Phage as agents of lateral gene transfer
Carlos Canchaya, Ghislain Fournous, Sandra Chibani-Chennoufi, Marie-Lise Dillmann and Harald Brüssow*

When establishing lysogeny, temperate phages integrate their genome as a prophage into the bacterial chromosome. Prophages thus constitute in many bacteria a substantial part of laterally acquired DNA. Some prophages contribute lysogenic conversion genes that are of selective advantage to the bacterial host. Occasionally, phages are also involved in the lateral transfer of other mobile DNA elements or bacterial DNA. Recent advances in the field of genomics have revealed a major impact by phages on bacterial chromosome evolution.

Addresses
Nestlé Research Centre, CH-1000 Lausanne 26, Vers-chez-les-Blanc, Switzerland
*e-mail: harald.bruessow@rdls.nestle.com

Lateral gene transfer
With about 100 sequenced genomes of bacteria in the public database and many more to come, genomics has changed our understanding of microbiology. In fact, the genomes of bacteria are remarkably fluid. A substantial part of the bacterial DNA is not transferred from the parental cell to its descendent (‘vertical’ transfer), but is acquired horizontally by transformation, conjugation or transduction (‘lateral’ transfer) [1**]. The replacement of a tree-like by a web-like representation of the phylogenetic relationship between bacteria is a visual expression of this change in perception of microbial evolution. An important element of mobile DNA is bacteriophages. Infection of a bacterial cell with a temperate phage (Figure 1a,b) can have two outcomes: multiplication of the phage with concomitant lysis of the bacterial host (Figure 1c) or lysogenization, (i.e. integration of the phage DNA into the bacterial chromosome as a prophage, Figure 1d). Bacterial genomics revealed that lysogeny is more the rule than the exception; many bacteria even contain multiple prophages (Figure 2a). Some temperate phages carry in their genomes extra genes that change the phenotype of the bacterial host (‘lysogenic conversion genes’, LCG) (Figure 2b). There is increasing evidence from bacterial pathogens that lysogeny is a motor of short-term bacterial evolution.

Phages as gene-transfer particles
Tailed phages are the most efficient gene-transfer particles developed in evolution. They represent densely compacted phage DNA [2] encased in a protective protein shell (the phage head) [3]. To this remarkable DNA storage device is added an equally efficient DNA transfer device, the phage tail and its associated fibres (Figure 1a). This structure assures both the specific recognition of the appropriate host cell and the guided injection of the phage DNA into the bacterial cell ([4,5], Figure 1b).

Some bacteria have learned to use phages for their own purposes. In Pseudomonas aeruginosa, two phage-tail gene-clusters have developed into bacteriocins [6]. The defective Bacillus subtilis prophage PBSX has maintained the capacity to build a size-reduced phage head into which 13 kb fragments of random bacterial DNA are packaged. A prophage remnant of Rhodobacter capsulatus acts as a gene-transfer agent for random 4.5 kb fragments of bacterial DNA in bacteria-controlled DNA exchange between cells in the stationary phase [7]. Prophage-like elements from Mycobacterium tuberculosis encode active integration/excision systems [8].
virulent phages. Prophage protein protects the lysogenic cell against superinfection with bacteria. Infection results in the multiplication of the phage and the lysis of the cell. EM thin section shows how phage Sfi21 adsorbs to its bacterial host, its head, a non-contractile tail and a single tail fibre are clearly visible.

A particularly interesting case is the 15 kb-long pathogenicity island SaPI1 from Staphylococcus aureus encoding the toxin Tst involved in toxic shock. In cells infected with S. aureus phage 80x, SaPI1 is excised from the chromosome, it replicates autonomously and interferes with phage growth by directing the encapsidation of its own DNA into specially tailored small phage 80x heads commensurate with its size. Upon phage-mediated transfer to a recipient organism, SaPI1 integrates by means of its own integrase [9**].

**Specialised transduction**

Resolvase-type integrases from phages of Gram-positive bacteria have no requirements for cofactors facilitating their integration into heterologous hosts [10]. If a prophage is imprecisely excised from the heterologous host, small segments of flanking bacterial DNA can be copackaged with the phage DNA and transferred to the original host (‘specialized transduction’). In accordance with this model, prophages from low GC content Gram-positive bacteria frequently contain extra genes in the vicinity of attR, the right attachment site (Figure 2b). Sometimes these genes differ in GC-content from the surrounding DNA and suggest a phage-mediated gene transfer from a rare heterologous host differing in GC content ([11], Figure 3a). In the case of pathogenic bacteria, these extra genes frequently encoded important virulence factors like bacterial toxins [12*,13*,14**] (Figure 2b). These extra genes were also observed in commensals and free-living bacteria and belonged to the few prophage genes expressed in the lysogenic state ([15], Figure 3b); only in a few cases did database matches suggest a physiological role for these extra genes (Figure 3c).

LCG were also identified in prophages from Gram-negative bacteria (Figure 4a). Some of them were located at the prophage genome ends (e.g. O serotype-converting enzymes were found near attL, the left attachment site, in several prophages [16,17]). However, the majority of the extra genes or ‘morons’ (for more DNA) were detected in the centre of the prophage genomes. Preferred insertion sites for LCG were located downstream of the Q anti-terminator, the lysis and the N antiterminator genes [18]. They tend to represent transcription units with their own promoters and terminators that are regulated independently from the rest of the prophage [19,20]. Some of the LCG were shown to respond to environmental cues [21,22]. In fact, when bacteria were grown under conditions that mimicked pathological conditions [23], or when they were grown in infected animals [24], prophage genes belonged to the most prominent genes of the entire bacterial chromosome that changed the expression level.

**Generalised transduction**

Phages such as Salmonella phage P22 or coliphage Mu occasionally commit the error to package even a headfull of bacterial DNA instead of phage DNA. Upon infection of the next host, this bacterial DNA can be incorporated into the bacterial chromosome (‘generalised transduction’). Despite the interest in gene flux in the environment, sparked by the discussion of the risks associated with the release of genetically modified microorganisms, only a few recent reports have investigated generalised transducing phages in terrestrial habitats (in Streptomyces and Listeria) [25,26]. One technical report addressed the problem of PCR-detection of phage-encapsidated bacterial DNA when working with uncultivable bacteria and their phages [27].

By contrast, phage ecology and phage-mediated DNA transfer became a focus in marine microbiology [28]. Researchers realised that viruses (most of them probably...
Phages from *Streptococcus pyogenes* encode many potential virulence factors. (a) Prophages are visualised as red boxes on the circular genome maps of four sequenced *S. pyogenes* strains representing three different M types. (b) Partial gene maps of the indicated *S. pyogenes* prophages covering the genome region between the lysis module (violet) to the right attachment site attR. Grey arrows represent genes of undetermined function. The prophages are noted in clockwise order as they appear in the indicated genome. Candidate lysogenic conversion genes are marked in red and are annotated: mf, mitogenic factors; sdn, streptodornase; sla, streptococcal phospholipase A2, spe, streptococcal pyrogenic exotoxins; ssa, streptococcal superantigen.

Prophages from *Streptococcus pyogenes* encode many potential virulence factors. (a) Prophages are visualised as red boxes on the circular genome maps of four sequenced *S. pyogenes* strains representing three different M types. (b) Partial gene maps of the indicated *S. pyogenes* prophages covering the genome region between the lysis module (violet) to the right attachment site attR. Grey arrows represent genes of undetermined function. The prophages are noted in clockwise order as they appear in the indicated genome. Candidate lysogenic conversion genes are marked in red and are annotated: mf, mitogenic factors; sdn, streptodornase; sla, streptococcal phospholipase A2, spe, streptococcal pyrogenic exotoxins; ssa, streptococcal superantigen.

Phages outnumber bacteria in the open ocean by a factor of ten [29]. In view of the large volume of the world’s oceans and the high titre of phage particles of $10^7$/ml of seawater, phages particles are the most abundant biological entities on earth [30]. If one anticipates a transduction frequency of $10^{-8}$/PFU per plaque forming unit for marine phages [31], it was calculated that phage-mediated gene transfer takes place at the incredible rate of about 20 million billion times per second in the oceans [1**]. However, the genomics of the predominant marine bacteria and their phages is still in its infancy. Only a handful of marine phages have been sequenced [32] from the 400–7000 viral types estimated in 100 litre water samples [33].

**New insights from prophage genomics**

The impact of phage-mediated lateral gene transfer can easily be read from the published bacterial genome
sequences [34*]: two-thirds of the sequenced low GC Gram-positive bacteria and $\gamma$-Proteobacteria (Gram-negative bacteria) contained identifiable prophages. Many bacteria were polylysogenic (contained multiple prophages, Figure 2a). Prophage DNA represented up to 16% of the chromosomal DNA (Escherichia coli O157 strain Sakai with 18 prophages) [35*]. Theoretical reasoning on the basis of Darwinian evolution predicted aspects of an arms race and of mutualism in the genetic interaction of phage and bacterial genomes [36*,37*]. Cooperation (mutualism) was demonstrated by the observation of many virulence factors encoded by prophages from pathogens [18], including prominent examples such as the cholera toxin from Vibrio cholerae or the shiga-like toxin from enterohaemorrhagic E. coli (see also Update). The arms race aspect of prophage genomics was also documented: most prophages from sequenced bacterial genomes showed inactivating point mutations, inactivating DNA insertion (often transposases) or progressive DNA deletion leading to defective prophages, prophage remnants and isolated prophage genes in bacterial genomes. A recurring observation was isolated phage integrase genes in bacterial genomes suggesting that these phage recombination genes involved in lateral gene transfer are of selective value to the bacterial host. Notably, several pathogenicity islands were flanked by direct repeats, the presence of phage integrase and integration into tRNA genes [38]. It is tempting to speculate that some pathogenicity islands have recruited the integration system from decaying prophages to achieve mobility.
Selected examples: *Streptococcus*, *Salmonella* and *Xylella*

*Streptococcus pyogenes* strains belonging to different M serotypes and associated with different pathologies (M1, wound infection; M3, toxic shock; M18, rheumatic fever) were closely related at the DNA sequence level [11,13**,14**]. The alignment of the genomes showed only a few gaps and, remarkably, prophages accounted for the majority of the variation in gene content between the strains. The two M3 strains differed from each other only by genome translocations and inversions. Some of them were flanked by prophages. The *S. pyogenes* prophages belonged to Sfi21-, Sfi11- and r1t-like Siphoviridae that represent the majority of prophages currently described in low GC-content Gram-positive bacteria [37*]. It was proposed that the recent emergence of highly virulent strains was the result of the sequential acquisition of three prophages with their specific LCG (superantigens, toxins and secreted enzymes) over the past decades. Microarray hybridization showed that prophage DNA represented a major part of the *S. pyogenes* strain-specific DNA.

*Salmonella enterica* serovars Typhimurium and Typhi share a closely related genome, although they differ substantially with respect to their prophages (Figure 4). Most of the larger gaps in the alignment of the two chromosomes were explained by prophage insertion. Microarray analysis [39] and further genome sequences
demonstrated that prophages are also a major contributor to the genetic differences between Salmonella strains belonging to the same serovar. The Typhimurium prophages belonged to the P2-like genus of Myoviridae and the lambda-like genus of Siphoviridae (Figure 4a). The Typhi prophages were distant lambda relatives and also a hybrid Mu/P2 prophage was observed (Figure 4b). Animal experiments with Salmonella deletion mutants demonstrated that prophages are not ephemeral selfish DNA that litters the bacterial chromosome, but contributors of numerous virulence factors (Figure 4a). In fact, the variable assortment of prophages was interpreted as a transferable repertoire of pathogenicity determinants in Salmonella [41].

The alignment of the genomes from two different pathovars of the plant pathogen Xylella fastidiosa revealed three chromosomal regions that were translocated and inverted, but otherwise they shared 98 per cent of the genes [42**]. The Temecula strain contained six prophages none of which shared sequence relatedness with the five prophages from the 9a5c strain. One prophage in each strain resembled filamentous phages. Prophage genes were again the major contributor to the strain-specific genes. The three chromosomal rearrangements were all flanked at one border by a phage integrase gene (see also Update).

**Lateral gene transfer between phages**

Virulence genes were apparently transferred between phages belonging to different phage groups [43] or infecting different bacterial species [37*] thereby increasing the lateral spread of these genes in bacteria. Sequencing data from coliphages and dairy phages also demonstrated that large phage gene clusters were transferred between distinct groups of phages [44] confirming tenets of the classical modular theory of phage evolution. The strikingly different GC-content of the left and right arm of phage lambda suggests the heterologous origin of this reference phage (Figure 4a). The mosaic character of phages was greater in Gram-negative than in Gram-positive bacteria [45]. In some lambdoid coliphages short conserved sequences were identified at the boundaries of functional modules. This suggested homologous recombination as the driving force for lateral gene transfer between phages [46]. However, the comparison of other lambdoid coliphage genomes suggested that non-homologous recombination occurs everywhere and the observed order in phage genome organisation is the consequence of selection forces eliminating all non-viable recombinants [47]. Recent sequencing data identified hybrids between phage genera (Figure 4b), phage families [17,48] and even temperate and virulent phages [34*]. This abundant lateral gene transfer between previously well-defined phage groups now poses a major dilemma for phage taxonomy and ideas on phage evolution [49].

**Conclusions and outlook**

Prophages contribute a substantial share of the mobile DNA of their bacterial hosts and seem to influence the short-term evolution of pathogenic bacteria. Automated methods for systematic investigation of prophages and other mobile DNA elements in the available 100 bacterial genome sequences will be necessary to understand their role in bacterial genome evolution. In the past, prophages were mainly investigated as the simplest model systems in molecular biology. Now it is increasingly realised that phage research will be instrumental in the understanding of bacterial abundance in the environment. One can predict that phage research will impact diverse areas such as geochemistry and medicine. Success will largely depend on integrative multidisciplinary approaches in a field that has, until recently, been dominated by reductionist thinking.

**Update**

Recent work has demonstrated that a prophage from a Lactobacillus oral-cavity commensal contains candidate lysogenic conversion genes near both prophage genome ends which are sequence-related to mf2 and mf4 from S. pyogenes prophages (see Figure 3b) [50]. In addition, microarray analysis demonstrated that 50% of the strain-specific DNA from Lactobacillus gut commensal is represented by prophage DNA [51].

In E. coli O157, induction of the prophage is required for toxin synthesis and release. Toxin synthesis is secondarily amplified by phage infection of non-toxigenic intestinal E. coli commensals, representing a new strategy of bacterial pathogenesis [52].

**Acknowledgements**

We thank the Swiss National Science Foundation for providing financial support for Carlos Canchaya and Sandra Chibani-Chennoufi (research grant 5002-057832). The paper is dedicated to Margret Beck-Brüssow for her patient and graceful support of the senior author.

**References and recommended reading**

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- or outstanding interest


Phage as agents of lateral gene transfer Canchaya et al. 423


A fascinating story of the genetic interaction between a pathogenicity island and a superinfecting phage. This remarkable article blurs the distinction between two different forms of mobile DNA (phage and genomics island). In fact, in this particular case evolution uses the phage for the benefit of the bacterial cell as the phage helps to disseminate the pathogenicity island DNA between different Staphylococcus aureus cells. The relationship between the superinfecting phage and the resident pathogenicity island somewhat resembles that of coliphage P2 and its satellite P4.


The sequences of a community- and hospital-acquired meticillin-resistant Staphylococcus aureus strain were compared. The combination of different allelic forms of genomic islands and prophages provide the genetic basis that determines the pathogenicity of this important human pathogen.


Phage-like elements and insertion sequences were the major sources of variation between the genomes from Streptococcus pyogenes strains involved in two distinct pathologies, wound infections and rheumatic fever. The prophages encode several secreted proteins involved in human–bacterium interaction, including the scarlet fever toxin. Phage DNA also represents the major source of genetic variability between S. pyogenes isolates belonging to the same M type.


Phages act as the vector for lateral gene transfer between bacterial genomes, in this case of O-serotype modifying enzymes. Horizontal gene transfer also occurs between two or more phages as demonstrated by a chimera phage genome combining structural genes from two different families of tailed phages, Siphoviridae and Myoviridae.


A demonstration that prophage DNA is the major source of DNA sequence diversification of bacterial genomes was presented by Desiere et al. – of DNA by a non-specific process. The prophage encodes the shiga-like toxin that is responsible for the clinical symptomatology of this enterohemorrhagic E. coli strain. Lambda-like prophages were apparently propagated by a copy-paste-like mechanism within the bacterial genome, resulting in a bacterial genome containing 18 prophages.


A stimulating theoretical argument that the constant influx of foreign DNA into bacterial genomes (including prophages) is balanced by a constant loss of DNA by a non-specific DNA deletion process. The reasoning leads to similar predictions to the theoretical arguments on phage–bacterium interaction, presented by Desiere et al. (2001) [37].

37. Desiere F, McShan WM, van Sinderen D, Ferretti JJ, Brüssow H: Comparative genomics reveals close genetic relationships between phages from dairy bacteria and pathogenic Streptococci: evolutionary implications for prophage–host interactions. Virology 2001, 283:325-341. On the basis of simple Darwinian reasoning, prophage–bacterium genome interaction was predicted to reflect elements of a genetic war between a parasite and its prey in the form of an arms race where measures of one protagonist are answered by counter-measures of the other protagonist. At the same time a certain degree of mutualism was predicted in this relationship, where the expression of lysogenic conversion genes from prophages was postulated to increase the fitness of the lysogenic bacterium.


40. Deng W, Liou SR, Plunkett G III, Mayhew GF, Rose DJ, Burland V, Kodoyianni V, Schwartz DC, Blattner FR: Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18. J Bacteriol 2003, 185:2330-2337. Prophage genes contribute importantly to the isolate-specific gene complement when the genomes from two Salmonella enterica strains belonging to the same Typhi serovar were compared.


42. Van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, Camargo LE, da Silva AC, Moon DH, Takita MA, Lemos EG et al.: Comparative analyses of the complete genome sequences of Pierce’s disease and citrus variegated chlorosis strains of Xylella fastidiosa. J Bacteriol 2003, 185:1018-1026. Two strains of Xylella fastidiosa, which are linked to different plant pathologies, share 98% of their genes. Genomic differences are limited to phage-associated genome rearrangements and deletions. Prophages also accounted for a substantial part of the strain-specific DNA. There is now increasing evidence (e.g. with the two M3 S. pyogenes strains) that prophages affect the overall chromosomal architecture of bacterial chromosomes and so are subsequently more flexible than anticipated.


