what remains after the prophages have decayed.

PHAGE-MEDIATED GENE TRANSFER AND INACTIVATION

Phage-mediated horizontal gene transfer occurs via transduction or lysogenic conversion. The global rate of phage-mediated genetic modification in bacteria has been estimated as being up to 20×10^{15} gene transfer events per s (39). In addition, bacterial gene disruption can occur by prophage integration into the bacterial genome.

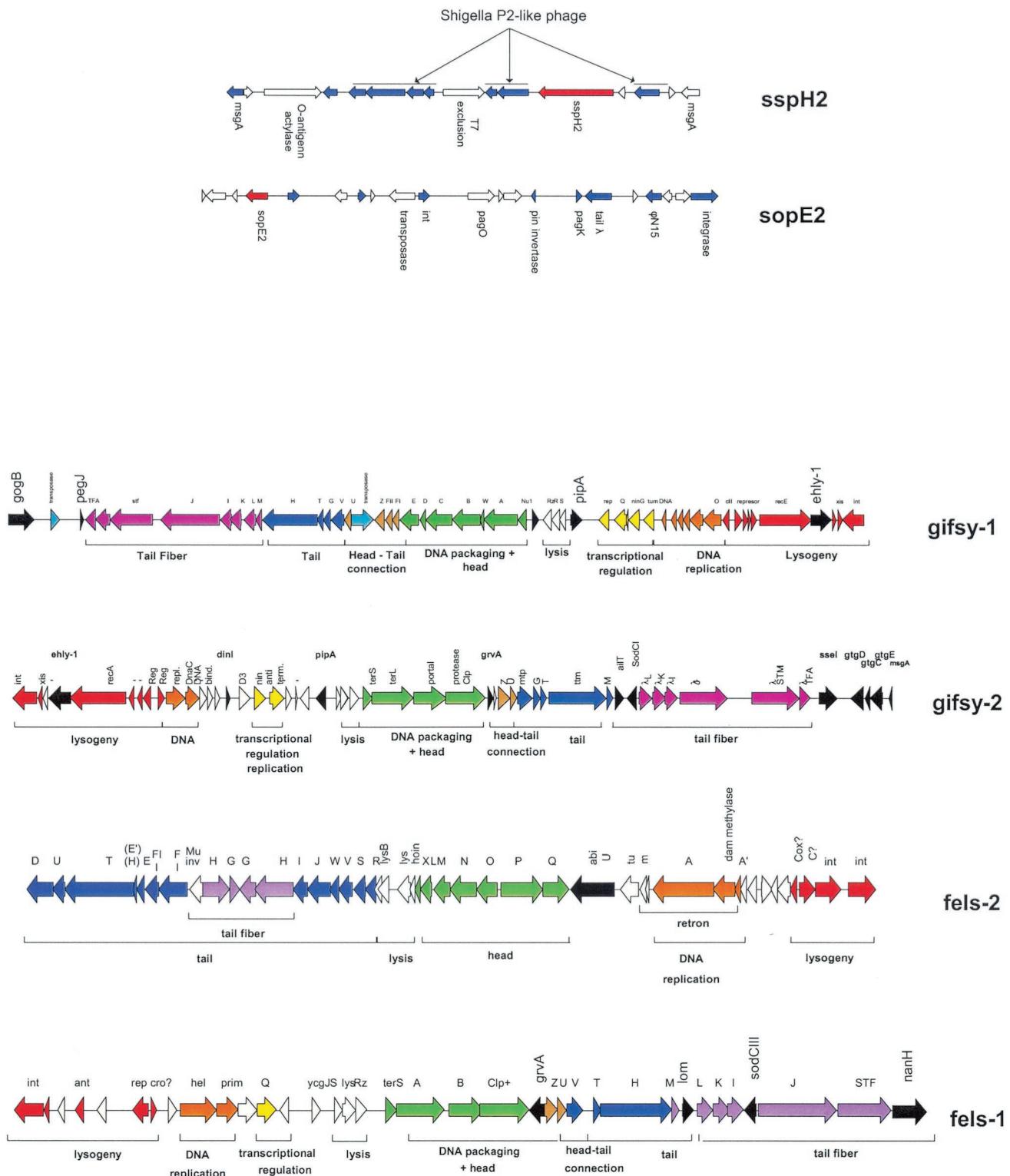


FIG. 10. Prophages and phage remnants in the *S. enterica* serovar Typhimurium strain LT2. (Top) Prophage remnants encoding the type III effector proteins SspH2 and SopE2. Phage-like genes in the vicinity of *sspH2* and *sopE2* are highlighted in blue. Selected genes are annotated. (Bottom) Genome maps of the four LT2 prophages. The modular structure of the prophages is indicated by the color code and the annotated brackets under the prophage genome maps. Proven or suspected fitness and virulence genes are colored in black.

Transduction

Generalized transduction is a “sloppy” feature observed with many bacteriophages. After the empty phage heads are completed, the phage DNA must be packaged. This process is quite accurate; however, DNA fragments of the host genome are packaged instead of the phage DNA at a finite frequency. This results in fully functional phage particles, which can attach and deliver the packaged DNA into suitable bacteria. Due to the absence of the phage DNA, this does not harm the bacterium. Instead, the injected foreign bacterial DNA can be incorporated into the genome. This is a typical example of phage-mediated horizontal gene transfer. Transducing phages have been observed in many bacteria including *Salmonella* spp. (200), *Streptomyces* spp. (38), and *Listeria* spp. (113).

Lysogenic Conversion

The acquisition of prophages would be an irrelevant process for the evolution of pathogenic bacteria if phages did not transfer useful genes to the lysogen. Some phage genes are known to increase the survival fitness of lysogens. The phage repressor and superinfection exclusion functions confer selective advantage to the lysogen by providing immunity against lytic infection. This was illustrated in a recent study of different *S. enterica* serovar Typhimurium strains harboring different sets of prophages. Lysogens like these often release low titers of phage (10^2 to 10^4 CFU/ml). These initially low titers of phage were sufficient to kick off an efficient decimation of a competing nonlysogenic strain (23). There are also some classical data that *E. coli* cells containing prophages (lambda, Mu, P2, and even cryptic prophages) grow quicker than nonlysogenic *E. coli* strains (68a, 68b, 141a).

It has been quite an exciting discovery that phages can also play an important role in the emergence of pathogens. This was recognized relatively early for toxins of *Corynebacterium diphtheriae* (diphtheria), *Clostridium botulinum* (botulism), *Streptococcus pyogenes* (scarlet fever), *Staphylococcus aureus* (food poisoning), and *E. coli* (Shiga toxin), which are all phage encoded. As far as we know, these genes do not play a role in the life cycle of the phages. The list of phage-encoded fitness factors is rapidly growing and now involves a wide range of different genes (Table 1). These factors include ADP-ribosyl transferase toxins, superantigens, lipopolysaccharide-modifying enzymes, type III effector proteins, detoxifying enzymes, hydrolytic enzymes, and proteins conferring serum resistance. In exceptional cases, phage tail genes seem to have developed dual functions and also serve as adhesion proteins for bacterial host attachment (*Streptococcus mitis* *pblA* and *pblB* genes). Many more prophage genes with sequence links to potential virulence factors were observed in the bioinformatic analysis of the genomes from bacterial pathogens, but experimental evidence for their role in bacterial pathogenicity is still largely lacking.

Another important lead was provided by the early observation that the toxin genes in *Corynebacterium* and *S. pyogenes* were located next to the phage attachment site (see Fig. 16). This specific location led to the hypothesis that these prophage genes actually represent bacterial genes that were acquired by a faulty excision process from a previous bacterial host. Some-

times these toxin genes still showed a clearly distinct G+C content, pointing to an unusual bacterial host as source for this DNA (78). However, to our knowledge, in no case was such a scenario demonstrated by direct experimental evidence. Fitness factors and “extra” genes of no known phage function are a relatively consistent finding in prophages from low-G+C-content gram-positive bacteria. “Extra” genes are not restricted to prophages from pathogenic bacteria but are also found in free-living bacteria and in bacterial commensals, where they are frequently located near the right attachment site, i.e., between the phage lytic gene and the transition zone from phage into bacterial DNA. Transcription studies of free-living dairy bacteria (*Streptococcus thermophilus* and *Lactococcus lactis*) (25, 224) and gut commensals (*Lactobacillus plantarum*) demonstrated that “extra” genes belong to the few constitutively transcribed prophage genes (225). Combined bioinformatic and transcription analysis revealed further extra genes between the phage repressor and the integrase gene near the left attachment site (226). In addition, isolated transcribed “extra” genes were found in the central part of the genome from prophages of low-G+C-content gram-positive bacteria (e.g., tRNA genes) (227).

In prophages from gram-negative bacteria, “extra” genes were also identified near both prophage DNA ends. Examples are O-antigen-modifying enzymes inserted between the phage integrase gene and the left attachment site in bacteriophage P22 (222) or genes inserted downstream of the phage tail genes (26). The location at both prophage ends has an intrinsic logic: an extension of these prophage transcripts would run into the bacterial DNA or as an antimessenger into the prophage DNA, in both cases preventing an accidental induction of the prophage via transcription of prophage DNA. Another frequent location for extra genes in lambdoid prophages from gram-negative bacteria was next to the N- or Q-like antiterminator genes in the middle of the prophage genome (26). In these locations, accidental induction of the prophage is a real risk. Therefore, these extra genes were postulated to be flanked by an independent promoter and terminator structure. For this independently transcribed gene cassette, the Pittsburgh phage group coined the expression “moron” (more DNA [107]). No explicit predictions were made for the possible origin of these morons. Overall, the hot spots for insertion of “extra” genes are located at strategic positions, which minimize interference with the functions required during lysogeny and for lytic infection.

Gene Disruption

Gain of virulence genes is not the only mechanism by which pathogenicity develops. Pathogenic bacteria also develop from commensal bacteria by loss of genes. *Shigella* is a prominent example of the gain of virulence by loss of *E. coli*-specific genes, namely flagellar genes and *cadA* (5, 144). The latter encodes an enzyme required for cadaverine synthesis, and cadaverine has been shown to inhibit the *Shigella* enterotoxin function. On a smaller scale, prophages can cause single-gene loss when they integrate disruptively into host genes. A prominent locus of prophage integration involves the tRNA genes (41a), and many phage attachment sites carry the necessary DNA sequences to reconstitute a functional tRNA after pro-

phage integration into the tRNA gene. Apparently, loss of tRNA genes could lead to fitness loss of the lysogen. However, not all prophage integration events reconstitute a functional tRNA gene. For example, two distinct *Lactobacillus* prophages share a nearly identical integrase gene but are located in two different tRNA genes at opposite locations on the bacterial genome. One prophage reconstitutes a functional tRNA, while the other apparently uses a secondary attachment site in a distinct tRNA gene of the same cell, which is disrupted by the integration event (226). Interestingly, both of these prophages, and many others, carry multiple tRNA genes that are transcribed from the prophage. This constellation might compensate for tRNA gene inactivation by prophages.

In other cases, prophages integrate into protein-encoding genes and lysogenization is linked to the loss of a protein function (negative lysogenic conversion phenotype). Well-characterized cases are the lipase- and the β -toxin-negative phenotype due to the integration of prophage L54a and phi13, respectively, into the *S. aureus* genome (54). Even when prophages integrate into intergenic DNA, the ca. 40 kb of prophage DNA might disrupt the coordinated transcription of genes belonging to the same transcription unit (45). In conclusion, there are ample possibilities to explain how prophage integration can disrupt or modulate bacterial gene expression and thereby alter bacterial fitness or virulence.

LYSOGENIC CONVERSION: ROLE OF PROPHAGE MORONS

Morons of Temperate Phages, Prophages, and Phage Remnants in Pathogenic Bacteria

Morons are thought to enhance phage replication in times when the temperate phage is residing as a prophage in the chromosome of a bacterium. The effect is indirect since the moron-encoded functions enhance the fitness of the lysogen and improve the fitness of the phage only passively via its propagation with its host bacterium (67, 107). This hypothesis provides the theoretical framework for phage-mediated horizontal transfer of fitness factors between bacteria. It is thought that this mechanism is the most important one by which phages affect the evolution of pathogenic bacteria.

For the purpose of this review, we use the term "moron" for all extra genes present in prophage genomes which do not have a phage function but (may) act as fitness factors for the lysogen. Table 1 provides an overview of the known and putative morons of phages and phage-like elements found in pathogenic bacteria. Often, there is quite a complex interplay between phage and bacterial functions in order to express the moron properly and to provide a selective benefit for the lysogen.

Examples of Distribution and Function of Morons

To provide a benefit for the lysogens, a moron must fulfill several requirements.

(i) The moron has to be useful in the ecological niche of the lysogen. Clearly, a host cell invasion factor would be of little use for an extracellular pathogen. A moron can provide a benefit in three ways: it enhances the fitness of the lysogen in

the bacterium's "old" ecological niche (in this case, it would allow the lysogen to outcompete the other inhabitants of this niche); its function allows the bacterium to counteract a sudden change occurring in its old niche (the former competitors are wiped out by the environmental change and the lysogen can proliferate); and its function allows the bacterium to conquer a new niche.

(ii) The expression of the moron function must be coordinated with the functions of the host bacterium. If the timing of moron expression is not well controlled, it cannot provide a benefit.

(iii) In some cases the moron function relies on the proper function of and interaction with other bacterial factors. This situation is somewhat similar to that in operons encoding entire metabolic pathways (138): one function makes sense only in the presence of the others. For example, in gram-negative bacteria, extracellular enzymes or toxins requiring a specialized export apparatus will be functional only if the lysogen provides the proper transport systems.

Prophages in *S. enterica* serovar Typhimurium. Recently *S. enterica* serovar Typhimurium has emerged as an excellent example of the integration of moron function into the biology of the host bacterium. *S. enterica* serovar Typhimurium is the causative agent of *Salmonella* food poisoning. A wide variety of animals can be infected with *S. enterica* serovar Typhimurium when they ingest contaminated water or foodstuff. Infected animals excrete *S. enterica* serovar Typhimurium in the feces. The fecal bacteria contaminate water and foodstuff, where they can persist—and multiply under proper growth conditions—for days or weeks (239). Therefore, *S. enterica* serovar Typhimurium is adapted to two principally different ecological settings: niches in the environment and the intestines of different animal species (239).

More than 200 different *S. enterica* serovar Typhimurium strains have been identified. Typically, these strains cause epidemics for periods of about one decade, one specific strain dominates and causes a large percentage of all animal or human infections, while the other *S. enterica* serovar Typhimurium strains are isolated only rarely. Finally, the incidence of this epidemic strain declines and a new epidemic strain takes over. The exact reasons for this are unknown, but it is thought that slight strain-specific differences in virulence might play a role. This hypothesis is based on the observation that different *S. enterica* serovar Typhimurium strains have different combinations of fitness factor-encoding mobile genetic elements and phages. The recent identification of morons encoding fitness factors in several of the *S. enterica* serovar Typhimurium prophages has led to the hypothesis that phage-mediated reassortment of virulence factors and fitness factors is a key driving force in the optimization of the *Salmonella*-host interaction and the emergence of new epidemic clones (81, 156). Some of these moron-encoded functions have been studied in detail. These are discussed to illustrate how moron function can be precisely integrated with the functions of the host bacterium. A selection of *S. enterica* serovar Typhimurium fitness factors is given in Table 2.

Several *S. enterica* serovar Typhimurium prophages encode so-called type III effector proteins (Table 2). Type III effector proteins are a specific class of virulence factors which are injected by the bacterium into animal cells via a type III se-

TABLE 2. Selection of *S. enterica* serovar Typhimurium fitness factors

Fitness factor	Location ^a	Function
Type III effector proteins		
SipA,B,C,D	Chr. (SPI-1)	SPI-1 TTSS effector + transport, actin manipulation
SptP	Chr. (SPI-1)	SPI-1 TTSS effector, tyrosine phosphatase, GAP ^b
AvrA	Chr. (SPI-1)	SPI-1 TTSS effector, NF-κB signaling
SopB	Chr. (SPI-5)	SPI-1 TTSS effector, inositol phosphate phosphatase
SopA	Chr.	SPI-1 TTSS effector, function unclear
SlrP	Chr.	SPI-1 TTSS effector, function unclear
SopD	Chr.	SPI-1 TTSS effector, function unclear
SopE	SopEΦ	SPI-1 TTSS effector, exchange factor for Cdc42, Rac1
SopE2	Phage remn.	SPI-1 TTSS effector, exchange factor for Cdc42
SspH1	GIFSY-3	SPI-1/2 TTSS effector, NF-κB signaling
SseA-G	Chr. (SPI-2)	SPI-2 TTSS effectors and transport proteins
SsaB	Chr. (SPI-2)	SPI-2 TTSS effector, function unclear
SseJ	Chr.	SPI-2 TTSS effector, function unclear
SifA	Chr.	SPI-2 TTSS effector, vesicle traffic
SseI	GIFSY-2	SPI-2 TTSS effector, function unclear
SspH2	Phage remn.	SPI-2 TTSS effector, function unclear
SopD2	Chr.	SPI-2 TTSS effector, function unclear
SifB	Chr.	SPI-2 TTSS effector, function unclear
GogB	GIFSY-1	Possible SPI-2 TTSS effector, function unclear
Others		
SpvB	Plasmid	ADP ribosyltransferase for actin
SodCI	GIFSY-2	Superoxide dismutase
SodCII	Chr.	Superoxide dismutase
SodCIII	Fels-1	Superoxide dismutase
NanH	Fels-1	Neuraminidase, role in virulence unclear
FliC	Chr.	Flagella, chemotaxis and inflammation
PefA	Plasmid	Plasmid-encoded fimbriae, adhesion
FimA	Chr.	Type I fimbriae
LpfA	Chr.	Long polar fimbriae
LPS	Chr./phage	Lipopolysaccharide, principal outer membrane component, protection, inflammation

^a Chr., chromosome; remn., remnant.^b GAP, GTPase activating protein.

cretion system (TTSS; Fig. 11B). Inside the animal cell, these type III effector proteins manipulate signal transduction pathways. In this way, the bacterium can manipulate the host cell response for its own benefit. TTSS are found in several gram-negative bacteria living in close association with plants or animals (88, 119). Future work will have to address whether any of these bacteria besides *Salmonella* spp. may also harbor phages encoding type III effector proteins.

TTSS are of key importance in *Salmonella* pathogenesis (87). *S. enterica* serovar Typhimurium encodes two TTSS (not counting the flagellar apparatus). The TTSS encoded in *Salmonella* pathogenicity island 1 (SPI-1 TTSS [Fig. 11A]) is active during the early, gut-associated stages of the infection (234). In contrast, the SPI-2 TTSS is essential for the survival of *S. enterica* serovar Typhimurium inside phagocytic cells during later stages of systemic (typhoid-like) disease (98). Expression of the SPI-1 and SPI-2 TTSS is strictly and inversely regulated. The SPI-1 TTSS is expressed when the bacteria grow in nutrient-rich high-osmolarity environments thought to simulate the intestinal lumen. SPI-2 is expressed when the bacteria replicate intracellularly in a vacuole with low ion concentrations and acidic pH (98).

How does the SPI-1 TTSS induce diarrhea? The type III effector protein cocktail injected into gut cells via the SPI-1 TTSS is very complex and consists of at least 11 different proteins including SopB, SopE2, AvrA, SipB, SptP, SipC,

SipA, SspH1, SopD, SlrP, SopA, and, in some strains, SopE (87). Each of these effector proteins seems to manipulate a specific signaling pathway of the mammalian cell, which together provoke a strong intestinal inflammation. It is becoming increasingly clear that *Salmonella* requires a fine-tuned delivery of the effector proteins into the intestinal cells of the infected animal. A similar picture emerges for the SPI-2 effector proteins. Overall, it is clear that acquisition of new type III effector proteins can alter virulence of *Salmonella* spp. Specificity for certain host animals might also be an issue, because signaling pathways are thought to be wired slightly different in every animal. Many prophages have been identified in *Salmonella* spp., and more are identified as additional *Salmonella* genomes are sequenced. Most of them belong to the P2 family (SopEΦ, Fels-1, and Fels-2) or the lambda family (GIFSY-1, GIFSY-2, GIFSY-3, and P22), and many carry morons encoding TTSS effector proteins and other fitness factors (Fig. 10). These are discussed in detail, below.

(i) **The *sopE* moron of SopEΦ.** SopEΦ is a P2-like phage. It has been identified and sequenced in *S. enterica* serovar Typhimurium strain DT49/DT204. This strain caused a major epidemic in England and the former German Democratic Republic in the 1970s and 1980s (156). In its tail and tail fiber region, SopEΦ carries a moron containing *sopE* (101, 157, 177). In SopEΦ lysogens, SopE is part of the effector protein cocktail that is injected into cells of the mammalian gut by the

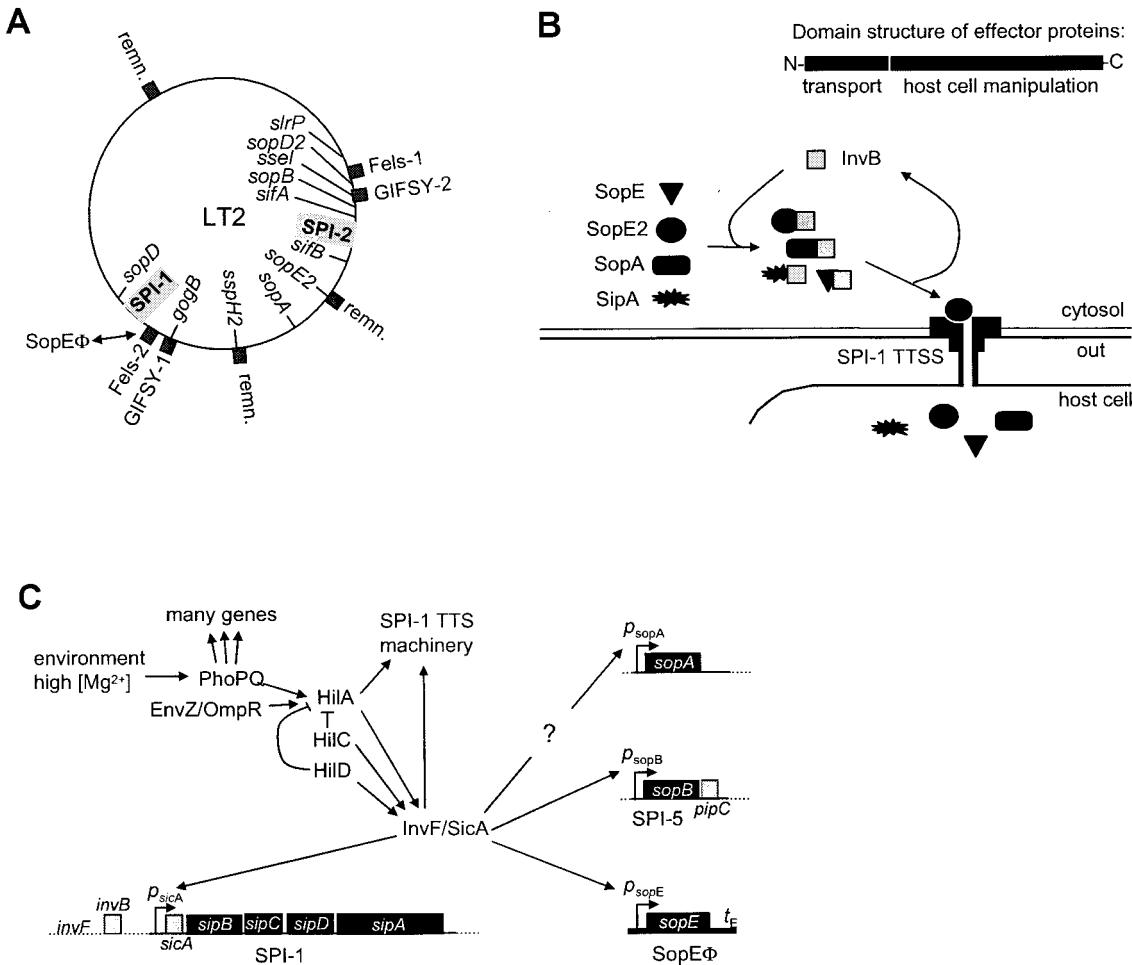


FIG. 11. *sopE* moron of *S. enterica* serovar Typhimurium. (A) Chromosomal map of *S. enterica* serovar Typhimurium LT2. The chromosomal locations of the SPI-1 and SPI-2 TTSS, the type III effector proteins (inside the circle), and of the prophages and phage remnants (remn.) (outside the circle) are shown. (B) Role of the chaperone InvB in the transport of the type III effector proteins SopE, SopE2, SopA, and SipA. (Inset) General domain structure of type III effector proteins. N terminus, chaperone interaction and transport via the TTSS; rest of the effector protein, domains for manipulation of signaling cascades in eukaryotic cells. (C) Outline of the regulation of SPI-1 TTSS effector protein expression in *S. enterica* serovar Typhimurium.

SPI-1 TTSS (101, 242) (Fig. 11B). It has been speculated that lysogenic conversion with SopEΦ was an important step in the emergence of this epidemic strain (156).

SopE is a so-called G-nucleotide exchange factor for RhoGTPases, central switches of mammalian cell physiology. For this reason, it is not surprising that SopE has profound effects once it is delivered into animal cells (100). In order to function, the expression of *sopE* must be tightly controlled and coregulated with the SPI-1 TTSS and the other effector proteins. This has been studied in considerable detail in the natural lysogen *S. enterica* serovar Typhimurium SL1344 (Fig. 11C).

sopE expression depends on the SPI-1-encoded proteins SicA and InvF (Fig. 11C) (70, 127, 218). The InvF/SicA complex is a positive regulator for the *sicAsipBCDA* operon, *sopB* and *sopE* (57–59, 70, 218). This ensures coexpression of *sopE* with the other type III effector proteins.

Recent work on the transport of SopE via the SPI-1 TTSS has identified a second mechanism ensuring proper functional

integration of the *sopE* moron. After production in the bacterial cytosol, SopE binds to the SPI-1 encoded chaperone InvB (69, 114). Type III secretion chaperones such as InvB are small acidic proteins which bind to the N-terminal region of the effector proteins (55, 175, 208). This interaction is required for proper recognition and transport by the TTSS (19). Recent work on the chaperone InvB has revealed that in addition to SopE, it is the chaperone for the effector proteins SipA, SopE2, and SopA (Fig. 11B) (31, 69, 69a, 114).

In conclusion, the proper function of the *sopE* moron in *S. enterica* serovar Typhimurium is ensured in two ways: *sopE* is coregulated with other SPI-1 effector proteins by a common regulator, and its proper recognition and transport via the SPI-1 TTSS is ensured by the SPI-1-encoded type III chaperone InvB (Fig. 11). This is quite an interesting situation because neither the regulators nor the chaperone are encoded by SopEΦ; *sopE* travels without this “dead freight.”

(ii) Other *S. enterica* serovar Typhimurium prophages: Fels-1, Fels-2, Gifsy-1, Gifsy-2, Gifsy-3, E34, and ST64B.

Besides SopEΦ, two other P2-like prophages (Fels-1 and Fels-2) have been identified in *S. enterica* serovar Typhimurium strains (2, 81) and several other *Salmonella* spp. also harbor Fels-1- or Fels-2-like prophages (117, 184, 194).

Fels-2 carries one moron of unknown function (*abiU* [26]). Three morons (*sodCIII*, *nanH*, and *grvA*) have been spotted in Fels-1 from *S. enterica* serovar Typhimurium LT-2: *sodCIII* encodes a superoxide dismutase (81). Superoxide dismutases are thought to protect bacteria from oxygen radicals, which are produced by macrophages (64, 74). *nanH* encodes a neuraminidase/sialidase, NanH. There is only circumstantial evidence for a role of *nanH* in *Salmonella* virulence: neuraminic acid-containing oligosaccharides are present on many animal cell surface glycoproteins and glycolipids, and neuraminidases are found in a variety of pathogenic bacteria such as *V. cholerae* and *Clostridium perfringens* (44, 123). However, the function of neuraminidases in virulence has remained enigmatic. Similarly, it has remained unclear if *nanH* can contribute to *Salmonella* virulence. The function of the *grvA* moron of Fels-1 has also remained unknown. Considering this large arsenal of putative fitness factor-encoding morons, it seems a bit surprising that lysogenic conversion of a *fels-1*-lacking *S. enterica* serovar Typhimurium mutant strain with Fels-1 did not enhance mouse virulence (81). Each of the morons might make only a small contribution. However, one also has to consider that *S. enterica* serovar Typhimurium can infect host animals as different as reptiles, birds, and mammals. It is conceivable that the morons enhance *S. enterica* serovar Typhimurium fitness only in very specific host animals and not in laboratory mice. Yet another possibility is that research has focused too much on the *S. enterica* serovar Typhimurium-animal interaction and that some of the morons might actually provide a selective advantage in the second niche of *S. enterica* serovar Typhimurium: the environment.

The lambdoid GIFSY-1 prophage is present in several *S. enterica* serovar Typhimurium strains. GIFSY-1 affects mouse virulence in certain *S. enterica* serovar Typhimurium strains lacking GIFSY-2 (79). The exact reasons for this are still unclear, but one can speculate that one of the GIFSY-1 morons might be involved. The *gogB* moron of GIFSY-1 (Fig. 10 and 11A) is expressed under control of the master regulatory system *ssrAB* of the SPI-2 TTSS (81). Future work will have to confirm whether GogB is an effector protein of the SPI-2 TTSS. The GIFSY-1 *gogA* moron is similar to PipA of *S. enterica* serovar Dublin and 98% identical to the *gtgA* moron of GIFSY-2, but so far it has no known function (81). Finally, GIFSY-1 includes the *ehly-1* moron (26) and the *gipA* moron, which encodes a putative transposase which enhances growth or survival in the Peyer's patches of the murine small intestine (207).

The lambdoid GIFSY-2 prophage has been identified in *S. enterica* Typhimurium strains LT2 and ATCC 14028 (Fig. 10 and 11A). Curing these strains from GIFSY-2 reduces virulence (79, 80). GIFSY-2 carries a multitude of known or putative fitness factors (Fig. 10). The periplasmic CuZn-superoxide dismutase SodCI was the first GIFSY-2 virulence factor described (79). Interestingly, most if not all *Salmonella* strains already encode such an enzyme (via the *sodCII* gene) in the chromosome. So why should the GIFSY-2-encoded enzyme enhance the fitness of a lysogen? Gene expression analyses

have revealed that the chromosomal *sodCII* gene is abundantly expressed in stationary-phase cultures in rich broth (RpoS control [74]) but only scarcely expressed when the bacteria are growing inside phagocytes or an infected spleen (220). In contrast, *sodCI* is expressed abundantly and protects the *S. enterica* serovar Typhimurium bacteria in this niche. In other words, *sodCI* moron expression requires correct timing to confer a selective advantage to GIFSY-2 lysogens. However, the regulatory cascades are unknown. The GIFSY-2 moron *gtgE* also contributes to *S. enterica* serovar Typhimurium virulence (111). Western blot analyses revealed that it is expressed during intracellular growth as well as in nutrient broth (221). The *grvA* moron carries a highly interesting gene, which has been classified as an "avirulence" factor (112). Its sequence is not homologous to any other known gene. Overexpression and deletion of *grvA* increased virulence (112). The exact mechanism underlying this observation is still unclear. Nevertheless, this illustrates that proper regulation of moron expression is of great importance. The *sseI/gtgB/srfH* moron encodes a type III effector protein for the SPI-2 TTSS (150). It is expressed intracellularly under the control of the master regulator of the SPI-2 TTSS (221, 243). This is somewhat reminiscent of the *sopE* moron described above and should guarantee proper integration of this moron function with SPI-2. However, *sseI* mutants did not show a virulence defect in mice (111). Other GIFSY-2 morons have sequence similarity to known or suspected fitness factors including *ailT*, *pipA*, *msgA*, and *gtgCD* (Table 1). So far, their function has not been studied in detail.

GIFSY-3 harbors two morons with possible fitness function (81, 150). The first, *pagJ*, was first identified by virtue of its regulation by the PhoPQ system (13, 93), a key regulator of *Salmonella* virulence functions (Fig. 11C) (92). The molecular function of *pagJ* is unclear. The second moron contains *sspH1*, a type III effector protein, which inhibits NF-κB-dependent production of inflammatory signals by infected animal cells (99). The C-terminal domain of SspH1, which manipulates cell signaling, harbors a leucine-rich sequence motif and is similar to type III effector proteins from various pathogenic bacteria and to SlrP and SspH2 (see below) from *S. enterica* serovar Typhimurium (150). The N-terminal region of SspH1 shows similarity to an even larger family of type III effector proteins from *S. enterica* serovar Typhimurium (SlrP, SspH1, SspH2, SseJ, SseI, SifA, SifB). This common N-terminal region harbors the signals for chaperone binding and transport via TTSS (55, 150). Probably, these type III effector proteins use the same chaperone for transport. This chaperone has not yet been identified but could ensure the function of *sspH1* and its coordination with the TTSS and the other effectors (Fig. 11B).

Regulation of *sspH1* expression is quite unusual for a type III effector protein. It is not under control of the "master regulators" for either SPI-1 or SPI-2 expression. In fact, SspH1 can be transported via both TTSS (150). Therefore, the *sspH1* moron might represent an evolutionary intermediate in the optimal integration of a moron into the functional networks of the host bacterium. Alternatively, *sspH1* might confer some (so far unknown) advantage during the extracellular and intracellular stages of the disease. In spite of these *pagJ* and *sspH1* morons, a GIFSY-3-cured strain did not show a virulence defect in mice (81). As discussed above, this does not mean that GIFSY-3 cannot enhance the fitness of certain *Salmonella*

strains if one looks at the right ecological niche (i.e., the right animal species). In fact, some data indicate that SspH1 contributes to virulence in calves (151).

Other phages found in *S. enterica* serovar Typhimurium strains include E34, ST64T, and ST64B (Table 1) (159, 160). It is safe to assume that this is only the tip of the iceberg. The *Salmonella* strains analyzed so far carry on average between two and six prophages, and the number of *Salmonella* strains is enormous: there are at least 200 different *S. enterica* subspecies 1 serovar Typhimurium strains, and Typhimurium is just 1 of more than 2,000 different subspecies 1 serovars. In addition, many strains seem to carry some prophage remnants. It will be a formidable task for the future to devise efficient strategies to identify these phages and analyze them.

(iii) Phage remnants encoding SopE2 and SspH2. Two further type III effector proteins (SopE2 and SspH2) of *S. enterica* serovar Typhimurium are encoded by phage remnants (Fig. 10). SopE2 is 70% identical to SopE and was detected in every analyzed *Salmonella* strain (155, 185). It contributes to intestinal inflammation in cows and mice (98a, 249, 250). SopE2 is coregulated with the SPI-1 TTSS (8, 209). The regulator has not been identified. SopE2 requires the same type III chaperone (InvB) for transport as SopE does (69), and we can assume that functional integration of *sopE2* works along the same lines as described above for the *sopE* moron (Fig. 11B).

SspH2 belongs to the same type III effector protein family as SspH1 does (see the discussion of GFSY-3 above). Expression of *sspH2* is controlled by the SPI-2 master regulator *ssrAB* (150), and it is translocated into host cells via the SPI-2 TTSS. Disruption of either *sspH1* or *sspH2* alone did not have much of an effect, but an *sspH1 sspH2* double mutant showed reduced virulence in a calf infection model (150, 151). This illustrates a common problem encountered in *Salmonella* virulence studies. Often the bacteria have several redundant proteins to achieve one specific task (e.g., SspH1 and SspH2; SopE1 and SopE2). It turns out that one has to identify and inactivate all the redundant proteins before one can observe a virulence phenotype for a specific fitness factor.

In conclusion, the morons of the *Salmonella* prophages and phage remnants encode products with a variety of functions. Even though we have learned quite a bit about the regulation and functional integration of the phage morons encoding type III effector proteins, only little is known about the other morons. Nevertheless, the diversity of phages and moron functions among *S. enterica* serovar Typhimurium strains is striking. The diversity of moron functions indicates that many bacterial properties can be optimized to yield a selective advantage for the lysogen and to establish a prophage-lysogen coevolution. Interestingly, many *Salmonella* phage morons just slightly modify the fitness factor repertoire of the lysogen; i.e., they add one more type III effector protein to the cocktail injected into host cells via the SPI-1 or SPI-2 TTSS, or they encode one additional periplasmic superoxide dismutase besides the chromosomally encoded enzyme. In most cases, lysogenic conversion leads to an incremental change in the fitness of a given *Salmonella* strain. Nevertheless, this small competitive edge might be decisive in determining the next emerging epidemic clone.

It is quite striking that many prophages and phage remnants of *Salmonella* spp. encode type III effector proteins. After

lysogenic conversion, these effector protein-encoding morons can modify fitness only if the lysogen has a TTSS, a suitable chaperone, and the correct regulators. None of these functions are encoded by the prophage. These restrictions represent a significant functional barrier, which limits the spread of the prophage between *Salmonella* spp. Most probably, these morons would be functional in only very few other bacterial species. This should result in an intimate coevolution of these type III effector-encoding phages with *Salmonella* spp. It will be an important task to devise experimental strategies to study this coevolution in detail.

***V. cholerae* toxin moron.** *Vibrio cholerae*, the causative agent of Asiatic cholera, is a gram-negative bacterial species. Of the >100 known *Vibrio* serogroups, the two toxigenic serogroups "classical" O1 and O139 have been associated with epidemic cholera. In the environment, *V. cholerae* is found in marine and estuarine habitats in association with copepods, planktonic species, insects, and water plants (12). Therefore, *V. cholerae* switches between two different ecological niches: the aquatic environment and mammalian and human hosts. *V. cholerae* enters its host via the oral route, colonizes the small intestine, and secretes "exoproteins," which cause severe secretory diarrhea.

The two human-pathogenic *V. cholerae* serogroups (O1 and O139) have evolved by sequential acquisition of two key fitness factors: the toxin-coregulated pilus (TCP) and cholera toxin (CT) (Fig. 12A) (27, 28). Both are encoded by phages or phage-like elements (129, 233). Extensive epidemiological studies support the notion that the evolution of new toxigenic *V. cholerae* clones (especially O139 strains) is still going on at breathtaking speed and includes repeated rearrangements of the CTXΦ element (76).

CT, the principal virulence factor of *V. cholerae*, is encoded by the bacteriophage CTXΦ (233). CT is expressed in the host intestine as a classical AB toxin, and human volunteer studies using purified CT demonstrated that it accounts for the induction of diarrhea (141). The B subunit of CT binds to enterocytes and transports the catalytic A subunit into the host cell cytoplasm. There, the A subunit triggers signaling cascades leading to rapid chloride and water efflux into the intestinal lumen, causing watery diarrhea, the hallmark of epidemic cholera. This diarrhea is thought to enhance fecal shedding and dissemination from host to host (see also below). In addition, CT might enhance bacterial survival in the intestine.

TCP is critical for intestinal colonization (149). It is a type IV bundle-forming pilus, whose major subunit (TcpA) was identified in a screen for secreted virulence factors which are coregulated with CT (214, 215). It is expressed in the human intestine and belongs to the major antigens in human infections (18, 97). The genetic element encoding TCP (also termed VPI for "*V. cholerae* pathogenicity island") has been described as the genome of a filamentous phage (VPIΦ or TCPΦ [129]), but the phage nature has been disputed recently (63, 77). Beside TCP, this locus encodes accessory colonization factors and the ToxT, TcpP, TcpH, and TcpI regulatory proteins (Fig. 12B).

Proper regulation of virulence factor expression is important, as suggested by a number of mutational analyses, which demonstrated that certain regulators are required for virulence. Recently, it was discovered that TCP and CT expression

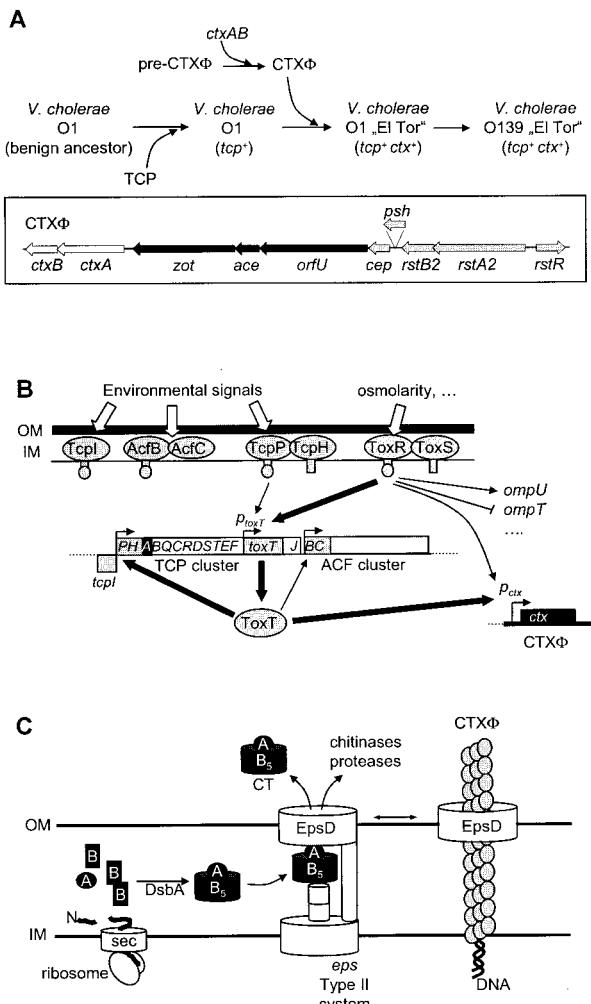


FIG. 12. CT moron. (A) Model for the evolution of toxigenic *V. cholerae* strains. (B) Regulation of *ctx* expression. (C) Assembly and secretion of CT and CTXΦ.

are controlled by at least three different quorum-sensing systems present in *V. cholerae* (154). The importance of regulation is also illustrated by the observation that *V. cholerae* isolated from fresh feces is more infectious than bacteria grown under laboratory conditions (148).

The key virulence factors required for the host phase of the *V. cholerae* life cycle (TCP and CT) are under the control of master regulators (ToxR, AphA, and AphB) encoded in the ancestral *Vibrio* chromosome and by newly acquired DNA (VPI [TcpP and ToxT] [139, 140] and RS1 [RstC] [61] [Fig. 12B]). ToxR is regarded as the central virulence regulator, and a set of about 60 genes (termed the ToxR regulon) involved in colonization, toxin production, and bacterial survival within the host are coregulated in response to external stimuli such as temperature, pH, and osmolarity (18, 202). ToxR belongs to an interesting family of transmembrane proteins with cytoplasmic transcription factor domains. ToxR has presumably evolved to control outer membrane (OM) synthesis in response to osmolarity. Later, expression of CT, TCP, and the regulators encoded in the TCP element (Fig. 12B) were plugged into this

preexisting ToxR regulatory system. In any event, expression of the virulence factors is controlled by fairly complex signaling cascades in modern epidemic *V. cholerae* strains, and key regulators are not encoded by the prophage.

CT has to be secreted into the intestinal lumen in order to exert its function. As with most other AB toxins, this occurs via type II secretion. Type II secretion is a multistep pathway, and transport across the inner membrane (IM) and (OM) occurs in separate steps (Fig. 12C). The A and B subunits of CT are synthesized as precursor proteins. These are transported into the periplasm and proteolytically processed by the ubiquitous Sec system. With the help of DsbA, the subunits are correctly folded and assembled into the AB₅ complex, which is subsequently transported across the OM via the *eps* type II secretion system (198). Besides CT, the *eps* type II secretion system transports chitinases and proteases. Type II secretion systems are widely distributed: However, CT export is specific for the *eps* type II system and does not function in other species (e.g., *Pseudomonas aeruginosa*, *Klebsiella oxytoca*, and *Erwinia chrysanthemi*) bearing different type II secretion systems (152). In contrast, environmental *Vibrio* spp. can secrete the CT B subunit, presumably via the *eps* type II system (152). It is thought that the *eps* type II system is shared by all *Vibrio* spp. and causes the secretion of proteins (chitinases and proteases) common to all vibrios. CT evolved to be specifically and efficiently transported by this system (198). In other words, the CT moron of CTXΦ can be functional and provide a selective advantage only in vibrios. Only here are the proper regulators and transport systems available.

Intriguingly, the filamentous phage CTXΦ, which does not encode its own OM pore, also requires one component of the *eps* system for its escape from the bacterium (secretion [Fig. 12C]). In *epsD* mutants, the phage cannot be assembled (62). This suggests that not only CTXΦ-encoded virulence factors but also phage viability require vibrios as the host bacteria. This is supported by the observation of lysogenic conversion of CT⁻ *Vibrio mimicus* strains by CTXΦ. The lysogens obtained in this study produced viable CTXΦ and secreted biologically active CT (75). In addition, integration of CTXΦ into the *Vibrio* chromosome requires host functions: the host-encoded recombinases XerC and XerD (118). It is unclear whether XerCD-like recombinases from other bacterial species could also foster CTXΦ integration. If not, it would be another mechanism for functional restriction of CTXΦ to *Vibrio* spp.

Shiga toxin-producing *E. coli*. *E. coli* is a gram-negative enterobacterium which diverged from the *Salmonella* lineage about 100 million years ago. Many contemporary *E. coli* strains are benign intestinal commensals; but a few strains have evolved into pathogens. These pathogenic *E. coli* strains are classified as enteropathogenic *E. coli*, enterohemorrhagic *E. coli*, Shiga-toxin producing *E. coli* (STEC), enterotoaggregative *E. coli*, enterotoxigenic *E. coli*, enteroinvasive *E. coli*, or uropathogenic *E. coli* based on their repertoires of virulence factors and the most common diseases associated with them. The vast majority of pathovar-specific virulence factors are encoded in horizontally acquired DNA fragments: pathogenicity islands, transposons, plasmids, and, last but not least, phages. The STEC strain O157:H7 is a prominent example of the last. During the past 20 years, *E. coli* strain O157:H7 has evolved from a clinical novelty, first described in 1982, to a global

public health concern (147). O157:H7 is a mucosal pathogen that produces several virulence factors, with the principal one being a prophage-encoded Shiga toxin (Stx), an AB-type toxin that inhibits protein synthesis. A second important virulence factor found in O157:H7 strains is the TTSS. It is encoded in a pathogenicity island called the locus of enterocyte effacement, which is located adjacent to prophage 933L (179). This locus is responsible for a specific pathological change in the infected intestinal mucosa, called the attaching and effacing lesion.

STEC causes diarrheal disease as well as more severe clinical manifestations, including hemorrhagic colitis and hemolytic-uremic syndrome. It is thought that the Shiga toxin Stx released by bacteria residing in the intestinal lumen is responsible for all these symptoms. The toxin traverses the intestinal epithelial barrier, enters the bloodstream, and damages vascular cells of the colon, the kidneys, and the central nervous system.

STEC strains are commonly found in the intestine of cattle and other ruminants, where they are associated with several O serotypes, including O157 (131). Cattle are the principal environmental reservoir of STEC. However, only a subgroup of STEC within the bovine reservoir is capable of causing disease (16). This might be linked to differences in the Stx expression characteristics (193).

The available genetic evidence suggests that O157:H7 is a group of closely related strains, which have emerged during the past 50 years from the enteropathogenic *E. coli* strain O55:H7 by a small number of genetic events (201). Key events in this process were the replacement of the *rfb* gene region (213) followed by the sequential acquisition of first bacteriophage stx2 and then phage stx1 (Fig. 13A). The O157 strains are not monolithic. Some geographical variation was documented between O157 strains from Europe, the United States, and Australia (130a). Variation occurred through the emergence of regional subclones showing distinct genetic polymorphisms. Genome diversity occurred through random drift and bacteriophage-mediated events. Many strains possess only the stx2 prophage. The *rfb* region encodes the enzymes necessary for the synthesis of the O side chains of the bacterial lipopolysaccharides. The lipopolysaccharide is the dominant and highly variable surface molecule under strong immune selection pressure, and the *rfb* locus is the target for frequent recombination and horizontal gene transfer.

Two bacteriophages encode the Shiga toxins Stx1 and Stx2, respectively. The stx2 phages sp5 and 933W from the two sequenced O157 strains Sakai and EDL933, respectively, are closely related at the DNA sequence level, share the genome organization of phage lambda (Fig. 13B), and are integrated into the same chromosomal site, *wrbA* (143, 180).

In contrast, the two stx1 phages sp15 and 993V from the two sequenced O157 strains share DNA sequence identity only across the nonstructural genes. Prophage sp15 closely resembles phage lambda over the structural genes (246). In comparison, *E. coli* strain Morioka V526 contained two nearly identical stx1 and stx2 prophages, which are closely related to sp5. The two Morioka stx prophages differed only over the region surrounding the distinct stx genes (199). The Stx2-encoding bacteriophage P27 from a clinical *E. coli* isolate in Germany clearly differed from the corresponding prophages in the Sakai

and EDL933 strains and was integrated into a different *E. coli* gene, *yecE* (191). The P27 genome organization corresponded to a mixture of modules from a siphovirus (λ) and myovirus (Mu). A similar hybrid phage organization was previously observed in a serotype-converting prophage from *Shigella* (3, 4). Overall, the stx phages provide strong evidence for the shuffling of phage modules and morons between phages from different *E. coli* strains.

Most differences between the genomes of the pathogenic *E. coli* strain O157:H7 Sakai and the laboratory strain K-12 (21, 171) are due to prophages (Sp1 to Sp18 in Sakai), prophage remnants resulting from phage genome truncations (e.g., e14, Rac, and Qin in K-12), and more distantly phage-related mobile DNA elements (SpLE1 to SpLE1 in Sakai) (Fig. 13C; see also Fig. 2B). A particularly striking feature was the sheer number of 18 prophages in the Sakai strain, covering P2-like, P4-like, Mu-like, and a large number of lambda-like prophages. Only two lambda-like prophages (sp5 and sp15) carried proven virulence factors (stx1 and stx2), but many other prophages contained potential virulence factors (e.g., sp4 *sodC*; sp6 candidate cytotoxin) (171). Cytotoxic distending toxin (124) was detected in O157 strains, where it was flanked by phage genes related to phage P2 and lambda, suggesting another toxin-carrying prophage in STEC (122).

Ohnishi et al. (172) analyzed the whole-genome structure of eight O157 strains by whole-genome PCR scanning. While the chromosomal DNA was conserved between the strains, a high level of variation was observed for the prophages. The stx2 prophages differed widely between these strains and were integrated at distinct chromosomal sites. In conclusion, prophages are thus the most dynamic genetic elements in this recently emerged *E. coli* lineage. The data suggest that prophages from contemporary O157 strains are the result of multiple, independent prophage acquisition events and that many modular exchanges between prophages must have occurred in O157 over the past few decades (see also Fig. 7).

A crucial issue for the understanding of the prophage-bacterium-host interaction is the regulation of the expression of the lysogenic conversion genes. The location of virulence genes in lambdoid coliphages next to the lambda *Q* or lambda *N* gene homologues provided first hints (Fig. 13B and 14A). These proteins are the phage lambda antitermination proteins and act in a cascade. The *N* protein associates with the RNA polymerase, and, assisted by several bacterially encoded Nus factors, it allows the polymerase to transcribe from the early *p_R* promoter in the genetic switch region into DNA replication genes and then the *Q* gene. The *Q* protein, being itself an antiterminator, allows then transcription to continue into late genes (lysis and structural genes [Fig. 14A]). *Q* binds DNA in the *p_{R'}* promoter located directly downstream of the *Q* gene (for a review, see reference 86).

In the majority of STEC strains, the *stx* genes are located downstream of the *Q* gene (Fig. 14A). Deletion of the *Q-p_{R'}* region in the Stx2-encoding prophage ϕ 361 abrogated the basic level of Stx2 toxin production by the lysogen. This result was somewhat surprising since previous studies had identified functional promoters immediately upstream of *stx2* (211a) that were not touched by the deletion. Mitomycin-C treatment led to high toxin production in the lysogen while the mutant showed only an increase to basic toxin levels (232). Apparently,

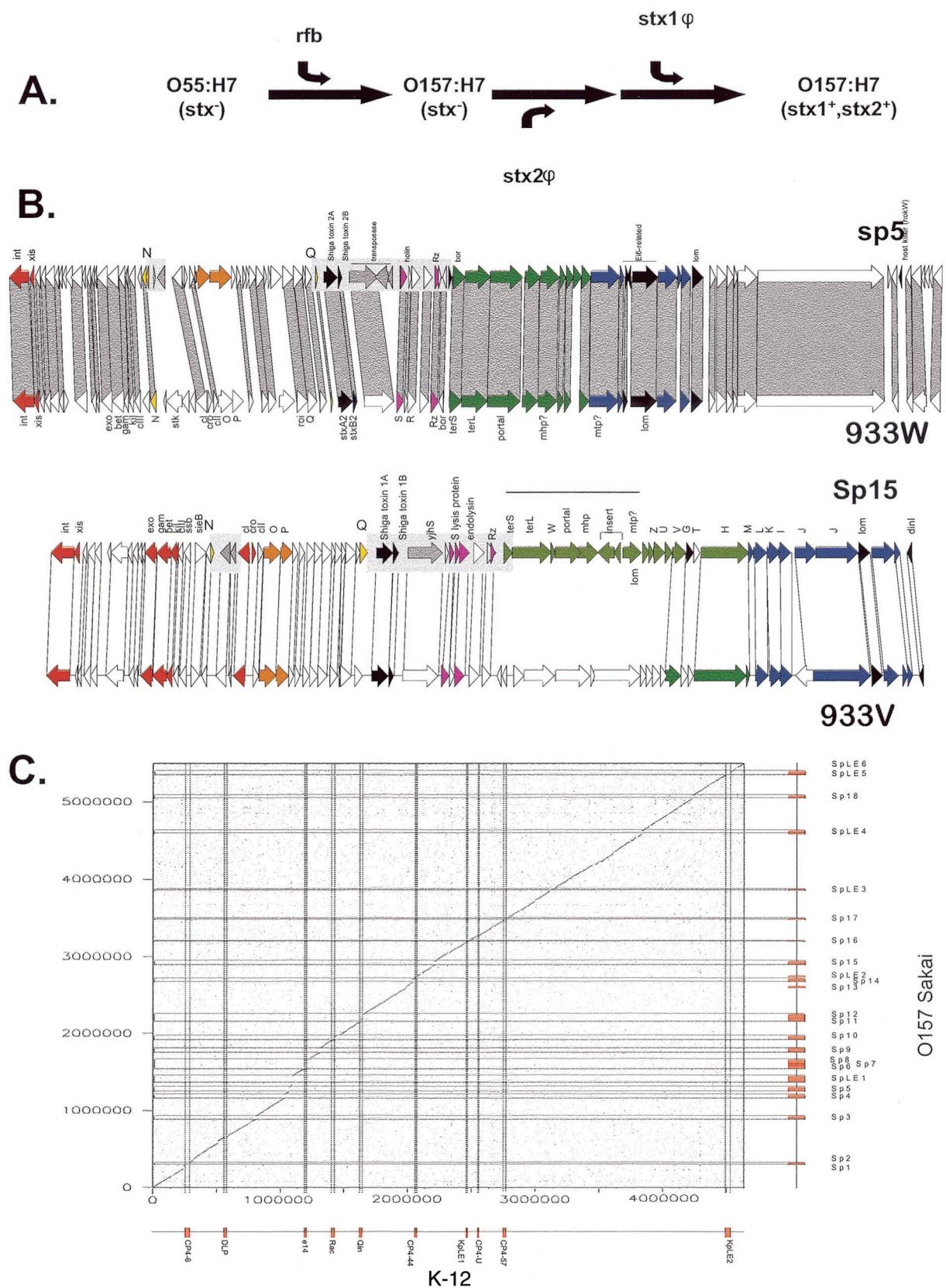


FIG. 13

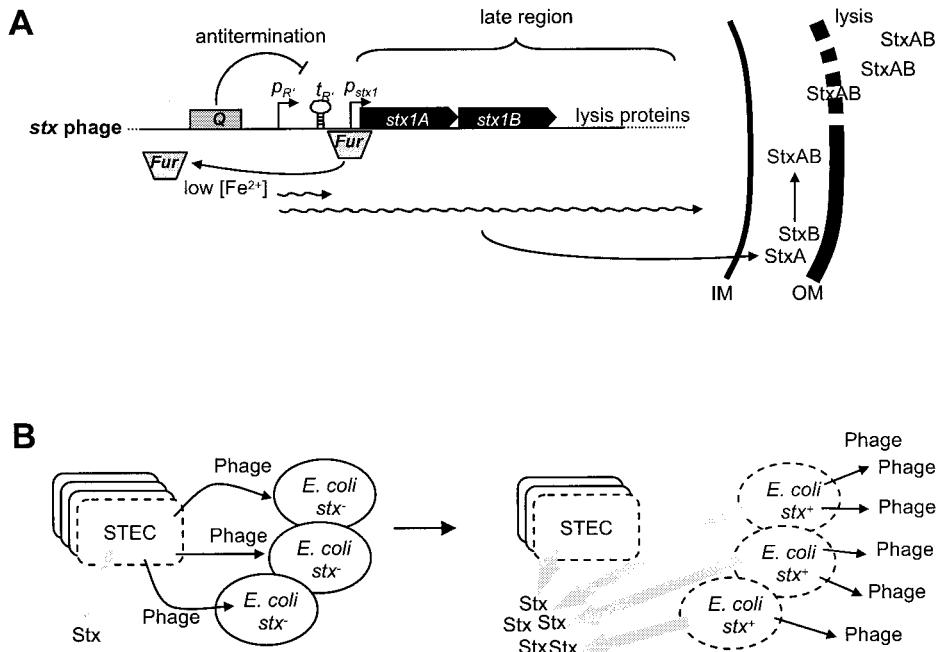


FIG. 14. Expression of Shiga toxin by STEC and amplification by recruitment of bystander bacteria. (A) Model for the regulation of *stx* expression in phage H19B. $P_{R'}$ is the promoter for the *stxAB* late-gene operon. Normally, transcription is terminated by $t_{R'}$. The antiterminator Q allows the transcription to override $t_{R'}$ and expression of the entire operon. A second Fur-regulated promoter might exist. The A and B subunits carry secretion signals and are assembled to form the mature AB toxin in the periplasm. Release from the periplasm occurs on cell lysis. IM, inner membrane; OM, outer membrane. (B) STEC converts nontoxicogenic *E. coli* strains present in the gut flora. This leads to amplification of toxin and phage release and ensures that most STEC bacteria survive.

in this prophage the Q function is not only necessary for transcription of the late phage genes but also important for *stx2* transcription (Fig. 14A). The effect was also reproduced *in vivo* (89). Substantial amounts of Stx2 were measured in the stools of mice inoculated with the wild-type lysogen but not in those inoculated with the deletion mutant. In this lysogen the *stx* gene is apparently entirely under the control of the phage, which is surprising since phage induction leads to the death of the lysogen and should thus reduce and not enhance the fitness of the lysogen. However, comparable amounts of wild-type and mutant cells were recovered from the gut (232).

There is precedent to this situation in bacteriocin-producing bacteria. Bacteriocins are molecules that are lethal to other bacteria of the same or closely related species. They are produced by a subpopulation of the bacteria harboring that gene, and these bacteria are lysed in the process of releasing the bacteriocin. The death of a minority of a bacterial population therefore contributes to the survival advantage of the majority. Notably, there are also examples where bacteriocins were derived from prophages; e.g., the R- and F-type pyocins from

Pseudomonas aeruginosa correspond to phage tail gene clusters from a P2- and a lambda-like prophage, respectively (165).

A functional promoter, p_{Stx1} , was identified directly upstream of the *Stx1* coding sequence. The activity of this promoter is regulated by the environmental iron concentration by a mechanism involving the iron-dependent Fur transcriptional repressor, which is thought to bind to a site near p_{Stx1} (40, 41). In this scenario, the toxin is under the control of the bacterial cell. This hypothesis has an inherent logic. The lysogen, or, better, a minority of the lysogens, expresses Stx only in case of low iron concentration, which is a typical growth-limiting factor for intestinal bacteria. Teleologically, the expression of the toxin leads to intestinal hemorrhage, which then liberates iron from the blood cells released into the gut and causes resumed growth of the bacteria and hence Stx downregulation.

Prophage H-19B revealed a complex network of transcriptional regulation of the *stx1* genes (168, 231). Stx1 concentrations were increased fivefold over basic levels by growth in low-iron medium. Mitomycin C induced the prophage and led to a remarkable 70-fold increase in the Stx1 titer. Deletion of

FIG. 13. Analysis of the sequenced *E. coli* O157:H7 strains. (A) Likely sequence of events that transformed *E. coli* O55:H7 into *E. coli* O157:H7. The sequential acquisition of O-antigen modifying genes and two prophages is indicated along the arrows representing the time line. (B) Comparative prophage maps of the Stx2-encoding prophages sp5 (Sakai) and 933W (EDL933) (top) and of the Stx1-encoding prophages sp15 (Sakai) and 933V (EDL933) (bottom). Selected genes are annotated above and below the gene map with their putative function. (C) Dot plot alignment of *E. coli* strains K-12 and O157 Sakai. The positions of the prophage and prophage-like elements are indicated by red boxes at the axes, and they are located by thin lines.

Q and *pstx1* had no effect on this titer, while deletion of *N* compromised but did not abolish the Stx1 production. Only a combined deletion of several phage genes including *N* and *Q* reduced Stx1 to basic levels.

There was another important difference between Fur-regulated and mitomycin-induced Stx1 production. In low-iron media, Stx1 remained intracellular while mitomycin induction resulted in fast release of Stx1 into the supernatant. Stx1 has no export system in STEC and is released by the cells via lysis due to prophage induction (Fig. 14A). Interestingly, the phage lysis cassette is directly downstream of the *stx1* genes and is cotranscribed with them.

These observations are interesting for several reasons. First, the phage seems to compete with the bacterium for *stx1* regulation. Second, it suggests that Stx1 can be released only from lysing cells. Based on a number of recent reports, STEC might have found an elegant way around this suicidal Stx1 production (Fig. 14B). When *E. coli* containing a genetically labeled H-19B prophage was introduced into mice, infectious virions are produced within the host intestine. The released phage was capable of converting other *E. coli* strains within the gastrointestinal tract (1). Apparently, specific mammalian host signals induce the Stx-encoding prophages. Candidate mammalian inducers were identified for prophage ϕ 361: coculture of the STEC strain with human neutrophils and the neutrophile product hydrogen peroxide induced Stx2 production (230). In vitro experiments verified that phage and toxin production by STEC was amplified in the presence of susceptible *E. coli* cells (89). When 37 intestinal *E. coli* isolates were individually incubated with the STEC test strain, 3 strains produced significantly more toxin in a coculture with STEC than did STEC alone, while cocultivation with one strain led to significantly less toxin production. This toxin amplification by a susceptible *E. coli* strain and toxin decrease by a resistant *E. coli* strain was also demonstrated to occur in mouse intestine inoculated jointly with STEC and an *E. coli* tester strain. In this way, the pathogenic bacteria pass the burden of toxin production to the harmless bystander member of the intestinal flora. This model has fascinating evolutionary and clinical implications. On the clinical side, Stx seems to play a role in the development of hemorrhagic lesions, but even in a single outbreak not all subjects excreting STEC strains experienced diarrhea and only a minority of the symptomatic individuals developed severe sequelae like hemolytic-uremic syndrome (163). It was recently suggested that the virulence of the isolates could be related to the clonal variability of the induced Stx prophages (podovirus versus siphovirus) (163). Alternatively, the susceptibility of the resident intestinal *E. coli* population could be a determining factor for clinical complications associated with STEC infections. Prophage induction was also obtained by exposure to antibiotics (89). This could contribute to antibiotic-induced exacerbation of STEC infection in mice (251) and in humans (241).

From an evolutionary viewpoint, *stx* genes may represent a recent acquisition by STEC strains, which have not yet recruited a protein secretion system for a nonsuicidal production of the toxin. Via the infection of bystander cells with the Stx-producing prophage, STEC can deploy the toxin without automatically lysing itself. In this way, even the possession of a toxin without a secretion system can confer a selective advan-

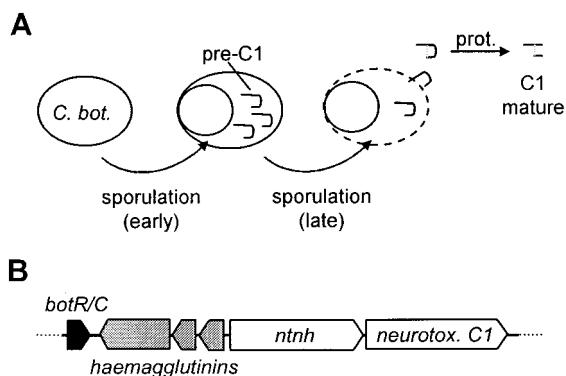


FIG. 15. C1 phage of *C. botulinum*. (A) Neurotoxin release during sporulation. (B) Organization of the neurotoxin (neurotox.) C1 locus.

tage. The selective advantage is deduced from the wide distribution of Stx prophages and high sequence conservation of the two *stx* gene cassettes in STEC. However, the teleological interpretation of the STEC data should be treated with caution, since in cattle, which are the natural reservoir for STEC, the strains are carried asymptotically in the intestine despite some (although more variable) Stx production. Virulence studies have been performed mostly in acute-infection models (see, e.g., reference 210), and it has remained unclear whether Stx toxins may help to shape the long-term coexistence of STEC with the bovine host. Even though the exact mechanism by which Stx toxins can provide a selective advantage has remained unclear, the Stx phages provide a formidable example of the rapid exchange of moron cassettes between phages from different *E. coli* strains.

***C. botulinum*.** Clostridia are strictly anaerobic gram-positive bacteria which are ubiquitous in the environment. These organisms produce extremely resistant spores which sporulate under anaerobic conditions.

C. botulinum strains were originally defined by their ability to produce one of the closely related but antigenically distinct members (A, B, C1, D, E, F, or G) of the botulinum neurotoxin family. Later, it turned out that the botulinum neurotoxin producers are quite heterogeneous and belong to different groups of strains (groups I to IV) and even different species (*C. butyricum* and *C. baratii*) (90, 94, 145, 146).

Human botulism is caused by the consumption of toxin-contaminated food. In other cases, the bacteria replicate within the human gut or sometimes in infected wounds, where they release the toxin in situ. The botulinum neurotoxins are expressed as ca. 150-kDa precursors lacking classical signal peptides (Fig. 15A). In vitro, release of the 150-kDa precursor from the cell is observed as a result of sporulation. Surprisingly, the exact mechanism of toxin release from the bacterium has not been studied in more detail. Once outside the cell, the botulinum neurotoxins are cleaved by bacterial or host proteases to yield the active toxin, consisting of 100-kDa H (heavy) and 50-kDa L (light) subunits. Often, the botulinum neurotoxins are complexed with hemagglutinins and Ntnh. This protects the toxin and facilitates its absorption through the gastric mucosa (96, 183). The "complexing" proteins are encoded in the same gene cluster as the corresponding botulinum neurotoxins (Fig. 15B).

The H subunit targets the toxin to neuronal tissue, mediates neuron binding, and delivers the L subunit into the neuronal cell. The L subunit is a metalloprotease, which cleaves protein components of the neuroexocytosis apparatus. This leads to irreversible blockade of acetylcholine secretion at neuromuscular synapses, resulting in flaccid paralysis.

The botulinum neurotoxins A, B, and F are encoded in the chromosome, while G is plasmid encoded (253) and C1 and D (and possibly E [252]) are encoded by prophages (CE β and DE β [11, 71, 72]). Surprisingly, sequence analysis of the botulinum neurotoxin C1 and D loci has remained restricted to the direct vicinity of the neutotoxin genes. There is a transcriptional start site 100 nucleotides upstream and a rho-independent terminator downstream of the toxin gene. Genes for both classes of "complexing" proteins and a gene for a regulator (*botR/C*) were also present in the phage (103, 217). Overall, the genetic organization of the botulinum neurotoxin C1 and D loci suggests that most regulators and cofactors required for proper toxin expression are encoded by the phage. Therefore, lysogenic conversion of a new strain or even another bacterial species should lead to "toxin conversion," provided that the recipient strain can sporulate or release the toxin in some other way. Surprisingly little work has been done in recent years to further explore this important aspect of CE β and DE β biology.

The *C. botulinum* lysogens can be cured easily, and cultures of the C1 and D toxin-producing strains release significant amounts of phage. The cured strains can be readily relysogenized, and curing-reinfection cycles may also occur in nature. This is backed by the identification of nontoxigenic or "low-toxin-producing" derivatives of certain *C. botulinum* "toxotypes" (73).

In addition, some phages carrying botulinum neurotoxins C1 and D were found to also encode a C3 toxin, an ADP-ribosyltransferase for Rho GTPases (104, 162, 181, 182). This supports the notion that the neurotoxin phages in *C. botulinum* behave like most other toxin-encoding phages in terms of monon exchange and variability.

Overall, much groundbreaking work has been done on the structure and the toxic mechanism of botulinum neurotoxins and ways to prevent and cure the disease. In sharp contrast, the complete sequences of the phages themselves are still not available in the public database. This is quite astonishing, considering that CE β and DE β were among the first toxin-converting phages ever discovered. Future work will have to address the genetic characteristics of these phages, their evolution, and also the issues of toxin gene expression and release of the toxin from the bacterium.

S. pyogenes. *S. pyogenes* has fascinated medical microbiologists for nearly a century. It is a protean pathogen, and humans are its only reservoir. One-third of all humans are colonized with *S. pyogenes*. The bacteria are commonly found in the throat and on the skin. *S. pyogenes* strains harbour a large variety of fitness factors in the chromosome as well as on mobile genetic elements (Fig. 16A). Sequence analysis of 10 chromosomally encoded fitness factors (extracellular proteins) revealed a striking correlation with the M serotypes of these strains (192). These observations suggest that the *S. pyogenes* strains (M serotypes) have evolved as distinct pathovars, harboring a specific set of chromosomally encoded fitness factors. These data fit with the epidemiological observation of an as-

sociation between the M serotype and certain pathological conditions in humans and mice (158). On the other hand, a detailed analysis of DNA sequence polymorphism of the set of chromosomal virulence genes studied by Reid et al. (192) revealed numerous likely horizontal transfer events between the different M strains, and the authors suspected that generalized phage transduction was responsible for at least some of these gene transfers.

S. pyogenes strains are polylysogenic organisms whose prophages constitute about 10% of the total genome of sequenced strains (up to two-thirds of the strain-specific genes [Fig. 1]). The prophages encode a wide variety of putative and established virulence factors (Table 3). Recently the sequences of four *S. pyogenes* strains representing three different M serotypes were published (14, 78, 164, 203). Each strain contained three to six seemingly complete prophage sequences closely related to phages known from the dairy field (67). Only one of the three prophages in strain SF370 could be induced by mitomycin C treatment. The two noninducible prophages contained stop codons within essential genes (67). In contrast, all five prophages from strain MGA315 could be induced (10) and thus could actively participate in gene transfer between the strains. The comparative genomics of these prophages was recently reviewed (46) and is not repeated here. In this review, we focus on the expression and role of the phage-encoded fitness factors in disease.

The vast majority of *S. pyogenes* prophages encode one or two likely or experimentally proven virulence or fitness factors between the lysis cassette and the right phage attachment site (Fig. 16B). Similar to the situation in *Salmonella*, the coexistence of multiple fitness factor-encoding prophages (polylysogeny) provides the opportunity for fast reassortment of fitness factors. Already the presently sequenced *S. pyogenes* phages allow a substantial permutation of virulence factor combinations. This permutation might account for the temporal and geographical variability and the distinct disease pathologies seen between clinical isolates. Diseases range from mild pharyngitis to life-threatening toxic shock syndromes, covering many distinct pathological entities associated with otherwise very similar strains (9).

Lytic induction of *S. pyogenes* prophages has been studied in considerable detail. In bacterial growth media, phages were either spontaneously released from the lysogenic *S. pyogenes* strain MGAS315 (two prophages) or induced by more or less physiological stimuli. Hydrogen peroxide, produced in vivo by attacking phagocytes, induced three prophages, while mitomycin C, which causes damage to the bacterial DNA, induced all five prophages, although with variable efficiency (10). Mitomycin C might not be a natural inducer, but this type of induction of prophages is interesting for the understanding of paradoxical effects of antibiotics on bacterial infections. Fluoroquinolone antibiotics, which poison DNA topoisomerase II, induce not only Shiga toxin-encoding bacteriophages in *E. coli* (see above) but also a mitogen-encoding bacteriophage in *Streptococcus canis* (121). *S. canis* is a commensal bacterium in dogs and may cause various opportunistic infections. Notably, after the introduction of fluoroquinolones into canine veterinary medicine, the incidence streptococcal toxic shock syndrome and necrotizing fasciitis, clinically very similar to the corre-

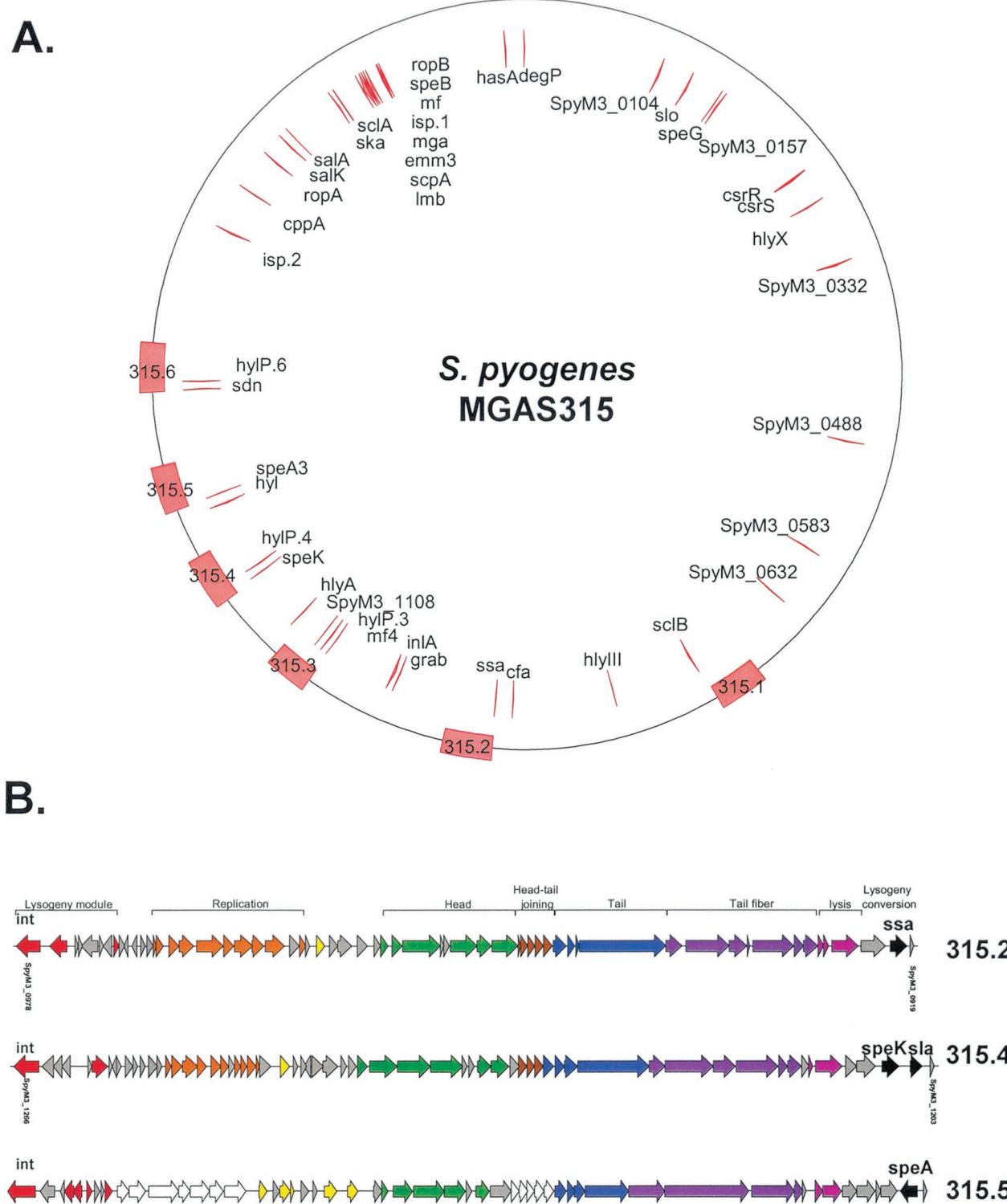


FIG. 16. *S. pyogenes* strain MGAS315 and its prophage- and chromosome-encoded virulence factors. (A) Genome map of the M3 *S. pyogenes* strain. The positions of the prophages are indicated on the outer circle by red boxes and their names. Proven and suspected virulence factors are located as red arrows on the inner circle and marked with their gene annotation. The origin of replication is located at the top. (B) Genome maps of three prophages. The modular structure of the prophage genome and the function of the modules are indicated. The virulence genes are marked in black (for details, see Table 3).

TABLE 3. Putative and proven virulence genes in *S. pyogenes* MGAS315^a

Fitness factor	Open reading frame no. ^b
Proteinase	
C3 degrading proteinase	1598
Exfoliative toxin-like protein	0632
IdeS/Sib38	0583
Streptokinase A	1698
C5a peptidase	1726
SpeB/Cysteine proteinase	1742
Serine proteinase	1864
Adhesin	
GRAB	1032
SclA	1702-3
SclB	0738
M protein	1727
Fibronectin-binding protein F2-like	0104
Internalin-like protein	1035
Laminin-binding protein	1725
Hemolysin ^c	
SLO	0130
SLS	0488
CAMP	0905
Hemolysin (<i>hylA1</i>)	1153
Hemolysin homolog (<i>hylX</i>)	0276
Hemolysin homolog (<i>hylIII</i>)	0815
Capsule synthesis and degradation	
Hyaluronate synthase (<i>hasA</i>)	1851
Hyaluronidase	1294
Hyaluronidase	1101, 1214, 1418
DNase	
Mitogenic factor (<i>mf</i>)	1745
Mitogenic factor-like	1095
Sda	1409
Superantigen	
SpeA	1301
SpeG	0155
SpeK	1205
SpeL	1205
SSA	0920
Immunoreactive antigen	
Immunogenic secreted protein	1562
Myosin crossreactive antigen	0332
Immunogenic secreted protein	1731
Regulators	
SpeB protease transcription regulator (<i>rgg</i>)	1744
RopA (trigger factor)	1634
M protein transacting positive regulator	1728
CsrR/S (CovR/S)	0244-5
RofA	0157
Bacteriocin	
Salivaricin precursor	1652
SalK homolog	1646

^a Modified from reference 164 with permission.^b Phage-encoded genes are in bold.^c SLO, streptolysin; SLS, streptolysin S; CAMP, CAMP factor.

sponding syndromes caused by *S. pyogenes* in humans, increased steadily (186).

Less artificial stimuli can also induce *S. pyogenes* prophages. Coculture of *S. pyogenes* with human pharyngeal cells induced the phage and the production of the phage-encoded strepto-

coccal pyrogenic exotoxin SpeC, Spd1, and a number of other bacterial proteins. A low-molecular-weight nonproteinaceous factor from the supernatant of the pharyngeal cells was the inducing factor (33, 34). Spd1 and SpeC are encoded by adjacent genes on the same prophage. Virtually identical adjacent genes (*speC* and *mf2*; Mf2 and Spd1 are 98% identical) were detected in a prophage from the sequenced M1 strain.

Spd1 belongs to a family of DNases found in a variety of *S. pyogenes* strains. Some of these proteins also function as bacterial superantigens (see below). Spd1 showed DNase activity but no superantigen activity (34). It has a leader sequence and is secreted via the Sec pathway. The authors suggested that in vivo streptococcal phage induction occurs in the pharynx, where other strains of *S. pyogenes* are likely to be present. Spd1 is secreted just before lysis of the induced lysogenic bacterial cell and could thus digest the bacterial DNA, which is spilled out after lysis. This would reduce the DNA-mediated viscosity of the tissue fluid and facilitate the spread of the released phage to the next target bacterium. Lysogenic conversion of Tox⁻ *S. pyogenes* in the pharyngeal flora by phages released from Tox⁺ lysogens or free phage was demonstrated in mice (32). Another function of the phage DNase is to liquefy pus, which contains substantial amounts of DNA from dying leukocytes. DNase would thus also increase the spread of the bacteria on the pharyngeal cells. Such a dual function of phage enzymes was also proposed for the phage hyaluronidase. This enzymatic activity is associated with a phage tail fiber protein. The enzyme apparently helps the phage to cross the hyaluronic acid-containing capsule surrounding *S. pyogenes* during the adsorption process. From the fact that patients mount an immune response to this phage enzyme (95), it was concluded that it might also assist bacteria in spreading along the connective tissue planes, which also contain hyaluronic acid.

The second phage protein induced by coculture with pharyngeal cells was SpeC, a member of a growing family of streptococcal and staphylococcal superantigens. These proteins simultaneously bind major histocompatibility complex (MHC) class II molecules and specific variable regions of T-cell receptors. In contrast to normal antigens presented by MHC class II, which activate 0.001 to 0.0001% of all T cells, superantigens activate up to 20% of all T cells. This results in massive proliferation and subsequent release of inflammatory cytokines. These factors are thought to cause the high fever and shock or autoimmune sequelae in some patients with streptococcal infections who develop acute rheumatic fever (ARF).

Epidemiologically, ARF is associated with M18 *S. pyogenes* serotypes. In the sequenced M18 strain, one prophage encodes the SpeL and SpeM superantigens, which are sequence-related to the SpeC and SpeK superantigens (203). SpeL and SpeM cause the proliferation of blood cells, are pyrogenic, and, in combination with endotoxin, are lethal in rabbits. They stimulated T cells with three specific β-chains. Their transcription was enhanced in exponentially growing cells, and all ARF patients showed significantly elevated antibody titers to both proteins (204).

Prophage 315.4 from the sequenced M3 strain had *speK* and *sla* genes in the lysogenic conversion region. The recombinant Sla protein had phospholipase A2 activity and structural similarity to a snake venom toxin (14). It is the target for the

antibody response in infected patients and may account for bleeding disorders, which accompany some invasive *S. pyogenes* infections. The authors proposed a model where the contemporary highly virulent M3 strains are clonal and are the result of the sequential acquisition of three prophages (ϕ 315.5, acquired circa 1920; ϕ 315.2, acquired circa 1940, and ϕ 315.4, acquired circa 1980), leading to the constellation of three particular superantigens (SpeA3, SSA, and SpeK) combined with Sla (9). However, there is only circumstantial evidence, but no experimental proof, for contribution of these prophage borne morons to the high virulence of this strain.

A basic problem with all studies of *S. pyogenes* virulence is that most of the time, these bacteria behave as commensals and do not cause disease. For this reason, much effort was spent in correlating virulence factor (i.e., superantigen) secretion with certain disease phenotypes. In 2000, Kotb and colleagues observed that the expression of the phage-encoded SpeA was either very low or undetectable in about half of the clinical M1 *S. pyogenes* isolates (51). When these isolates were introduced into Teflon tissue chambers within mice for 5 days, the expression of SpeA was turned on and concomitantly the expression of SpeB was downregulated (130). SpeB is a chromosomally encoded secreted cysteine protease. Interestingly, the same authors had previously observed an inverse relationship between SpeB expression and disease severity in clinical M1 *S. pyogenes* infections (128). Isolates recovered from the chambers continued to produce SpeA for extended passages in vitro, suggesting a stable genetic switch for SpeA expression. In cases where in vitro SpeA expression was finally downregulated, SpeB expression was again turned on. Electrophoretic two-dimensional gel analysis of the secreted M1 *S. pyogenes* proteome, coupled with matrix-assisted laser desorption-ionization/time-of-flight mass spectroscopy, revealed that expression of active SpeB caused the degradation of the vast majority of the secreted bacterial proteins, including several known virulence factors (6). Deletion of the *speB* gene or addition of a cysteine protease inhibitor inactivating SpeB yielded cells that revealed more than 150 spots in the secreted proteome, including the prophage-encoded Sda (streptodornase). A complex secreted proteome was also reported for the strain recovered from the mouse tissue chamber. The proteome included SpeA, representing a clear case of in vivo up-regulation of a phage-encoded virulence protein.

The regulation of gene expression in *S. pyogenes* grown under different conditions was the focus of a series of studies from the Musser laboratory. At the extremes of its habitats, *S. pyogenes* has to adapt to a range of temperatures extending from about 25°C on the superficial skin to 40°C or more in deep tissue infections. In microarray analysis, globally 9% of the genes from the M1 strain SF370 were differentially transcribed by organisms grown at 29°C compared with 37°C (205). Genes from mobile DNA (mainly phages) belonged to the most prominently up-regulated genes at 29°C followed by transport and binding protein genes (Fig. 1B) (45a, 46, 164).

During *S. pyogenes*-phagocyte interaction, about 16% of the bacterial genes were differentially transcribed. The largest fraction (69 genes) represented hypothetical up-regulated genes. The next most prominent group comprised prophages (23 genes). Since most of the prophage genes showed increased transcription, the authors suggested that these phage genes

play a role in the host-pathogen interaction (229). The expression pattern of putative virulence and regulatory genes in *S. pyogenes* strains recovered from pediatric patients presenting with pharyngitis was investigated. Notably, the most prominently up-regulated gene (with a 60-fold increase in transcript level) was the prophage *sda* gene (228).

Two regulators were identified that controlled the prophage moron expression. One was the two-component regulator CovR/S. When it was inactivated, a mucoid colony phenotype was observed in *S. pyogenes*, associated with overexpression of the hyaluronic acid capsule. The phage-encoded DNase was more abundant in the strain (91) since it was relieved from repressor binding to the *sda* promoter (153). The second identified transcriptional regulator was Rgg, which controlled the expression of the prophage-encoded genes *sda* and *mf3* (52).

However, the exact role of the up-regulated phage genes in *S. pyogenes* cells making contact with the mammalian cells is not yet clear: do they directly benefit the propagation of the phage, or do they benefit the lysogen via a bacteriocin-like effect or the expression of phage-encoded fitness factors? In summary, it is fair to say that perhaps more than in any other system, the recent evolution of *S. pyogenes* has been guided by bacteriophages. Even when making contact with the mammalian host, *S. pyogenes* strains not only alter their gene expression pattern but—by lysogenization of bystander cells—also alter the genomes of commensal *S. pyogenes* (32). In addition, prophages change the bacterial genome while residing silently in the chromosome by serving as anchor points for homologous recombination leading to bacterial chromosome rearrangements independent of lysogenization. In contrast, the actual role of the prophage morons in the emergence of new epidemic strains with different disease characteristics is mainly speculative so far, and there is no direct evidence from animal experiments.

***S. aureus*.** *S. aureus* strains are gram-positive cocci which can grow under aerobic and anaerobic conditions. Their natural habitat is the nose and the skin of warm-blooded animals. A large fraction of the human population is colonized with this bacterium. These colonizing *S. aureus* strains cause disease only rarely. Nevertheless, *S. aureus* is one of the most frequent causes of bacterial infection in humans. Specific antibiotic-resistant strains cause epidemics in hospital settings, and one can distinguish between these “nosocomial” infections and community-acquired infections, which are caused by a much more diverse group of strains with different properties. *S. aureus* can cause a wide range of diseases ranging from toxicoses such as food poisoning to invasive diseases. Many skin infections such as furunculosis, staphylococcal scalded skin syndrome, and wound infections are caused by this bacterium. *S. aureus* strains encode a large variety of secreted toxins, and these toxins (Fig. 17A) (Table 4) are responsible for most of the clinical symptoms associated with the infections.

The production of some *S. aureus* toxins was first linked to lysogeny over 40 years ago: a phage could convert nontoxicogenic strains to alpha-hemolysin production (20). However, it took more than 20 years after that discovery until a toxin gene was located on a staphylococcal phage by molecular means. The staphylococcal enterotoxin A gene, *sea*, was mapped near the attachment site of the temperate phage PS42-D (15). Southern hybridizations revealed that the *sea* genes in staphylococcal

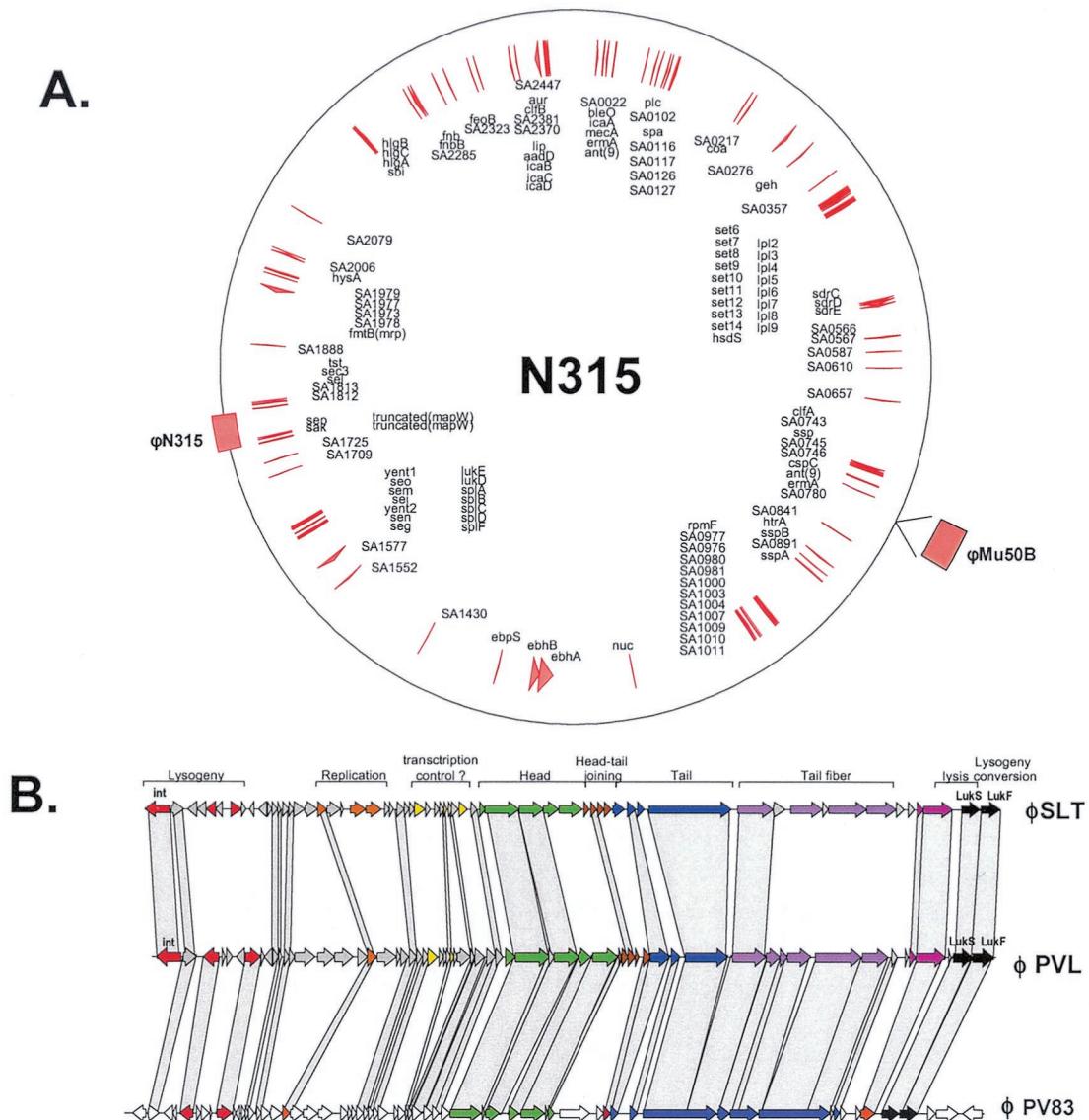


FIG. 17. *S. aureus* strain N315 and its prophage- and chromosome-encoded virulence factors. (A) Genome map of the N315 strain. The positions of the φN315 prophage and of the φMu50B prophage present in a closely related *S. aureus* strain are indicated. A selection of proven and suspected virulence factors are indicated by red arrows on the inner circle (for details, see Table 4). The origin of replication is located at the top. (B) Genome maps of three prophages from different *S. aureus* strains. The modular structure of the prophage genome is indicated. The virulence genes are indicated by black arrows.

strains were associated with a family of phages rather than with one particular phage. The potential medical importance of staphylococcal phage-encoded toxins has recently motivated a number of phage sequencing projects. Phage PVL encodes a bicomponent cytotoxin, the Panton-Valentine leukocidin. The two toxin genes *lukS* and *lukF* were located between the phage lysin gene and the right attachment site (126). The two toxins assemble into pore-forming transmembrane complexes and lyse their target cells, human polymorphonuclear leukocytes (82). Phosphorylation of LukS by protein kinase A was found to be required for the leukocytolytic activity (167). Very similar toxin genes were found at the same location in a morphologically and molecularly distinct *S. aureus* phage, SLT (Fig. 17B). Sequence comparison suggested that the entire region sur-

rounding the attachment site was the result of a modular exchange between the two phages (166). *S. aureus* prophage PV83 also encodes a leukocidin, this time a *lukM-lukF* gene combination (Fig. 17B). As in other staphylococcal phages, these toxin genes were located between the lysin gene and the right attachment site. The toxin genes in PV83 are flanked by a transposase gene, suggesting that this gene cassette was derived from a mobile DNA element (254). The different *S. aureus* phages showed a patchwork pattern of relatedness (46), as predicted by the modular theory of phage evolution (24).

The exfoliative toxin is an extracellular protein that underlies the scalded skin syndrome associated with *S. aureus* infections (133). The causative proteins are ETA and ETB. The former is encoded by prophage ETA, and the latter is encoded

TABLE 4. Putative and proven fitness factors identified in *S. aureus* N315^a

Product	Open reading frame	Gene name
Exoenzymes		
Phosphatidylinositol phosphodiesterase	SA0091	<i>plc</i>
Staphylocoagulase	SA0222	<i>coa</i>
Lipase	SA0309	<i>geh</i>
Nuclease	SA0746	
Cysteine protease	SA0900	<i>sspB</i>
V8 protease	SA0901	<i>sspA</i>
Thermonuclease	SA1160	<i>nuc</i>
Serine protease	SA1627–31	<i>spl4–F splF</i>
Staphylokinase	SA1758	<i>sak</i>
Hyaluronidase	SA2003	<i>hysA</i>
Zinc metalloproteinase aureolysin	SA2430	<i>aur</i>
Triacylglycerol lipase	SA2463	<i>lip</i>
Toxins		
Probable exotoxins	SA0382–93	<i>set6–set14</i>
Alpha-hemolysin	SA1007	
Leukotoxins	SA1637, SA1638	<i>lukD, lukE</i>
Enterotoxins	SA1642–48	<i>seg, sen, ent2, ent1, sei, sem, seo</i>
Enterotoxin P	SA1761	<i>sep</i>
Enterotoxin A	SAV19	<i>sea</i>
Leukocidin LukM ?	SA1813	
Enterotoxins	SA1816, SA1817	<i>sel, sec3</i>
Toxic shock syndrome toxin 1	SA1819	<i>tst</i>
Delta-hemolysin	SAS065	<i>hld</i>
Gamma-hemolysin components	SA2207–9	<i>hlgA, hlgC, hlgB</i>
Adhesins		
Fibrinogen-binding proteins	SA0742–3	<i>clfA, B</i>
Extracellular matrix-binding proteins	SA0519–21	<i>sdrC–sdrE</i>
Elastin-binding protein	SA1267, SA1268	<i>ebhA, ebhB</i>
Fibronectin-binding proteins	SA1312	<i>ebpS</i>
Intercellular adhesion proteins	SA2290, SA2291	<i>fnbB, fnbA</i>
	SA2459–62	<i>icaA–icaD</i>
Others^c		
MHC class II-like protein	SA0102	
IgG-binding protein A	SA0107	<i>spa</i>
IgG-binding protein SBI	SA2206	<i>sbi</i>

^a Modified from reference 132 with permission.^b Genes located on pathogenicity islands are in bold; those on prophages are in bold and italics.^c IgG, immunoglobulin G.

by a large plasmid. Again, the *eta* gene is located downstream of the phage lysis gene. Notably, the ETA protein was found by Western blot analysis in the supernatant from the parental lysogen and in two *S. aureus* strains lysogenized with ETA phage. Its toxic activity was demonstrated in mouse experiments (245). Interestingly, ETA production was not stimulated by mitomycin C induction, suggesting that it is constitutively produced from its own promotor and neither regulated by the preceding lysis gene cassette nor up-regulated with phage replication during prophage induction (see above).

The recent sequencing of several *S. aureus* strains confirmed and extended the observations from the phage-sequencing projects. The bacterial genome projects were motivated by the increasing antibiotic resistance demonstrated by clinical *S. aureus* isolates. Methicillin-resistant *S. aureus* is now the main etiological agent of nosocomial infection, and vancomycin, the only antibiotic effective against it, is no longer effective against all *S. aureus* isolates. The methicillin-resistant strain N315 isolated in 1982 differed from the vancomycin-resistant strain Mu50 isolated in 1997 by less than 4% at the nucleotide level

(132). Most of the differences in genome structure were due to the insertion of Mu50-specific mobile genetic elements, including a Mu50-specific prophage. Two phages are very similar between the two strains: φN315 and φMu50A (Fig. 6A). The two prophages carry known virulence factors: a gene encoding enterotoxin P (the *sep* gene), a superantigen involved in the symptoms of food poisoning, and a gene encoding staphylokinase (the *sak* gene), suspected to be involved in the proteolytic destruction of host tissue. In addition, an M-like protein fragment is encoded by a gene preceding *sep*. The virulence genes flank the phage lysis cassette on both sides. However, the two prophages are not identical. Especially over the lysogeny and early genes, the two prophages differed in numerous small modular exchanges (Fig. 6A).

One major difference between the two sequenced *S. aureus* strains was prophage φMu50B. Strain N315 showed no prophage at this position. φMu50B shares segments of sequence relatedness with phage ETA, especially over the late genes, but lacks the ET toxin gene (Fig. 6B). In its place, it carries four genes of unknown function. A potential virulence gene was

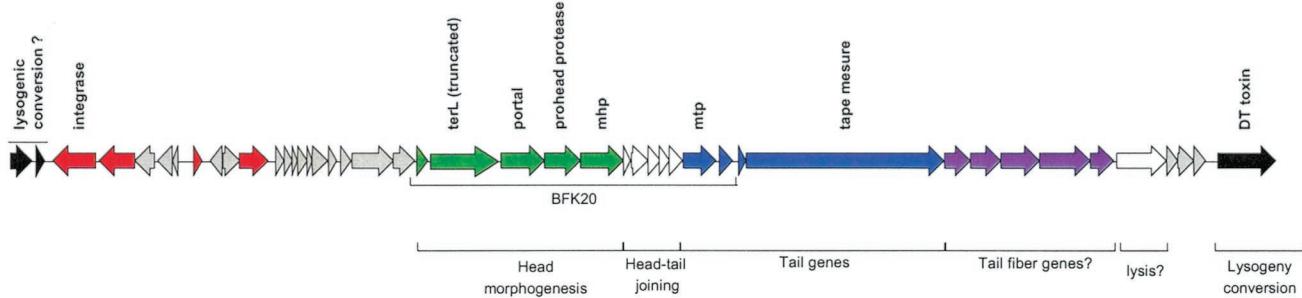


FIG. 18. Prophage genome map of the *C. diphtheriae* prophage encoding DT. The map indicates the modular structure of the corynephage by a color code and the annotated brackets. The BFK20 bracket marks a region of high similarity to a *Brevibacterium* phage. Selected genes are annotated.

identified upstream of the integrase gene; it had similarity to a gene from the staphylococcal pathogenicity island SaPI1. Notably, both prophages were flanked near their integrase gene by pathogenicity or genomic islands, suggesting a potential link between the acquisition of the two types of mobile DNA elements in strain Mu50.

A relationship between the two types of mobile DNA element was demonstrated in *S. aureus* with the 15 kb-long pathogenicity island SaPI1 and the temperate staphylococcal phage 80 α (142, 197). SaPI1 carries the gene for the toxic shock syndrome toxin 1 (TSST-1) and an integrase gene. It is flanked by a 17-bp repeat, and it is excised and circularized and replicates autonomously. Notably, SaPI1 interferes with phage 80 α growth and is encapsidated into special small phage 80 α heads commensurate with its smaller DNA size. On phage-mediated transfer to a recipient organism, SaPI1 integrates at a specific attachment site via the SaPI1-encoded integrase. However, as with the P4 satellite phage and its P2 helper phage, phage 80 α provides functions for excision, replication, and encapsidation of this pathogenicity island. This peculiar link with a phage ensures mobility to the pathogenicity island and may be responsible for the spread of TSST-1 among *S. aureus* strains.

Also a community-acquired *S. aureus* isolate, strain MW2, was sequenced (7). It differed from strain N315 by numerous insertions, deletions, and gene replacements. The most obvious differences were near the origin of replication, including the DNA element encoding methicillin resistance (SCCmec). Further differences between the strains were linked to other mobile DNA elements: prophages, transposons, and a number of small genomic islands. MW2 contains two prophages: ϕ Sa2 and ϕ Sa3. ϕ Sa2 resembles *S. aureus* phage ϕ 12, but also carries the *lukS* and *lukF* genes in a constellation identical to that phage in SLT. ϕ Sa3 closely resembles phage PVL over most of their genomes, but the two phages differed in their content of virulence genes. Sa3 contains, in addition to *sak* and *sea* around the lysin gene, two new enterotoxin gene alleles, *seg2* and *sek2*. The latter genes are encoded between repressor and integrase genes in the lysogeny module. Prophage Sa3ms from an unfinished *S. aureus* genome sequenced at the Sanger Centre differed from Sa3 at only 14 positions (211). The extensive patchwise sequence similarities between *S. aureus* phages suggested that multiple recombination events between N315-, PVL-, PV83-, Sa3-, and ϕ 13-like phages had occurred in the evolutionary history of these phages. In the Sa3ms lysogen, a

single 1- and 1.7-kb long transcript was detected with a *sak* and *sea* probe, respectively. Previously, a promoter was demonstrated upstream of *sea* (22). Mitomycin C induction led to a marked increase in the number of these mRNAs, and transcription of two higher-molecular-weight forms of mRNA was observed that covered both genes. If the replication of the prophage was prevented by a mutation, no increase in transcription was observed, suggesting that the augmented transcription was the direct result of the increased phage DNA copy number. Also, the transcription of the *seg2* and *sek2* genes was increased by mitomycin C (211). This result was surprising since these genes were cotranscribed from a promoter upstream of the *cI*-like repressor gene. This constellation ensures constitutive expression during lysogeny in coliphages, but repression via the Cro repressor occurs after prophage induction.

A comparison of the different *S. aureus* prophages revealed that the toxin genes are mobile DNA elements of their own and suggested that they are not stably associated with an individual prophage (7). Microarray analysis also demonstrated an extensive variation in gene content among different strains of *S. aureus*, with 22% of the genome comprising dispensable genetic material (84). A total of 18 large regions of difference were identified, 10 of which encode virulence factors or antibiotic resistance genes. Apparently, lateral gene transfer has played a fundamental role not only in the evolution of *S. aureus* prophages but also in that of their hosts.

***C. diphtheriae*.** *C. diphtheriae* is a strictly human-adapted gram-positive bacterium. It can cause local infections of the tonsils, pharynx, nose, and conjunctiva and systemic intoxications when the released toxin destroys the parenchyma of the heart, liver, kidneys, or adrenal glands. The diphtheria toxin (DT) is the major virulence factor of this pathogen, and the DT gene is carried by a family of closely related bacteriophages (Fig. 18).

DT is a classical AB toxin. The precursor-protein has a signal sequence for Sec-mediated protein secretion. Later, the pre-toxin is proteolytically processed to yield the mature AB toxin. DT is one of the best-characterized bacterial toxins, and we possess X-ray crystallographic structures for free DT and DT complexed with different substrates and a detailed knowledge of the mode of action of this toxin (reviewed in reference 115). The A subunit of DT is an ADP-ribosyltransferase which covalently modifies the elongation factor EF-2, thereby inhibiting chain elongation during protein synthesis. Detailed

knowledge is also available for the cellular receptor of DT and the way in which DT is processed after entry into the cell. This knowledge has allowed the establishment of a transgenic mouse model for diphtheria, even though mice are naturally resistant to this disease (50).

Diphtheria research has a long history. The discovery of DT goes back to Roux and Yersin in 1888. Subsequently, the work performed by Behring and Kitasato in 1891 and the discovery of the DT-overproducing strain PW8 in 1896 laid the ground for antitoxin therapy of and active immunization against this upper respiratory tract infection, respectively. The link of DT to lysogeny was established in the 1950s, when the nonlysogenic *C. diphtheriae* strains C4 and C7 were shown to become toxicogenic after infection with the *tox*⁺ corynephage beta but not with the *tox*-lacking corynephage gamma. Work in the late 1960s established that DT expression was achieved from the integrated prophage as well as from an extrachromosomal replicating or nonreplicating phage DNA, suggesting that DT expression was not regulated with other phage genes. In the mid-1970s a restriction map of phage beta was elaborated which showed that the *tox* gene was located next to the phage attachment site, making its independent regulation plausible. Earlier work had already established that *tox*⁺ lysogens produce DT only in iron-depleted media (174). The regulation occurred at the level of transcription. In the mid-1980s, a promoter directly upstream of the *tox* gene was characterized; this was followed by the discovery of a diphtheria toxin repressor (DtxR) that binds to an operator directly upstream of the *tox* gene. DtxR functions as an iron-dependent global regulator in *C. diphtheriae*. Currently, at least 18 DtxR binding sites are known to occur in *C. diphtheriae*, and they affect the expression of about 40 genes. There are fascinating analogies between the iron-regulated Fur repressor in *E. coli* and its effect on Shiga toxin-encoding prophages and DtxR in *C. diphtheriae* on DT-encoding prophages, despite the wide phylogenetic distance separating these bacteria (gram-negative proteobacteria versus high-G+C-content gram-positive bacteria). A recent review provides a detailed list of the classical papers on DT briefly mentioned in the preceding paragraph (115).

C. diphtheriae phages have been poorly investigated. Most toxicogenic *C. diphtheriae* strains contain DNA sequences related to phage beta, but the *tox* gene was also found to be associated with the distinct phages δ and ω (37, 190). Interestingly, the *tox* corynephage γ also contained the *tox* gene. However, a 1.5-kb IS-like DNA element inserted near the 5' end of the coding sequence prevented *tox* translation (140a). The recently published *C. diphtheriae* genome sequence of a bacterial strain from the current diphtheria epidemic in Eastern Europe (49) provided the first complete sequence of a *tox*⁺ phage. Surprisingly, its genome organization resembled that of phages found in low-G+C-content gram-positive bacteria. In fact, several proteins still demonstrated weak sequence links with this group of phages. However, its closest sequence matches were with structural genes of *Brevibacterium* phage BFK20. The *tox* gene was found at the right prophage genome end. Interestingly, two additional candidate lysogenic conversion genes were detected at the left genome end between phage integrase and the attachment site, a tRNA gene (Fig. 18). Both prophage genome ends showed a decreased G+C content. A tRNA gene was also found downstream of the *tox* gene, followed by what

could be a further mobile DNA element. Overall, the bacterial genome sequence showed 13 regions with local anomalies in nucleotide composition. Seven of them were flanked by tRNA genes, and none were present in the two sequenced environmental corynebacteria. Many genes that could contribute to the pathogenicity of *C. diphtheriae* are found in these genomic islands. Six regions contained phage-related genes. Unlike its closest sequenced pathogenic relative, *Mycobacterium tuberculosis*, *C. diphtheriae* appears to have recently acquired many genes necessary for survival, attachment, and virulence in the host. *M. tuberculosis* contains only a few small prophage remnants that still have functions related to the excision of these elements (17), and these were not associated with recognizable virulence factors. This difference may be a reflection of the different ecology: *M. tuberculosis* is a predominantly intracellular pathogen and thus has less opportunity for genetic exchanges than does the extracellular *C. diphtheriae*.

Fitness Cassettes in Nonpathogenic Bacteria?

Pathogenic bacteria cannot be sharply defined, as demonstrated by opportunistic pathogens that cause disease only in compromised patients (e.g., *P. aeruginosa* lung infection in cystic fibrosis patients). Bacteria of the normal human flora are another interesting case. For example, lactobacilli represent the dominant bacterial biota in the human vagina, where they mediate an important protective function by preventing the growth of pathogenic bacteria and yeasts via their production of lactic acid and hydrogen peroxide. In addition, lactobacilli are part of the commensal bacteria along the alimentary tract. However, in a few patients, lactobacilli are also found associated with disease processes. These observations led some researchers to postulate gradual transitions from bacterial symbionts via commensals to pathogens (116). It is interesting that certain aspects of the prophage-host interaction described above for pathogenic bacteria are also found in lactobacilli. For example, it was demonstrated for a *Lactobacillus* gut commensal that about half of the isolate-specific DNA was represented by prophage DNA (226). The locations of these nonphage extra genes were at the same positions as lysogenic conversion genes in prophages from pathogenic staphylococci and streptococci, i.e., downstream of the phage lyisin and phage repressor genes. In fact, some of the extra genes even showed sequence similarity to candidate lysogenic conversion genes in prophages from *S. pyogenes*, which inhabits the same ecological niche. Furthermore, in a *Lactobacillus* commensal from the oral cavity, the prophage genome segments near both attachment sites were transcribed (225). Transcribed extra genes of prophages from a *Lactobacillus* gut commensal showed sequence links to anonymous genes from SaPI1 (227). These extra genes were located between phage repressor and integrase genes. These observations lend support to the notion that the virulence genes in prophages from pathogenic bacteria are only a special case of a much broader phenomenon. According to that hypothesis, prophages encode fitness factors for their lysogenic hosts irrespective of the ecological niche inhabited by the lysogen (46, 67). Support for this generalization came from transcription studies of prophages from still another ecological group of bacteria, i.e., free-living bacteria like the milk bacteria *Streptococcus thermophilus* (224) and *Lactococcus lactis* (25).

Notably, a lysogenic conversion phenotype was observed in *S. thermophilus*: lysogens containing the prophage TP-J34 showed homogeneous growth in suspension, while the prophage-cured cell aggregated (169). This change in growth properties could be of selective value for this milk organism.

Evidence for Increased or Modulated Virulence after Phage Conversion

In a number of bacterial pathogens, the key virulence factor is prophage encoded. Curing the prophage from these pathogens alleviates virulence, and benign relatives do not carry the respective prophage. In these cases, it has often been possible to provide direct experimental evidence of a key role for lysogenic conversion in the emergence of virulent bacterial strains. This is the litmus test for the evolutionary function of prophage-encoded fitness factors, because it requires "improving" the properties of a bacterium which has already been optimized by evolution.

There are a number of classic examples. It is long known that lysogenic conversion can convert nontoxigenic *Clostridium* spp. into toxigenic *C. botulinum* strains expressing the botulinum neurotoxin (71, 72). Curing of the phage resulted in loss of virulence. Similarly, lysogenic conversion of nontoxigenic, non-virulent *C. diphtheriae* strains with DT phages transformed these strains into toxin-producing, highly virulent strains (85, 219). Also, lysogenic conversion of nontoxigenic *E. coli* strains by Shiga toxin-encoding phages from *E. coli* O157:H7 converted them to toxigenic strains. Last but not least, work on CTXΦ from *V. cholerae* has demonstrated that lysogenic conversion of several environmental *Vibrio* spp. yielded toxin-producing and -secreting lysogens (76, 152). In all these cases, the phage encodes the key virulence factor of the bacterial pathogen. This has made it quite easy to demonstrate a significant phenotypic change in the lysogen.

The second group of pathogenic bacteria employs a multitude of virulence factors to damage the host and cause disease. Here, it has proven much more difficult to demonstrate clear-cut selective advantages of strains obtained by lysogenic conversion with a toxin-encoding phage. For example, several superantigens, including SpeA and SpeC, have been identified in highly virulent strains of *S. pyogenes* (discussed above). They can transform nontoxigenic *S. pyogenes* strains into toxin-expressing strains. Also, toxins recovered from the supernatants of these strains cause typical symptoms when injected into the skin of laboratory animals. However, there are (to our knowledge) no experimental data showing that lysogenic conversion really changes the virulence characteristics of the lysogen. Moreover, disruption of *speA* in a toxigenic *S. aureus* M1 strain did not affect virulence in a mouse model and epidemiologic analyses did not reveal a strong correlation between the presence of SpeAC and specific streptococcal diseases (206, 247, 248). It is conceivable that SpeA and SpeC play an important role in the pathogen-host interaction that is simply masked by the activities of all the other virulence factors expressed by *S. pyogenes*. However, one cannot rule out that association of SpeAC-encoding prophages with certain serotypes may simply reflect host specificity of these phages and that the different virulence phenotypes of the different serotypes are attributable mainly to chromosomally encoded virulence factors. Yet an-

other possibility is that prophage-encoded fitness factors increase only the colonization and persistence capacity of the strain, not the virulence. Successful colonizers may simply have a better chance to disseminate and cause infection. This scenario is supported by the observation that the clones with the specific prophage combination found in the highly virulent strains are also now found to be dominant among the colonizers of the oral cavity (120). *S. aureus* is another example of a bacterial pathogen using a multitude of virulence factors. Nevertheless, it has been possible to demonstrate altered virulence characteristics of certain *S. aureus* strains after lysogenic conversion with ΦETA (133). *eta* *S. aureus* mutant strains are transformed into ETA-producing strains by lysogenic conversion with this phage. Skin infection experiments with suckling mice demonstrated that these new lysogens have gained the ability to induce a specific type of skin exfoliation (245).

S. enterica serovar Typhimurium also belongs to this second group of pathogens. In many cases it has been impossible to detect phenotypic changes after lysogenic conversion with phages encoding fitness or virulence factors. For example, lysogenic conversion with Fels-1, which carries the morons *sodCIII*, *nanH*, and *grvA*, did not enhance virulence in an intraperitoneal mouse infection (81). In other cases, it has at least been possible to detect reduced virulence in strains cured from certain prophages. For example, curing GIFSY-1 and GIFSY-2 from *S. enterica* serovar Typhimurium significantly reduced virulence in a mouse model (79). It was possible to demonstrate that lysogenic conversion of the *S. enterica* serovar Typhimurium with phage SopEΦ led to a slight but significant increase in virulence (249). In tissue culture invasion experiments, SopEΦ lysogens are twofold more invasive than the nonlysogenic parental strain (S. Mirold and W. D. Hardt, unpublished data). In addition, the lysogen is slightly but significantly more virulent in bovine infection assays than the parental strain (249). Control experiments with a SopEΦ derivative carrying a disrupted *sopE* moron demonstrated that the increased virulence was really attributable to the moron. This supports the notion that lysogenic conversion with SopEΦ expands the host cell-signaling capacity of *S. enterica* serovar Typhimurium and suggests that lysogenic conversion with SopEΦ has given the epidemic strain DT49/DT204 a competitive edge in its "traditional" niche (discussed above).

Overall, there is substantial experimental evidence available demonstrating that lysogenic conversion can really alter the virulence characteristics of a lysogen. Conceptually this is quite important, because it provides a solid basis for the hypothesis that fitness factor-encoding phages engage in a coevolution with their host bacteria. Nevertheless, it is quite clear that the experimental evidence is limited to a few examples. For the vast majority of prophages known today, much work remains to be done.

CONCLUSION AND OUTLOOK

With respect to prophages, there are two groups of bacterial pathogens. On one side are pathogens in which no prophages have been detected so far. However, in the majority of bacterial pathogens, prophages and phage-like elements are present. Some of these bacteria (*V. cholerae*, STEC, *C. diphtheriae*, and *C. botulinum*) depend on a specific prophage-

encoded virulence or fitness factors to cause a specific disease. Others (*S. aureus*, *S. pyogenes*, and *S. enterica* serovar Typhimurium) harbor a multitude of prophages, and each phage-encoded virulence or fitness factor makes an incremental contribution to the fitness of the lysogen; as discussed in detail, this latter group of pathogens provides ample evidence for pro-phage-lysogen coevolution.

The prophage-lysogen coevolution observed in the second group has resulted in several interesting phenomena. First, different strains of the same bacterial species generally harbor distinct sets of prophages. Often, these prophages behave like "swarms" of related prophages (Fig. 7; 13B; and 17B). This points to a fast prophage diversification. The diversification seems to be fueled by the frequent transfer of phage material, especially between the different strains of the same species but also by recombination with genomic DNA material and occasional acquisition of mobile DNA elements from more distantly related species and gene cassettes from unrelated phages infecting distinct bacterial hosts. In addition to changing themselves, bacteriophages contribute to the diversification of the bacterial genome architecture (Fig. 1, 2, and 4). In many cases, they actually represent a large fraction of the strain-specific DNA sequences. In addition, they can serve as anchoring points for genome inversions.

The acquisition and evolution of more and more phages in the bacterial genome is apparently kept at bay by frequent loss of functional prophages (curing) and stepwise prophage degradation. The latter process is suggested by prophages inactivated by point mutations, incomplete pieces of prophages, and single phage genes (mostly integrases and possibly morons).

The current evidence suggests that protection from phage lysis via superinfection exclusion functions and the postulated selective advantage conferred by moron-encoded factors provided the functional link between the evolution of prophages and the lysogens. The morons have some very interesting properties. Often, they are transferred within the prophage "swarm" present in the same bacterial species and occasionally even between nonrelated phages. The distribution of a specific moron seems frequently to be restricted to one specific species or closely related species. It is not clear whether this observation suffers from an observation bias. Functional restrictions are certainly important: in many cases, morons would not function within the genetic apparatus of a distantly related receptor cell. This argument has two aspects: one deals with the compatibility of the genes for transcription, translation, and control, and the other deals with the integration of the new virulence factors into the existing network of finely tuned virulence factors.

Not every prophage in a bacterial pathogen harbors fitness factor-encoding morons. Furthermore, morons are also found at chromosomal sites (e.g., *sopE* in *Salmonella* spp. or Shiga toxin in STEC and *S. flexneri*), suggesting that morons have a "life of their own" and might be regarded as selfish genetic elements sui generis that just exploit phages to attain mobility. Such a hypothesis is not unprecedented, since we know of other selfish DNA elements, like type II introns, that have invaded a number of phage systems. However, these systems are frequently associated with H-N-H endonucleases, which confer mobility to the introns. In contrast, morons generally lack mobility genes and therefore depend on other mobile-

DNA-like phages for their dissemination. Some morons are flanked by transposase genes in their vicinity, pointing to alternative methods of mobility. More work is required to prove the "selfish" nature of morons and to decipher the molecular mechanisms driving moron transfer among phages and between phages and chromosomal sites.

What is the origin of morons? In many cases, distantly related genes are found at chromosomal sites in other bacteria. Therefore, many morons were possibly acquired by horizontal gene transfer from an ancestral source. Possibly, polyvalent phages contributed to this horizontal transfer. However, the high rate of phage recombination and moron transfer makes it impossible to reconstruct these events.

The contemporary prophages found in the sequenced genomes are the result of endless recombinations. Similarly, the genomes of the host bacterial species are in a constant (but somewhat slower) flux. Therefore, the >200 phage and >200 prophage sequences available in the current database are nothing but a snapshot documenting only a few of the possible combinations and configurations. The dynamics of these processes, their function in the evolution of bacterial pathogens, and their impact on animal evolution are just beginning to be explored. Integrative approaches are required to understand how phages, bacterial pathogens, and animal populations interact and coevolve. Microbes will, over the next decade, be perceived as members of a given environment, and after the first phase of sequencing of individual genomes and the laboratory analysis of single strains, we will see interacting genomes and microbes in the complex fabric of their natural environment. This will pose new analytical challenges that will probably again transform microbiology as genomics did over the last decade.

ACKNOWLEDGMENTS

We are grateful to Cosima Pelludat for critical reading of the manuscript and to Anne Bruttin for help with the editorial work.

Work in our laboratories is supported by the Swiss National Science Foundation (grants SNF 3100A0-100175/1 to W.D.H. and S.N.F. 5002-057832 to H.B.).

REFERENCES

1. Acheson, D. W., J. Reidl, X. Zhang, G. T. Keusch, J. J. Mekalanos, and M. K. Waldor. 1998. In vivo transduction with shiga toxin 1-encoding phage. *Infect. Immun.* **66**:4496-4498.
2. Affolter, M., C. Parent-Vaugeois, and A. Anderson. 1983. Curing and induction of the Fels 1 and Fels 2 prophages in the Ames mutagen tester strains of *Salmonella typhimurium*. *Mutat. Res.* **110**:243-262.
3. Allison, G. E., D. Angeles, N. Tran-Dinh, and N. K. Verma. 2002. Complete genomic sequence of SFV, a serotype-converting temperate bacteriophage of *Shigella flexneri*. *J. Bacteriol.* **184**:1974-1987.
4. Allison, G. E., D. C. Angeles, P. Huan, and N. K. Verma. 2003. Morphology of temperate bacteriophage SFV and characterisation of the DNA packaging and capsid genes: the structural genes evolved from two different phage families. *Virology* **308**:114-127.
5. Al Mamun, A. A., A. Tominaga, and M. Enomoto. 1997. Cloning and characterization of the region III flagellar operons of the four *Shigella* subgroups: genetic defects that cause loss of flagella of *Shigella boydii* and *Shigella sonnei*. *J. Bacteriol.* **179**:4493-4500.
6. Aziz, R. K., M. J. Pabst, A. Jeng, R. Kansal, D. E. Low, V. Nizet, and M. Kotb. 2004. Invasive M1T1 group A *Streptococcus* undergoes a phase-shift in vivo to prevent proteolytic degradation of multiple virulence factors by SpeB. *Mol. Microbiol.* **51**:123-134.
7. Baba, T., F. Takeuchi, M. Kuroda, H. Yuzawa, K. Aoki, A. Oguchi, Y. Nagai, N. Iwama, K. Asano, T. Naimi, H. Kuroda, L. Cui, K. Yamamoto, and K. Hiramatsu. 2002. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* **359**:1819-1827.
8. Bakshi, C. S., V. P. Singh, M. W. Wood, P. W. Jones, T. S. Wallis, and E. E. Galov. 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.
9. Banks, D. J., S. B. Beres, and J. M. Musser. 2002. The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. *Trends Microbiol.* **10**:515–521.
 10. Banks, D. J., B. Lei, and J. M. Musser. 2003. Prophage induction and expression of prophage-encoded virulence factors in group A *Streptococcus* serotype M3 strain MGAS315. *Infect. Immun.* **71**:7079–7086.
 11. Barksdale, L., and S. B. Arden. 1974. Persisting bacteriophage infections, lysogeny, and phage conversions. *Annu. Rev. Microbiol.* **28**:265–299.
 12. Barua, D., and W. B. Greenough. 1992. *Cholera*. Plenum Medical Book Co., New York, N.Y.
 13. Belden, W. J., and S. I. Miller. 1994. Further characterization of the PhoP regulon: identification of new PhoP-activated virulence loci. *Infect. Immun.* **62**:5095–5101.
 14. Beres, S. B., G. L. Sylva, K. D. Barbian, B. Lei, J. S. Hoff, N. D. Mammarella, M. Y. Liu, J. C. Smoot, S. F. Porcella, L. D. Parkins, D. S. Campbell, T. M. Smith, J. K. McCormick, D. Y. Leung, P. M. Schlievert, and J. M. Musser. 2002. Genome sequence of a serotype M3 strain of group A *Streptococcus*: phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc. Natl. Acad. Sci. USA* **99**:10078–10083.
 15. Betley, M. J., and J. J. Mekalanos. 1985. Staphylococcal enterotoxin A is encoded by phage. *Science* **239**:185–187.
 16. Beutin, L., D. Geier, H. Steinruck, S. Zimmermann, and F. Scheutz. 1993. Prevalence and some properties of verotoxin (Shiga-like toxin)-producing *Escherichia coli* in seven different species of healthy domestic animals. *J. Clin. Microbiol.* **31**:2483–2488.
 17. Bibb, L. A., and G. F. Hatfull. 2002. Integration and excision of the *Mycobacterium tuberculosis* prophage-like element, phiRv1. *Mol. Microbiol.* **45**: 1515–1526.
 18. Bina, J., J. Zhu, M. Dziejman, S. Faruque, S. Calderwood, and J. Mekalanos. 2003. ToxR regulon of *Vibrio cholerae* and its expression in vibrios shed by cholera patients. *Proc. Natl. Acad. Sci. USA* **100**:2801–2806.
 19. Birtalan, S. C., R. M. Phillips, and P. Ghosh. 2002. Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens. *Mol. Cell* **9**:971–80.
 20. Blair, J. E., and M. Carr. 1961. Lysogeny in staphylococci. *J. Bacteriol.* **82**:984–993.
 21. 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.
 22. Borst, D. W., and M. J. Betley. 1994. Promoter analysis of the staphylococcal enterotoxin A gene. *J. Biol. Chem.* **269**:1883–1888.
 23. Bossi, L., J. A. Fuentes, G. Mora, and N. Figueroa-Bossi. 2003. Prophage contribution to bacterial population dynamics. *J. Bacteriol.* **185**:6467–6471.
 24. Botstein, D. 1980. A theory of modular evolution for bacteriophages. *Ann. N. Y. Acad. Sci.* **354**:484–490.
 25. Boyce, J. D., B. E. Davidson, and A. J. Hillier. 1995. Identification of prophage genes expressed in lysogens of the *Lactococcus lactis* bacteriophage BK5-T. *Appl. Environ. Microbiol.* **61**:4099–4104.
 26. Boyd, E. F., and H. Brüssow. 2002. Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol.* **10**:521–529.
 27. Boyd, E. F., A. J. Hellperrn, and M. K. Waldor. 2000. Molecular analyses of a putative CTXφ precursor and evidence for independent acquisition of distinct CTXφs by toxigenic *Vibrio cholerae*. *J. Bacteriol.* **182**:5530–5538.
 28. Boyd, E. F., and M. K. Waldor. 2002. Evolutionary and functional analyses of variants of the toxin-coregulated pilus protein TepA from toxigenic *Vibrio cholerae* non-O1/non-O139 serogroup isolates. *Microbiology* **148**: 1655–1666.
 29. Breitbart, M., I. Hewson, B. Felts, J. M. Mahaffy, J. Nulton, P. Salamon, and F. Rohwer. 2003. Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* **185**:6220–6223.
 30. Breitbart, M., P. Salamon, B. Andresen, J. M. Mahaffy, A. M. Segall, D. Mead, F. Azam, and F. Rohwer. 2002. Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci. USA* **99**:14250–14255.
 31. Bronstein, P. A., E. A. Miao, and S. I. Miller. 2000. InvB is a type III secretion chaperone specific for SspA. *J. Bacteriol.* **182**:6638–6644.
 32. Brody, T. B., and V. A. Fischetti. 2003. In vivo lysogenic conversion of Tox(−) *Streptococcus pyogenes* to Tox(+) with lysogenic streptococci or free phage. *Infect. Immun.* **71**:3782–3786.
 33. Brody, T. B., V. Pancholi, and V. A. Fischetti. 2001. Induction of lysogenic bacteriophage and phage-associated toxin from group A streptococci during coculture with human pharyngeal cells. *Infect. Immun.* **69**:1440–1443.
 34. Brody, T. B., V. Pancholi, and V. A. Fischetti. 2002. The in vitro interaction of *Streptococcus pyogenes* with human pharyngeal cells induces a phage-encoded extracellular DNase. *Infect. Immun.* **70**:2805–2811.
 35. Brüssow, H., and F. Desiere. 2001. Comparative phage genomics and the evolution of *Siphoviridae*: insights from dairy phages. *Mol. Microbiol.* **39**: 213–223.
 36. Brüssow, H., and R. W. Hendrix. 2002. Phage genomics: small is beautiful. *Cell* **108**:13–16.
 37. Buck, G. A., R. E. Cross, T. P. Wong, J. Loera, and N. Groman. 1985. DNA relationships among some tox-bearing corynebacteriophages. *Infect. Immun.* **49**:679–684.
 38. Burke, J., D. Schneider, and J. Westpheling. 2001. Generalized transduction in *Streptomyces coelicolor*. *Proc. Natl. Acad. Sci. USA* **98**:6289–6294.
 39. Bushman, F. 2002. Lateral DNA transfer: mechanisms and consequences. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 40. Calderwood, S. B., F. Auclair, A. Donohue-Rolfe, G. T. Keusch, and J. J. Mekalanos. 1987. Nucleotide sequence of the Shiga-like toxin genes of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **84**:4364–4368.
 41. Calderwood, S. B., and J. J. Mekalanos. 1987. Iron regulation of Shiga-like toxin expression in *Escherichia coli* is mediated by the fur locus. *J. Bacteriol.* **169**:4759–4764.
 - 42a. Campbell, A. 1992. Chromosomal insertion sites for phages and plasmids. *J. Bacteriol.* **174**:7495–7499.
 42. Campbell, A., and D. Botstein. 1983. Evolution of the lambdoid phages, p. 365–380. In J. R. W. Hendrix, F. W. Stahl, and R. A. Weisberg (ed.), *Lambda II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 43. Campo, N., M. J. Dias, M. L. Daveran-Mingot, P. Ritzenthaler, and P. Le Bourgeois. 2004. Chromosomal constraints in Gram-positive bacteria revealed by artificial inversions. *Mol. Microbiol.* **51**:511–522.
 44. Canard, B., and S. T. Cole. 1989. Genome organization of the anaerobic pathogen *Clostridium perfringens*. *Proc. Natl. Acad. Sci. USA* **86**:6676–6680.
 45. Canchaya, C., F. Desiere, W. M. McShan, J. J. Ferretti, J. Parkhill, and H. Brüssow. 2002. Genome analysis of an inducible prophage and prophage remnants integrated in the *Streptococcus pyogenes* strain SF370. *Virology* **302**:245–258.
 - 45a. Canchaya, C., G. Fournous, and H. Brüssow. 2004. The impact of prophages on bacterial chromosomes. *Mol. Microbiol.* **53**:9–18.
 46. Canchaya, C., C. Proux, G. Fournous, A. Bruttin, and H. Brüssow. 2003. Prophage genomics. *Microbiol. Mol. Biol. Rev.* **67**:238–276.
 47. Casjens, S. 2003. Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.* **49**:277–300.
 48. Casjens, S., G. F. Hatfull, and R. W. Hendrix. 1992. Evolution of dsDNA tailed-bacteriophage genomes. *Semin. Virol.* **3**:383–397.
 49. Cerdeno-Tarraga, A. M., A. Efstratiou, L. G. Dover, M. T. Holden, M. Pallen, S. D. Bentley, G. S. Besra, C. Churcher, K. D. James, A. De Zoysa, T. Chillingworth, A. Cronin, L. Dowd, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Moule, M. A. Quail, E. Rabinowitz, K. M. Rutherford, N. R. Thomson, L. Unwin, S. Whitehead, B. G. Barrell, and J. Parkhill. 2003. The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. *Nucleic Acids Res.* **31**:6516–6523.
 50. Cha, J. H., M. Y. Chang, J. A. Richardson, and L. Eildes. 2003. Transgenic mice expressing the diphtheria toxin receptor are sensitive to the toxin. *Mol. Microbiol.* **49**:235–240.
 51. Chatellier, S., N. Ihendyane, R. G. Kansal, F. Khambaty, H. Basma, A. Norrby-Teglund, D. E. Low, A. McGeer, and M. Kotb. 2000. Genetic relatedness and superantigen expression in group A streptococcus serotype M1 isolates from patients with severe and nonsevere invasive diseases. *Infect. Immun.* **68**:3523–3534.
 52. Chaussee, M. S., G. L. Sylva, D. E. Sturdevant, L. M. Smoot, M. R. Graham, R. O. Watson, and J. M. Musser. 2002. Rgg influences the expression of multiple regulatory loci to coregulate virulence factor expression in *Streptococcus pyogenes*. *Infect. Immun.* **70**:762–770.
 53. Clark, A. J., W. Inwood, T. Cloutier, and T. S. Dhillon. 2001. Nucleotide sequence of coliphage HK620 and the evolution of lambdoid phages. *J. Mol. Biol.* **311**:657–679.
 54. Coleman, D., J. Knights, R. Russell, D. Shanley, T. H. Birkbeck, G. Dougan, and I. Charles. 1991. Insertional inactivation of the *Staphylococcus aureus* beta-toxin by bacteriophage phi 13 occurs by site- and orientation-specific integration of the phi 13 genome. *Mol. Microbiol.* **5**:933–939.
 55. Cornelis, G. R., and F. Van Gijsegem. 2000. Assembly and function of type III secretory systems. *Annu. Rev. Microbiol.* **51**:735–774.
 56. d'Adda di Fagagna, F., G. R. Weller, A. J. Doherty, and S. P. Jackson. 2003. The Gam protein of bacteriophage Mu is an orthologue of eukaryotic Ku. *EMBO Rep.* **4**:47–52.
 57. Darwin, K. H., and V. L. Miller. 1999. InvF is required for expression of genes encoding proteins secreted by the SPI1 type III secretion apparatus in *Salmonella typhimurium*. *J. Bacteriol.* **181**:4949–4954.
 58. Darwin, K. H., and V. L. Miller. 2000. The putative invasion protein chaperone SicA acts together with InvF to activate the expression of *Salmonella typhimurium* virulence genes. *Mol. Microbiol.* **35**:949–960.
 59. Darwin, K. H., and V. L. Miller. 2001. Type III secretion chaperone-dependent regulation: activation of virulence genes by SicA and InvF in *Salmonella typhimurium*. *EMBO J.* **20**:1850–1862.
 60. Daubin, V., N. A. Moran, and H. Ochman. 2003. Phylogenetics and the cohesion of bacterial genomes. *Science* **301**:829–832.
 61. Davis, B. M., H. H. Kimsey, A. V. Kane, and M. K. Waldor. 2002. A satellite

- phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. *EMBO J.* **21**:4240–4249.
62. Davis, B. M., E. H. Lawson, M. Sandkvist, A. Ali, S. Sozhamannan, and M. K. Waldor. 2000. Convergence of the secretory pathways for cholera toxin and the filamentous phage, CTXphi. *Science* **288**:333–335.
 63. Davis, B. M., and M. K. Waldor. 2003. Filamentous phages linked to virulence of *Vibrio cholerae*. *Curr. Opin. Microbiol.* **6**:35–42.
 64. De Groot, M. A., U. A. Ochsner, M. U. Shiloh, C. Nathan, J. M. McCord, M. C. Dinaner, S. J. Libby, A. Vazquez-Torres, Y. Xu, and F. C. Fang. 1997. Periplasmic superoxide dismutase protects *Salmonella* from products of phagocyte NADPH-oxidase and nitric oxide synthase. *Proc. Natl. Acad. Sci. USA* **94**:13997–14001.
 65. Deng, W., V. Burland, G. Plunkett III, A. Boutin, G. F. Mayhew, P. Liss, N. T. Perna, D. J. Rose, B. Mau, S. Zhou, D. C. Schwartz, J. D. Fetherston, L. E. Lindler, R. R. Brubaker, G. V. Plano, S. C. Straley, K. A. McDonough, M. L. Nilles, J. S. Matson, F. R. Blattner, and R. D. Perry. 2002. Genome sequence of *Yersinia pestis* KIM. *J. Bacteriol.* **184**:4601–4611.
 66. Deng, W., S. R. Liou, G. Plunkett III, G. F. Mayhew, D. J. Rose, V. Burland, V. Kodeyanni, D. C. Schwartz, and F. R. Blattner. 2003. Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18. *J. Bacteriol.* **185**:2330–2337.
 67. Desiere, F., W. M. McShan, D. van Sinderen, J. J. Ferretti, and H. Brüssow. 2001. Comparative genomics reveals close genetic relationships between phages from dairy bacteria and pathogenic streptococci: evolutionary implications for prophage-host interactions. *Virology* **288**:325–341.
 68. Doherty, A. J., and S. P. Jackson. 2001. DNA repair: how Ku makes ends meet. *Curr. Biol.* **11**:920–924.
 - 68a. Edlin, G., L. Lin, and R. Kudmar. 1975. Lambda lysogens of *E. coli* reproduce more rapidly than non-lysogens. *Nature* **255**:735–737.
 - 68b. Edlin, G., L. Lin, and R. Bitner. 1977. Reproductive fitness of P1, P2, and Mu lysogens of *Escherichia coli*. *J. Virol.* **21**:560–564.
 69. Ehrbar, K., A. Friebel, S. I. Miller, and W. D. Hardt. 2003. Role of the *Salmonella* pathogenicity island 1 (SPI-1) protein InvB in type III secretion of SopE and SopE2, two *Salmonella* effector proteins encoded outside of SPI-1. *J. Bacteriol.* **185**:6950–6967.
 - 69a. Ehrbar, K., S. Hapfelmeyer, B. Stecher, and W. D. Hardt. 2004. InvB is required for type III-dependent secretion of SopA in *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* **186**:1215–1219.
 70. Eichelberg, K., and J. E. Galan. 1999. Differential regulation of *Salmonella typhimurium* type III secreted proteins by pathogenicity island 1 (SPI-1)-encoded transcriptional activators InvF and HilA. *Infect. Immun.* **67**:4099–4105.
 71. Eklund, M. W., F. T. Poysky, and S. M. Reed. 1972. Bacteriophage and the toxigenicity of *Clostridium botulinum* type D. *Nat. New Biol.* **235**:16–17.
 72. Eklund, M. W., F. T. Poysky, S. M. Reed, and C. A. Smith. 1971. Bacteriophage and the toxigenicity of *Clostridium botulinum* type C. *Science* **172**:480–482.
 73. Eklund, M. W., V. R. Dowell, and Joint United States-Japan Cooperation on the Development and Utilization of Natural Resources Program, Toxic Micro-Organisms Panel. 1987. Avian botulism: an international perspective. Charles C. Thomas, Springfield, Ill.
 74. Fang, F. C., M. A. DeGroot, J. W. Foster, A. J. Baumler, U. Ochsner, T. Testerman, S. Bearson, J. C. Giard, Y. Xu, G. Campbell, and T. Laessig. 1999. Virulent *Salmonella typhimurium* has two periplasmic Cu,Zn-superoxide dismutases. *Proc. Natl. Acad. Sci. USA* **96**:7502–7507.
 75. Faruque, S. M., M. M. Rahman, Asadulghani, K. M. Nasirul Islam, and J. J. Mekalanos. 1999. Lysogenic conversion of environmental *Vibrio mimicus* strains by CTXφ. *Infect. Immun.* **67**:5723–5729.
 76. Faruque, S. M., D. A. Sack, R. B. Sack, R. R. Colwell, Y. Takeda, and G. B. Nair. 2003. Emergence and evolution of *Vibrio cholerae* O139. *Proc. Natl. Acad. Sci. USA* **100**:1304–1309.
 77. Faruque, S. M., J. Zhu, Asadulghani, M. Kamruzzaman, and J. J. Mekalanos. 2003. Examination of diverse toxin-coregulated pilus-positive *Vibrio cholerae* strains fails to demonstrate evidence for *Vibrio* pathogenicity island phage. *Infect. Immun.* **71**:2993–2999.
 78. Ferretti, J. J., W. M. McShan, D. Ajdic, D. J. Savic, G. Savic, K. Lyon, C. Primeaux, S. Sezate, A. N. Suvorov, S. Kenton, H. S. Lai, S. P. Lin, Y. Qian, H. G. Jia, F. Z. Najar, Q. Ren, H. Zhu, L. Song, J. White, X. Yuan, S. W. Clifton, B. A. Roe, and R. McLaughlin. 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci. USA* **98**:4658–4663.
 79. Figueroa-Bossi, N., and L. Bossi. 1999. Inducible prophages contribute to *Salmonella* virulence in mice. *Mol. Microbiol.* **33**:167–176.
 80. Figueroa-Bossi, N., E. Coissac, P. Netter, and L. Bossi. 1997. Unsuspected prophage-like elements in *Salmonella typhimurium*. *Mol. Microbiol.* **25**:161–173.
 81. Figueroa-Bossi, N., S. Uzzau, D. Maloriol, and L. Bossi. 2001. Variable assortment of prophages provides a transferable repertoire of pathogenic determinants in *Salmonella*. *Mol. Microbiol.* **39**:260–272.
 82. Finck-Barbancon, V., G. Duportail, O. Meunier, and D. A. Colin. 1993. Pore formation by a two-component leukocidin from *Staphylococcus aureus* within the membrane of human polymorphonuclear leukocytes. *Biochim. Biophys. Acta* **1182**:275–282.
 83. Finck-Barbancon, V., J. Goranson, L. Zhu, T. Sawa, J. P. Wiener-Kronish, S. M. Fleiszig, C. Wu, L. Mende-Mueller, and D. W. Frank. 1997. ExoU expression by *Pseudomonas aeruginosa* correlates with acute cytotoxicity and epithelial injury. *Mol. Microbiol.* **25**:547–557.
 84. Fitzgerald, J. R., D. E. Sturdevant, S. M. Mackie, S. R. Gill, and J. M. Musser. 2001. Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. *Proc. Natl. Acad. Sci. USA* **98**:8821–8826.
 85. Freeman, V. J. 1951. Studies on the virulence of bacteriophage infected strains of *Corynebacterium diphtheriae*. *J. Bacteriol.* **61**:675–688.
 86. Friedman, D. I., and D. L. Court. 1995. Transcription antitermination: the lambda paradigm updated. *Mol. Microbiol.* **18**:191–200.
 87. Galan, J. E. 2001. *Salmonella* interactions with host cells: type III secretion at work. *Annu. Rev. Cell Dev. Biol.* **17**:53–86.
 88. Galan, J. E., and A. Collmer. 1999. Type III secretion machines: bacterial devices for protein delivery into host cells. *Science* **284**:1322–1328.
 89. Gamage, S. D., J. E. Strasser, C. L. Chalk, and A. A. Weiss. 2003. Non-pathogenic *Escherichia coli* can contribute to the production of Shiga toxin. *Infect. Immun.* **71**:3107–3115.
 90. Gimenez, J. A., M. A. Gimenez, and B. R. DasGupta. 1992. Characterization of the neurotoxin isolated from a *Clostridium baratii* strain implicated in infant botulism. *Infect. Immun.* **60**:518–522.
 91. Graham, M. R., L. M. Smoot, C. A. Migliaccio, K. Virtaneva, D. E. Sturdevant, S. F. Porcella, M. J. Federle, G. J. Adams, J. R. Scott, and J. M. Musser. 2002. Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and in vivo infection modeling. *Proc. Natl. Acad. Sci. USA* **99**:13855–13860.
 92. Groisman, E. A. 2001. The pleiotropic two-component regulatory system PhoP-PhoQ. *J. Bacteriol.* **183**:1835–1842.
 93. Gunn, J. S., W. J. Belden, and S. I. Miller. 1998. Identification of PhoP-PhoQ activated genes within a duplicated region of the *Salmonella typhimurium* chromosome. *Microb. Pathog.* **25**:77–90.
 94. Hall, J. D., L. M. McCroskey, B. J. Pincomb, and C. L. Hatheway. 1985. Isolation of an organism resembling *Clostridium baratii* which produces type F botulinal toxin from an infant with botulism. *J. Clin. Microbiol.* **21**:654–655.
 95. Halperin, S. A., P. Ferrieri, E. D. Gray, E. L. Kaplan, and L. W. Wannamaker. 1987. Antibody response to bacteriophage hyaluronidase in acute glomerulonephritis after group A streptococcal infection. *J. Infect. Dis.* **155**:253–261.
 96. Hambleton, P. 1992. Clostridium botulinum toxins: a general review of involvement in disease, structure, mode of action and preparation for clinical use. *J. Neurol.* **239**:16–20.
 97. Hang, L., M. John, M. Asaduzzaman, E. A. Bridges, C. Vanderspurt, T. J. Kirn, R. K. Taylor, J. D. Hillman, A. Progulske-Fox, M. Handfield, E. T. Ryan, and S. B. Calderwood. 2003. Use of in vivo-induced antigen technology (VIAT) to identify genes uniquely expressed during human infection with *Vibrio cholerae*. *Proc. Natl. Acad. Sci. USA* **100**:8508–8513.
 98. Hansen-West, I., and M. Hensel. 2001. *Salmonella* pathogenicity islands encoding type III secretion systems. *Microbes Infect.* **3**:549–559.
 - 98a. Hapfelmeyer, S., K. Ehrbar, B. Stecher, M. Barthel, M. Kremer, and W. D. Hardt. 2004. Role of the *Salmonella* pathogenicity island 1 effector proteins SipA, SopB, SopE, and SopE2 in *Salmonella enterica* subspecies 1 serovar Typhimurium colitis in streptomycin-pretreated mice. *Infect. Immun.* **72**:795–809.
 99. Haraga, A., and S. I. Miller. 2003. A *Salmonella enterica* serovar Typhimurium-translocated leucine-rich repeat effector protein inhibits NF-κB-dependent gene expression. *Infect. Immun.* **71**:4052–4058.
 100. Hardt, W. D., L. M. Chen, K. E. Schuebel, X. R. Bustelo, and J. E. Galan. 1998. *S. typhimurium* encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells. *Cell* **93**:815–826.
 101. Hardt, W. D., H. Urlaub, and J. E. Galan. 1998. A substrate of the centosome 63 type III protein secretion system of *Salmonella typhimurium* is encoded by a cryptic bacteriophage. *Proc. Natl. Acad. Sci. USA* **95**:2574–2579.
 102. Hauser, A. R., P. J. Kang, and J. N. Engel. 1998. PepA, a secreted protein of *Pseudomonas aeruginosa*, is necessary for cytotoxicity and virulence. *Mol. Microbiol.* **27**:807–818.
 103. Hauser, D., M. W. Eklund, P. Boquet, and M. R. Popoff. 1994. Organization of the botulinum neurotoxin Cl gene and its associated non-toxic protein genes in *Clostridium botulinum* C 468. *Mol. Gen. Genet.* **243**:631–640.
 104. Hauser, D., M. Gibert, M. W. Eklund, P. Boquet, and M. R. Popoff. 1993. Comparative analysis of C3 and botulinic neurotoxin genes and their environment in *Clostridium botulinum* types C and D. *J. Bacteriol.* **175**:7260–7268.
 105. Hendrix, R. W. 2002. Bacteriophages: evolution of the majority. *Theor. Popul. Biol.* **61**:471–480.
 106. Hendrix, R. W., G. F. Hatfull, and M. C. Smith. 2003. Bacteriophages with tails: chasing their origins and evolution. *Res. Microbiol.* **154**:253–257.

107. Hendrix, R. W., J. G. Lawrence, G. F. Hatfull, and S. Casjens. 2000. The origins and ongoing evolution of viruses. *Trends Microbiol.* **8**:504–508.
108. Hendrix, R. W., M. C. Smith, R. N. Burns, M. E. Ford, and G. F. Hatfull. 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl. Acad. Sci. USA* **96**:2192–2197.
109. Hershey, A. D. 1971. The bacteriophage lambda. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
110. Hilbi, H., G. Segal, and H. A. Shuman. 2001. Icm/dot-dependent upregulation of phagocytosis by *Legionella pneumophila*. *Mol. Microbiol.* **42**:603–617.
111. Ho, T. D., N. Figueroa-Bossi, M. Wang, S. Uzzau, L. Bossi, and J. M. Slauch. 2002. Identification of GtGE, a novel virulence factor encoded on the Gifsy-2 bacteriophage of *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* **184**:5234–5239.
112. Ho, T. D., and J. M. Slauch. 2001. Characterization of grvA, an antivirulence gene on the gifsy-2 phage in *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* **183**:611–620.
113. Hodgson, D. A. 2000. Generalized transduction of serotype 1/2 and serotype 4b strains of *Listeria monocytogenes*. *Mol. Microbiol.* **35**:312–323.
114. Ho Lee, S., and J. E. Galan. 2003. InvB is a type III secretion-associated chaperone for the *Salmonella enterica* effector protein SopE. *J. Bacteriol.* **185**:7279–7284.
115. Holmes, R. K. 2000. Biology and molecular epidemiology of diphtheria toxin and the tox gene. *J. Infect. Dis.* **181**(Suppl. 1):S156–S167.
116. Hooper, L. V., and J. I. Gordon. 2001. Commensal host-bacterial relationships in the gut. *Science* **292**:1115–1118.
117. Hoyer, L. L., A. C. Hamilton, S. M. Steenbergen, and E. R. Vimr. 1992. Cloning, sequencing and distribution of the *Salmonella typhimurium* LT2 sialidase gene, *nanH*, provides evidence for interspecies gene transfer. *Mol. Microbiol.* **6**:873–884.
118. Huber, K. E., and M. K. Waldor. 2002. Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* **417**:656–659.
119. Hueck, C. J. 1998. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.* **62**:379–433.
120. Ikebe, T., A. Wada, Y. Inagaki, K. Sugama, R. Suzuki, D. Tanaka, A. Tamaru, Y. Fujinaga, Y. Abe, Y. Shimizu, and H. Watanabe. 2002. Dissemination of the phage-associated novel superantigen gene *spcL* in recent invasive and noninvasive *Streptococcus pyogenes* M3/T3 isolates in Japan. *Infect. Immun.* **70**:3227–3233.
121. Ingrey, K. T., J. Ren, and J. F. Prescott. 2003. A fluoroquinolone induces a novel mitogen-encoding bacteriophage in *Streptococcus canis*. *Infect. Immun.* **71**:3028–3033.
122. Janka, A., M. Bielaszewska, U. Dobrindt, L. Greune, M. A. Schmidt, and H. Karch. 2003. Cytolytic distending toxin gene cluster in enterohemorrhagic *Escherichia coli* O157:H- and O157:H7: characterization and evolutionary considerations. *Infect. Immun.* **71**:3634–3638.
123. Jermyn, W. S., and E. F. Boyd. 2002. Characterization of a novel *Vibrio* pathogenicity island (VPI-2) encoding neuraminidase (NanH) among toxicogenic *Vibrio cholerae* isolates. *Microbiology* **148**:3681–3693.
124. Johnson, W. M., and H. Lior. 1987. Response of Chinese hamster ovary cells to a cytolytic distending toxin (CDT) of *Escherichia coli* and possible misinterpretation as heat-labile (LT) enterotoxin. *FEMS Microbiol. Lett.* **43**:19–23.
125. Juhala, R. J., M. E. Ford, R. L. Duda, A. Youlton, G. F. Hatfull, and R. W. Hendrix. 2000. Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *J. Mol. Biol.* **299**:27–51.
126. Kaneko, J., T. Kimura, S. Narita, T. Tomita, and Y. Kamio. 1998. Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage φPVL carrying Panton-Valentine leukocidin genes. *Gene* **215**:57–67.
127. Kaniga, K., J. C. Bossio, and J. E. Galan. 1994. The *Salmonella typhimurium* invasion genes *invF* and *invG* encode homologues of the AraC and PuuD family of proteins. *Mol. Microbiol.* **13**:555–568.
128. Kansal, R. G., A. McGeer, D. E. Low, A. Norrby-Teglund, and M. Kotb. 2000. Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive group A streptococcal infection cases. *Infect. Immun.* **68**:6362–6369.
129. Karaolis, D. K., S. Somara, D. R. Maneval, Jr., J. A. Johnson, and J. B. Kaper. 1999. A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* **399**:375–379.
130. Kazmi, S. U., R. Kansal, R. K. Aziz, M. Hooshdaran, A. Norrby-Teglund, D. E. Low, A. B. Halim, and M. Kotb. 2001. Reciprocal, temporal expression of SpeA and SpeB by invasive M1T1 group A streptococcal isolates in vivo. *Infect. Immun.* **69**:4988–4995.
- 130a. Kim, J., J. Nietfeld, J. Ju, J. Wise, N. Fegan, P. Demarchelier, and A. K. Benson. 2001. Ancestral divergence, genome diversification, and phylogeographic variation in subpopulations of sorbitol-negative, β-glucuronidase-negative enterohemorrhagic *Escherichia coli* O157. *J. Bacteriol.* **183**:6885–6897.
131. Kobayashi, H., J. Shimada, M. Nakazawa, T. Morozumi, T. Pohjanvirta, S. Pelkonen, and K. Yamamoto. 2001. Prevalence and characteristics of Shiga toxin-producing *Escherichia coli* from healthy cattle in Japan. *Appl. Environ. Microbiol.* **67**:484–489.
132. Kuroda, M., T. Ohta, I. Uchiyama, T. Baba, H. Yuzawa, I. Kobayashi, L. Cui, A. Oguchi, K. Aoki, Y. Nagai, J. Lian, T. Ito, M. Kanamori, H. Matsumaru, A. Maruyama, H. Murakami, A. Hosoyama, Y. Mizutani-Uti, N. K. Takahashi, T. Sawano, R. Inoue, C. Kaito, K. Sekimizu, H. Hirakawa, S. Kuwara, S. Goto, J. Yabuzaki, M. Kanehisa, A. Yamashita, K. Oshima, K. Furuya, C. Yoshino, T. Shiba, M. Hattori, N. Ogasawara, H. Hayashi, and K. Hiramatsu. 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* **357**:1225–1240.
133. Ladha, S., C. L. Joannou, D. P. Lochrie, R. W. Evans, and S. M. Poston. 1999. Clinical, microbial, and biochemical aspects of the exfoliative toxins causing staphylococcal scalded-skin syndrome. *Clin. Microbiol. Rev.* **12**:224–242.
134. Lawrence, J. G. 1997. Selfish operons and speciation by gene transfer. *Trends Microbiol.* **5**:355–359.
135. Lawrence, J. G., G. F. Hatfull, and R. W. Hendrix. 2002. Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. *J. Bacteriol.* **184**:4891–4905.
136. Lawrence, J. G., R. W. Hendrix, and S. Casjens. 2001. Where are the pseudogenes in bacterial genomes? *Trends Microbiol.* **9**:535–540.
137. Lawrence, J. G., and H. Ochman. 1997. Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* **44**:383–397.
138. Lawrence, J. G., and J. R. Roth. 1996. Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**:1843–1860.
139. Lee, S. H., S. M. Butler, and A. Camilli. 2001. Selection for *in vivo* regulators of bacterial virulence. *Proc. Natl. Acad. Sci. USA* **98**:6889–6894.
140. Lee, S. H., D. L. Hava, M. K. Waldor, and A. Camilli. 1999. Regulation and temporal expression patterns of *Vibrio cholerae* virulence genes during infection. *Cell* **99**:625–634.
- 140a. Leong, D., and J. R. Murphy. 1985. Characterization of the diphtheria *tox* transcript in *Corynebacterium diphtheriae* and *Escherichia coli*. *J. Bacteriol.* **163**:1114–1119.
141. Levine, M. M., J. B. Kaper, R. E. Black, and M. L. Clements. 1983. New knowledge on pathogenesis of bacterial enteric infections as applied to vaccine development. *Microbiol. Rev.* **47**:510–550.
- 141a. Lin, L., R. Bitner, and G. Edlin. 1977. Increased reproductive fitness of *Escherichia coli* lambda lysogens. *J. Virol.* **21**:554–559.
142. Lindsay, J. A., A. Ruzin, H. F. Ross, N. Kurepina, and R. P. Novick. 1998. The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*. *Mol. Microbiol.* **29**:527–543.
143. Makino, K., K. Yokoyama, Y. Kubota, C. H. Yutsudo, S. Kimura, K. Kurokawa, K. Ishii, M. Hattori, I. Tatsuno, H. Abe, T. Iida, K. Yamamoto, M. Onishi, T. Hayashi, T. Yasunaga, T. Honda, C. Sasakawa, and H. Shinagawa. 1999. Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak. *Genes Genet. Syst.* **74**:227–239.
144. Maurelli, A. T., R. E. Fernandez, C. A. Bloch, C. K. Rode, and A. Fasano. 1998. "Black holes" and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **95**:3943–3948.
145. McCroskey, L. M., C. L. Hatheway, L. Fenicia, B. Pasolini, and P. Aureli. 1986. Characterization of an organism that produces type E botulinal toxin but which resembles *Clostridium butyricum* from the feces of an infant with type E botulism. *J. Clin. Microbiol.* **23**:201–202.
146. McCroskey, L. M., C. L. Hatheway, B. A. Woodruff, J. A. Greenberg, and P. Jurgenson. 1991. Type F botulism due to neurotoxigenic *Clostridium baratii* from an unknown source in an adult. *J. Clin. Microbiol.* **29**:2618–2620.
147. Mead, P. S., and P. M. Griffin. 1998. *Escherichia coli* O157:H7. *Lancet* **352**:1207–1212.
148. Merrell, D. S., S. M. Butler, F. Qadri, N. A. Dolganov, A. Alam, M. B. Cohen, S. B. Calderwood, G. K. Schoolnik, and A. Camilli. 2002. Host-induced epidemic spread of the cholera bacterium. *Nature* **417**:642–645.
149. Merrell, D. S., D. L. Hava, and A. Camilli. 2002. Identification of novel factors involved in colonization and acid tolerance of *Vibrio cholerae*. *Mol. Microbiol.* **43**:1471–1491.
150. Miao, E. A., and S. I. Miller. 2000. A conserved amino acid sequence directing intracellular type III secretion by *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. USA* **97**:7539–44.
151. Miao, E. A., C. A. Scherer, R. M. Tsolis, R. A. Kingsley, L. G. Adams, A. J. Baumler, and S. I. Miller. 1999. *Salmonella typhimurium* leucine-rich repeat proteins are targeted to the SPI1 and SPI2 type III secretion systems. *Mol. Microbiol.* **34**:850–864.
152. Michel, L. O., M. Sandkvist, and M. Bagdasarian. 1995. Specificity of the protein secretory apparatus: secretion of the heat-labile enterotoxin B subunit pentamers by different species of gram– bacteria. *Gene* **152**:41–45.
153. Miller, A. A., N. C. Engleberg, and V. J. DiRita. 2001. Repression of virulence genes by phosphorylation-dependent oligomerization of CsrR at target promoters in *S. pyogenes*. *Mol. Microbiol.* **40**:976–990.
154. Miller, M. B., K. Skorupski, D. H. Lenz, R. K. Taylor, and B. L. Bassler.

2002. Parallel quorum sensing systems converge to regulate virulence in *Vibrio cholerae*. *Cell* **110**:303–314.
155. Mirold, S., K. Ehrbar, A. Weissmüller, R. Prager, H. Tschäpe, H. Russmann, and W. D. Hardt. 2001. *Salmonella* host cell invasion emerged by acquisition of a mosaic of separate genetic elements, including *Salmonella* pathogenicity island 1 (SPI1), SPI5, and *sopE2*. *J. Bacteriol.* **183**:2348–2358.
156. Mirold, S., W. Rabsch, M. Rohde, S. Stender, H. Tschape, H. Russmann, E. Igwe, and W. D. Hardt. 1999. Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic *Salmonella typhimurium* strain. *Proc. Natl. Acad. Sci. USA* **96**:9845–9850.
157. Mirold, S., W. Rabsch, H. Tschape, and W. D. Hardt. 2001. Transfer of the *Salmonella* type III effector *sopE* between unrelated phage families. *J. Mol. Biol.* **312**:7–16.
158. Miyoshi-Akiyama, T., J. Zhao, K. Kikuchi, H. Kato, R. Suzuki, M. Endoh, and T. Uchiyama. 2003. Quantitative and qualitative comparison of virulence traits, including murine lethality, among different M types of group A streptococci. *J. Infect. Dis.* **187**:1876–1887.
159. Mmolawa, P. T., H. Schmieger, and M. W. Heuzenroeder. 2003. Bacteriophage ST64B, a genetic mosaic of genes from diverse sources isolated from *Salmonella enterica* serovar Typhimurium DT 64. *J. Bacteriol.* **185**:6481–6485.
160. Mmolawa, P. T., H. Schmieger, C. P. Tucker, and M. W. Heuzenroeder. 2003. Genomic structure of the *Salmonella enterica* serovar Typhimurium DT 64 bacteriophage ST64T: evidence for modular genetic architecture. *J. Bacteriol.* **185**:3473–3475.
161. Molmeret, M., O. A. Alli, S. Zink, A. Flieger, N. P. Cianciotto, and Y. A. Kwak. 2002. *icmT* is essential for pore formation-mediated egress of *Legionella pneumophila* from mammalian and protozoan cells. *Infect. Immun.* **70**:69–78.
162. Moriishi, K., B. Syuto, M. Saito, K. Oguma, N. Fujii, N. Abe, and M. Naiki. 1993. Two different types of ADP-ribosyltransferase C3 from *Clostridium botulinum* type D lysogenized organisms. *Infect. Immun.* **61**:5309–5314.
163. Muniesa, M., M. de Simon, G. Prats, D. Ferrer, H. Panella, and J. Jofre. 2003. Shiga toxin 2-converting bacteriophages associated with clonal variability in *Escherichia coli* O157:H7 strains of human origin isolated from a single outbreak. *Infect. Immun.* **71**:4554–4562.
164. Nakagawa, I., K. Kurokawa, A. Yamashita, M. Nakata, Y. Tomiyasu, N. Okahashi, S. Kawabata, K. Yamazaki, T. Shiba, T. Yasunaga, H. Hayashi, M. Hattori, and S. Hamada. 2003. Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res.* **13**:1042–1055.
165. Nakayama, K., K. Takashima, H. Ishihara, T. Shimomiya, M. Kageyama, S. Kanaya, M. Ohnishi, T. Murata, H. Mori, and T. Hayashi. 2000. The R-type pyocin of *Pseudomonas aeruginosa* is related to P2 phage, and the F-type is related to lambda phage. *Mol. Microbiol.* **38**:213–231.
166. Narita, S., J. Kaneko, J. Chiba, Y. Piemont, S. Jarraud, J. Etienne, and Y. Kamio. 2001. Phage conversion of Panton-Valentine leukocidin in *Staphylococcus aureus*: molecular analysis of a PVL-converting phage, φSLT. *Gene* **268**:195–206.
167. Nariya, H., A. Nishiyama, and Y. Kamio. 1997. Identification of the minimum segment in which the threonine246 residue is a potential phosphorylated site by protein kinase A for the LukS-specific function of staphylococcal leukocidin. *FEBS Lett.* **415**:96–100.
168. Neely, M. N., and D. I. Friedman. 1998. Functional and genetic analysis of regulatory regions of coliphage H-19B: location of Shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. *Mol. Microbiol.* **28**:1255–1267.
169. Neve, H., W. Freudenberg, F. Diestel-Feddersen, R. Ehlert, and K. J. Heller. 2003. Biology of the temperate *Streptococcus thermophilus* bacteriophage TP-J34 and physical characterization of the phage genome. *Virology* **315**:184–194.
170. Nilsson, A. S., J. L. Karlsson, and E. Haggard-Ljungquist. 2004. Site-specific recombination links the evolution of P2-like coliphages and pathogenic enterobacteria. *Mol. Biol. Evol.* **21**:1–13.
171. Ohnishi, M., K. Kurokawa, and T. Hayashi. 2001. Diversification of *Escherichia coli* genomes: are bacteriophages the major contributors? *Trends Microbiol.* **9**:481–485.
172. Ohnishi, M., J. Terajima, K. Kurokawa, K. Nakayama, T. Murata, K. Tamura, Y. Ogura, H. Watanabe, and T. Hayashi. 2002. Genomic diversity of enterohemorrhagic *Escherichia coli* O157 revealed by whole genome PCR scanning. *Proc. Natl. Acad. Sci. USA* **99**:17043–17048.
173. Okamoto, K., J. A. Mudd, and J. Marmor. 1968. Conversion of *Bacillus subtilis* DNA to phage DNA following mitomycin C induction. *J. Mol. Biol.* **34**:429–437.
174. Pappenheimer, A. M. 1938. Diphtheria toxin. II. The action of ketene and formaldehyde. *J. Biol. Chem.* **125**:201–208.
175. Parsot, C., C. Hamiaux, and A. L. Page. 2003. The various and varying roles of specific chaperones in type III secretion systems. *Curr. Opin. Microbiol.* **6**:7–14.
176. Paul, J. H., M. B. Sullivan, A. M. Segall, and F. Rohwer. 2002. Marine phage genomics. *Comp. Biochem. Physiol. Ser. B* **133**:463–476.
177. Pedulla, M. L., M. E. Ford, J. M. Houtz, T. Karthikeyan, C. Wadsworth, J. A. Lewis, D. Jacobs-Sera, J. Falbo, J. Gross, N. R. Pannunzio, W. Brucker, V. Kumar, J. Kandasamy, L. Keenan, S. Bardarov, J. Kriakov, J. G. Lawrence, W. R. Jacobs, Jr., R. W. Hendrix, and G. F. Hatfull. 2003. Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**:171–182.
178. Pelludat, C., S. Mirold, and W. D. Hardt. 2003. The SopEd phage integrates into the *ssrA* gene of *Salmonella enterica* serovar Typhimurium A36 and is closely related to the Fels-2 prophage. *J. Bacteriol.* **185**:5182–5191.
179. Perna, N. T., G. F. Mayhew, G. Posfai, S. Elliott, M. S. Donnenberg, J. B. Kaper, and F. R. Blattner. 1998. Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli* O157:H7. *Infect. Immun.* **66**:3810–3817.
180. Plunkett, G., III, D. J. Rose, T. J. Durfee, and F. R. Blattner. 1999. Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J. Bacteriol.* **181**:1767–1778.
181. Popoff, M., P. Boquet, D. M. Gill, and M. W. Eklund. 1990. DNA sequence of exoenzyme C3, an ADP-ribosyltransferase encoded by *Clostridium botulinum* C and D phages. *Nucleic Acids Res.* **18**:1291.
182. Popoff, M. R., D. Hauser, P. Boquet, M. W. Eklund, and D. M. Gill. 1991. Characterization of the C3 gene of *Clostridium botulinum* types C and D and its expression in *Escherichia coli*. *Infect. Immun.* **59**:3673–3679.
183. Popoff, M. R., and J.-C. Marvaud. 1999. Structural and genomic features of clostridial neurotoxins, p. 174–201. Academic Press Ltd., London, United Kingdom.
184. Porwollik, S., R. M. Wong, and M. McClelland. 2002. Evolutionary genomics of *Salmonella*: gene acquisitions revealed by microarray analysis. *Proc. Natl. Acad. Sci. USA* **99**:8956–8961.
185. Prager, R., S. Mirold, E. Tietze, U. Strutz, B. Knüppel, W. Rabsch, W. D. Hardt, and H. Tschäpe. 2000. Prevalence and polymorphism of genes encoding translocated effector proteins among clinical isolates of *Salmonella enterica*. *Int. J. Med. Microbiol.* **290**:605–617.
186. Prescott, J. F., C. W. Miller, K. A. Mathews, J. A. Yager, and L. DeWinter. 1997. Update on canine streptococcal toxic shock syndrome and necrotizing fasciitis. *Can. Vet. J.* **38**:241–242.
187. Proux, C., D. van Sinderen, J. Suarez, P. Garcia, V. Ladero, G. F. Fitzgerald, F. Desiere, and H. Brüssow. 2002. The dilemma of phage taxonomy illustrated by comparative genomics of Sfi21-like *Siphoviridae* in lactic acid bacteria. *J. Bacteriol.* **184**:6026–6036.
188. Pukatzki, S., R. H. Kessin, and J. J. Mekalanos. 2002. The human pathogen *Pseudomonas aeruginosa* utilizes conserved virulence pathways to infect the social amoeba *Dictyostelium discoideum*. *Proc. Natl. Acad. Sci. USA* **99**:3159–3164.
189. Ramirez, M., E. Severina, and A. Tomasz. 1999. A high incidence of pro-phage carriage among natural isolates of *Streptococcus pneumoniae*. *J. Bacteriol.* **181**:3618–3625.
190. Rappuoli, R., J. L. Michel, and J. R. Murphy. 1983. Restriction endonuclease map of corynebacteriophage omega ctox+ isolated from the Park-Williams no. 8 strain of *Corynebacterium diphtheriae*. *J. Virol.* **45**:524–530.
191. Recktenwald, J., and H. Schmidt. 2002. The nucleotide sequence of Shiga toxin (Stx) 2e-encoding phage φP27 is not related to other Stx phage genomes, but the modular genetic structure is conserved. *Infect. Immun.* **70**:1896–1908.
192. Reid, S. D., N. M. Green, J. K. Buss, B. Lei, and J. M. Musser. 2001. Multilocus analysis of extracellular putative virulence proteins made by group A *Streptococcus*: population genetics, human serologic response, and gene transcription. *Proc. Natl. Acad. Sci. USA* **98**:7552–7557.
193. Ritchie, J. M., P. L. Wagner, D. W. Acheson, and M. K. Waldor. 2003. Comparison of Shiga toxin production by hemolytic-uremic syndrome-associated and bovine-associated Shiga toxin-producing *Escherichia coli* isolates. *Appl. Environ. Microbiol.* **69**:1059–1066.
194. Roggentin, P., R. Schauer, L. L. Hoyer, and E. R. Vimr. 1993. The Slalidase superfamily and its spread by horizontal gene transfer. *Mol. Microbiol.* **9**:915–921.
195. Rohwer, F. 2003. Global phage diversity. *Cell* **113**:141.
196. Rohwer, F., and R. Edwards. 2002. The phage proteomic tree: a genome-based taxonomy for phage. *J. Bacteriol.* **184**:4529–4535.
197. Ruzin, A., J. Lindsay, and R. P. Novick. 2001. Molecular genetics of SaPI1—a mobile pathogenicity island in *Staphylococcus aureus*. *Mol. Microbiol.* **41**:365–377.
198. Sandkvist, M. 2001. Biology of type II secretion. *Mol. Microbiol.* **40**:271–283.
199. Sato, T., T. Shimizu, M. Watarai, M. Kobayashi, S. Kano, T. Hamabata, Y. Takeda, and S. Yamasaki. 2003. Genome analysis of a novel Shiga toxin 1 (Stx1)-converting phage which is closely related to Stx2-converting phages but not to other Stx1-converting phages. *J. Bacteriol.* **185**:3966–3971.
200. Schicklmaier, P., and H. Schmieger. 1995. Frequency of generalized transducing phages in natural isolates of the *Salmonella typhimurium* complex. *Appl. Environ. Microbiol.* **61**:1637–1640.
201. Shaikh, N., and P. I. Tarr. 2003. *Escherichia coli* O157:H7 Shiga toxin-encoding bacteriophages: integrations, excisions, truncations, and evolutionary implications. *J. Bacteriol.* **185**:3596–3605.
202. Skorupski, K., and R. K. Taylor. 1997. Control of the ToxR virulence

- regulon in *Vibrio cholerae* by environmental stimuli. *Mol. Microbiol.* **25**: 1003–1009.
203. Smoot, J. C., K. D. Barbian, J. J. Van Gompel, L. M. Smoot, M. S. Chaussee, G. L. Sylva, D. E. Sturdevant, S. M. Ricklefs, S. F. Porcella, L. D. Parkins, S. B. Beres, D. S. Campbell, T. M. Smith, Q. Zhang, V. Kapur, J. A. Daly, L. G. Veasy, and J. M. Musser. 2002. Genome sequence and comparative microarray analysis of serotype M18 group A *Streptococcus* strains associated with acute rheumatic fever outbreaks. *Proc. Natl. Acad. Sci. USA* **99**:4668–4673.
 204. Smoot, L. M., J. K. McCormick, J. C. Smoot, N. P. Hoe, I. Strickland, R. L. Cole, K. D. Barbian, C. A. Earhart, D. H. Ohlendorf, L. G. Veasy, H. R. Hill, D. Y. Leung, P. M. Schlievert, and J. M. Musser. 2002. Characterization of two novel pyrogenic toxin superantigens made by an acute rheumatic fever clone of *Streptococcus pyogenes* associated with multiple disease outbreaks. *Infect. Immun.* **70**:7095–7104.
 205. Smoot, L. M., J. C. Smoot, M. R. Graham, G. A. Somerville, D. E. Sturdevant, C. A. Migliaccio, G. L. Sylva, and J. M. Musser. 2001. Global differential gene expression in response to growth temperature alteration in group A *Streptococcus*. *Proc. Natl. Acad. Sci. USA* **98**:10416–10421.
 206. Sriskandan, S., M. Unnikrishnan, T. Krausz, and J. Cohen. 1999. Molecular analysis of the role of streptococcal pyrogenic exotoxin A (SPEA) in invasive soft-tissue infection resulting from *Streptococcus pyogenes*. *Mol. Microbiol.* **33**:778–790.
 207. Stanley, T. L., C. D. Ellermeier, and J. M. Slouch. 2000. Tissue-specific gene expression identifies a gene in the lysogenic phage Gifsy-1 that affects *Salmonella enterica* serovar Typhimurium survival in Peyer's patches. *J. Bacteriol.* **182**:4406–4413.
 208. Stebbins, C. E., and J. E. Galan. 2003. Priming virulence factors for delivery into the host. *Nat. Rev. Mol. Cell Biol.* **4**:738–743.
 209. Stender, S., A. Friebel, S. Linder, M. Rohde, S. Mirold, and W. D. Hardt. 2000. Identification of SopE2 from *Salmonella typhimurium*, a conserved guanine nucleotide exchange factor for Cdc42 of the host cell. *Mol. Microbiol.* **36**:1206–1221.
 210. Stevens, M. P., O. Marches, J. Campbell, V. Huter, G. Frankel, A. D. Phillips, E. Oswald, and T. S. Wallis. 2002. Intimin, tir, and Shiga toxin 1 do not influence enteropathogenic responses to Shiga toxin-producing *Escherichia coli* in bovine ligated intestinal loops. *Infect. Immun.* **70**:945–952.
 211. Sumby, P., and M. K. Waldor. 2003. Transcription of the toxin genes present within the staphylococcal phage φSa3ms is intimately linked with the phage's life cycle. *J. Bacteriol.* **185**:6841–6851.
 - 211a. Sung, L. M., M. P. Jackson, A. D. O'Brien, and R. K. Holmes. 1990. Transcription of the Shiga-like toxin type II and Shiga-like toxin type II variant operons of *Escherichia coli*. *J. Bacteriol.* **182**:6386–6395.
 212. Susskind, M. M., and D. Botstein. 1978. Molecular genetics of bacteriophage P22. *Microbiol. Rev.* **42**:385–413.
 213. Tarr, P. I., L. M. Schoening, Y. L. Yea, T. R. Ward, S. Jelacic, and T. S. Whittam. 2000. Acquisition of the *rfb-gnd* cluster in evolution of *Escherichia coli* O55 and O157. *J. Bacteriol.* **182**:6183–6191.
 214. Taylor, R. K., V. L. Miller, D. B. Furlong, and J. J. Mekalanos. 1986. Identification of a pilus colonization factor that is coordinately regulated with cholera toxin. *Ann. Scuola Collana Monogr.* **3**:51–61.
 215. Taylor, R. K., V. L. Miller, D. B. Furlong, and J. J. Mekalanos. 1987. Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci. USA* **84**:2833–2837.
 216. Tettelin, H., V. Masignani, M. J. Cieslewicz, J. A. Eisen, S. Peterson, M. R. Wessels, I. T. Paulsen, K. E. Nelson, I. Margarit, T. D. Read, L. C. Madoff, A. M. Wolf, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, R. T. Deboy, A. S. Durkin, J. F. Kolonay, R. Madupu, M. R. Lewis, D. Radune, N. B. Fedorova, D. Scanlan, H. Khouri, S. Mulligan, H. A. Carty, R. T. Cline, S. E. Van Aken, J. Gill, M. Scarselli, M. Mora, E. T. Iacobini, C. Brettoni, G. Galli, M. Mariani, F. Vegini, D. Maione, D. Rinaudo, R. Rappuoli, J. L. Telford, D. L. Kasper, G. Grandi, and C. M. Fraser. 2002. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. *Proc. Natl. Acad. Sci. USA* **99**:12391–12396.
 217. Tsuzuki, K., K. Kimura, N. Fujii, N. Yokosawa, T. Indoh, T. Murakami, and K. Oguma. 1990. Cloning and complete nucleotide sequence of the gene for the main component of hemagglutinin produced by *Clostridium botulinum* type C. *Infect. Immun.* **58**:3173–3177.
 218. Tucker, S. C., and J. E. Galan. 2000. Complex function for SicA, a *Salmonella enterica* serovar Typhimurium type III secretion-associated chaperone. *J. Bacteriol.* **182**:2262–2268.
 219. Uchida, T., D. M. Gill, and A. M. Pappenheimer, Jr. 1971. Mutation in the structural gene for diphtheria toxin carried by temperate phage. *Nat. New Biol.* **233**:8–11.
 220. Uzzau, S., L. Bossi, and N. Figueroa-Bossi. 2002. Differential accumulation of *Salmonella* [Cu, Zn] superoxide dismutases SodCI and SodCII in intracellular bacteria: correlation with their relative contribution to pathogenicity. *Mol. Microbiol.* **46**:147–156.
 221. Uzzau, S., N. Figueroa-Bossi, S. Rubino, and L. Bossi. 2001. Epitope tag-
 - ging of chromosomal genes in *Salmonella*. *Proc. Natl. Acad. Sci. USA* **98**:15264–15269.
 222. Vander, B. C., and A. M. Kropinski. 2000. Sequence of the genome of *Salmonella* bacteriophage P22. *J. Bacteriol.* **182**:6472–6481.
 223. Van Sluys, M. A., M. C. de Oliveira, C. B. Monteiro-Vitorelo, C. Y. Miyaki, L. R. Furlan, L. E. Camargo, A. C. da Silva, D. H. Moon, M. A. Takita, E. G. Lemos, M. A. Machado, M. I. Ferro, F. R. da Silva, M. H. Goldman, G. H. Goldman, M. V. Lemos, H. El Dorry, S. M. Tsai, H. Carrer, D. M. Carraro, R. C. de Oliveira, L. R. Nunes, W. J. Siqueira, L. L. Coutinho, E. T. Kimura, E. S. Ferro, R. Harakava, E. E. Kuramae, C. L. Marino, E. Giglioti, I. L. Abreu, L. M. Alves, A. M. da Amaral, G. S. Baia, S. R. Blanco, M. S. Brito, F. S. Cannavan, A. V. Celestino, A. F. da Cunha, R. C. Fenille, J. A. Ferro, E. F. Formighieri, L. T. Kishi, S. G. Leoni, A. R. Oliveira, V. E. Rosa, Jr., F. T. Sassaki, J. A. Sena, A. A. de Souza, D. Truffi, F. Tsukumo, G. M. Yanai, L. G. Zaros, E. L. Civerolo, A. J. Simpson, N. F. Almeida, Jr., J. C. Setubal, and J. P. Kitajima. 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. *J. Bacteriol.* **185**:1018–1026.
 224. Ventura, M., A. Bruttin, C. Canchaya, and H. Brüssow. 2002. Transcription analysis of *Streptococcus thermophilus* phages in the lysogenic state. *Virology* **302**:21–32.
 225. Ventura, M., C. Canchaya, M. Kleerebezem, W. M. de Vos, R. J. Siezen, and H. Brüssow. 2003. The prophage sequences of *Lactobacillus plantarum* strain WCFS1. *Virology* **316**:245–255.
 226. Ventura, M., C. Canchaya, D. Pridmore, B. Berger, and H. Brüssow. 2003. Integration and distribution of *Lactobacillus johnsonii* prophages. *J. Bacteriol.* **185**:4603–4608.
 227. Ventura, M., C. Canchaya, R. D. Pridmore, and H. Brüssow. 2004. The prophages of *Lactobacillus johnsonii* NCC 533: comparative genomics and transcription analysis. *Virology* **320**:229–242.
 228. Virtaneva, K., M. R. Graham, S. F. Porcella, N. P. Hoe, H. Su, E. A. Graviss, T. J. Gardner, J. E. Allison, W. J. Lemon, J. R. Bailey, M. J. Parnell, and J. M. Musser. 2003. Group A *Streptococcus* gene expression in humans and cynomolgus macaques with acute pharyngitis. *Infect. Immun.* **71**:2199–2207.
 229. Voyich, J. M., D. E. Sturdevant, K. R. Braughton, S. D. Kobayashi, B. Lei, K. Virtaneva, D. W. Dorward, J. M. Musser, and F. R. DeLeo. 2003. Genome-wide protective response used by group A *Streptococcus* to evade destruction by human polymorphonuclear leukocytes. *Proc. Natl. Acad. Sci. USA* **100**:1996–2001.
 230. Wagner, P. L., D. W. Acheson, and M. K. Waldor. 2001. Human neutrophils and their products induce Shiga toxin production by enterohemorrhagic *Escherichia coli*. *Infect. Immun.* **69**:1934–1937.
 231. Wagner, P. L., J. Livny, M. N. Neely, D. W. Acheson, D. I. Friedman, and M. K. Waldor. 2002. Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. *Mol. Microbiol.* **44**:957–970.
 232. Wagner, P. L., M. N. Neely, X. Zhang, D. W. Acheson, M. K. Waldor, and D. I. Friedman. 2001. Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain. *J. Bacteriol.* **183**:2081–2085.
 233. Waldor, M. K., and J. J. Mekalanos. 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**:1910–1914.
 234. Wallis, T. S., and E. E. Galyov. 2000. Molecular basis of *Salmonella*-induced enteritis. *Mol. Microbiol.* **36**:997–1005.
 235. Weller, G. R., and A. J. Doherty. 2001. A family of DNA repair ligases in bacteria? *FEBS Lett.* **505**:340–342.
 236. Weller, G. R., B. Kysela, R. Roy, L. M. Tonkin, E. Scanlan, M. Della, S. K. Devine, J. P. Day, A. Wilkinson, F. d'Adda di Fagagna, K. M. Devine, R. P. Bowater, P. A. Jeggo, S. P. Jackson, and A. J. Doherty. 2002. Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* **297**:1686–1689.
 237. Westmoreland, B. C., W. Szybalski, and H. Ris. 1969. Mapping of deletions and substitutions in heteroduplex DNA molecules of bacteriophage lambda by electron microscopy. *Science* **163**:1343–1348.
 238. Williamson, S. J., L. A. Houchin, L. McDaniel, and J. H. Paul. 2002. Seasonal variation in lysogeny as depicted by prophage induction in Tampa Bay, Florida. *Appl. Environ. Microbiol.* **68**:4307–4314.
 239. Winfield, M. D., and E. A. Groisman. 2003. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. *Appl. Environ. Microbiol.* **69**:3687–3694.
 240. Wommack, K. E., and R. R. Colwell. 2000. Viriplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* **64**:69–114.
 241. Wong, C. S., S. Jelacic, R. L. Habeeb, S. L. Watkins, and P. I. Tarr. 2000. The risk of the hemolytic-uremic syndrome after antibiotic treatment of *Escherichia coli* O157:H7 infections. *N. Engl. J. Med.* **342**:1930–1936.
 242. Wood, M. W., R. Rosqvist, P. B. Mullan, M. H. Edwards, and E. E. Galyov. 1996. SopE, a secreted protein of *Salmonella dublin*, is translocated into the target eukaryotic cell via a sip-dependent mechanism and promotes bacterial entry. *Mol. Microbiol.* **22**:327–338.
 243. Worley, M. J., K. H. Ching, and F. Heffron. 2000. *Salmonella* SsrB activates a global regulon of horizontally acquired genes. *Mol. Microbiol.* **36**:749–761.

244. Wyss, A. 2001. Paleontology. Digging up fresh clues about the origin of mammals. *Science* **292**:1496–1497.
245. Yamaguchi, T., T. Hayashi, H. Takami, K. Nakasone, M. Ohnishi, K. Nakayama, S. Yamada, H. Komatsuzawa, and M. Sugai. 2000. Phage conversion of exfoliative toxin A production in *Staphylococcus aureus*. *Mol. Microbiol.* **38**:694–705.
246. Yokoyama, K., K. Makino, Y. Kubota; M. Watanabe, S. Kimura, C. H. Yutsudo, K. Kurokawa, K. Ishii, M. Hattori, I. Tatsuno, H. Abe, M. Yoh, T. Iida, M. Ohnishi, T. Hayashi, T. Yasunaga, T. Honda, C. Sasakawa, and H. Shinagawa. 2000. Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic *Escherichia coli* O157:H7 strain derived from the Sakai outbreak. *Gene* **258**:127–139.
247. Yu, C. E., and J. J. Ferretti. 1989. Molecular epidemiologic analysis of the type A streptococcal exotoxin (erythrogenic toxin) gene (*speA*) in clinical *Streptococcus pyogenes* strains. *Infect. Immun.* **57**:3715–3719.
248. Yu, C. E., and J. J. Ferretti. 1991. Frequency of the erythrogenic toxin B and C genes (*speB* and *speC*) among clinical isolates of group A streptococci. *Infect. Immun.* **59**:211–215.
249. Zhang, S., R. L. Santos, R. M. Tsolis, S. Mirold, W.-D. Hardt, L. G. Adams, and A. J. Bäumler. 2002. Phage-mediated horizontal transfer of the *sopE1* gene increases enteropathogenicity of *Salmonella enterica* serotype Typhimurium for calves. *FEMS Microbiol. Lett.* **217**:243–247.
250. Zhang, S., R. L. Santos, R. M. Tsolis, S. Stender, W. D. Hardt, A. J. Bäumler, and L. G. Adams. 2002. The *Salmonella enterica* serotype Typhimurium effector proteins SipA, SopA, SopB, SopD, and SopE2 act in concert to induce diarrhea in calves. *Infect. Immun.* **70**:3843–3855.
251. Zhang, X., A. D. McDaniel, L. E. Wolf, G. T. Keusch, M. K. Waldor, and D. W. Acheson. 2000. Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice. *J. Infect. Dis.* **181**:664–670.
252. Zhou, Y., H. Sugiyama, and E. A. Johnson. 1993. Transfer of neurotoxicity from *Clostridium butyricum* to a nontoxicogenic *Clostridium botulinum* type E-like strain. *Appl. Environ. Microbiol.* **59**:3825–3831.
253. Zhou, Y., H. Sugiyama, H. Nakano, and E. A. Johnson. 1995. The genes for the *Clostridium botulinum* type G toxin complex are on a plasmid. *Infect. Immun.* **63**:2087–2091.
254. Zou, D., J. Kaneko, S. Narita, and Y. Kamio. 2000. Prophage, ϕ PV83-pro, carrying Panton-Valentine leukocidin genes, on the *Staphylococcus aureus* P83 chromosome: comparative analysis of the genome structures of ϕ PV83-pro, ϕ PVL, ϕ 11, and other phages. *Biosci. Biotechnol. Biochem.* **64**:2631–2643.