Prophages in Salmonella enterica: a driving force in reshaping the genome and physiology of their bacterial host?

Astrid Wahl, Aurelia Battesti, Mireille Ansaldi

To cite this version:

Astrid Wahl, Aurelia Battesti, Mireille Ansaldi. Prophages in Salmonella enterica: a driving force in reshaping the genome and physiology of their bacterial host?. Molecular Microbiology, Wiley, 2018, pp.1 - 14. 10.1111/mmi.14167. hal-01988228

HAL Id: hal-01988228
https://hal-amu.archives-ouvertes.fr/hal-01988228
Submitted on 21 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
MicroReview

Prophages in *Salmonella enterica*: a driving force in reshaping the genome and physiology of their bacterial host?

Astrid Wahl, Aurélia Battesti and Mireille Ansaldi*
Laboratoire de Chimie Bactérienne, UMR7283, Institut de Microbiologie de la Méditerranée, Centre National de la Recherche Scientifique, Aix-Marseille Université, Marseille, France.

Summary
Thanks to the exponentially increasing number of publicly available bacterial genome sequences, one can now estimate the important contribution of integrated viral sequences to the diversity of bacterial genomes. Indeed, temperate bacteriophages are able to stably integrate the genome of their host through site-specific recombination and transmit vertically to the host siblings. Lysogenic conversion has been long acknowledged to provide additional functions to the host, and particularly to bacterial pathogen genomes where prophages contribute important virulence factors. This review aims particularly at highlighting the current knowledge and questions about lysogeny in *Salmonella* genomes where functional prophages are abundant, and where genetic interactions between host and prophages are of particular importance for human health considerations.

Introduction
Bacteriophages are recognized as the most abundant biological entities on earth, participating to numerous biological cycles and constantly reshaping bacterial communities (Suttle, 2007; Brussaard et al., 2008). In all environments they outcompete the number of available hosts by one to several log10. Moreover, due to their propensity to lysogenize, i.e. become quiescent proviruses, temperate phages are recognized as essential drivers of bacterial genomes’ evolution (Roux et al., 2015a; 2015b; Casjens and Grose, 2016) (some detailed definitions can be found in the glossary).

This review aims at highlighting the contribution of prophage genes to the host physiology. To date, a lot of emphasis has been put on the identification and characterization of phage-encoded virulence factors in various pathogenic bacteria (Brüssow et al., 2004; Dearborn and Dokland, 2012; Rabinovich et al., 2012; Busby et al., 2012; Fortier and Sekulovic, 2013; Davies et al., 2016; Kraushaar et al., 2017). However, as it becomes clearer that bacterial genomes contain large amounts of DNA from (pro)phage origin, we want to stress that these horizontally acquired genes are important contributors to the genomes evolution and provide discrete adaptive physiological contributions such as increasing fitness under certain environmental conditions or providing non-obvious metabolic or signaling functions (D’Ari and Casadesús, 1998). We chose to focus on *Salmonella enterica* prophages for the following reasons: (i) it is a widespread enterobacteria displaying a broad host range, frequently carried by wild and domestic birds as well as rodents, and an animal and human pathogen, (ii) host-prophage interactions have been studied for many years and still lead to amazing pieces of work encompassing many topics, such as host-phage interactions, virulence, ecology and genome evolution.

Prophage abundance and integration sites in *S. enterica* genomes
The first prophages in *Salmonella* species have been identified in 1950, just before transduction has been discovered (Boyd, 1950; Zinder and Lederberg, 1952). Since the 1990s, the Bossi group has been a pioneer in *S. enterica* prophage research, which highlighted the diversity of the prophage repertoire of various strains

© 2018 The Authors. *Molecular Microbiology* Published by John Wiley & Sons Ltd
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
A. Wahl, A. Battesti and M. Ansaldi

Glossary

Defective: a defective prophage has lost part of its genome and can no longer produce viral particles, but may still be able to excise from the host chromosome.

Functional: a prophage is considered functional if it can resume a lytic cycle and re-infect naive cells.

Lysogeny: upon lysogenic infection by a temperate phage its genome is stably maintained as a prophage (most of the time integrated into the chromosome) in the host bacterium and vertically replicates with it. No progeny is produced until a lytic cycle resumes.

Lysogenic conversion: describes the phenotypic changes bacteria undergo upon infection by a temperate phage going through lysogeny.

Morons (“more on”): genes regulated independently from the rest of the prophage and conferring a fitness advantage under specific environmental conditions.

Polylysogeny: the hosting of multiple prophages in a single host genome.

Temperate bacteriophage: a phage that can multiply either lytically or through lysogeny.

Propagolysin: inherited form of a temperate bacteriophage; dormant form of the viral genome which replicates with the bacterial host genome.

Pseudolysogeny: a phage–host cell interaction in which the phage genome does not integrate into the host’s and forms an episomal form that can be still transmitted vertically. Pseudolysogeny applies to virulent and temperate phages and does not lead to the usual outcome lysis or lysogeny.

Pseudogenization: a gene becomes a pseudogene by accumulating mutations that hinder its correct transcription and/or translation.

Viral or strictly lytic bacteriophage: a phage that multiplies exclusively through a lytic cycle.

(Figueroa-Bossi et al., 1997; 2001; Figueroa-Bossi and Bossi, 1999; Bossi et al., 2003). Since then many more prophages have been identified, every time a set of new Salmonella genomes is sequenced. Dormant prophages are transmitted vertically along with bacterial cell division and can be induced under stressful conditions, such as DNA damages or in animal guts (Kim et al., 2014). They can also undergo spontaneous induction, which can increase the fitness of a given strain whenever in competition when entering a new niche (Bossi et al., 2003).

Salmonella enterica genomes also carry defective prophages that are no longer able to form infectious particles, meanwhile being present – and perhaps maintained – in the host chromosome (Cajens, 2003; Bobay et al., 2013). Different events can lead to prophage degradation including large genomic reduction, targeting by insertion sequences (IS) as well as point mutations (Bobay et al., 2014). However, when prophages are not too degraded it is possible to “resuscitate” them into fully functional prophages, meaning inducible and able to form infectious particles. Such reactivations of defective prophages involve either a temporal complementation by an infecting phage that provides the missing function, or a recombination event that allows a permanent complementation (Figueroa-Bossi and Bossi, 2004; De Paepe et al., 2014). The recombination events driven either by the host homologous recombinases or phage-encoded recombinases inside the host cells are causing pervasive mosaicism in phage genomes (Lopes et al., 2010; De Paepe et al., 2014; Menouni et al., 2015). The best known and long studied temperate phage infecting S. enterica is P22 (the λ equivalent paradigm in S. enterica) that was a key model for transduction discovery (Boyd, 1950; Zinder and Lederberg, 1952). However, P22 itself is not a common prophage in S. enterica genomes that contain in average 5.29 prophages representing around 3.52% of the total gene content and close to 30% of the accessory genome (based on 21 S. enterica genomes analyzed) (Bobay et al., 2013). In other words, these numbers show that polylsogeny, i.e. the hosting of multiple prophages by a single genome, is a very frequent event. Another striking point is that integration sites are highly conserved between the two closely related Escherichia coli and S. enterica species and even beyond (Bobay et al., 2013; Oliveira et al., 2017). Among favored integration sites are found all categories of non-translated RNA genes such as sRNA, tmRNA and tRNA, the latter being the most frequently targeted (Bobay et al., 2013). Other sites in the chromosome may be targeted as well, such as intergenic regions, while integration within protein-encoding genes is much less frequent. Even when integrating at the 3’ end of genes, the site-specific reaction involved in the integration process leads to the reconstitution of the targeted genes since the equivalent portion of the gene is provided on the phage genome, without affecting their function or expression (Argos et al., 1986). Alternatively, a prophage may disrupt a gene and therefore a cellular function. However, when the phage excises, the interrupted gene can be reconstituted and the host regains the lost function, a process called phage-driven regulatory switch or active lysogeny (Feiner et al., 2015). However, no such a switch has been experimentally described so far in S. enterica genomes.

The quasi-weekly release of new draft genomes from S. enterica prevents an accurate description of the prophage content as prophage description and annotation are not so obvious, even though facilitated by various softwares (Clokie and Kropinski, 2009). Indeed, the presence of multiple contigs may hinder the correct description of prophages as they frequently co-localize with
contig borders, impairing correct genome assembly and are sometimes interrupted by insertion sequences (IS). As a result, prophage predictions need to undergo expert manual curation.

A recent study based on public health surveillance in the UK highlighted that *S. enterica* Typhimurium causing invasive non-typhoidal salmonellosis in Africa carried a specific prophage as well as antibiotic resistance genes that are not found in the UK version of this lineage (Kintz *et al*., 2015; Owen *et al*., 2017; Ashton *et al*., 2017). As a consequence stably integrated prophages are useful tools as epidemiology markers in addition to CRISPR-Cas typing. However, as the latter were found to be poorly active and show a very slow spacer turnover, such typing should be restricted to anciently diverged strains (Touchon and Rocha, 2010).

**Lysogenic conversion**

The notion of lysogenic conversion, meaning the propensity of temperate phage undergoing lysogeny to contribute to the host physiology has been described and admitted for numerous years. However, a strong bias is observed in the literature toward lysogenic conversion aspects that contribute to the host virulence. As an example, the Gifsy-2 encoded superoxide dismutase SodC that obviously contribute to the establishment of *Salmonella* cells into the macrophage (Figueroa-Bossi and Bossi, 1999). Needless to say that *S. enterica* is an organism of choice for such contribution examples. However, one must consider that more subtle contributions do exist and pave the way for multiple interactions with the host genome as well as with the eukaryotic cells targeted by *S. enterica* or the microbiome encountered by the pathogen during its infectious journey in animals.

**Prophage induction and prophage gene expression under lysogenic conditions**

As in *Escherichia coli*, the repressor model is widespread in *Salmonella*’s Lambdoid prophages (Sauer *et al*., 1981; Campbell, 1994; Whipple *et al*., 1998). However, a striking and widely conserved feature is the involvement of antirepressor proteins in prophage induction. If most *S. enterica* prophages are induced by the activation of the SOS response due to DNA damaging factors (mitomycin C, UV or H$_2$O$_2$), the cleavage of the repressor is not the major outcome of the induction system. Indeed, it was shown for several *S. enterica* prophages that upon SOS response induction and LexA self-cleavage, an antirepressor protein (Ant), homologous to the Tum one in phage 186, is being made that inhibits the lytic repressor through protein-protein interaction (Shearwin *et al*., 1998; Lemire *et al*., 2011; Kim and Ryu, 2013). P22 also encodes such an antirepressor whose expression is negatively controlled by the Mnt repressor. Nevertheless, in this case, an ant mutant remains SOS inducible (Botstein *et al*., 1975; Levine *et al*., 1975). Interestingly, antirepressors are responsible for prophage induction crosstalk: a given prophage-encoded antirepressor was shown to counteract the action of a repressor from another prophage (Lemire *et al*., 2011) (Fig. 1 and Table 1 subitem 2.1). This prophage crosstalk has probably a role in prophage dissemination, as non-coordinated prophages could be lost upon massive host cell lysis.

![Fig. 1.](image-url) *Multiple interactions between prophages and Salmonella hosts*: The multiple host-prophage and prophage-prophage interactions depicted in the text are illustrated. For more details see Table 1.
provoked by a neighboring prophage undergoing induction. In contrast, a partial cell lysis is often provoked by spontaneous or uncompleted induction, which allows a non-induced prophage to remain in the bacterial population.

A number of inducing conditions or chemicals have been described that induce prophages in *Salmonella* through the SOS-response activation. However, induction driven by some of these should lead to preoccupation in the context of public health risks. Indeed, it has been shown recently that some prophages are induced by antibiotics widely used in agriculture, such as fluoroquinolones (Bearson and Brunelle, 2015) and carbadox (Bearson et al., 2014). As a result, these antibiotics that are most probably present in our daily food intake may have thus consequences on prophage induction in our gut. Also relevant to the physiology of *Salmonella* could be the induction by bile salts, which are encountered by this bacterium in its natural ecological niche (Hernández et al., 2012). In all cases, little is known about the *in vivo* consequences of these treatments. However, one should consider the demonstration of *in vivo* transfer of prophage-encoded virulence genes at loci of inflammation (Diard et al., 2017), which raises the question of the consequences of inducing antibiotics on the transmission of prophage-encoded genes.

A commonly admitted view is that prophage genes are under control either by the powerful phage repressor or by host factors that tend to limit the negative effects of horizontal gene acquisition. In the first case, thanks to the seminal work of Jacob and Monod “The operon model” (Lewis, 2011), the lysogeny repressor is known to control the expression of the lytic promoters (PL and PR) in the absence of induction. Host-controlled expression of exogenic, and prophage genes in particular, has been described in several bacterial models (Navarre et al., 2006; Cardinale et al., 2008). However, it is in *S. enterica* that the role of the nucleoid associated protein H-NS has been acknowledged as a genome sentinel (see below) (Navarre et al., 2006; Dorman, 2007; Navarre et al., 2007; Ali et al., 2013). Although there is no doubt about the extensive repression of prophage genes involved in lytic functions, we argue that it may have occulted the expression of those genes involved

**Table 1.** Host-prophage and prophage-prophage interactions in *S. enterica*.

<table>
<thead>
<tr>
<th>1. Interactions bacteria – prophage and prophage – bacteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.1 Host factors</strong></td>
</tr>
<tr>
<td>• Phase variation: Dam in conjunction with other regulator, such as OxyR</td>
</tr>
<tr>
<td>• H-NS (silencing of exogenic DNA as heterodimers with proteins from Hha family)</td>
</tr>
<tr>
<td>• InvF (transcriptional regulator from AraC/ XylS family; sopE)</td>
</tr>
<tr>
<td>• SsrAB (two-component system; gogB, sseI, stm2239)</td>
</tr>
<tr>
<td><strong>1.2 Prophage factors</strong></td>
</tr>
<tr>
<td>• ΦW104 (ryeA, ryeB)</td>
</tr>
<tr>
<td>• Gifsy1 (AntQ forms complex with bacterial RNA polymerase)</td>
</tr>
<tr>
<td><strong>2. Interactions prophage – (pro)phage</strong></td>
</tr>
<tr>
<td><strong>2.1 Cross-regulation</strong></td>
</tr>
<tr>
<td>Repressor – Antirepressor (Fels-2 and Gifsy prophages)</td>
</tr>
<tr>
<td><strong>2.2 Super-infection exclusion</strong></td>
</tr>
<tr>
<td>• SieA and SieB (P22)</td>
</tr>
<tr>
<td>• Repressor C2 in cytoplasm (phage carrier state of P22)</td>
</tr>
<tr>
<td><strong>3. Surface features encoded by prophages and phage remnants</strong></td>
</tr>
<tr>
<td><strong>3.1 Glucosylation</strong></td>
</tr>
<tr>
<td>• GtrABCa (phage remnant)</td>
</tr>
<tr>
<td>• GtrABC&lt;sub&gt;P22&lt;/sub&gt; (P22)</td>
</tr>
<tr>
<td><strong>3.2 O-acetylation</strong></td>
</tr>
<tr>
<td>• OafA (phage remnant)</td>
</tr>
<tr>
<td>• Protein similar to acyltransferase 3 (SPC-P1 prophage)</td>
</tr>
<tr>
<td>• GtrABC&lt;sub&gt;BTP1&lt;/sub&gt; (BTP1; GtrC containing an acyltransferase domain)</td>
</tr>
<tr>
<td><strong>3.3 Glycosidic bonds</strong></td>
</tr>
<tr>
<td>• phage beta polymerase changes to β-1,6 glycosidic linkage (<em>ε</em>15; new phage host range: <em>ε</em>34)</td>
</tr>
<tr>
<td><strong>3.4 Adhesion</strong></td>
</tr>
<tr>
<td>• tail-like protein STM2699 (Fels-2)</td>
</tr>
</tbody>
</table>
in lysogenic conversion and for which no obvious function in metabolic or pathogenic pathways have been identified. In order to find more of these genes whose expression remains significant under lysogenic conditions, one can look at the impressive amount of transcriptomic data available in the public databases and have a search in prophage-related regions. One such very helpful database dedicated to *S. enterica* is the SalCom compendium developed by and hosted in Jay Hinton’s lab (Kröger et al., 2013; Colgan et al., 2016) (http://bioinf.gen.tcd.ie/cgi-bin/salcom.pl?_HL).

**Contributions to phage resistance and physiology**

Being extremely abundant on earth while presenting a gene content very different from that currently available in the genomic databases, bacteriophage genomes are believed to contribute extensively to the genetic “dark matter” (Rinke et al., 2013; Roux et al., 2015b; Nadeem and Wahl, 2017). Indeed, even when focusing on a single bacterial host the gene repertoire carried by its infecting phages is highly diverse (Hatfull et al., 2010; Cressawn et al., 2011; Hatfull, 2015), and a large proportion of these genes remain annotated with unknown functions. A number of contributions to the host physiology might also benefit to the prophages themselves; among them are those preventing phage superinfection and/or recognition by the immune system from mammals.

**LPS modifications.** One particularity of phages infecting *Salmonella* strains is that they often recognize the lipopolysaccharide (LPS) or endotoxin consisting of a lipid and a polysaccharide composed of O-antigens, as receptor as opposed to *E. coli* bacteriophages that often target outer membrane proteins (Bertozzi Silva et al., 2016). LPS contributes to the structural integrity and negative charge of the outer membrane of Gram-negative bacteria and protects the membrane from certain kinds of chemicals, such as anionic bile salts or lipophilic antibiotics. Interestingly, the cognate *Salmonella* prophages frequently carry genes encoding proteins involved in O-antigen modifications (Fig. 1 and Table 1 subitem 3) (Duerr et al., 2009; Broadbent et al., 2010; Andres et al., 2013; Davies et al., 2013; Sun et al., 2014; Cota et al., 2015; Kintz et al., 2015).

One of the most prominent examples is bacteriophage P22 of *S. enterica* which recognizes O-antigen polysaccharides with its tailspike protein and when integrated as a lysogen in the bacterial chromosome provides itself a *gtrABC* operon for O-antigen glucosylation (Andres et al., 2010; 2013). Up to four *gtr* operons encoded on different prophages and responsible for distinct modifications can be found within one bacterial genome (Broadbent et al., 2010). The *gtrA* and *gtrB* genes encode conserved membrane proteins, the bactoprenol-linked glucosyl translocase or “flipppase” and the bactoprenol glucosyl transferase respectively. It is the third gene, *gtrC*, in the cluster that is variable and confers specificity since it encodes the glucosyl transferase attaching a glucose group at a distinct O-antigen position. These *gtr* operons are regulated according to phase variation (see the “Negative host-control of prophage genes” section). LPS is a potent activator of immune cells and therefore these temporarily changing surface modifications can shade the bacteria vis-à-vis the eukaryotic immune system (Duerr et al., 2009). Besides glucosylation, another possible LPS O-antigen modification is acetylation. An example is the gene *oatA* which is located on a phage remnant (at ~2.33 Mb on the *S. enterica* 4/74 chromosome) and codes for an integral membrane transacylase conferring the specific serotype 05 (Slauch et al., 1996). More recently, a gene coding for a protein similar to the acetyltransferase 3 of *Pseudomonas syringae* *py*. *Syringae* B728a has been identified on prophage SPC-P1, which was associated to increased virulence in *S. paratyphi* C (Zou et al., 2010). Another intriguing example is the *S. enterica* serovar Anatum specific phase ε15 which modulates glycosidic linkage of O-antigen by blocking the host α polymerase and producing its own β polymerase (changes from α-1,6 to β-1,6 glycosidic linkage). This modification restricts super-infection by ε15 itself and in turn allows infection by phage ε34 (Kropinski et al., 2007).

All these genetic features responsible for O-antigen modifications and prophage-encoded factors without being “real” virulence factors, may affect the resistance to the intestinal environment as well as to sur-infecting phages, and thus in sum the fitness and pathogenicity of the bacterial host.

**Superinfection exclusion.** Besides the LPS modifications, phages have acquired other ways to prevent superinfection by themselves and other phages. As examples we will mention here the genes *sieA* and *sieB* and the so-called phage carrier state of P22 (Fig. 1 and Table 1 subitem 2.2). The inner membrane protein SieA is seemingly responsible for blocking the phage DNA transfer across the membrane into the bacterial cytoplasm (Susskind et al., 1974; Susskind and Botstein, 1980). SieB, also encoded by *E. coli* phage λ, aborts the lytic development of other *Salmonella* phages - such as P22-like MG178 and MQ40 - by stopping RNA, DNA, and protein synthesis. P22 itself is not affected, since it produces an early uncharacterized escape factor (Susskind et al., 1974). Interestingly, both SieA and SieB inhibit infection of the same phages, including λ, although *Salmonella* is not a standard host for the latter (Ranade and Poteete, 1993).
Phage carrier cells were identified recently and are cells infected with P22 harboring an episomal form of P22 that is transmitted asymmetrically during division (Cenens et al., 2013a; 2013b). The daughter cell inheriting this episome enters lysogeny resulting in a chromosomally integrated prophage. The other daughter cell becomes P22-free, but intriguingly stays resistant to P22 infection in a transient way. The immunity factors responsible for this resistance are GtrABC, SieA and the repressor C2, which have been constitutively produced by the phage carrier cell. The P22-free daughter cells cytoplasmically inherit these immunity factors, which then dilute out upon subsequent cell divisions (Cenens et al., 2015). By conferring this temporary resistance to a bacterial subpopulation, phages might thus insure both vertical and horizontal transmission routes throughout an infected population while maintaining a bacterial population they can infect. This process, described as “host-farming” by A. Aertsen, allows P22 to cultivate susceptible non-immune cells as a prey reservoir. This is an example of how up-to-date single-cell studies contribute to the field and also shows that even the extensively studied phages such as P22 still reserves surprising new features to uncover.

Contribution of prophage gene products to virulence

This section will be mainly dedicated to recent or not yet reviewed examples of prophage-encoded factors that are involved in Salmonella virulence (Boyd and Brüssow, 2002; Boyd, 2012; Boyd et al., 2012). Salmonella displays full panoply of virulence factors permitting to adhere to and infect eukaryotic cells and survive within them, in particular the hostile microenvironment of macrophages.

Adhesion. Shah et al. have identified a prophage gene, gpE, coding for a putative tail-spike protein in SopEΦ in S. enterica LT2 that increased binding to epithelial cells (specifically via Spectrin1, an eukaryotic surface protein) and increased cell invasion (Shah et al., 2014; Fig. 1, Table 1 subitem 3.4). However, this prophage-encoded gene was only expressed when bacteria were exposed to a cold stress before the infection assay. This emphasizes the importance of storage conditions for Salmonella contaminated food (mostly eggs and poultry) that are stored at cold and then reheated exposing bacteria to a chain of stresses, which may in turn induce prophages genes and finally, increase bacterial virulence.

Host entry, manipulation and intracellular survival (virulence factors per se). One of the most impressive features of S. enterica virulence is the secretion of multiple effectors involved in virulence. The effectors are secreted via two different Type Three Secretion Systems (T3SS) encoded by the pathogenicity islands 1 and 2 (SPI1 and 2). SPI1 is activated when S. enterica is in contact with eukaryotic host cells, whereas SPI2 is expressed during the phagocytosis step (Kaur and Jain, 2012). Bacterial effectors are able to interfere and hijack the host signaling pathways. In addition, other SPI-encoded factors exist that facilitate bacterial survival, among which so-called anti-virulence factors, which deletions curiously confer more virulence to the bacteria.

Among the SPI1 secreted effectors, SopE is one of the best studied effector proteins and its secretion results in actin cytoskeleton rearrangements and stimulates membrane ruffling, promoting bacterial entry into non-phagocytic cells such as epithelial cells (reviewed in (Ehrbar and Hardt, 2005)). Its encoding gene has been originally identified on the SopEΦ prophage in S. enterica SL1344, but the sopE gene and a constant flanking sequence, called the SopE-cassette, is sporadically distributed in other lambdoid prophages of the Gifsy family among several Salmonella serovars, as well as on a P2-like prophage in S. typhi (Mirold et al., 2001; Bachmann et al., 2014). The SopE-cassette has most probably been transferred and integrated among these prophages by homologous recombination resulting in multiple sopE copies present in a single bacterial genome (Hoffmann et al., 2014). In the context of evolution, such a modular exchange mechanism could enhance effector protein diversity, since genes may duplicate and then potentially evolve to other functions. However, even the well-studied SopE virulence factor may still not have unveiled all its functions. Indeed, recently SopE has been found to be not only produced and secreted for entering the eukaryotic cell, but also during the intracellular state where it seems to participate to the formation of the early Salmonella-containing vacuole (SCV) (Vonaesch et al., 2014). The SCV is formed in order to create a replicative niche for the bacteria within the host cell. This double function of SopE during cell entry as well as during intracellular survival suggests that other prophage-encoded effectors may have additional functions for host-cell manipulation.

Gifsy1 prophage has been found to encode three genes involved in intra-cellular survival: gogB, sarA, and pagK2. The first gene, gogB, codes for an anti-inflammatory effector, which inhibits NFκB activation by interaction with host factors Skp1 and FBX022. It is thought that this anti-inflammatory effect limits tissue damage during longer term infection, while short-term inflammation enhances colonization in the intestine (Pilar et al., 2012).

The second gene, sarA, has been identified only very recently. SarA is mainly secreted by SPI2-encoded
T3SS, although there is also some translocation by SPI1 T3SS. It activates the eukaryotic transcription factor STAT3, which induces the transcription of IL-10 as well as of other anti-inflammatory factors. SarA is thus the first example of an effector that activates an anti-inflammatory pathway in the eukaryotic host cell (Jaslow et al., 2018). PagK2 is secreted in outer membrane vesicles and contributes to intracellular survival in macrophages through an unknown mechanism (Yoon et al., 2011). The anti-inflammatory effects of the T3SS effectors seem to be crucial at systemic sites later in infection when S. enterica must evade immunity and promote intracellular growth. Apparently, there is an evolutionary advantage to maintain gogB and sarA on the same prophage and the recent identification of a new prophage ST-1974 in S. enterica Enteritidis supports this idea (D’Alessandro et al., 2018). In this case, the two genes coding for anti-inflammatory functions, gogB and ssek3, are present on a single prophage. As mentioned above, gogB is encoded on Gifsy1 but can be found elsewhere on the chromosome (see below), while ssek3 has been previously identified on prophage ST64B (Brown et al., 2011). So, it seems that recombination events, similar to the above-mentioned modular exchange of the SopE-cassette, have taken place between these prophages. Interestingly, both GogB and SseK3 act on the same anti-inflammatory NFκB pathway. However, the SseK3 host targets remain to be identified (Yang et al., 2015).

Antivirulence. Some prophage-encoded genes confer an intriguing phenotype termed anti-virulence. The Gifsy2-encoded grvA gene is such an anti-virulence factor: in its absence, and in contrast to what one would expect with classical virulence genes, the bacterial host is more virulent than a wild type strain in competition assays in mice (Ho and Slauch, 2001). However, this phenotype is only observed when sodCI, a Gifsy2-encoded superoxide dismutase, is present as well. Thus, it is hypothesized that in a wild type situation, GrvA decreases the pathogenicity of the host probably by affecting resistance to toxic oxygen species via SodCI through an unknown mechanism. Another peculiar example of a factor that can be a virulence or an anti-virulence factor depending on the serovar type of its host, is bstA encoded on prophage BTP1. Indeed, it acts as virulence factor in S. enterica ST313, i.e. higher uptake in macrophages (Herrero-Fresno et al., 2014), while it was described as an anti-virulence factor (lower uptake) in S. enterica Dublin; however the molecular mechanisms underlying both phenotypes are not yet understood (Herrero-Fresno et al., 2018). A potential reason for this difference may be that, similarly to GrvA, another virulence factor is affected by BstA and is present in only one serovar. This highlights the possibility of a different output of prophage genes depending of the pre-existing bacterial regulatory networks. Currently, it is not fully understood why bacterial pathogens would possess these antivirulence genes and what the evolutionary advantage (for both the prophage and the host) might be. One may speculate that bacterial pathogens might evolve toward less virulence in order to ensure their own propagation by keeping the potential host in shape in a way resembling to the above described phage-carrier state. The fact that these factors are prophage-encoded might give an advantage to changing environmental niches of the mammal’s body, since prophages can be lost and acquired in only one recombination event in the gut and therefore provide a fast way of adaptation (Diard et al., 2017).

Host–prophage regulatory networks

As mentioned earlier, genes from phage origin represent a large part of S. enterica accessory genome. Some of these genes contribute to the host physiology and therefore need to be expressed at the right time and the right place. To this end, they become part of the bacterial regulatory network. How does acquisition of these new genes not disturb the normal bacterial functioning? How are they integrated into the host regulatory network? What potential benefit do they provide to the bacterial host? How can bacterial regulators modulate prophage behavior by modulating gene expression?

Negative host-control of prophage genes

Expression of new genes must not be detrimental for the bacterial host. Therefore, genes acquired by horizontal gene transfer (HGT) are generally first silenced before being integrated into the host regulatory network. The silencing of genes from foreign origin can, for example, occur via DNA modification or involve regulatory proteins that bind DNA to prevent transcription. One of the most studied DNA modifications is responsible for an epigenetic regulation called phase variation and occurs only in a small fraction of the bacterial population. This regulation relies on the methylation of deoxyadenosines by the Dam methylase (Deoxyadenosine methyltransferase) (Casadesús, 2016). The Dam enzyme recognizes and specifically modifies the 5′-GATC-3′ sequences; when these sequences are localized in a promoter region, methylation events can block the binding of transcriptional regulators and consequently modify gene expression. DNA methylation is involved in the silencing of genes localized on the Gifsy1, Fels1 and ST64B prophages in S. enterica SL1344 (Baldontin et al., 2006).
Strikingly, it negatively regulates most of the ST64B genes. This observation is in accordance with previous results showing that ST64B excision is inhibited by Dam regulation (Alonso et al., 2005). This has been attributed to the down regulation of two genes located on this prophage and coding for proteins involved in phage induction: the anti-repressor Sb41 and the replication protein Sb42. The bacterial regulator involved in this regulation and hindered in its function by the Dam methylation has not been identified to date. While the Dam-regulation observed for genes located on Gifsy1 and Fels1 prophages does not affect their excision, SopΦ prophage excision is favored by Dam methylation. However, the transcriptional regulator as well as the target genes responsible for this phenotype has not been described (Alonso et al., 2005).

Epigenetic regulation is also involved in the regulation of O-antigen glucosylation (see LPS modification section). Indeed, under lysogenic conditions expression of the P22 encoded gtrABC operon is regulated by Dam methylation and the bacterial regulator OxyR (Broadbent et al., 2010; Davies et al., 2013). The region upstream of the gtr operon contains several OxyR binding sites as well as several methylation sites. Depending on the methylation state, OxyR bind different sites and can act as an activator or a repressor of the system. OxyR binding to one site decreases the methylation by Dam on this site and thus increases its own binding. But this works also the other way around: increased methylation results in reduced OxyR binding, which favors methylation, etc. This confers heritability of the expression state to the system and only a part of the population is in an “ON” state, leading to an heterogeneous population (Broadbent et al., 2010; García-Pastor et al., 2018). This regulatory mechanism is thought to be conserved among the P22 temperate phage family and can prevent superinfection by the same or other phages that use similar O-antigen co-receptor during a limited time (Davies et al., 2013).

Another factor involved in the silencing of genes acquired by HGT, in addition to core genes, is the DNA binding protein H-NS (Lucchini et al., 2006; Navarre et al., 2006). By preferentially binding to AT rich sequences, this protein can discriminate between self and non-self. Interestingly, several studies suggest that H-NS dependent regulation would involve different mechanisms for ancestral genes or genes acquired by HGT (Vivero et al., 2008; Baños et al., 2009). It has been suggested that ancestral genes would be regulated directly by H-NS binding whereas the genes acquired by HGT would require the formation of heterodimers between H-NS and proteins belonging to the Hha family. What characteristics of the promoter are required to favor the binding of homodimers or heterodimers are not known. In the same order of idea, H-NS proteins encoded by conjugative plasmids have evolved to specifically regulate foreign genes (Baños et al., 2009; 2011). This could be due to structural differences between plasmid-encoded or chromosomally encoded H-NS leading to different affinity for promoter regions. Indeed, although the N- and C-terminal domains are conserved, the linker region presents some variability that could be responsible for this differential regulation. All these observations concern genes acquired by HGT in general, including genes from phage origin. However, it has been noticed that the GC content of prophages in the reference strain of S. enterica LT2 is similar to the average GC content of the genome (Navarre et al., 2007). Thus, we can wonder if the conclusions made above really apply to genes from prophage origin. Studies focusing specifically on the regulation of these genes are missing so far and need to be performed to answer this question. Ongoing work in our lab suggests however that H-NS regulates prophage genes that have not been identified by global approaches in S. enterica ST4/74 and that these regulations have consequences on prophage maintenance in the host chromosome (Wahl et al., unpublished).

Positive host-control of prophage genes

In addition to the negative regulation that we have just mentioned, some bacterial regulators also positively regulate genes from prophage origin. Surprisingly, if one looks at the different transcriptomic studies performed in S. enterica to define the targets of global regulator such as PhoP, SlyA, ArcA, FNR or RpoS, only a handful of prophage genes were identified (Navarre et al., 2005; Fink et al., 2007; Evans et al., 2011; Lévi-Meyrueis et al., 2015). Furthermore, the molecular mechanism(s) leading to these regulations or the consequences on bacterial physiology are rarely looked at. Not surprisingly, what has been mostly studied is the regulation of genes coding for proteins involved in virulence and host colonization. These studies have shown that genes under the control of bacterial regulators are morons, which defines genes regulated independently from the rest of the prophage and conferring an advantage (fitness effect under specific conditions such as virulence) to the host. Among them, are several effectors proteins secreted by T3SS. As mentioned above, S. enterica possesses two T3SS encoded by the pathogenicity island 1 and 2 (SPI1 and 2). Among the regulators known to control the expression of genes located on the SPI are the InvF transcriptional regulator belonging to the AraC/XyIIS family for SPI1, itself regulated by the master regulator HilD, and the SsrAB two-component system for SPI2. Although both regulators were initially thought to be only dedicated to the regulation of genes located on SPI1 and 2, they also
regulate expression of prophage-encoded effectors. Indeed, InvF regulates effectors from prophage origin secreted by SPI1, whereas SsrAB controls effectors that are SPI2-dependent.

For example, SopE is a virulence factor encoded on the SopEΦ prophage in S. enterica SL1344 and secreted by the SPI1 system. Consequently, sopE expression is regulated by InvF, in association with the chaperon protein SicA (Darwin and Miller, 2000). The role of SicA is probably indirect, by stabilizing or allowing the function of a so far unidentified transcriptional regulator involved in sopE regulation (Tucker and Galán, 2000).

Other prophage-encoded effectors are secreted by SPI2, and as a consequence, their expression depends on the SsrAB two-component system. Among them, GogB is encoded by the first gene of the Gifsy1 prophage in S. enterica SL1344. Interestingly it has been shown that gogB expression is independent of Gifsy1 prophage factors since it can be transferred by itself in the enteropathogenic E. coli strain E2348/69, expressed from its own promoter and secreted via the T3SS of its new host. This shows that gogB can be integrated easily into the host regulatory network (Coombes et al., 2005). The GC content of gogB shows differences with the GC content of Gifsy1 suggesting that this gene has been recently acquired by the prophage. Moreover, gogB can be found outside of Gifsy1 and is not always prophage-encoded, which further supports its transcriptional independence from the Gifsy1 prophage (Porwollik et al., 2002).

ssel is a gene located on the Gifsy2 prophage, encoding another T3SS effector. ssel expression is strongly activated by the direct binding of the phosphorylated form of SsrB in its promoter region (Worley et al., 2000; Feng et al., 2004). ssel expression is also regulated by the phosphorylated form of OmpR but it is not clear whether this regulation is direct or dependent on SsrB (Feng et al., 2004). Interestingly, the pseudogenization of ssel together with the higher expression of pgtE, encoding an outer membrane protein, allows S. enterica ST313 adaptation to cause systemic disease (Carden et al., 2017; Hammarläf et al., 2018). The increase in pgtE transcription is due to a single SNP in its promoter region. Further studies are required to understand how these changes in gene expression modify S. enterica ST313 behavior (Hammarläf et al., 2018).

SsrB also regulates genes in the phage remnant SPI12. Among the regulated genes STM2239 encodes a Q antiterminator protein that interacts with the RNA polymerase to facilitate the transcription of late promoters. The absence of STM2239 affects the fitness of the bacteria within the host. STM2239 allows the transcription of phage-encoded genes but also of bacterial genes involved in metabolic pathways including ribose modification and transport, acetyl coenzyme A synthesis and recycling as well as galactose metabolism. None of these regulations have been characterized further, but it has been speculated that some of them may be important for S. enterica fitness within the host (Tomljenovic-Berube et al., 2013).

Phage-controlled bacterial genes

Except for virulence, examples of bacterial processes under prophage control are scarce (Fig. 1 and Table 1 subitem 1.2). However, P22 offers a nice illustration of bacterial genes encoding proteins involved in metabolism and under the control of a regulator from phage origin. The dgo operon is involved in the uptake and metabolism of D-galactonate, an important carbon source during intracellular proliferation. In S. enterica LT2 strain, expression of the dgo operon is derepressed in the presence of Pid, a protein encoded on a moron locus in P22 (Cenens et al., 2013a; 2013b). This regulation only occurs when P22 undergoes pseudolysogeny, suggesting the existence of a dedicated genetic program in this condition.

Phage-dependent regulation can be conserved among closely related bacteria. It is the case for the pckA gene encoding a phosphoenolpyruvate carboxykinase required for gluconeogenesis. In E. coli, this gene is under the control of the CI repressor of the λ phage (Chen et al., 2005). Interestingly, pckA regulatory region is conserved among Enterobacteriaceae and contains sequences homologs to several phage operators, one of them being the binding site for the C2 repressor of P22. This shared regulation between several prophages could be part of an adaptive strategy to increase lysogens fitness by lowering their growth rate under glucose-limited conditions (Chen et al., 2005).

Prophages often integrate into tRNA encoding genes. One counter example is given by phage ΦW104 that integrates the host chromosome at a locus encoding the RyeA and RyeB sRNA located on the opposite DNA strand in S. enterica DT104 (Balbontín et al., 2008). The attachment site for ΦW104 is within the 23 last base pairs of ryeB and corresponds to an internal site in ryeA. Therefore, ΦW104 lysogenization modifies the 5’ portion of ryeA, leading to a decreased transcription of ryeA and an increased transcription of ryeB. This transcriptional regulation has probable physiological consequences on the bacterial host, by modifying the expression of RyeA and RyeB mRNA targets.

Finally, another example of host-gene regulation involves a gene located on the Gifsy1 prophage, isrK (Hershko-Shalev et al., 2016). This gene encodes an sRNA and a long polycistrionic mRNA comprising isrK, orf45, anrP and isrJ coding sequences. However, there is no translation observed from this mRNA unless IsrK sRNA is present. Indeed, IsrK sRNA binds next to the orf45 ribosome binding site and facilitates the binding
of the 30S ribosomal subunit to this site, leading to the translation of downstream sequences and the production of the AnrP protein. AnrP is an anti-repressor activating the transcription of phage-encoded genes. Among them, it activates the expression of the antQ gene coding for the AntQ anti-terminator that interacts and forms a stable complex with RNA polymerase. This leads to an aberrant transcriptional elongation, DNA damage and ultimately cell death (Hershko-Shalev et al., 2016).

All the above examples concern regulation of gene expression. However, phages can also influence bacterial physiology by other means. For example, the release of colicin 1b in S. enterica SL1344 depends on the lysis genes of the ST64B prophage (Nedialkova et al., 2015). Indeed, under specific conditions such as DNA damage or iron limitation, colicin 1b accumulates in the cell and needs the induction of ST64B lysis genes to be found in the extracellular medium. Interestingly, complex crosstalk between ST64B and other prophages present in that strain contributes to this regulation and need further characterization.

Conclusions

What is striking whenever considering and comparing different Salmonella genomes is the diversity of the prophage content as well as the diverse relationships these prophages engage with the host strains. As stated before, we think that contributions to virulence have been more studied and highlighted up to now than regulatory and metabolic interactions between the Salmonella host and its prophages. This bias seems largely due to the prevalent role of Salmonella species in public health threats. In addition, contrary to some previous statements, phage genomes rarely contain antibiotic resistance genes, and if antibiotic resistance transfer can be sometimes attributed to phages, it is more likely due to generalized transduction rather than lysogeny (Enault et al., 2017). Nevertheless, a new category of self-transferable plasmid-phages could change this view as some of them carry ATB resistance genes (SSU5 super-cluster) (Gilcrease and Casjens, 2018).

Apart from actual contributions to virulence, we foresee that many more interactions do exist, and, it seems that microbiologists have only just began to explore the expense of prophage–host interactions and their short and long-term effects on bacterial metabolism and evolution. More interactions certainly remain to be discovered such as the up-to-now neglected carrier state and its implications on the host metabolism (Cenens et al., 2015).

In a context of multidrug resistance spreading and recurrent warnings from WHO and other health authorities, the use of bacteriophages as therapeutic agents is coming back to the scene, not only as potent antimicrobials by themselves but also as synergistic or complementary agents in combination with antibiotics (Kamal and Dennis, 2015; Abedon, 2018). However, even though only virulent (or strictly lytic) phages are considered for therapeutic usage, one must be aware of the possible interference of prophages whenever considering phages as a treatment. As described above, prophages are important contributors of serotype conversion, and particularly in Salmonella species. The literature becomes quite abundant regarding Salmonella phage quests, but little is known about the consequences of poly-lysogeny, which can rapidly modify the bacterial surface, and therefore, the resistance to surinfesting phages on the efficacy of phage cocktails, particularly for those developed to treat swine and poultry or in the case of adjuvant in food processing (Wernicki et al., 2017). We suggest to systematically assessing the prophage content of the targeted strains to evaluate and adapt the composition of phage cocktails.

Acknowledgements

Work in our group is supported by the French Research National Agency grants ANR-16-CE12-0029-01 and ANR-17-CE11-0038-02, and institutional funding from the Centre National de la Recherche Scientifique and the Aix-Marseille Université. We are grateful to S. Champ, N. Ginet and the whole phage group at LCB for stimulating discussions and suggestions. We thank M. Touchon and E. Rocha for Salmonella genome analysis and to J.R. Fantino for his help in figure design. M.A. also thanks J. Casadesús who provided an inspiring environment while part of this review was written. Although literature citations have not been limited, we may have omitted some pieces of the abundant literature to be more focused, we apologize to those authors whose research is not included.

References


Brown, N.F., Coombes, B.K., Bishop, J.L., Wickham, M.E., Lowden, M.J., Gal-Mor, O. et al. (2011) *Salmonella* phage ST64B encodes a member of the SseK/NleB effector family. *PloS One, 6*, e17824.


foreign DNA with low GC content by the H-NS protein in Salmonella. Science, 313, 236–238.


