



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/181212/>

Version: Published Version

---

**Article:**

Ford, S., Moeskjær, S., Young, P. et al. (2021) Introducing a novel, broad host range temperate phage family infecting *Rhizobium leguminosarum* and beyond. *Frontiers in Microbiology*, 12 (11). 765271. ISSN: 1664-302X

<https://doi.org/10.3389/fmicb.2021.765271>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



# Introducing a Novel, Broad Host Range Temperate Phage Family Infecting *Rhizobium leguminosarum* and Beyond

Sam Ford<sup>1</sup>, Sara Moeskjær<sup>1</sup>, Peter Young<sup>2</sup>, Rosa I. Santamaría<sup>3</sup> and Ellie Harrison<sup>1\*</sup>

<sup>1</sup> Department of Animal and Plant Sciences, University of Sheffield, Sheffield, United Kingdom, <sup>2</sup> Department of Biology, University of York, York, United Kingdom, <sup>3</sup> Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, México City, Mexico

## OPEN ACCESS

### Edited by:

Heejoon Myung,  
Hankuk University of Foreign Studies,  
South Korea

### Reviewed by:

Jens Andre Hammerl,  
Bundesinstitut für Risikobewertung,  
Germany  
Robert Hertel,  
Brandenburg University of Technology  
Cottbus-Senftenberg, Germany

### \*Correspondence:

Ellie Harrison  
ellie.harrison@sheffield.ac.uk

### Specialty section:

This article was submitted to  
Phage Biology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 26 August 2021

**Accepted:** 13 October 2021

**Published:** 09 November 2021

### Citation:

Ford S, Moeskjær S, Young P,  
Santamaría RI and Harrison E (2021)  
Introducing a Novel, Broad Host  
Range Temperate Phage Family  
Infecting *Rhizobium leguminosarum*  
and Beyond.  
Front. Microbiol. 12:765271.  
doi: 10.3389/fmicb.2021.765271

Temperate phages play important roles in bacterial communities but have been largely overlooked, particularly in non-pathogenic bacteria. In rhizobia the presence of temperate phages has the potential to have significant ecological impacts but few examples have been described. Here we characterize a novel group of 5 *Rhizobium leguminosarum* prophages, capable of sustaining infections across a broad host range within their host genus. Genome comparisons identified further putative prophages infecting multiple *Rhizobium* species isolated globally, revealing a wider family of 10 temperate phages including one previously described lytic phage, RHEph01, which appears to have lost the ability to form lysogens. Phylogenetic discordance between prophage and host phylogenies suggests a history of active mobilization between *Rhizobium* lineages. Genome comparisons revealed conservation of gene content and order, with the notable exception of an approximately 5 kb region of hypervariability, containing almost exclusively hypothetical genes. Additionally, several horizontally acquired genes are present across the group, including a putative antirepressor present only in the RHEph01 genome, which may explain its apparent inability to form lysogens. In summary, both phenotypic and genomic comparisons between members of this group of phages reveals a clade of viruses with a long history of mobilization within and between *Rhizobium* species.

**Keywords:** prophage, temperate phage, rhizobia, phage ecology, phage evolution

## INTRODUCTION

Bacteriophages (phages) are ubiquitous in bacterial populations and contribute significantly to the ecology and evolution of microbial communities. Temperate phages—phages capable of integrating into the bacterial genome—in particular, occupy a multitude of roles in bacterial communities (Harrison and Brockhurst, 2017); like purely lytic phages they can infect and kill susceptible bacteria, contributing to population turnover and driving selection for resistance (Paterson et al., 2010). As integrated phages—“prophages”—they can provide competitive advantages to their hosts—which are known as “lysogens” (Brown et al., 2009); Lytic phage particles produced from the lysogen population at low rates infect and kill non-lysogens—likely to be unrelated competitors—and thus act as anti-competitor “weapons” during competition. Self-exclusion mechanisms encoded by phages mean that lysogens are themselves immune to infection and thus

benefit from the resulting competitive release (Brown et al., 2009; Burns et al., 2015). Phages also play a major role as agents of horizontal gene transfer (Harrison and Brockhurst, 2017) either through transduction—packaging and transporting host DNA within viral particles—or by carriage of bacterial genes on the phage genome. Genes encoded on temperate phages can have significant ecological effects, such as the shiga toxin genes encoded on STX phages that make *Shigella* bacteria highly virulent (Allison, 2007) or photosynthesis genes encoded on *Prochlorococcus* phages (Lindell et al., 2004). Thus temperate phages can be significant players in microbial ecosystems. For many bacterial groups, however, particularly in terrestrial environments, very little is known about the lives of these genetic symbionts.

Rhizobia—bacteria that form nitrogen-fixing mutualistic symbioses with legumes—are key members of the soil community. Rhizobia are characterized by their ability to form beneficial intracellular infections in legume hosts which form specialized root structures—“nodules”—to house their bacterial symbionts. Here atmospheric nitrogen is fixed and exchanged with the plant for sugars. Rhizobia-mediated nitrogen fixation is a major source of accessible nitrogen for the biosphere (Vitousek et al., 2002). The rhizobia symbiosis is highly dependent on mobile genetic elements. Rhizobia typically contain numerous plasmids and integrated elements, including those encoding genes required for legume symbiosis (Cavassim et al., 2020). Lytic rhizobia phages have been well described and evidence suggests that they may play a role in structuring rhizobia communities, and correspondingly, in interactions with their legume hosts. Rhizobia phages generally show specificity within their host clades (Shahaby et al., 2014) but can have broad infectivity within genera (Msimbira et al., 2016). Phage resistance and susceptibility can lead to differential nodulation success demonstrating a direct role in the important legume symbiosis (Evans et al., 1979; Hashem and Angle, 1988; Ahmad and Morgan, 1994). Lytic rhizobia phages have also been shown to be capable of efficient horizontal gene transfer by transduction (Buchanan-Wollaston, 1979; Finan et al., 1984). However, the presence and role of temperate phages in rhizobia is less well understood.

Currently, only a small number of prophages have been described across rhizobia clades; in *Rhizobium* (Halmillawewa et al., 2015), *Ensifer* (Dziewit et al., 2014; Decewicz et al., 2017), and *Bradyrhizobium* (Dhar et al., 2012). Here we describe a clade of novel temperate phages infecting *Rhizobium leguminosarum* strains nodulating clover from across Europe. Sequence comparisons show that this group is closely related to a lytic phage isolated in Mexico on *R. etli*, RHEph01 (Santamaría et al., 2014), as well as a range of putative prophages detectable within published *Rhizobium* genomes. Characterization of these phages showed that all prophages were “active”—able to form infectious phage particles and new lysogenic infections, with potential implications for host bacteria. Confirming the results of Santamaría et al. (2014), we find that related phage RHEph01 was unable to form lysogens, and genome analysis suggests that this may be an acquired trait rather than a loss of function through gene loss or disruption.

## MATERIALS AND METHODS

### General Bacterial and Phage Culturing

All bacterial culturing was performed in TY media (3 g/L yeast extract, 6 g/L tryptone, 0.5 g/L calcium chloride) in 6 ml culture volumes under shaken (180 rpm) conditions at 28°C. Colony forming units (CFUs) were enumerated on solid TY agar (13 g/L) and plaque forming units (PFUs) on soft TY agar overlays after cells were removed through filtration (using 0.45 μM syringe filters). Soft agar overlays consist of 5 ml of 6 g/L agar containing 100 μl of a growing culture of phage free *R. leguminosarum* bv. *trifolii* (*Rlt*) TRX19. Where stated phage lysis was induced with mitomycin C. A 1000X concentration of mitomycin C was added at a final concentration of 0.05 μg/ml to lysogenic cultures grown for 6 h. Cultures were then allowed to grow for a further 40 h before phages were harvested. Phage clones were isolated by “picking” a plaque with a sterile toothpick and inoculating into a culture of phage free *R. leguminosarum* bv. *trifolii* (*Rlt*) TRX19.

### Phage Isolation and Sequencing of vTRX32-1

The phage vB\_RleA\_TRX32-1 (hereafter shortened to vTRX32-1, accession number MW023914.1) was isolated through induction from lysogenic strain, *Rlt* TRX32. The strain was isolated in York, United Kingdom by Kumar et al. (2015) from a clover (*Trifolium repens*) nodule. vTRX32-1 was then cultured on phage-free strain *Rlt* TRX19. Phage clones were amplified from single plaques and DNA extracted from lytic phage particles following published methods (Jakočičuně and Moodley, 2018). Briefly, 1 ml of a filtered phage lysate was prepared first by digestion of any contaminating bacterial DNA and RNA with DNaseI and RNaseA, followed by lysis of the phage capsid by proteaseK digestion. DNA was then extracted using a DNeasy Blood and Tissue kit (Qiagen). Further purification was achieved using the DNA Clean and Concentrator kit (Zymo) following kit protocol. Phage DNA was sequenced on Illumina MiSeq platform. An initial alignment was performed using SPAdes creating one primary contig of 42.8 Kb (initial coverage of 29.27x). All reads were then realigned to this primary contig using BWA and SAMtools resulting in 88% of reads being mapped and a final mean coverage of 49.1x. The assembly pipeline is provided in **Supplementary Materials**. Open reading frames (ORFs) were predicted and annotated provisionally using the RAST pipeline (Aziz et al., 2008). The PHASTER webserver (Arndt et al., 2016) and blastp were then used to aid the annotation of the remaining hypothetical ORFs. tRNAs were identified using tRNAscan-SE 2.0 (Chan and Lowe, 2019). Host and phage promoter sequences were predicted using PhagePromoter (Sampaio et al., 2019).

A nucleotide Blast revealed closely related phage sequences in four published bacterial genomes belonging to an extensive (196 strain) collection of *Rlt* strains (Cavassim et al., 2020); SM48, SM95, SM113 and SM135A. These strains were kindly provided by Stig U. Andersen and phages induced as above. Phage clones were isolated from single plaques and verified using phage specific primers. Phages were named based on isolating strain as before, vB\_RleA\_vSM48-1, or “vSM48-1” from SM48 etc.

The blast search also revealed that these phages are related to the lytic phage, RHEph01, which was isolated on *R. etli* in Mexico and has been previously described (Santamaría et al., 2014; **Supplementary Table 1**). Phage RHEph01 was kindly provided by Victor Gonzalez for phenotypic comparison with the *Rlt* phages.

## Electron Microscopy

Phage lysates of vTRX32-1 and RHEph01 were prepared through lysis of the mutually susceptible host TRX19. 20  $\mu$ l of phage stocks were introduced to 1 ml culture in mid-exponential phase. Phages were recovered by filtration with a 0.45  $\mu$ m syringe filter. A 5  $\mu$ l spot of phage lysate was added to a copper/palladium grid, left for 1 min and excess removed. Phages were then washed twice in water and stained twice with 1% uranyl formate for 20 s. Grids were examined with a CM100 Transmission electron microscope.

## Phenotypic Comparisons Between Strains

Phage phenotypes were compared for host range, virulence and lytic growth dynamics. Host range was determined for the five *Rlt* phages on a range of hosts drawn from the same culture collections as the original lysogens as well as a number of non-*R. leguminosarum* type strains. The lytic phage RHEph01 was included for comparison. Phages were standardized to 10<sup>6</sup> PFU per ml and 10-fold serial dilutions spotted onto bacteria grown in soft TY agar overlays. Phages that showed a zone of partial (turbid) clearing at high phage titer after 1 day were graded as “weak.” Phages that produced a clear zone of clearing at high titer and individual plaques at lower densities at day 1 were graded as “strong.” Zones of clearing were inspected each day for a further 3 days. Where regrowth was observed within the high titer spots, bacteria from this turbid zone was streaked out to isolate individual colonies to check for lysogenization. Putative lysogens were confirmed through colony PCR targeting the phage DNA maturase gene (forward: GTCGAGTGCTTGACCTCCTC, reverse: ACCTCTTCTTGGTCGCTTCA) and experimentally by spotting supernatant onto susceptible host TRX19.

Replication dynamics of 3 phages, vTRX32-1, vSM113-2 and lytic phage RHEph01, were determined by one-step growth curves of lytic viral infections. vSM113-2 was selected as the most divergent phage from vTRX32-1. Exponentially growing cultures ( $n = 4$ ) were mixed with phages (100:1) and left to adsorb for 15 min. Unadsorbed phages were then removed and plated to determine the number of phage infections (total phage—unadsorbed). Infected cells were then diluted 10 fold and incubated at 28°C. Samples were taken every 30-min for 8 h and phage density enumerated by plating onto soft agar overlays. Phage growth curves were modeled using R package “GrowthRates” to estimate latent period (time between adsorption and the beginning of phage lysis) and burst size (maximum phages produced per infection).

Virulence of these 3 phages was estimated on susceptible *Rlt* host TRX19 as well as *R. etli* strain Bra5, the original isolating strain of RHEph01. Mid-exponential phase cultures of host strains were diluted 10 fold and combined in a 96

well plate with phage inoculants at a multiplicity of infection 0.1 or phage free TY for a phage- control. Three replicates of each phage x host combination were established and grown for 72 h at 28°C with shaking in an incubated plate reader measuring OD600 at 20 min intervals. Phage virulence was measured as reduction in bacterial growth (RBG) at 2-time points: at 10 h post-infection to estimate the initial impact of infection, and at 40 h, to estimate the overall impact of phage infection on growth (thus capturing both lysis and recovery following in emergence of resistance/lysogenization). RBG was estimated as  $1 - (\text{OD}_{\text{phage} + T_x} - \text{OD}_{\text{phage} + T_0}) / (\text{OD}_{\text{phage} - T_x} - \text{OD}_{\text{phage} - T_0})$ , where time = T. Phenotypic data can be found in **Supplementary Table 2**.

## Genome Annotation and Comparison

In addition to the isolated and previously described phages, four further, more distantly related, prophage sequences were identified in other *Rhizobium* strains through NCBI blastn using the vTRX32-1 genome and default settings: *R. anhuiense* CCBAU 23252, *R. loessense* CGMCC 1.3401, *Rhizobium* sp. NXC14 and *Rhizobiales* sp. AFS092257. These were included for comparison of the wider clade.

Provisional prophage sequences for the four *Rlt* prophages vSM48-1, vSM95-2, vSM113-1 and vSM135A-2, and for the four putative prophages from *Rhizobium* spp. sequences were obtained using PHASTER (Arndt et al., 2016)<sup>1</sup>. Prophage termini (attL and attR sites) and the corresponding integration site in the bacterial genome (attB) were identified based on their alignment to the circular vTRX32-1 and RHEph01 phage genomes and annotation by PHASTER (Arndt et al., 2016) and comparison to phage free *Rhizobium leguminosarum* sequences. A conserved terminally repeated 12 bp motif (“CAGATTTAGGTT”) was identified at the terminal ends of all prophages, which could also be found in phage free strains within a tRNA-leu ORF.

Pairwise blastp of RAST-predicted coding sequences and blastn of whole genomes were used to generate percent identity values between the genes/genomic regions of vTRX32-1 and those of its relatives. Average nucleotide identity was calculated using the VIRIDIC phage genome comparison tool (Moraru et al., 2020)<sup>2</sup>.

## Phylogenetic Analyses of Prophage/Phage

As phage genomes can be highly variable the phage phylogeny was generated using a 5 gene concatenated nucleotide sequence of the RNA polymerase, DNA polymerase, DNA helicase, Collar protein and Maturase B genes. Host phylogeny was produced based on core gene *rpoB*. A proteomic tree based on the RNA polymerase amino acid sequence was produced showing vTRX32-1 and its relatives in the context of other phages representing major *Autographiviridae* clades. Both protein and nucleotide sequences were produced in MEGAX using MUSCLE (following Kumar et al., 2015) using a maximum likelihood method (Tamura-Nei model), with 1,000 bootstrap replications.

<sup>1</sup><https://phaster.ca>

<sup>2</sup><http://rhea.icbm.uni-oldenburg.de/VIRIDIC>

Phage mobilization within bacterial lineages was assessed using empress (Santichaivekin et al., 2020) maximum parsimony reconciliations. Reduced phylogenies were produced as above using the 9 prophages and corresponding lysogen strains. Cost weightings for events were set as follows: cospeciation = 0, transfers = 1, duplications = 2 and loss = 1–2. Analysis of the “cost space” (the impact of cost values on the predicted outcomes) showed that outcomes were robust to changes in the weighting of duplication events. A value of 2 was chosen as all prophages in this group show integration into one conserved site, making duplication unlikely. The relative weighting of transfer and loss is less clear and strongly affects the outcome therefore loss was considered at two probability levels, firstly with equal probability with transfer (loss = 1) and secondly, as less likely than transfer (loss = 2). Where loss = 1, 20 maximum parsimony reconciliations (MPRs) were estimated and where loss = 2, 38 MPRs were estimated. These were clustered into 2 and 4 groups of MPRs, respectively, and a representative scenario of each cluster produced following program guidance.

## RESULTS AND DISCUSSION

### The vTRX32-1 Phage Family

The vTRX32-1 phage family represents a novel group of temperate phages isolated within *Rhizobium leguminosarum* bv. *trifolii* (*Rlt*) strains. Originally, vTRX32-1 (vB\_RleA\_TRX32-1) was isolated from strain TRX32 (Kumar et al., 2015) (from a clover nodule in York, United Kingdom), sequenced and visualized by transmission electron microscopy (Figure 1). Both sequence and morphology place this phage as a member of the *Autographiviridae*—related to *Podoviridae*. Subsequent screening of *Rlt* published sequences revealed four further lysogens containing closely related prophages; strains SM48 and SM135A (isolated from white clover nodules in Denmark) and SM95 and SM113 (isolated from white clover nodules in Rennes, France) (Cavassim et al., 2020; Supplementary Table 1). All five *Rlt* phages could be isolated by induction and formed plaques on isolating strain *Rlt* TRX19.

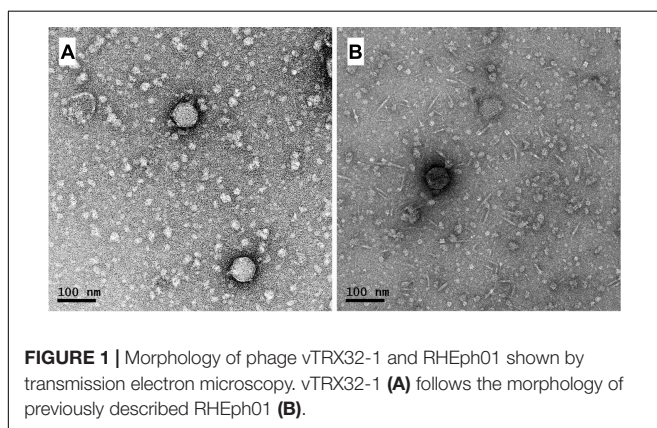
Blast searches revealed that these prophages were also related to the previously described lytic phage, RHEph01, isolated on

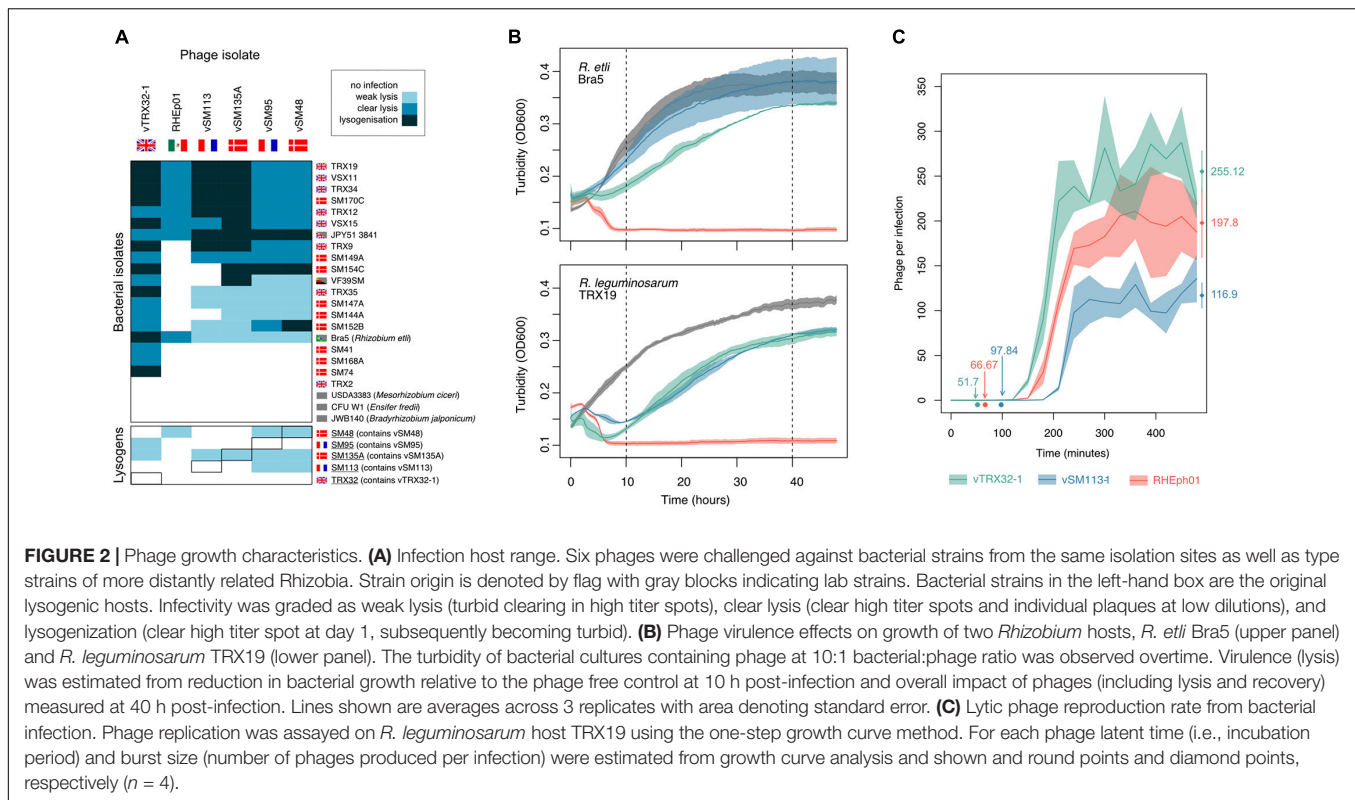
*R. etli* in Mexico (Santamaría et al., 2014) as well as four further putative prophages in more distantly related *Rhizobium* strains; *R. anhuiense* CCBAU 23252, *R. loessense* CGMCC 1.3401, *Rhizobium* sp. NXC14 and *Rhizobiales* sp. AFS092257 (Supplementary Table 1).

Patterns of relatedness between phages show that 8 phages—five *Rlt* phages, the NXC14 and CCBAU 23252 putative prophages and the RHEph01 lytic phage—form a clade of > 70% average nucleotide identity (ANI) (Supplementary Figure 1). RHEph01 is currently classified by ICTV as the sole member of the genus *Paadamvirus*, in the family *Autographiviridae* (International Committee on Taxonomy of Viruses [ICTV], 2021) and under current guidance this places these 7 additional phages within this genus. Within this group, the *Rlt* phages form a distinct, more closely related cluster (ANI 87.3–93.9%; Supplementary Figure 1). Under classification guidelines, however, these phages would each correspond to a separate species (< 97% ANI) (International Committee on Taxonomy of Viruses [ICTV], 2021).

### Variable Life Histories Across the vTRX32-1 Like *Rlt* Phages

Analysis of phage infection dynamics showed significant variation within this small group. The infectious host ranges of the five *Rlt* phages and the lytic phage RHEph01 were compared on a panel of natural isolates drawn from the same *Rlt* communities as vTRX32-1 and the vSM phages as well as several type strains from more distantly related species (Figure 2A). All 6 phages were able to infect a wide range of *R. leguminosarum* hosts as well as *R. etli* strain Bra5 indicating a relatively broad host range, i.e., spanning multiple *Rhizobium* species (Ross et al., 2016). No phages were able to infect strains from more distantly related genera. Within the host panel (primarily *Rlt* strains) phage RHEph01 had the most narrow host range—which very likely reflects local adaptation of the *Rlt* phages to European *R. leguminosarum* populations, compared with RHEph01 which was isolated in Mexico where *R. leguminosarum* is less dominant. RHEph01 showed no evidence of lysogeny in either species, confirming the findings of Santamaría et al. (2014) that this phage is purely lytic, at least under lab conditions. All five *Rlt* phages were able to form lysogens in new hosts, though lysogens were not detected for all hosts. Phage vTRX32-1 was by far the most infectious. Indeed, relatively few resistant *R. leguminosarum* isolates have been identified. The four vSM phages had a slightly reduced host range with many infections that produced zones of clearing but no visible individual plaques, designated as weak lysis. These phages divided into two patterns of host range—with vSM113 and vSM135A showing more evidence of lysogeny of strains compared to vSM95 and vSM48. These patterns do not follow either phage relatedness (Supplementary Figure 1) nor isolation environment, e.g., vSM95 and vSM113 were isolated from the same site in France yet show different patterns of infection. Multiple phages also showed weak but consistent lysis of lysogen strains containing either related phages or, in the case of vSM48 and vSM135A, their own lysogens. This may imply that superinfection immunity is impaired, or that the zone of growth





inhibition could be associated with induction triggered by high rates of secondary phage adsorption (Abeldon, 2019).

Phage virulence was assessed for 3 phages—vTRX32-1 (isolated as a prophage in United Kingdom), vSM113 (isolated as a prophage in France) and RHEph01 (isolated as a lytic phage in Mexico)—on growing cultures of mutually susceptible *R. leguminosarum*, TRX19, and *R. etli*, Bra5, hosts. At 10 h, where growth is most severely affected by phage killing (Figure 2B), virulence varied significantly between phages ( $F_{2, 12} = 10.736$ ,  $p = 0.00212$ ) but not bacterial host ( $F_{1, 12} = 4.005$ ,  $p = 0.0685$ ). All phages reduced host density at this time point but lytic phage RHEph01 was significantly more virulent than the two *Rlt* phages (contrasts $_{vSM113---RHEp01}$ :  $t = -4.352$ ,  $p = 0.0025$ , contrasts $_{vTRX32-1---RHEp01}$ :  $t = -3.067$ ,  $p = 0.0245$ ). No significant difference was observed between these two *Rlt* phages (contrasts $_{vSM113---vTRX32-1}$ :  $t = 1.219$ ,  $p = 0.465$ ).

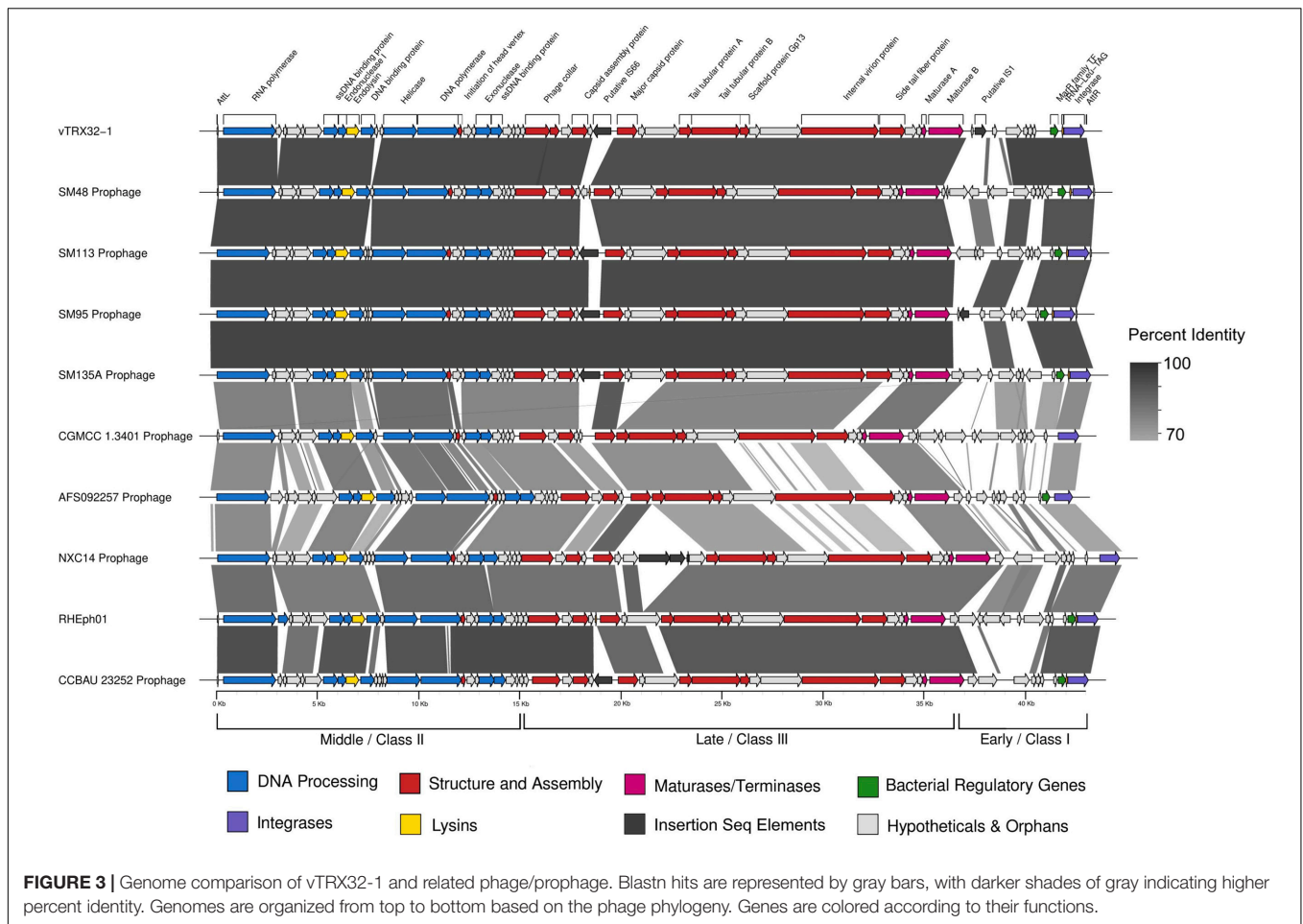
Following this initial drop in bacterial density in phage containing populations, bacteria infected with either of the *Rlt* phages recovered dramatically (Figure 2B). Populations infected with RHEph01, however, experienced no such rescue. Correspondingly, reduction in bacterial growth at 40 h differed significantly between phages ( $F_{2, 12} = 84.791$ ,  $p < 0.0001$ ), which was again driven by difference between RHEph01 and the *Rlt* phages (contrasts $_{vSM113---RHEp01}$ :  $t = -11.683$ ,  $p < 0.001$ ; contrasts $_{vTRX32-1---RHEp01}$ :  $t = -10.303$ ,  $p < 0.001$ ) with no difference between the vSM113 and vTRX32-1 (contrasts $_{vSM113---vTRX32-1}$ :  $t = 1.309$ ,  $p = 0.417$ ). PCR of colonies isolated from a subset of strains confirmed that recovered bacteria infected with both *Rlt* phages had become

lysogens. In contrast RHEph01 effectively cleared the population without any appearance of resistance emerging, again confirming that it is unable—or unlikely—to form lysogens.

The dynamics of lytic phage replication were further analyzed using a one step growth curve of these three strains on the mutually susceptible host TRX19. In comparison with other phages, the *Rlt* phages have relatively long latent periods and large burst sizes. The latent period—the time between adsorption and lysis when lytic phage particles are assembled—was relatively long; from 50 min in vTRX32-1 to over 90 min in the case of vSM113. In comparison, in *E. coli* phages latent periods can be as short as 20 min (Topka et al., 2018). All 3 phages were able to replicate as lytic phages on the host and produced large burst sizes of up to  $255 \pm$  PFUs per cell in the case of vTRX32-1 (Figure 2C). Phages differed significantly in both growth characteristics (burst size:  $F_{2, 9} = 6.522$ ,  $p = 0.018$ , latent period:  $F_{2, 9} = 38.21$ ,  $p < 0.0001$ ). vSM113 exhibited the least efficient lysis—both in terms of longest latent period and smallest burst size—while vTRX32-1 exhibited the most productive infections with burst size being significantly larger compared with vSM113-1 (contrasts $_{vTRX32-1---vSM113-1}$ :  $t = 3.594$ ,  $p = 0.014$ ) and latent period shorter than both phages (contrasts $_{vTRX32-1---vSM113-1}$ :  $t = -8.567$ ,  $p < 0.001$ , contrasts $_{vTRX32-1---RHEp01}$ :  $t = -2.781$ ,  $p = 0.051$ ).

## Comparative Genomics of the vTRX32-1 Like Phages and Wider Relatives

In addition to the five vTRX32-1-like *Rlt* phages and RHEph01, we identified four related prophages within published sequences

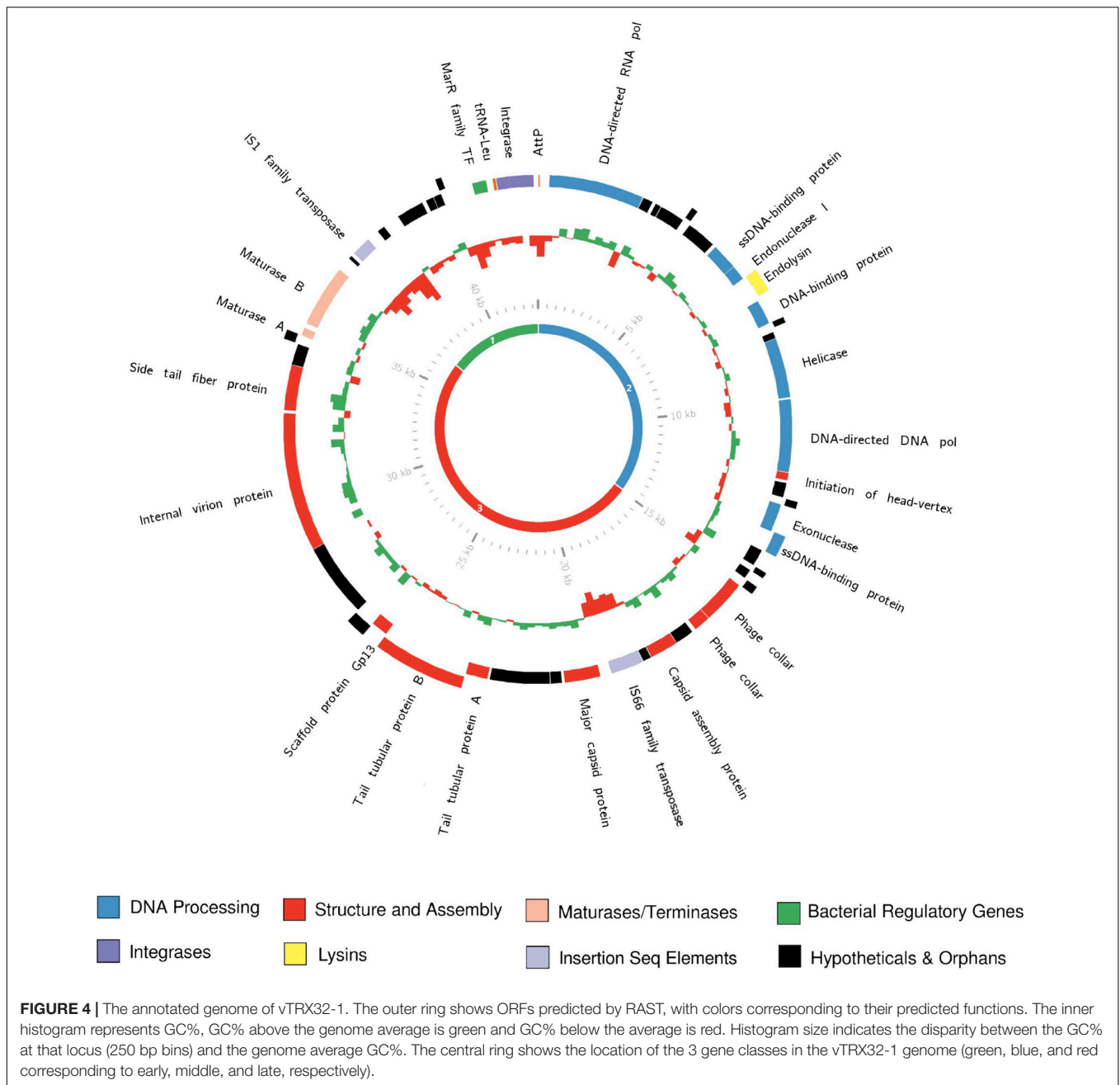


of different *Rhizobium* species and included these putative phages for a comparison of the wider clade. Initial comparison between phage genomes shows that genome architecture across this group of 10 phages is well conserved; they are between 42 and 45 kb long, have an average GC content of 58.2–59.5% and contain between 52 and 55 predicted open reading frames each. In the vTRX32-1 genome 25 genes were annotated, 23 of which were present in all close relatives. tRNA-SCAN-SE identified a single tRNA-Leu in all pro/phage genomes, and additional tRNAs in 4 of the 10 genomes.

All prophages share the same attachment site. A conserved terminally repeated 12 bp motif (“CAGATTAGGTT”) was identified at the terminal ends of all prophages—annotated by PHASTER as the attL/attR sites. Phages sequenced in their circular form, vTRX32-1 and RHEph01, both carry this repeated motif in the same position. In *R. leguminosarum* strains which lack the prophage this 12 bp motif is present within a tRNA-Leu indicating that this is the site of prophage integration (attB; see **Supplementary Figure 2**). tRNA sequences are used by a range of other mobile genetic elements for site-specific integration (Reiter et al., 1989), including other rhizophages (Semsey et al., 2002; Halmillawewa et al., 2016), retrotransposons (Mularoni et al., 2012), and cyanopodoviruses (Sullivan et al., 2005). This includes

closely related pelagibacter viruses in the *HTVC019Pvirus* genus (Zhao et al., 2019). The tRNA-Leu carried within the phage genome lacks the 12 bp motif, presumably replacing function of the chromosomal tRNA-Leu interrupted through integration but preventing “super-integration”—integration of additional phages into the phage itself. It is common for tRNA integrating phages to encode a replacement tRNA, although this is more often in the form of a partial tRNA sequence which completes one half of the disrupted gene (Halmillawewa et al., 2016).

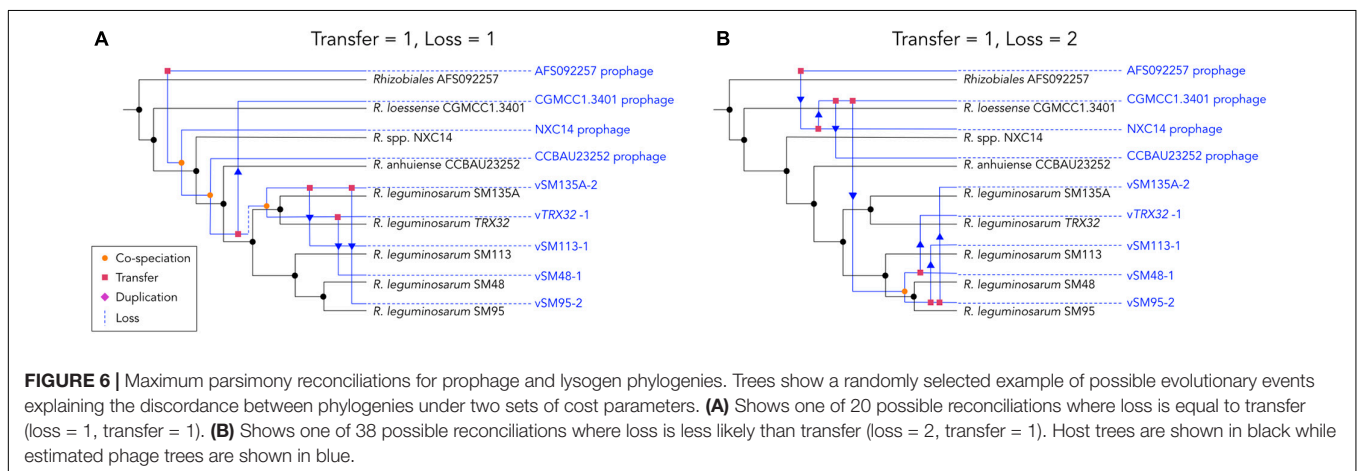
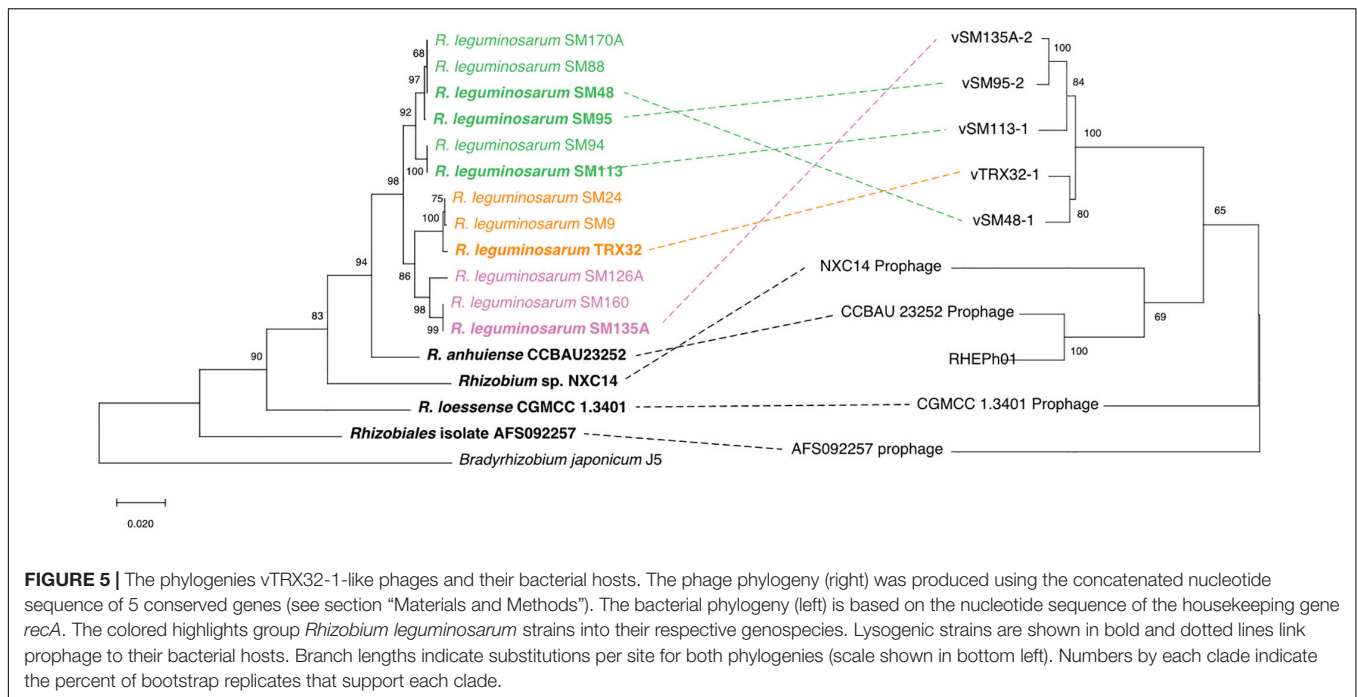
Genome analysis revealed that this group of *Rhizobium* phages have compact genomes typical of the *Autographiviridae* and related (and better described) *Podoviridae*, with similar genome size, synteny and core gene content (**Figure 3**). As is typical for mobile genetic elements such as prophages, roughly half of genes are of unknown function. Distribution of hypothetical genes as well as genome conservation was uneven across the genome. Based on homology to type phage, *E. coli* bacteriophage T7, genes can be grouped into early, middle and late clusters—named on the basis of their transcription order (also referred to as the class I, II and III genes) (Dunn and Studier, 1983; Lindell et al., 2007). Regions of the genome showing homology to genes from middle and late regions, involved in DNA processing and phage structure, respectively, are characterized by high levels of



synteny and nucleotide homology. Genes in these regions contain predominantly phage-like predicted promoters and can be well annotated by homology with other known core phage genes.

In contrast, the region downstream of the maturase B gene, which contains genes with homology to early region genes, is highly variable in both nucleotide identity (Figure 3) and ORF presence/absence (Supplementary Figure 3). The majority of ORFs called within this region are of unknown function, based on both homology to published sequences or identifiable domains. In contrast to middle and late genes, predicted promoters for this region were typical of bacteria, rather than phage genes. Finally, the early region also has significantly lower GC% ( $54.32 \pm 0.65\%$ )

than the rest of the genome ( $59.95 \pm 0.24\%$ ), ( $t_9 = 29.67$ ,  $p < 0.00001$ ; Figure 4), characteristic of genes recently acquired by horizontal gene transfer. Early region genes, as the name implies, are transcribed first upon entry to the cell. In closely related phages, such as T7, these genes encode proteins that must act fast to overcome host defenses and takeover host replicative machinery (Studier, 1972; Dunn and Studier, 1983). Thus, variation in this region could reflect selection for host specificity and/or infectivity driving divergence in this region, which has been shown to drive recombination between phages (De Paepe et al., 2014). This pattern of conservation and variability is common amongst close relatives within the *Autographiviridae*

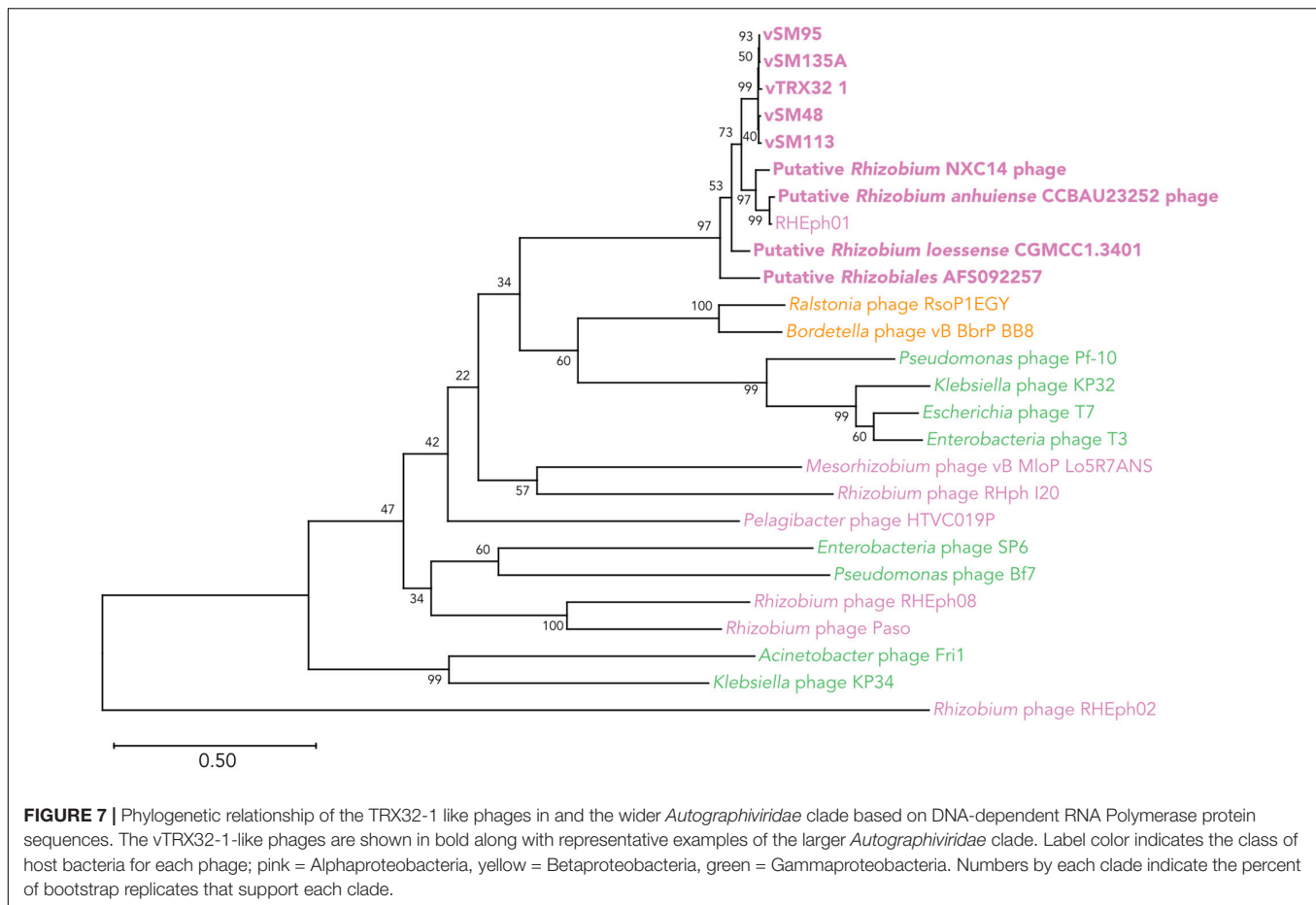


(Zhu et al., 2010; Labrie et al., 2013; Zhao et al., 2013; Hamdi et al., 2016; Kawasaki et al., 2016).

## Evidence of Horizontal Gene Transfer Within Phage Genomes

Phage genomes often feature significant mosaicism (Juhala et al., 2000) and are themselves subject to integration of other mobile elements. Three insertion sequences, in the form of transposases, were identified across the 10 phages. Firstly, a putative IS66 element (Machida et al., 1984) was found in 4/5 of the *Rlt* phages, as well as a putative prophage detected in *R. anhuiense* CCBAU23252. IS66 elements were originally isolated in the *Agrobacterium tumefaciens* [now reclassified as *Rhizobium radiobacter* (Young et al., 2001)] Ti plasmid (Machida et al., 1984), and are common throughout *Rhizobium* species

(Rochepeau et al., 1997; González et al., 2006). In addition to IS66 elements both the vTRX32-1 and vSM95-2 genomes contain an additional gene annotated as an IS1 transposase within the variable early region. Finally, an IS21 transposase is present in the genome of a putative prophage identified in *Rhizobium* spp. NXC14. This IS element is identical to IS21 transposase domains found on NXC14 plasmids, pRspNXC14a and pRspNXC14c, as well as being closely related to other insertion sequences on a variety of *Rhizobium* replicons, making it likely that it was acquired from co-habiting plasmids. Like the early region genes, these regions are notable for their low GC% content (Figure 4). IS elements are known to be highly active within *Rhizobium* (Hernandez-Lucas et al., 2006), and are thought to contribute to *Rhizobium* genome evolution, creating their characteristic mosaic structure (González et al., 2003; Hernandez-Lucas et al., 2006). The presence of IS elements within phage genomes suggests that



prophages can act as a conduit for dissemination of IS elements throughout *Rhizobium*.

In contrast to plasmids where they are common, the presence of insertion sequences is unusual in phages, and in temperate phages specifically (Leclercq and Cordaux, 2011). One survey of phage genomes found 92% of active temperate phages (i.e., excluding immobile/partial prophages) had no detectable insertion elements (Leclercq and Cordaux, 2011). This is likely due to the strict genome size constraints imposed by phage capsid size, the density of essential genes in phage genomes, and the huge population sizes of active phages resulting in a strong selection pressure against insertion element retention. In the *R. leguminosarum* phages, for which lysogen strains were available, the presence of insertion elements does not impair the ability of the phages to produce functional infective phage particles as all prophages tested could be induced. In addition to the IS66 element phages vTRX32-1 and vSM95-1 also carry an additional transposase, yet both have been shown to form efficient lytic infections, suggesting the phages are capable of accommodating additional novel DNA into their genome without impairing the packaging of phage DNA.

In addition to these insertion elements, four of the 10 phages carry unique genes that do not appear on other phages in the clade suggesting they have been acquired by horizontal

gene transfer. Interestingly these genes are found in the same location—directly downstream of the RNA polymerase gene. In two cases functional information indicates that these additional genes may be adaptive. Firstly, the lytic phage RHEph01 contains a Bro-domain putative antirepressor. Antirepressors are involved in prophage lysis-lysogeny decisions, which may provide an explanation for the lytic nature of this phage. When integrated into the bacterial genome, most prophages express a repressor protein which silences phage gene transcription and prevents the phage excising, producing infectious virus particles and lysing the host (Bednarz et al., 2014). Antirepressors, such as Bro domain proteins represent one mechanism for relieving repression of phage genes when entering lysis (Lemire et al., 2011) and have been identified in numerous prophages (Botstein et al., 1975; Farlow et al., 2018). Interestingly, however, the absence of this gene on the 9 other phage genomes suggests that it has been acquired independently and is not required to moderate the lysogeny-lysis switch. While antirepressors are typically associated with integrated prophages exiting lysogeny, it is possible that the presence of this gene could explain the inability of RHEph01 to form stable lysogenic infections. For example, overactive expression of the antirepressor gene may inhibit repressor activity entirely, preventing stable lysogeny from establishing. Further work will be needed to confirm the

role of this gene in the conversion of phage RHEph01 to a purely lytic lifestyle.

Secondly *Rhizobiales* sp. AFS092257 prophage contains a putative cytosine-methyltransferase. Orphan DNA-methyltransferases have been identified on numerous phage genomes and have been linked to protection against host restriction enzymes used to defend hosts from phage infection (Murphy et al., 2013, 2014). DNA-methyltransferases methylate restriction sites within the phage genome, preventing host restriction enzymes from cleaving phage DNA (Murphy et al., 2013, 2014). For example a methyltransferase encoded on the genome of the *Bacillus subtilis* prophage SP $\beta$  becomes active after the prophage is induced, methylating newly replicated phage DNA at BsuR restriction sites and as a result protecting them from the BsuR endonuclease (Trautner et al., 1980). It may be that the AFS092257 prophage cytosine-methyltransferase performs a similar protective function.

Finally vTRX32-1, and the prophage identified in *R. anhuiense* CCBAU23252 each carry a unique ORF at the same site as these two genes, with no homology to each other and no functional domains which could be identified. The complete lack of homology between genes at this site suggest this location is a hotspot for gene gain and loss.

## Relatedness and Evidence for Transfer of *Rhizobium* Phages

Patterns of relatedness between phages were estimated using a 5-concatenated-gene phylogeny and compared to their host strain phylogeny (Figure 5). The phage phylogeny based on core genes follows the general pattern observed from ANI calculated across the whole phage genomes (Supplementary Figure 1), with prophages isolated from *Rlt* strains (vTRX32-1, vSM48, vSM95, vSM113, and vSM135A) being most closely related to each other and the five additional phages more distantly related. Within the host tree 7 additional *Rlt* strains were included which lack any prophages at the attB site demonstrating that lysogenic strains are nested within closely related but non-lysogenic strains. Indeed, the four SM-strain lysogens (Cavassim et al., 2020) were identified within a wider strain collection of 196 fully *Rlt* sequenced strains. Detailed assessment of these genomes (through both blastn and inspection of the highly conserved insertion site) found no intact or degraded vTRX32-1-like phages and attB insertion sites were intact.

Analysis of co-phylogeny was performed for the 9 prophage and lysogen matched pairs using empress (Santichaivekin et al., 2020), which calculates maximum parsimony reconciliations (MPRs) for patterns of relatedness between 2 groups of species (e.g., hosts and symbionts). Relative costs of four possible evolutionary events were set as: cospeciation = 0, transfer = 1, duplication = 2 and loss was modeled at two levels, 1 and 2. Where the likelihood of loss events was equal to transfer events (loss = 1) reconstruction estimated that the pattern of phage distribution within the host phylogeny could be explained by 4 cospeciation events, 2 losses and 4 transfer events (Figure 6A). Where loss events are less likely than transfer events (loss = 2) reconstruction estimates that the pattern of distribution is

predominantly explained by transfers events (7) rather than cospeciation events (1) (Figure 6B). Under both scenarios the distribution of phages and hosts is not significantly different from that of random trees (loss = 1,  $P < 0.168$ , loss = 2,  $P = 0.693$ ), which indicates that processes such as transfer rather than simple cospeciation is the dominant process explaining phage distribution. This is supported by the relative rarity of phages in closely related populations and lack of evidence for degraded phage genomes within the insertion site.

Comparison of phage RNA polymerase sequences places the vTRX32-1-like group within the newly demarcated family *Autographiviridae*, previously a subfamily of the *Podoviridae* (Walker et al., 2019). Phages from this clade infect a wide diversity of bacteria spanning Alphaproteobacteria (including *Rhizobium*), Betaproteobacteria, and Gammaproteobacteria (Figure 7).

## CONCLUSION

The vTRX32-1 like phages are a novel group of temperate phages infecting *Rhizobium leguminosarum*, closely related to the previously described lytic phage RHEph01—suggesting this phage is ancestrally temperate. In addition we identified a wider group of phage genomes within other *Rhizobium* species indicating that these phages form part of a group of temperate phages which have a long evolutionary history within *Rhizobium* spp.

Phenotypic demonstrate that the vTRX-1-like phages are inducible, widely infectious, able to form new lysogenic partnerships. Intriguingly, no degraded versions of these phages were identified within 192 other *Rlt* strains suggesting both that they are being maintained by active mobilization and are not under selection for degradation and loss of function at least in these populations. This supports the findings from the phylogenetic distribution of the wider phage group, which implies a significant contribution of horizontal transmission to phage distribution over longer evolutionary timescales. We also identified instances of horizontal gene transfer within the phage genomes including insertion sequences—which are known to play a role in *Rhizobium* evolution—and several unique genes which are mostly likely to have been acquired following divergence between phages.

Temperate phages are relatively unstudied in terrestrial species such as *Rhizobium*, with very few published examples (Halmillawewa et al., 2016). In both pathogenic and marine ecosystems temperate phages play a significant role in bacterial ecology and evolution. The presence of this active group of temperate phages suggests they may have a role to play in *Rhizobia* communities.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/genbank/>, MW023914.1.

## AUTHOR CONTRIBUTIONS

SM and EH contrived and conducted experiments and analysis for the study. PY, SM, and RS provided strains and advised on analysis. EH awarded the funding. All authors contributed to the preparation of the manuscript.

## FUNDING

This work was funded by an NERC independent fellowship NE/P017584/1 and BBSRC standard grant BB/R014884/1.

## ACKNOWLEDGMENTS

We would like to thank Stig U. Andersen and Victor Gonzalez for providing strains from their collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.765271/full#supplementary-material>

**Supplementary Figure 1** | VIRIDIC genome comparisons for *Rhizobium* phages. Grid shows pairwise comparisons for each of the 10 phage genomes. Squares in

the top right show ANI values for whole phage genomes with intensity of color denoting intergenomic similarity. Squares in the bottom left half show the genome length ratio (numbered) and the proportion of sequence aligned for each comparison (shading). Highlighted squares represent the 5 phages isolated from *Rhizobium leguminosarum* (green box) and the phages which share > 70% ANI, and are therefore likely members of the *Paadmavirus* genus (pink).

**Supplementary Figure 2** | Illustrating the integration of the vTRX32-1 like phages into the *Rhizobium* genome (based on genome comparisons of vTRX32-1 and vRHEph01 with closely related prophage). The tRNA-Leu in the *Rhizobium* genome contains a 12 bp motif also carried by the phage (red), this acts as the region of homology (attB/attP site) for phage integration. In the process of integration the bacterial tRNA-Leu (yellow) is interrupted and replaced by a phage encoded tRNA-Leu (yellow, hashed) in the phage genome.

**Supplementary Figure 3** | Percent identity between amino acid sequences predicted by RAST in the vTRX32-1 genome and those of related phage/prophage. Genes (X-axis) are ordered from the highest percent identity (left) to the lowest (right), phage/prophage genomes (Y-axis) are ordered similarly from top to bottom. Colored highlights adjacent to gene names indicate predicted class membership (yellow = early/class I, blue = middle/class II, red = late/class III). Pairwise blastP was used to generate percent identity values and only hits with an E-value less than 1E-4 were included.

**Supplementary Table 1** | Details of phages and putative phages identified as related to vTRX32-1.

**Supplementary Table 2** | Annotated data sets from the 3 phenotypic assays; phage virulence, phage replication dynamics (one step growth curve) and host range. Details are included in each dataframe.

**Supplementary Data Sheet 1** | Bioinformatics pipeline for phage genome assembly and provisional annotation.

## REFERENCES

- Abedon, S. T. (2019). Look Who's talking: T-Even phage lysis inhibition, the granddaddy of virus-virus intercellular communication research. *Viruses* 11:951. doi: 10.3390/v111100951
- Ahmad, M. H., and Morgan, V. (1994). Characterization of a cowpea (*Vigna unguiculata*) rhizobiophage and its effect on cowpea nodulation and growth. *Biol. Fertil. Soils* 18, 297–301. doi: 10.1007/BF00570632
- Allison, H. E. (2007). Stx-phages: drivers and mediators of the evolution of STEC and STEC-like pathogens. *Future Microbiol.* 2, 165–174. doi: 10.2217/17460913.2.2.165
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75
- Bednarz, M., Halliday, J. A., Herman, C., and Golding, I. (2014). Revisiting bistability in the lysis/lysogeny circuit of bacteriophage lambda. *PLoS One* 9:e100876. doi: 10.1371/journal.pone.0100876
- Botstein, K., Lew, K. K., Jarvik, V., and Swanson, C. A. (1975). Role of antirepressor in the bipartite control of repression and immunity by bacteriophage P22. *J. Mol. Biol.* 91, 439–462. doi: 10.1016/0022-2836(75)90271-5
- Brown, S. P., Fredrik Inglis, R., and Taddei, F. (2009). Evolutionary ecology of microbial wars: within-host competition and (incidental) virulence. *Evol. Appl.* 2, 32–39. doi: 10.1111/j.1752-4571.2008.00059.x
- Buchanan-Wollaston, V. (1979). Generalized transduction in *Rhizobium leguminosarum*. *J. Gen. Microbiol.* 112, 135–142. doi: 10.1099/00221287-112-1-135
- Burns, N., James, C. E., and Harrison, E. (2015). Polylysogeny magnifies competitiveness of a bacterial pathogen in vivo. *Evol. Appl.* 8, 346–351. doi: 10.1111/eva.12243
- Cavassim, M. I. A., Moeskjær, S., Moslemi, C., Fields, B., Bachmann, A., Vilhjálmsson, B. J., et al. (2020). Symbiosis genes show a unique pattern of introgression and selection within a *Rhizobium leguminosarum* species complex. *Microb. Genom.* 6:e000351. doi: 10.1099/mgen.0.000351
- Chan, P. P., and Lowe, T. M. (2019). tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* 1962, 1–14. doi: 10.1007/978-1-4939-9173-0\_1
- De Paepe, M., Hutinet, G., Son, O., Amarir-Bouhram, J., Schbath, S., and Petit, M.-A. (2014). Temperate phages acquire DNA from defective prophages by relaxed homologous recombination: the role of Rad52-like recombinases. *PLoS Genet.* 10:e1004181. doi: 10.1371/journal.pgen.1004181
- Decewicz, P., Radlinska, M., and Dziewit, L. (2017). Characterization of *Sinorhizobium* sp. LM21 prophages and virus-encoded DNA methyltransferases in the light of comparative genomic analyses of the sinorhizobial virome. *Viruses* 9:161. doi: 10.3390/v9070161
- Dhar, B., Jaiswal, S. K., Kumar, B., and Vaishampayan, A. (2012). Characteristics and host range nodulation of a lysogenic *Bradyrhizobium* strain of pigeonpea (*Cajanus cajan* L.) isolated from an Indian soil. *Biol. Fertility Soils* 49, 343–349. doi: 10.1007/s00374-012-0728-728
- Dunn, J. J., and Studier, F. W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* 166, 477–535. doi: 10.1016/S0022-2836(83)80282-4
- Dziewit, L., Oscik, K., Bartosik, D., and Radlinska, M. (2014). Molecular characterization of a novel temperate sinorhizobium bacteriophage, ÖLM21, encoding DNA methyltransferase with CcrM-like specificity. *J. Virol.* 88, 13111–13124. doi: 10.1128/JVI.01875-14
- Evans, J., Barnet, Y. M., and Vincent, J. M. (1979). Effect of a bacteriophage on colonisation and nodulation of clover roots by paired strains of *Rhizobium trifolii*. *Can. J. Microbiol.* 25, 974–978. doi: 10.1139/m79-149
- Farlow, J., Bolkvadze, D., Leshkasheli, L., Kusradze, I., Kotorashvili, A., Kotaria, N., et al. (2018). Genomic characterization of three novel Basilisk-like phages infecting *Bacillus anthracis*. *BMC Genomics* 19:685. doi: 10.1186/s12864-018-5056-4
- Finan, T. M., Hartweg, E., LeMieux, K., Bergman, K., Walker, G. C., and Signer, E. R. (1984). General transduction in *Rhizobium meliloti*. *J. Bacteriol.* 159, 120–124. doi: 10.1128/jb.159.1.120-124.1984

- González, V., Bustos, P., Ramírez-Romero, M. A., Medrano-Soto, A., Salgado, H., Hernández-González, I., et al. (2003). The mosaic structure of the symbiotic plasmid of *Rhizobium etli* CFN42 and its relation to other symbiotic genome compartments. *Genome Biol.* 4:R36. doi: 10.1186/gb-2003-4-6-r36
- González, V., Santamaría, R. I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., et al. (2006). The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc. Natl. Acad. Sci. U S A.* 103, 3834–3839. doi: 10.1073/pnas.0508502103
- Halmillawewa, A. P., Restrepo-Córdoba, M., Perry, B. J., Yost, C. K., and Hynes, M. F. (2016). Characterization of the temperate phage vB\_RleM\_PPF1 and its site-specific integration into the *Rhizobium leguminosarum* F1 genome. *Mol. Genet. Genomics* 291, 349–362. doi: 10.1007/s00438-015-1113-8
- Halmillawewa, A. P., Restrepo-Córdoba, M., Yost, C. K., and Hynes, M. F. (2015). Genomic and phenotypic characterization of *Rhizobium gallicum* phage vB\_RglS\_P106B. *Microbiology* 161, 611–620. doi: 10.1099/mic.0.000022
- Hamdi, S., Rousseau, G. M., Labrie, S. J., Kourda, R. S., Tremblay, D. M., Moineau, S., et al. (2016). Characterization of five podoviridae phages infecting *Citrobacter freundii*. *Front. Microbiol.* 7:1023. doi: 10.3389/fmicb.2016.01023
- Harrison, E., and Brockhurst, M. A. (2017). Ecological and evolutionary benefits of temperate phage: what does or doesn't kill you makes you stronger. *Bioessays* 39:1700112. doi: 10.1002/bies.201700112
- Hashem, F. M., and Angle, J. S. (1988). Rhizobiophage effects on *Bradyrhizobium japonicum*, nodulation and soybean growth. *Soil Biol. Biochem.* 20, 69–73. doi: 10.1016/0038-0717(88)90128-9
- Hernandez-Lucas, I., Ramirez-Trujillo, J. A., Gaitan, M. A., Guo, X., Flores, M., Martinez-Romero, E., et al. (2006). Isolation and characterization of functional insertion sequences of rhizobia. *FEMS Microbiol. Lett.* 261, 25–31. doi: 10.1111/j.1574-6968.2006.00319.x
- International Committee on Taxonomy of Viruses [ICTV] (2021). *International Committee on Taxonomy of Viruses*. Available online at: <https://ictv.global/taxonomy/> (accessed July 9, 2021).
- Jakočiūnė, D., and Moodley, A. (2018). A rapid bacteriophage DNA extraction method. *Methods Protoc.* 1:27. doi: 10.3390/mps1030027
- Juhala, R. J., Ford, M. E., Duda, R. L., Youlton, A., Hatfull, G. F., and Hendrix, R. W. (2000). Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdaoid bacteriophages. *J. Mol. Biol.* 299, 27–51. doi: 10.1006/jmbi.2000.3729
- Kawasaki, T., Narulita, E., Matsunami, M., Ishikawa, H., Shimizu, M., Fujie, M., et al. (2016). Genomic diversity of large-plaque-forming podoviruses infecting the phytopathogen *Ralstonia solanacearum*. *Virology* 492, 73–81. doi: 10.1016/j.virol.2016.02.011
- Kumar, N., Lad, G., Giuntini, E., Kaye, M. E., Udomwong, P., Shamsani, N. J., et al. (2015). Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*. *Open Biol.* 5:140133. doi: 10.1098/rsob.140133
- Labrie, S. J., Frois-Moniz, K., Osburne, M. S., Kelly, L., Roggensack, S. E., Sullivan, M. B., et al. (2013). Genomes of marine cyanopodoviruses reveal multiple origins of diversity. *Environ. Microbiol.* 15, 1356–1376. doi: 10.1111/1462-2920.12053
- Leclercq, S., and Cordaux, R. (2011). Do phages efficiently shuttle transposable elements among prokaryotes? *Evolution* 65, 3327–3331. doi: 10.1111/j.1558-5646.2011.01395.x
- Lemire, S., Figueroa-Bossi, N., and Bossi, L. (2011). Bacteriophage crosstalk: coordination of prophage induction by trans-acting antirepressors. *PLoS Genet.* 7:e1002149. doi: 10.1371/journal.pgen.1002149
- Lindell, D., Jaffe, J. D., Coleman, M. L., Futschik, M. E., Axmann, I. M., Rector, T., et al. (2007). Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* 449, 83–86. doi: 10.1038/nature06130
- Lindell, D., Sullivan, M. B., Johnson, Z. I., Tolonen, A. C., Rohwer, F., and Chisholm, S. W. (2004). Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc. Natl. Acad. Sci. U S A.* 101, 11013–11018. doi: 10.1073/pnas.0401526101
- Machida, Y., Sakurai, M., Kiyokawa, S., Ubasawa, A., Suzuki, Y., and Ikeda, J. E. (1984). Nucleotide sequence of the insertion sequence found in the T-DNA region of mutant Ti plasmid pTiA66 and distribution of its homologues in octopine Ti plasmid. *Proc. Natl. Acad. Sci. U S A.* 81, 7495–7499. doi: 10.1073/pnas.81.23.7495
- Moraru, C., Varsani, A., and Kropinski, A. M. (2020). VIRIDIC-A novel tool to calculate the intergenomic similarities of prokaryote-infecting viruses. *Viruses* 12:1268. doi: 10.3390/v12111268
- Msimbira, L. A., Jaiswal, S. K., and Dakora, F. D. (2016). Identification and characterization of phages parasitic on bradyrhizobia nodulating groundnut (*Arachis hypogaea* L.) in South Africa. *Appl. Soil Ecol.* 108, 334–340. doi: 10.1016/j.apsoil.2016.09.010
- Mularoni, L., Zhou, Y., Bowen, T., Gangadharan, S., Wheelan, S. J., and Boeke, J. D. (2012). Retrotransposon Ty1 integration targets specifically positioned asymmetric nucleosomal DNA segments in tRNA hotspots. *Genome Res.* 22, 693–703. doi: 10.1101/gr.129460.111
- Murphy, J., Klumpp, J., Mahony, J., O'Connell-Motherway, M., Nauta, A., and van Sinderen, D. (2014). Methyltransferases acquired by lactococcal 936-type phage provide protection against restriction endonuclease activity. *BMC Genomics* 15:831. doi: 10.1186/1471-2164-15-831
- Murphy, J., Mahony, J., Ainsworth, S., Nauta, A., and van Sinderen, D. (2013). Bacteriophage orphan DNA methyltransferases: insights from their bacterial origin, function, and occurrence. *Appl. Environ. Microbiol.* 79, 7547–7555. doi: 10.1128/AEM.02229-13
- Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., et al. (2010). Antagonistic coevolution accelerates molecular evolution. *Nature* 464, 275–278. doi: 10.1038/nature08798
- Reiter, W. D., Palm, P., and Yeats, S. (1989). Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. *Nucleic Acids Res.* 17, 1907–1914. doi: 10.1093/nar/17.5.1907
- Rochepeau, P., Fellay, R., Jabbouri, S., Perret, X., and Broughton, W. J. (1997). Region II of *Rhizobium* sp. NGR234 inhibits nodulation of *Medicago sativa* by R. meliloti nodIJ and nodQ1 mutants. *Mol. Plant. Microbe. Interact.* 10, 978–983. doi: 10.1094/MPMI.1997.10.8.978
- Ross, A., Ward, S., and Hyman, P. (2016). More is better: selecting for broad host range bacteriophages. *Front. Microbiol.* 7:1352. doi: 10.3389/fmicb.2016.01352
- Sampaio, M., Rocha, M., Oliveira, H., and Dias, O. (2019). Predicting promoters in phage genomes using PhagePromoter. *Bioinformatics* 35, 5301–5302. doi: 10.1093/bioinformatics/btz580
- Santamaría, R. I., Bustos, P., Sepúlveda-Robles, O., Lozano, L., Rodríguez, C., Fernández, J. L., et al. (2014). Narrow-host-range bacteriophages that infect *Rhizobium etli* associate with distinct genomic types. *Appl. Environ. Microbiol.* 80, 446–454. doi: 10.1128/AEM.02256-13
- Santichaivekin, S., Yang, Q., Liu, J., Mawhorter, R., Jiang, J., Wesley, T., et al. (2020). eMPress: a systematic copyphylogeny reconciliation tool. *Bioinformatics* 37, 2481–2482. doi: 10.1093/bioinformatics/btaa978
- Semsey, S., Blaha, B., Köles, K., Orosz, L., and Papp, P. P. (2002). Site-specific integrative elements of rhizobiophage 16-3 can integrate into proline tRNA (CGG) genes in different bacterial genera. *J. Bacteriol.* 184, 177–182. doi: 10.1128/JB.184.1.177-182.2002
- Shahaby, A. F., Alharthi, A. A., El-Tarras, A. E., and Others. (2014). Molecular characterization of rhizobiophages specific for *Rhizobium* sp. *Sinorhizobium* sp., and *Bradyrhizobium* sp. *Int. J. Curr. Microbiol. Appl. Sci.* 3, 155–171.
- Studier, F. W. (1972). Bacteriophage T7. *Science* 176, 367–376. doi: 10.1126/science.176.4033.367
- Sullivan, M. B., Coleman, M. L., Weigele, P., Rohwer, F., and Chisholm, S. W. (2005). Three *Prochlorococcus cyanophage* genomes: signature features and ecological interpretations. *PLoS Biol.* 3:e144. doi: 10.1371/journal.pbio.0030144
- Topka, G., Bloch, S., Nejman-Faleńczyk, B., Gąsior, T., Jurczak-Kurek, A., Necel, A., et al. (2018). Characterization of bacteriophage vB-EcoS-95, isolated from urban sewage and revealing extremely rapid lytic development. *Front. Microbiol.* 9:3326. doi: 10.3389/fmicb.2018.03326
- Trautner, T. A., Pawlek, B., Günther, U., Canosi, U., Jentsch, S., and Freund, M. (1980). Restriction and modification in *Bacillus subtilis*: identification of a gene in the temperate phage SP beta coding for a BsuR specific modification methyltransferase. *Mol. Gen. Genet.* 180, 361–367. doi: 10.1007/BF00425849
- Vitousek, P. M., Cassman, K., Cleveland, C., Crews, T., Field, C. B., Grimm, N. B., et al. (2002). "Towards an ecological understanding of biological nitrogen fixation," in *The Nitrogen Cycle at Regional to Global Scales*, eds E. W. Boyer and R. W. Howarth (Dordrecht: Springer Netherlands), 1–45. doi: 10.1007/978-94-017-3405-9\_1

- Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Dempsey, D. M., Dutilh, B. E., et al. (2019). Changes to virus taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses (2019). *Arch. Virol.* 164, 2417–2429. doi: 10.1007/s00705-019-04306-w
- Young, J. M., Kuykendall, L. D., Martínez-Romero, E., Kerr, A., and Sawada, H. (2001). A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *Int. J. Syst. Evol. Microbiol.* 51, 89–103. doi: 10.1099/00207713-51-1-89
- Zhao, Y., Qin, F., Zhang, R., Giovannoni, S. J., Zhang, Z., Sun, J., et al. (2019). Pelagiphages in the Podoviridae family integrate into host genomes. *Environ. Microbiol.* 21, 1989–2001. doi: 10.1111/1462-2920.14487
- Zhao, Y., Temperton, B., Thrash, J. C., Schwalbach, M. S., Vergin, K. L., Landry, Z. C., et al. (2013). Abundant SAR11 viruses in the ocean. *Nature* 494, 357–360. doi: 10.1038/nature11921
- Zhu, J., Rao, X., Tan, Y., Xiong, K., Hu, Z., Chen, Z., et al. (2010). Identification of lytic bacteriophage Mmp1, assigned to a new member of T7-like phages infecting *Morganella morganii*. *Genomics* 96, 167–172. doi: 10.1016/j.ygeno.2010.06.001
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Ford, Moeskjær, Young, Santamaría and Harrison. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.