The ISME Journal (2017) 11, 1988–1996
www.nature.com/ismej

OPEN

ORIGINAL ARTICLE

Comparative transcriptomics analyses reveal the conservation of an ancestral infectious strategy in two bacteriophage genera

Bob G Blasdel[1,5], Anne Chevallereau[2,3,5], Marc Monot[4], Rob Lavigne[1] and Laurent Debarbieux[2]

[1]*Laboratory of Gene Technology, Department of Biosystems, Leuven, Belgium;* [2]*Institut Pasteur, Molecular Biology of the Gene in Extremophiles Unit, Department of Microbiology, Paris, France;* [3]*Université Paris Diderot, Sorbonne Paris Cité, Cellule Pasteur, Paris, France and* [4]*Institut Pasteur, Laboratoire Pathogenèse des bactéries anaérobies, Département de Microbiologie, Paris, France*

Although the evolution of tailed bacteriophages has increasingly been better understood through comparisons of their DNA sequences, the functional consequences of this evolution on phage infectious strategies have remained unresolved. In this study, we comprehensively compared the transcriptional strategies of two related myoviruses, PAK_P3 and PAK_P4, infecting the same *Pseudomonas aeruginosa* host strain. Outside of the conservation of their structural clusters, their highly syntenic genomes display only limited DNA similarity. Despite this apparent divergence, we found that both viruses follow a similar infection scheme, relying on a temporal regulation of their gene expression, likely involving the use of antisense transcripts, as well as a rapid degradation of 90% of the host non-ribosomal mRNA, as previously reported for PAK_P3. However, the kinetics of the mRNA degradation is remarkably faster during PAK_P4 infection. Moreover, we found that each virus has evolved specific adaptations, as exemplified by the distinct patterns of their core genes expression as well as the specific manipulation of the expression of iron-related host genes by PAK_P4. This study enhances our understanding of the evolutionary process of virulent phages, which relies on adjusting globally conserved ancestral infection mechanisms.

The ISME Journal (2017) **11,** 1988–1996; doi:10.1038/ismej.2017.63; published online 12 May 2017

Introduction

Viruses that infect bacteria, also known as bacteriophages (phages), have crucial roles in a variety of environments including the renewal of bacterial populations, recycling biochemical resources and horizontal gene transfer, as reflected by their abundance and diversity. Multiple comparative genomic analyses have suggested that related phage genomes can be clustered based on their sequence similarity (Kwan *et al.*, 2005; Kropinski *et al.*, 2007; Marinelli *et al.*, 2012; Grose and Casjens, 2014). However, several studies have shown that bacteria and their phages rapidly coevolve, promoting genetic diversity in both populations (Marston *et al.*, 2012; Williams, 2013; Koskella and Brockhurst, 2014). This genetic diversity is further stimulated by horizontal gene transfer (HGT) events, which in some cases can make it difficult to reconstruct evolutionary history of phages. Consequently, many studies have aimed at understanding how different phage clusters are related to each other or how phage populations are structured (based on comparative (meta)genomics) (Hatfull *et al.*, 2006; Kwan *et al.*, 2006; Grose and Casjens, 2014; Pope *et al.*, 2015; Roux *et al.*, 2015).

Apart from host-range and defense systems evolution, no mechanistic studies have yet experimentally addressed the evolution of phage infection strategies themselves, leaving unresolved questions on the consequences of divergent evolution of phage genomic sequences on the functions of those sequences. In particular, it has not yet been directly investigated how conserved the infection strategy is during the evolution of virulent phages, which can only be weakly predicted using bioinformatics tools. Following the isolation and genomic characterization of several phages infecting the same *Pseudomonas aeruginosa* host strain, an opportunistic pathogen widely present in the environment, we previously found that, although their morphologies were nearly

Correspondence: R Lavigne, Laboratory of Gene Technology, Department of Biosystems, KU Leuven Kasteelpark Arenberg 21, Box 2462, Heverlee 3001, Belgium.
E-mail: rob.lavigne@kuleuven.be
or L Debarbieux, Unité Biologie Moléculaire du Gène chez les Extrêmophiles, Département de Microbiologie. Institut Pasteur, 25 rue du Dr Roux, 75015 Paris, France.
E-mail: laurent.debarbieux@pasteur.fr
[5]These authors contributed equally to this work.
Received 25 October 2016; revised 22 February 2017; accepted 22 March 2017; published online 12 May 2017

identical, their genomic contents were divergent enough to support their classification within two new genera among *Myoviridae*, namely *Kpp10virus* and *Pakpunavirus* (Henry *et al.*, 2015; Krupovic *et al.*, 2016). Moreover, an additional study analyzing phages belonging to these genera pointed out that their genomes are typically composed of strongly conserved regions (0.1% single-nucleotide polymorphism (SNP) frequency) alternating with heterogeneous regions (up to 20% SNP frequency, scars of HGT and recombination events; Essoh *et al.*, 2015). Overall, our previous comparative genomic analysis revealed that phages belonging to these distinct genera differ by their mean genome length, their number of predicted open reading frames (ORFs) and tRNA, and their GC content, but they also share homologous proteins (about 25%), most of which encoding structural components (Henry *et al.*, 2015). Despite this significant genomic divergence, they display a strong synteny and the reconstruction of their phylogeny, based on four conserved proteins, revealed that they likely share a relatively recent common ancestor (Henry *et al.*, 2015).

By selecting PAK_P3 and PAK_P4 as representatives of *Kpp10virus* and *Pakpunavirus* genera respectively, we aimed to investigate whether the transcriptional strategy is conserved during evolution of phages by performing a transcriptomic-based comparative analysis (RNA-sequencing). Our investigation was driven by the following questions. Do phages that have a common origin and subsequently diverge, still share common mechanisms to infect their host? How similar are the host transcriptional responses to infections by two different phages?

We found that the global infection process is conserved (temporal regulation of gene expression, production of antisense transcripts, rapid degradation of host mRNA) with few phage specific tweaks (specific manipulation of host gene expression, custom patterns of phage core gene expression). Altogether, these results suggest that, despite genomic divergence, phages are bound to their ancestral mode of action.

Materials and methods

Strains and growth conditions

P. aeruginosa strain PAK (Takeya and Amako, 1966) was grown in LB medium supplemented with 10 mM $CaCl_2$ at 37 °C unless stated otherwise. For RNA-seq experiments, cells were infected with bacteriophage PAK_P4 (Henry *et al.*, 2013) using a multiplicity of infection (MOI) of 25 to ensure the synchronicity of the infection (95% of the bacterial population killed after 5 min phage-introduction). Bacteriophage PAK_P3 was isolated and described by Morello *et al.* (2011). Updated phage genome sequences are available in Genbank and can be accessed via accession numbers NC_022970 (PAK_P3), NC_022986 (PAK_P4). Updated annotations of strain PAK genome are available via GenBank accession number CP020659.

Adsorption assay and one-step growth experiment

PAK_P4 adsorption assays (MOI 10^{-3}) and one-step growth experiments (MOI 10^{-1}) were performed as described in Chevallereau *et al.* (2016). Briefly, a liquid culture of strain PAK was infected at indicated MOI and incubated at 37 °C during 5 min with agitation to allow phage particles to adsorb. The culture was diluted (1/1000) and two 100 µl samples were collected every 2 min and either kept on ice until titration, or mixed with $CHCl_3$. For each sample, free phage titer (samples with $CHCl_3$) as well as the number of free phages and infective centers (samples without $CHCl_3$) were recorded to calculate eclipse and latency periods, respectively.

Dependence to host RNA polymerase assay

The experiment was performed according to the methods described by Ceyssens *et al.* (2014). The strain PAK was grown in LB medium until it reached OD_{600nm} ~0.3, then the culture was split with one half supplemented with 100 µg ml^{-1} rifampicin (Rif) and second half with the same volume of water. Ten minutes later, phages were added at MOI of 1. Samples were taken immediately ($t = 0$), 15, 60 and 150 min post infection and phages were titered.

Whole transcriptome sequencing and analysis

RNA-sequencing analysis was performed as described in Chevallereau *et al.* (2016). Over 12 million 75 bp reads mapping to non-ribosomal regions were obtained from each library with the exception of one early sample (2.4 million reads). After trimming, sequencing reads were aligned separately to both the phage and host genomes with the CLC Genomics workbench v7.5.1 (QIAGEN Bioinformatics, Aarhus, Denmark) These alignments were then summarized into count tables of Total Gene Reads that map to phage or host gene features respectively. Each statistical comparison presented was performed using the DESeq2 (Love *et al.*, 2014) R/Bioconductor package to normalize host transcript populations to host transcript populations, or phage to phage, before testing for differential expression.

RNA-seq data have been deposited in NCBI-GEO with accession no. GSE86022.

RNA-seq coverage visualization is available through the COV2HTML software at https://mmonot.eu/COV2HTML/visualisation.php?str_id=-36 for a comparison of the host (0 min/13 min) and https://mmonot.eu/COV2HTML/visualisation.php?str_id=-38 for a comparison of the phage (3.5 min/13 min).

Results

PAK_P3 and PAK_P4 share common life-history traits

To compare the molecular mechanisms of infection used by virulent phages that recently diverged, we chose two phages infecting the same *P. aeruginosa* host, strain PAK. These two myoviruses, PAK_P3 and

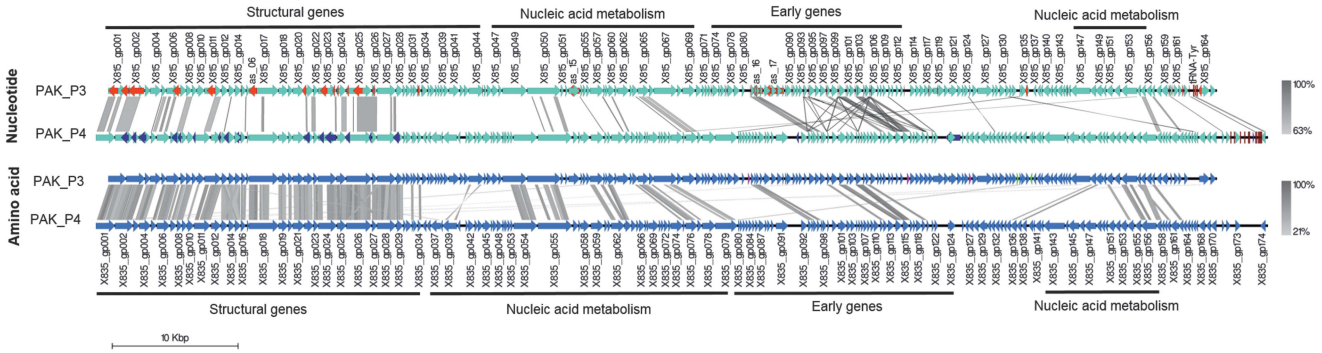

Figure 1 Genomic comparison of phages PAK_P3 and PAK_P4. Upper panel: nucleotide comparison. ORFs are depicted as cyan arrows, antisense RNAs are represented as red and dark blue arrows, and small RNAs (ncRNA, tRNA) appear as red squares. Lower panel: translated nucleotide comparison. Coding sequences are colored in blue.

PAK_P4, respectively belong to the newly defined genera *Kpp10virus* and *Pakpunavirus* (Henry *et al.*, 2015) and display distinct host ranges (Saussereau *et al.*, 2014). PAK_P4 shares an average 76% (±9%) identity over 13% of PAK_P3 genome, in other words, only 9% of the nucleotide sequence is identical between the two phages (Supplementary Table 1; Figure 1). Sixty-nine homologous coding sequences, hereafter referred as core genes, have been defined (that is, ~40% of the proteomic content of the two phages, according to CoreGenes analysis with default parameters (Mahadevan *et al.*, 2009)). The sequence identity for each pair of homologous predicted proteins ranges from 24 to 85% (with a mean of 49 ± 14%). Most of these homologous proteins belong to the structural (49%) and nucleic acid metabolism (32%) modules, while 19% of them are encoded by early-expressed ORFs (defined in the third paragraph) which are generally considered to be the least conserved phage proteins, often responsible for the subversion of host metabolism (Calendar, 2006) (Supplementary Table 1; Figure 1). It has to be noted that among the coding sequences that do not meet the criterions required to be considered as homologous (that is, blastp score <75), some appear to have the same predicted functions. For example, the putative ribonucleoside-reductase beta subunits (PAK_P3_gp069 and PAK_P4_gp078) share only ~30% of similarity (~17% of identity) in their protein sequences. Nevertheless, most of the predicted ORFs could not be associated with a putative function.

Altogether, these observations raise the question of whether the infectious mechanisms and strategies used by PAK_P3 and PAK_P4 are conserved, given their divergent evolution. To answer this, we first compared their growth parameters. PAK_P4 adsorbs onto the strain PAK within 4.8 ± 1.7 min ($k_a = 2.4 \times 10^{-9} \pm 1.4 \times 10^{-9}$ ml min^{-1}) and rapidly assembles new functional virions (eclipse period: 13.3 ± 2.0 min), although they are not immediately released (latency period: 18.2 ± 2.9 min). PAK_P4 infected cells eventually produce an average of 13 ± 5 progeny phages within a mean infection cycle duration as short as 21.4 ± 1.8 min (Supplementary Figure 1). In comparison with PAK_P3, PAK_P4 displays a smaller burst size and a longer latency period (Chevallereau *et al.*, 2016). These results hint that, if both phages have largely similar infection strategies (as shown below), they may have evolved unique processes for enacting them that may be variably effective. To investigate this further, we studied the transcriptomic dynamics of both phages.

PAK_P4 and PAK_P3 rapidly overwhelm the transcriptional environment of the host by efficiently hijacking bacterial RNA polymerase

In a previous study, we investigated the dynamics of host strain PAK transcriptome throughout PAK_P3 infection focusing on early (3.5 min) and late (13 min) stages of infection (Chevallereau *et al.*, 2016). As PAK_P4 has a comparable eclipse period as PAK_P3, we performed a transcriptomic analysis using the same time points. While PAK_P3 induces a progressive depletion of host transcripts that eventually represent only 13% of non-ribosomal RNAs in the cell 13 min post infection, PAK_P4 is much more rapid as only 20% of the total transcripts matches the host genome as soon as 3.5 min post infection (Figure 2). Both phages shut down host cell transcription but following different kinetics. We hypothesize that they may have a different control of the machineries dedicated to (i) degrade host mRNAs and/or (ii) catalyze the transcription of viral genes.

The rapid and overwhelming replacement of host transcripts with viral transcripts could indicate that both phages rely on their own transcription apparatus even though they are not predicted to encode a viral RNA polymerase (RNAP) based on sequence homology searches. Therefore, we experimentally assessed the dependency of PAK_P3 and PAK_P4 infections on host RNAP using rifampicin (Rif), which is known to inactivate this enzyme. We found that addition of Rif abolished PAK_P3 and PAK_P4 amplification while it did not affect phage PhiKZ transcription, as previously shown (Supplementary Figure 2; Ceyssens *et al.*, 2014). This indicates that both phages appear to

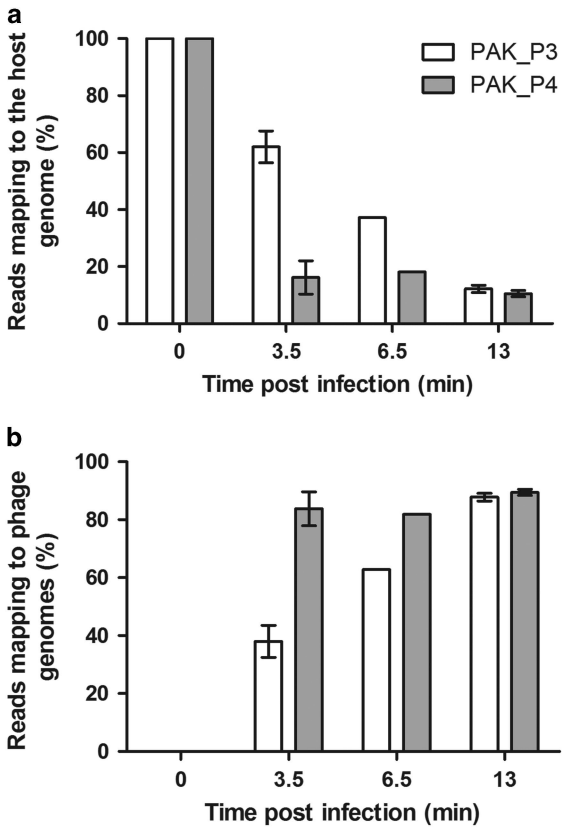

Figure 2 PAK_P4 takes over the host cell transcription, faster than PAK_P3, over the course of infection. Three independent biological replicates of RNA extracts were harvested from PAK_P4 and PAK_P3 infected cells at 0 min, 3.5 min and 13 min post infection, as well as a single sample collected at 6.5 min and all were subsequently sequenced. Percentages of reads mapping to phage genomes (**a**) and to the host genome (**b**) over the course of PAK_P4 infection (gray bars) or PAK_P3 infection (white bars) are plotted. The 6.5 min time point was not included in further analyses as it was performed as a single replicate.

efficiently hijack the host transcription machinery to complete their infectious cycle.

The transcriptional strategy of PAK_P4 is broadly similar to PAK_P3

PAK_P4 temporally regulates its transcription over the course of infection, as previously found for PAK_P3 (Chevallereau *et al.*, 2016). Most of the gene features encoding unique hypothetical proteins (that is, with low or no homology to any other amino acid sequence in databases) are expressed during the early stage of infection (3.5 min). This 'early region' encompasses gene features encoding gp080 through gp124, while the 'late region' encompasses genes encoding gp001 to gp034. These gene features, which are massively expressed 13 min post infection, primarily encode the predicted phage structural proteins (Supplementary Table 2; Figure 3). PAK_P4 also produces antisense RNAs (asRNAs) that appear to be acting as *cis*-encoded asRNAs since most of them (15 out of 17) are encoded within structural genes (Supplementary Table 2). These asRNAs are significantly more expressed (2 to 3-fold change) 3.5 min after infection compared with late infection stage (13 min). We hypothesize that the transcription of these non-coding RNAs act as repressors to attenuate the expression of the structural genes during the early stage of infection. Since both phage strategies display similar characteristics, our data support an overall conservation of regulatory mechanisms to control viral gene expression, but does not exclude specific differences (see below).

Core genes are similarly expressed in both infections but display extensive phage specific adjustment

To determine more precisely how the regulation of phage gene expression is adjusted, we compared the differential expression of the set of 69 core genes. While most conserved gene features still undergo the same temporal regulation, the strength of this differential expression appears to have changed between the two viruses (Figure 4; Supplementary Table 3). For instance, PAK_P3 tends to differentially express gp002-gp006 more drastically than PAK_P4, while PAK_P4 favors the differential expression of gp007-gp010. Strikingly, one gene is strongly upregulated during PAK_P4 late infection while it is not temporally regulated in PAK_P3. The corresponding hypothetical protein (gp014 in PAK_P4) possesses a wide range of homologs in other unrelated phages and in some bacteria. We also found that genes encoding DNA-metabolism related proteins as well as early genes are overall more strongly temporally regulated in PAK_P3 than they are in PAK_P4.

This comparative analysis of gene expression highlights that evolution has readily selected for small modifications of the mechanisms regulating a variety of specific systems, potentially having a major impact on the way the virus interacts with its host. To address this, we compare how PAK_P3 and PAK_P4 manipulate host transcription and investigate whether the host similarly responds to both phages.

PAK_P4 specifically evolved mechanisms to modulate the expression of genes involved in iron uptake

Phage PAK_P4 strongly upregulates the expression of three host operons predicted to be related to siderophore synthesis and transport, which are not upregulated during PAK_P3 infection (Figure 5; Supplementary Table 4). Indeed, *PAK_4094-4104* (mean $Log_2(FC) = 6.1 \pm 1.1$) and *PAK_4106-09* (mean $Log_2(FC) = 6.3 \pm 0.4$) are respectively homologous to the *pchEFG* and *pchABCD* operons of *P. aeruginosa* PAO1 (Winsor *et al.*, 2016), which are involved in pyochelin biosynthesis, one of the major siderophores of *P. aeruginosa* while *PAK_4090-93* (mean $Log_2(FC) = 5.4 \pm 0.5$) is homologous to *fpt* system, encoding the corresponding siderophore receptor. Interestingly, another upregulated operon, *PAK_1971-78* (mean $Log_2(FC) = 3.0 \pm 0.3$) is homologous

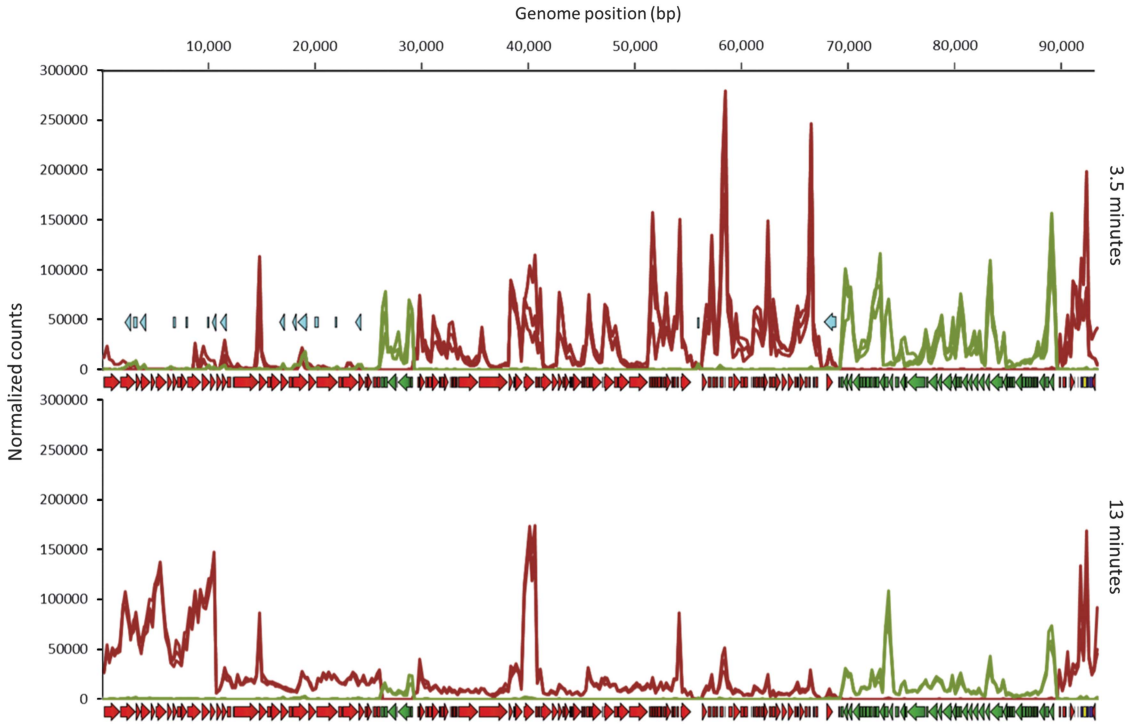

Figure 3 PAK_P4 Transcription map. Transcript densities for early and late infections, normalized by the Total Count of reads that align to the phage genome for each sample. Read density and coding sequence annotations on the Watson and Crick strands are highlighted in red and green respectively while antisense RNA, tRNA and other non-coding RNA annotations are depicted in light blue, dark blue and yellow.

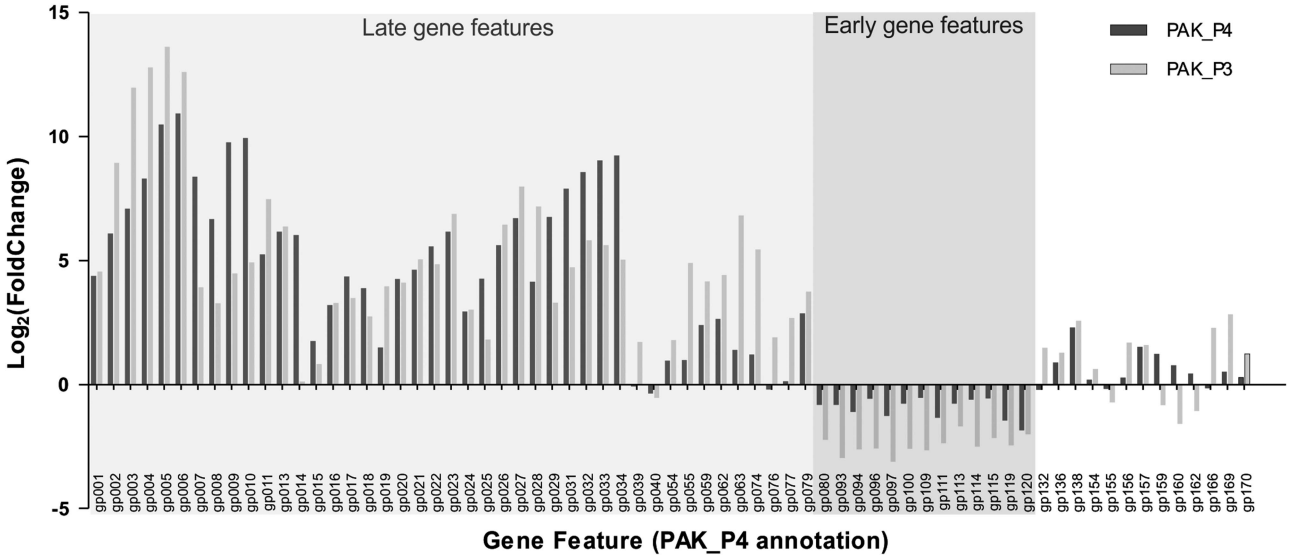

Figure 4 Comparison of PAK_P3 and PAK_P4 homologous genes expression. Phage homologous genes are listed and their differential expression (late versus early) during PAK_P3 and PAK_P4 infections are indicated as Log_2(fold change) values in light and dark gray, respectively.

to the *pvdIJD* operon (Winsor *et al.*, 2016), encoding non-ribosomal peptide synthetases involved in the biosynthesis of the other major *P. aeruginosa* siderophore pyoverdin. Moreover, *PAK_4645-51* and *PAK_4652*, respectively homologous to the *phuS-TUVW* operon and *phuR* receptor gene (Winsor *et al.*, 2016), are also upregulated, although to a lesser extent (mean $Log_2(FC) = 2.3 \pm 1.0$). These genes encode an ABC transporter involved in heme uptake (Ochsner *et al.*, 2000). Finally, 14 other genes putatively involved in iron acquisition (mostly TonB-dependent receptors) are upregulated upon infection (mean $Log_2(FC) = 2.0 \pm 0.4$) while *PAK_4711*, a Fur-like protein (ferric uptake regulator), which represses the expression of genes involved in iron uptake, storage and metabolism in the presence

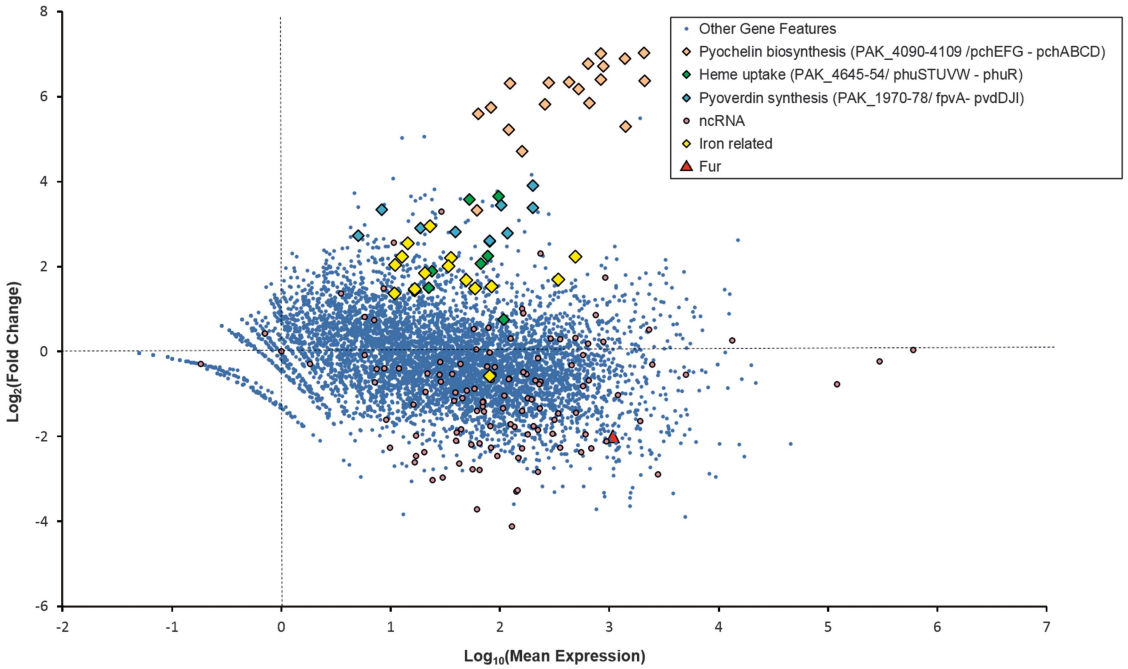

Figure 5 PAK_P4 specifically modulates transcript abundance of many host gene features. An MA plot depicting a differential expression analysis of host gene features demonstrates how the host transcriptional content changes from pre-infection exponential growth ($t=0$) to late infection ($t=13$ min). Notably, this analysis was performed after normalizing the host read counts to host read counts, ignoring the rapidly accumulating phage transcripts in a way that artificially enriches host transcripts in late infection relative to the total transcripts in the cell. This method allows specific effects on host transcription to be tested for independently of the global replacement of host RNAs by phage RNA.

of sufficient iron (Hantke, 2001), is downregulated (Figure 5; Supplementary Table 4).

As this upregulation of iron uptake-related genes may be the downstream consequence of a general activation of stress responses, we investigated the level of expression of alternative sigma factors which are acknowledged to orchestrate such responses. The only sigma factor that is meaningfully transcriptionally upregulated is *algU*, which may reflect an increase of AlgU-associated bacterial stress response. We subsequently looked at the expression of *algU* regulon that typically includes genes involved in alginate, peptidoglycan and LPS biosynthesis, and found that none of them were upregulated upon PAK_P4 infection (Supplementary Table 4). Overall, the differential expression of iron-related genes does not appear to be part of a more general stress response.

Altogether, these results indicate that, unlike PAK_P3, which only triggers the specific and meaningful upregulation of a single operon (Chevallereau *et al.*, 2016), PAK_P4 has specifically evolved processes leading to the complex manipulation of iron homeostasis in the host cell. We next investigate whether we could detect common host transcripts variations to both phage infections which would reflect a conserved host response against this group of viruses.

A common host response is elicited during infection by PAK_P3 and PAK_P4

We identified 55 host coding sequences upregulated in response to both PAK_P3 and PAK_P4 infection while 445 were downregulated (Supplementary Table 5). While the latter may mainly be the result of host mRNA degradation, the former could more likely represent a transcriptional general host response to phage infection, suggesting that some mechanisms for manipulating host transcription may have been conserved. Specifically, infections by both PAK_P3 and PAK_P4 prompt the marked transcriptional upregulation of a predicted P2-like prophage (Figure 6; Supplementary Table 5). However, like PAK_P3, PAK_P4 transcripts globally replace all host mRNAs faster than the increased prophage specific transcription can compete with (Chevallereau *et al.*, 2016). Moreover, one operon linked to RNA processing (*PAK_4493-4499*), which was previously found to be specifically upregulated during PAK_P3 infection, is also upregulated during PAK_P4 infection, albeit more weakly (mean $Log_2(FC)=6.4$ vs mean $Log_2(FC)=2.5$, respectively; Figure 6; Supplementary Table 5). The level of upregulation of this operon between both phages also inversely correlates with the extent and speed of RNA degradation. Additionally, we found that both infections also specifically deplete Hfq transcripts, similar to PhiKZ infected PAO1 cells (Ceyssens *et al.*, 2014; Figure 6 and Supplementary Table 5). Future experiments will address the biological consequences of these transcriptional modifications.

Discussion

While in cellular organisms, the increasing amount of genomic information is strengthening comparative

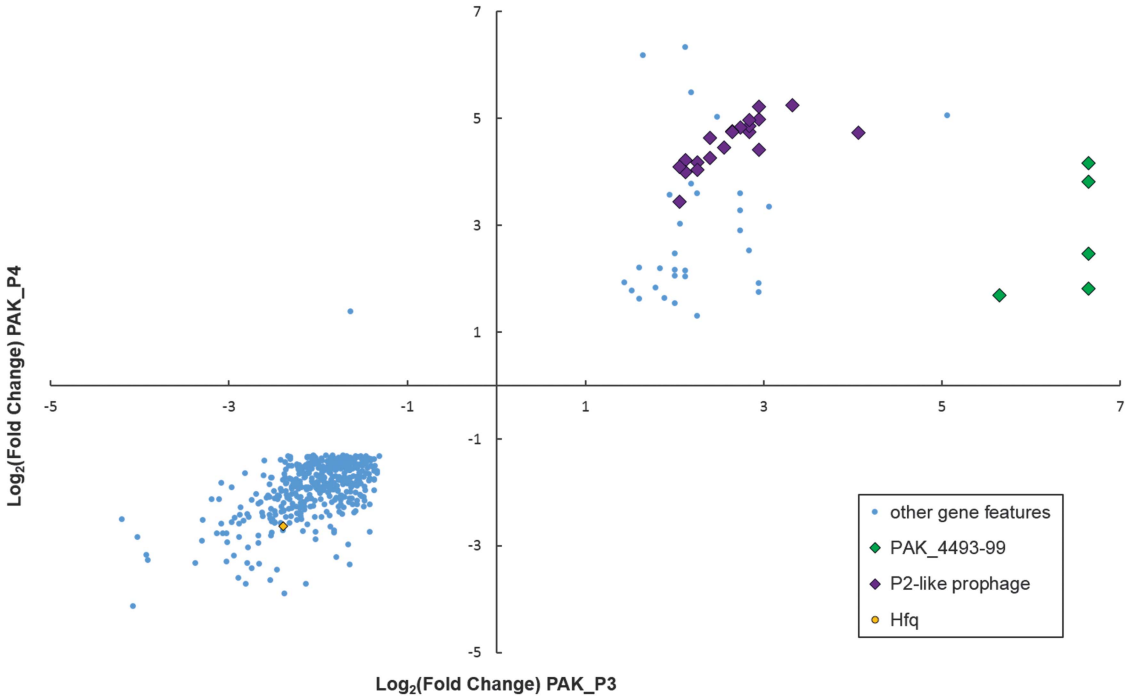

Figure 6 Host strain specifically upregulates the expression of one operon involved in RNA processing in response to both phage infections. Host genes significantly ($P<0.05$) differentially expressed ($|\mathrm{Log_2(FC)}|>1.3$) between $t=0$ and $t=13$ min upon both PAK_P3 and PAK_P4 infections were listed and their $\mathrm{Log_2(FC)}$ values upon each infection are plotted.

analyses delineating evolutionary relationships between cellular lineages, the lack of universal genes has hampered such analyses in viruses. In addition, the diversity of phage sequences is such that most genomes contain numerous ORFs displaying no homology to any other sequences in databanks. Therefore, reconstruction of viral evolution cannot be based solely on genomic sequences and requires further information such as elucidation of mechanistic processes (Forterre *et al.*, 2014). In this context, we have assessed whether two phages (PAK_P3 and PAK_P4), with divergent proteomic contents and infecting the same host strain, conserved the infectious strategy of their common ancestor during their evolution.

First, we observed that both phages share similar infection parameters although PAK_P4 displays a smaller burst size and a longer latent period than PAK_P3. This would not be predicted by the optimal lysis timing model which suggests that longer latency periods result in larger burst sizes (Wang *et al.*, 1996), and requires an alternative explanation. Comparison of expressions of putative endolysin-coding genes (*PAK_P4gp29* vs *PAK_P3gp28*), as well as the four downstream ORFs (holins and/or spanins candidates), revealed that these genes are more efficiently transcribed in PAK_P4, which also appears contradictory with the longer latency period. However, PAK_P4 encodes an asRNA (*PAK_P4as15*), not present in PAK_P3, matching the endolysin gene that could act as a translational repressor, preventing a rapid accumulation of endolysins and delaying lysis. With a supplementary level of lysis control, it was anticipated that PAK_P4 should be able to produce higher number of progeny than PAK_P3. However, as it is not the case, we hypothesized that both phages may rely on distinct infection mechanisms, consistent with their largely divergent protein content, that could explain the observed differences in latency period and burst size. The consequences of these adjustments in the efficiency of each phage may seem of little importance but actually appear to have much broader consequences outside the test tube. This is exemplified by the distinct ability of PAK_P3 to more efficiently treat murine pneumonia compared with PAK_P4 (Henry *et al.*, 2013).

Therefore, we further explored phage molecular lifecycles and found that they both massively degrade host RNA with little manipulation of host gene expression, while relying on the host RNAP which is hijacked to efficiently transcribe phage gene features throughout infection. However: (i) the kinetics of RNA degradation are different and (ii) each phage co-opts the host transcription in a unique way. Remarkably, among host gene features differentially expressed in common between the two phage infections, the *PAK_4493-4499* operon encoding an RtcAB system was previously proposed to be upregulated by PAK_P3 to use the RNA 3'-phosphate cyclase RtcA to prime host transcripts for adenylation as part of a degradation pathway (Chevallereau *et al.*, 2016). However, upon PAK_P4 infection, the upregulation of this operon is much weaker while the degradation of host transcripts is more extensive. Thus, the upregulation of this operon could instead be a reflection of the host attempting to respond to

the extensive RNA damage imposed by both phages, which is more weakly active and thus more weakly protective in PAK_P4 infected cells, which may explain the different kinetics of RNA degradation upon both infections.

As a result, we anticipated that co-option of nucleotide metabolism must be central in PAK_P4 infection process, as it has been shown for other phages, including PAK_P3 (Chevallereau *et al.*, 2016; De Smet *et al.*, 2016). The strong and extensive increased transcription of iron acquisition genes upon PAK_P4 infection supports this hypothesis as increased concentrations of intracellular iron would supply cofactors for viral auxiliary metabolic enzymes, eventually favoring nucleotide turnover.

Alternatively, this increased expression of iron uptake genes could potentially be explained by the recent 'Ferrojan Horse hypothesis', proposing that phage tail fibers are bound to iron ions which allow them to use siderophores as receptors and consequently take advantage of bacterial iron uptake mechanisms (Bonnain *et al.*, 2016). However, no HxH motifs, known to facilitate iron binding, have been found in PAK_P4 putative tail fibers and, more generally, PAK_P4 does not possess more HxH motifs throughout its proteome than PAK_P3, which does not up-regulate iron-related genes (14 motifs versus 10, respectively).

In addition to studying direct phage-host interaction, we compared how these phages regulate their own gene expression, which appears to be broadly conserved and thus inherited from their common ancestor. Notably, recent work by Doron *et al.* has elegantly determined that the Syn9 phage produces a nearly identical transcriptional program when infecting three evolutionarily divergent hosts. In contrast, host response to phage infection was found to be host-specific (Doron *et al.*, 2016). Similar results were obtained by Howard-Varona *et al.* using the φ38:1 phage on two *Cellulophaga baltica* hosts (Howard-Varona *et al.*, 2016) indicating that neither of these two phages adjust their transcriptional program to different hosts. We have now added another piece to the puzzle by conversely studying the infection processes of two divergent phages on one common host.

We found that both phages display similar temporal regulatory schemes of their syntenic gene sets. In addition, this regulation appears to partially rely on the production of antisense transcripts, indicating that this mechanism probably already existed in the common ancestor. However, beside the genes unique to each phage (expected to specifically modulate each phage infection), we noted that differences in the transcriptional program of conserved genes (core genes) also impact infection cycles, as exemplified by the endolysin. These observations highlight the extensive panel of solutions that phages have developed to orchestrate bacterial infections. In addition, such transcriptomics studies revealing infection strategies could help implementing phage taxonomy. Indeed, including genome transcription patterns add a novel layer of properties that could be considered to cluster and differentiate phage species and genera.

In summary, even though they have almost identical morphologies and are able to infect the same host, both phages display substantial variability of their proteomic content, especially within early proteins. Thus, it was unclear whether such related phages would share common mechanisms of action. We showed that the global subversion scheme underlying infection cycle is conserved between phages deriving from a common ancestor although with distinct evolutionary trajectories. We hypothesize that this common mechanism is most likely mediated by the set of the 13 homologous phage early proteins. More intuitive is the observation that both phages have evolved unique processes to infect one particular host, probably explaining why they display distinct killing efficiency. These processes may also indicate that both phages adapted to different ecological niches (for example, low iron environment for PAK_P4).

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This research was supported by the Geconcerteerde Onderzoeks Actie grant 'Phage Biosystems' from the KU Leuven. (http://www.kuleuven.be/onderzoek/kernprojecten/goa.htm). BB has a PhD scholarship within the framework of an Onderzoeks Toelage grant of the KU Leuven. AC was supported by a PhD fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche (ED N°516 B3MI/ ED N°157 BioSPC Paris Diderot Université). RL and LD are members of the « PhageBiotics research community », supported by the FWO Vlaanderen. We deeply thank Damien Mornico (Hub Bioinformatique et Biostatistique, Institut Pasteur – C3BI, USR 3756 IP CNRS) for his invaluable help to sucessfully update the annotations of strain PAK.

References

Bonnain C, Breitbart M, Buck KN. (2016). The Ferrojan horse hypothesis: iron-virus interactions in the ocean. *Front Mar Sci* **3**: 82.

Calendar R. (2006). *The Bacteriophages*. 2nd edn. Oxford University Press: Oxford; New York.

Ceyssens PJ, Minakhin L, Van den Bossche A, Yakunina M, Klimuk E, Blasdel B *et al.* (2014). Development of giant bacteriophage ϕKZ is independent of the host transcription apparatus. *J Virol* **88**: 10501–10510.

Chevallereau A, Blasdel BG, De Smet J, Monot M, Zimmermann M, Kogadeeva M *et al.* (2016). Next-generation '-omics' approaches reveal a massive alteration of host RNA metabolism during bacteriophage

infection of *Pseudomonas aeruginosa*. *PLoS Genet* **12**: e1006134.

De Smet J, Zimmermann M, Kogadeeva M, Ceyssens PJ, Vermaelen W, Blasdel B *et al.* (2016). High coverage metabolomics analysis reveals phage-specific alterations to Pseudomonas aeruginosa physiology during infection. *ISME J* **10**: 1823–1835.

Doron S, Fedida A, Hernandez-Prieto MA, Sabehi G, Karunker I, Stazic D *et al.* (2016). Transcriptome dynamics of a broad host-range cyanophage and its hosts. *ISME J* **10**: 1437–1455.

Essoh C, Latino L, Midoux C, Blouin Y, Loukou G, Nguetta SP *et al.* (2015). Investigation of a large collection of *Pseudomonas aeruginosa* bacteriophages collected from a single environmental source in Abidjan, Cote d'Ivoire. *PLoS One* **10**: e0130548.

Forterre P, Krupovic M, Prangishvili D. (2014). Cellular domains and viral lineages. *Trends Microbiol* **22**: 554–558.

Grose JH, Casjens SR. (2014). Understanding the enormous diversity of bacteriophages: the tailed phages that infect the bacterial family Enterobacteriaceae. *Virology* **468-470**: 421–443.

Hantke K. (2001). Iron and metal regulation in bacteria. *Curr Opin Microbiol* **4**: 172–177.

Hatfull GF, Pedulla ML, Jacobs-Sera D, Cichon PM, Foley A, Ford ME *et al.* (2006). Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet* **2**: e92.

Henry M, Lavigne R, Debarbieux L. (2013). Predicting in vivo efficacy of therapeutic bacteriophages used to treat pulmonary infections. *Antimicrob Agents Chemother* **57**: 5961–5968.

Henry M, Bobay LM, Chevallereau A, Saussereau E, Ceyssens PJ, Debarbieux L. (2015). The search for therapeutic bacteriophages uncovers one new subfamily and two new genera of Pseudomonas-infecting Myoviridae. *PLoS One* **10**: e0117163.

Howard-Varona C, Roux S, Dore H, Solonenko NE, Holmfeldt K, Markillie LM *et al.* (2016). Regulation of infection efficiency in a globally abundant marine Bacteriodetes virus. *ISME J* **11**: 284–295.

Koskella B, Brockhurst MA. (2014). Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol Rev* **38**: 916–931.

Kropinski AM, Sulakvelidze A, Konczy P, Poppe C. (2007). Salmonella phages and prophages–genomics and practical aspects. *Methods Mol Biol* **394**: 133–175.

Krupovic M, Dutilh BE, Adriaenssens EM, Wittmann J, Vogensen FK, Sullivan MB *et al.* (2016). Taxonomy of prokaryotic viruses: update from the ICTV bacterial and archaeal viruses subcommittee. *Arch Virol* **161**: 1095–1099.

Kwan T, Liu J, DuBow M, Gros P, Pelletier J. (2005). The complete genomes and proteomes of 27 Staphylococcus aureus bacteriophages. *Proc Natl Acad Sci USA* **102**: 5174–5179.

Kwan T, Liu J, Dubow M, Gros P, Pelletier J. (2006). Comparative genomic analysis of 18 Pseudomonas aeruginosa bacteriophages. *J Bacteriol* **188**: 1184–1187.

Love MI, Huber W, Anders S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Mahadevan P, King JF, Seto D. (2009). CGUG: in silico proteome and genome parsing tool for the determination of 'core' and unique genes in the analysis of genomes up to ca. 1.9 Mb. *BMC Res Notes* **2**: 168.

Marinelli LJ, Fitz-Gibbon S, Hayes C, Bowman C, Inkeles M, Loncaric A *et al.* (2012). Propionibacterium acnes bacteriophages display limited genetic diversity and broad killing activity against bacterial skin isolates. *MBio* **3**: e00279–12.

Marston MF, Pierciey FJ Jr, Shepard A, Gearin G, Qi J, Yandava C *et al.* (2012). Rapid diversification of coevolving marine Synechococcus and a virus. *Proc Natl Acad Sci USA* **109**: 4544–4549.

Morello E, Saussereau E, Maura D, Huerre M, Touqui L, Debarbieux L. (2011). Pulmonary bacteriophage therapy on Pseudomonas aeruginosa cystic fibrosis strains: first steps towards treatment and prevention. *PLoS One* **6**: e16963.

Ochsner UA, Johnson Z, Vasil ML. (2000). Genetics and regulation of two distinct haem-uptake systems, phu and has, in Pseudomonas aeruginosa. *Microbiology* **146**: 185–198.

Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG *et al.* (2015). Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *Elife* **4**: e06416.

Roux S, Hallam SJ, Woyke T, Sullivan MB. (2015). Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *Elife* **4**: e08490.

Saussereau E, Vachier I, Chiron R, Godbert B, Sermet I, Dufour N *et al.* (2014). Effectiveness of bacteriophages in the sputum of cystic fibrosis patients. *Clin Microbiol Infect* **20**: O983–O990.

Takeya K, Amako K. (1966). A rod-shaped Pseudomonas phage. *Virology* **28**: 163–165.

Wang I-N, Dykhuizen DE, Slobodkin LB. (1996). The evolution of phage lysis timing. *Evolut Ecol* **10**: 545–558.

Williams HT. (2013). Phage-induced diversification improves host evolvability. *BMC Evol Biol* **13**: 17.

Winsor GL, Griffiths EJ, Lo R, Dhillon BK, Shay JA, Brinkman FS. (2016). Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. *Nucleic Acids Res* **44**: D646–D653.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/4.0/

© The Author(s) 2017

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)