Functional ecology of bacteriophages in the environment

Richard J Puxty[1] and Andrew D Millard[2]

Check for updates

Bacteriophages are as ubiquitous as their bacterial hosts and often more abundant. Understanding how bacteriophages control their bacterial host populations requires a number of different approaches. Bacteriophages can control bacterial populations through lysis, drive evolution of bacterial immunity systems through infection, provide a conduit for horizontal gene transfer and alter host metabolism by carriage of auxiliary metabolic genes. Understanding and quantifying how bacteriophages drive these processes, requires both technological developments to take measurements *in situ*, and laboratory-based studies to understand mechanisms. Technological advances have allowed quantification of the number of infected cells *in situ*, revealing far-lower levels than expected. Understanding how observations in laboratory conditions relate to what occurs in the environment, and experimental confirmation of the predicted function of phage genes from observations in environmental omics data, remains challenging.

Addresses
[1] University of Warwick, School of Life Sciences, Gibbet Hill Road, Coventry CV4 7AL, United Kingdom
[2] University of Leicester, Dept of Genetics and Genome Biology, University Road, Leicester, United Kingdom

Corresponding authors: Puxty, Richard J (R.Puxty@warwick.ac.uk), Millard, Andrew D (adm39@le.ac.uk)

Introduction

At the most basic level, bacteriophages (phages) are simply nucleic acids encased in a protein capsid [1]. However, these small biological entities are machiavellian masters of the microbial world. The sheer abundance of phages means the lysis of their bacterial host has an important role in the ecology of some environments. For example, in the Oceans, where they are estimated to kill 20% of bacteria per day and divert the flow of carbon and other nutrients from higher trophic levels, via the viral shunt (see [2] for review). Phages also have other roles by acting as mediators of lateral gene transfer, driving the evolution of bacterial immune systems, altering host metabolism and shaping bacterial populations through their continual interaction. Within this review, we will discuss recent advances in viromics that are further informing our understanding of phage diversity. We further review recent lab studies on model systems that have pertinent implications for phage-functional ecology, and pose important outstanding questions for phage ecologists.

The study of viruses through viromics, be that from the identification of viruses from bulk metagenomes or enrichment of virions before sequencing, has revolutionised our understanding of phage diversity compared with culture-based approaches [3]. Vastly expanding our understanding of phage diversity and distribution of phages in multiple environments. Long-read sequencing technologies have shown between 22% and 160% more viral Operational Taxonomiv Units (vOTUs) can be detected when long reads are used compared with short reads alone [4–6], thus, we are likely missing much of the diversity of even the most abundant phages in many environments. As the requirement for large quantities of DNA for long-read sequencing is overcome by technological approaches of using virION [6,7] or Multiple displacement amplification [5,8], further phage diversity will likely be uncovered in already well-sampled environments.

While viromics has produced new data, it is culture-based approaches that unveil further biases in viromic datasets. Several culture-based studies have identified the modification of phage DNA makes it recalcitrant to standard DNA library preparation techniques [20–22] and thus will be missing from most viromes. How much we are missing is yet to be quantified. In a study that used metatranscriptomic data and DNA metagenomes to identify dsDNA phages, ~50% of phages were not found in the DNA fraction [9]. What proportion of those phages only detected in the metatranscriptomic were because of DNA modification, preventing detection in DNA samples is unknown. But, it does suggest we could still be missing a large proportion of dsDNA and the use of multiple omic technologies will provide better estimates of the diversity of phages within an environment.

The recent utilisation of transcriptomic datasets [9–11] and specific sequencing of virions containing RNA [12,13] has vastly expanded the diversity of RNA viruses, with a doubling in the number of phyla [11]. Of genome-sequenced phage isolates, only ∼0.1% are from RNA phages. To understand the role of RNA phages in the environment, efforts are needed to bring more into culture to develop model systems.

Population control

Despite several decades of work [14], methods to directly assess the rates of stress imposed by phages have been challenging. Phages span from being obligately virulent (i.e. antagonistic), through to temperate, which in many cases may confer a fitness advantage (i.e. mutualistic). For virulent phages, it has been relatively straightforward to enumerate free-phage particles from samples. More challenging has been converting these numbers to mortality rates of susceptible hosts. This requires knowing who infects whom and the number of cells infected at any time. The iPolony method allows direct quantification of the number of infected cells, and thus, deducing of mortality rates [15]. When applied to marine cyanobacteria, these data suggest relatively low levels of infection *in situ* (< 1% of cells), despite high abundances of free phages within samples. This is a conceptual challenge for phage ecologists to balance apparent infection rates with those predicted from encounter theory. Three mechanisms have been proposed to explain this paradox, namely, that a large fraction of free phage is not infectious, lower rates of adsorption *in situ* than measured in the laboratory or high rates of resistant cells in the environment.

The relative contributions of these processes to the observed discrepancy are currently unclear. Laboratory experiments show antagonistic coevolution between host and phage is rapid [16], but how quickly these scenarios play out in nature remains poorly resolved. In one study, *Pseudomonas fluorescens* clones isolated from soil are broadly resistant to phages isolated from the past and similarly 'future' phages are broadly infectious against past bacteria [17]. Thus, at least phenotypically, timescales were over several weeks. Meanwhile, wild strains of *Flavobacterium columnare* and their phages undergo antagonistic coevolution within a ten-year time series [18]. In both these examples, infectivity and resistance increase linearly over time, indicative of 'arms-race' dynamics. In contrast, time-series infection networks of *Vibrio crassostreae* and their phages in an oyster farm setting show that phages are 'locally-adapted' in time [19]. That is, they broadly infect host clones isolated from the same time but are less likely to infect hosts from the past or the future. This argues for a 'negative-frequency dependent selection' mechanism operating *in situ* [16].

With respect to temperate phages, around two-thirds of genomes from isolated cultures have at least one prophage-like signal (Thomas Sicheritz-Ponten *pers. comms.*). This may depend on environmental contexts, with reports of only 18% of genomes from marine bacteria containing a prophage [20]. Conflicting data from single-amplified genomes from a sample of > 5000 marine bacteria show all genomes contained at least one prophage-like signal [21]. However, it is currently unclear whether these prophage prediction tools are accurate [21,22]. A major focus of the community should be validation of these tools in the lab. Nevertheless, these results pose interesting questions for phage ecologists. For example, why are not all bacteria lysogens? Why are some bacteria more susceptible to prophage than others? Recent work described fitness costs and benefits to prophage carriage in varying environmental contexts. In competition assays with prophage-free strains, lysogens have either a fitness cost or benefit in ordinary laboratory conditions [23]. The fitness peaks then shift depending on certain environmental conditions (e.g. the presence of antibiotics). Moreover, direct effect on fitness also depends on the genes encoded by the prophage in addition to prophage carriage itself [23].

Beyond the laboratory, researchers have sought to correlate metadata with observations of prophage carriage to better understand the evolutionary basis for lysogeny. The Piggyback-the-Winner hypothesis stipulates that a temperate mode is selected for when host population density is high [24]. Correlational analysis supports this theory, yet there is some doubt over the validity of these correlations in providing a mechanistic explanation between cell density and the lysis-lysogeny decision [25]. This mechanism has been explored using modelling approaches [26], which have recently been adapted to suggest a non-linear trend between cell density and frequency of lysogeny [27], whereby lysogeny is promoted at both high and low cell densities. This relationship forms through interactions of cell density, fitness benefits to lysogenisation and energetic status of the bacterial cell.

It is clear that experimental studies on the rates of population control by phages are disparate, have produced contrasting results and have not provided a unifying theory of phage control of bacterial populations. Indeed, one may not exist. However, future experimental approaches should build on the time-series studies and integrate isolation, genomics, estimates of the frequency of infected cells and modelling.

Evolution of bacterial immunity

High-throughput, forward genetic screens are expanding the diversity of described bacterial immune systems (for review see [28]). Yet, perhaps understudied, is the

relevance of such systems for phage-host interactions in the environment. That is, can we predict who infects whom from the complement of immune and immune evasion systems? In the case of clustered regularly interspaced short palindromic repeats (CRISPR), decades of research has unmasked the myriad ways to achieve CRISPR immunity (for review see [29]). Thus, there have been many attempts to use the sequence information contained in spacers to predict phage-host relationships [30–32]. Yet these attempts are conceptually flawed, not least because the spacer contains a record of past infections, not current ones. More importantly, in recent years, there has been a rapid expansion in the catalogue of ways that phages subvert CRISPR [33–36]. Interestingly, many of these anti-CRISPR systems are encoded by phages or plasmids [37,38]. This extends to many other phage defense systems. For instance, a functional genetic screen of *Escherichia coli* strains discovered an overwhelming propensity of novel defense systems within intact prophages and other mobile elements [39]. Therefore, phages may be important in controlling bacterial populations indirectly through phage–phage antagonism.

It is also true that in non-CRISPR systems, phages evolve rapid counter-defence. Phage T4 rapidly evolves segmental amplification mutations to evade the ToxIN toxin–anti-toxin defence system [40]. Interestingly, these amplifications result in deletions elsewhere in the phage chromosome, presumably to maintain packaging fidelity. These deletions happen to be regions encoding other counter-defence systems relevant when infecting other hosts. Thus, this experiment provides the first mechanistic explanation of why arms-races may come at a cost and therefore not continue *ad infinitum* [16].

Time-series metagenomics of a biological wastewater treatment plant has sought to resolve how bacterial immune systems shape the dynamics of phage-host interactions *in situ* [31]. Using CRISPR as a model, it is clear that CRISPR spacers are extremely dynamic in time, despite microbial communities remaining relatively constant. This implicates CRISPR in being important for structuring phage-host networks at fine-scale taxonomic resolution (i.e. strain level). Similar conclusions are drawn when attempting to model the temporal behaviour of host and virus immunity [41] and are likely to extend to other defense systems beyond CRISPR.

Active metabolism

The advent of genome sequencing revealed phages could potentially alter the metabolism of their host by the carriage of genes termed auxiliary metabolic genes (AMGs) (see [42] for a review). With increased (meta)-genomic data, the number of putative AMGs involved in nitrogen and phosphorus cycles has increased [43–45].

Whether all putative phage AMGs are real and not misidentification of phage contigs or assembly artefacts, cannot be fully ruled out until they are observed in phage isolates. Greater confidence in predictions occurs when related AMGs are repeatedly found on predicted phage contigs (containing phage hallmark genes) from multiple environments. Such is the case for AMGs associated with dissimilatory sulphur metabolism, that have been found on ~200 phage contigs from a variety of environments [46]. Comparison of the abundance of phage:host transcription ratios of these key AMGs, demonstrated they are also active and expressed at high levels [46]. While transcriptomics clearly suggests these genes are important, without cultured phage-host systems, it is difficult to decipher the mechanism or consequence of phages altering host sulphur metabolism.

The very nature of metagenomics and the limited number of phages in cultures, means testing the function of most AMGs remains elusive. However, the use of heterologous expression systems can be used to infer the function of AMGs. One such example is the expression of a phage-encoded CAZyme that cleaves the β-1,4-linked mannose units in galactomannan and glucomannan [47]. Suggesting phages may contribute to the breakdown of plant-derived polymers. While the function of one phage CAZyme has been confirmed, a plethora of different CAZyme types have been identified in a range of environments, including mangroves [48], and agroecosystems used for rice cultivation [49]. With both the abundance and type of phage-encoded CAZymes differing between bulk soil and the rhizosphere, and the type of cropping system used for cultivation of rice [49,50]. However, if the CAZymes are all functional and how much they contribute to the breakdown of complex carbon and release simple sugars into the environment as predicted remains unknown. Without cultured phage-host systems, the prediction of such contributions to the carbon cycle remains difficult to quantify.

Even when phage-host systems are available, it is not a simple task to determine the functional role of AMGs. The study of photosynthetic genes found in marine cyanophages, exemplifies how confirmation of prediction function is crucial. With the discovery of *psbAD* in cyanophages, it was hypothesised that cyanophages carry these genes to allow a repair cycle to operate in PSII, maintaining host photosynthetic activity and accompanying oxygen evolution [51]. Utilising a model phage-host system found the proposed hypothesis is only partly true. While phage versions of the proteins D1 and D2 are produced and the photosynthetic electron transport function of the host is maintained [52–54]. Unexpectedly, the fixation of CO_2 was halted during cyanophage infection, decoupling the light and dark reactions of photosynthesis [54]. As a consequence of this decoupling, cyanophages are estimated to result in

reduced CO_2 fixation on a global scale of between 0.02 and 5.39 Pg C per year [54].
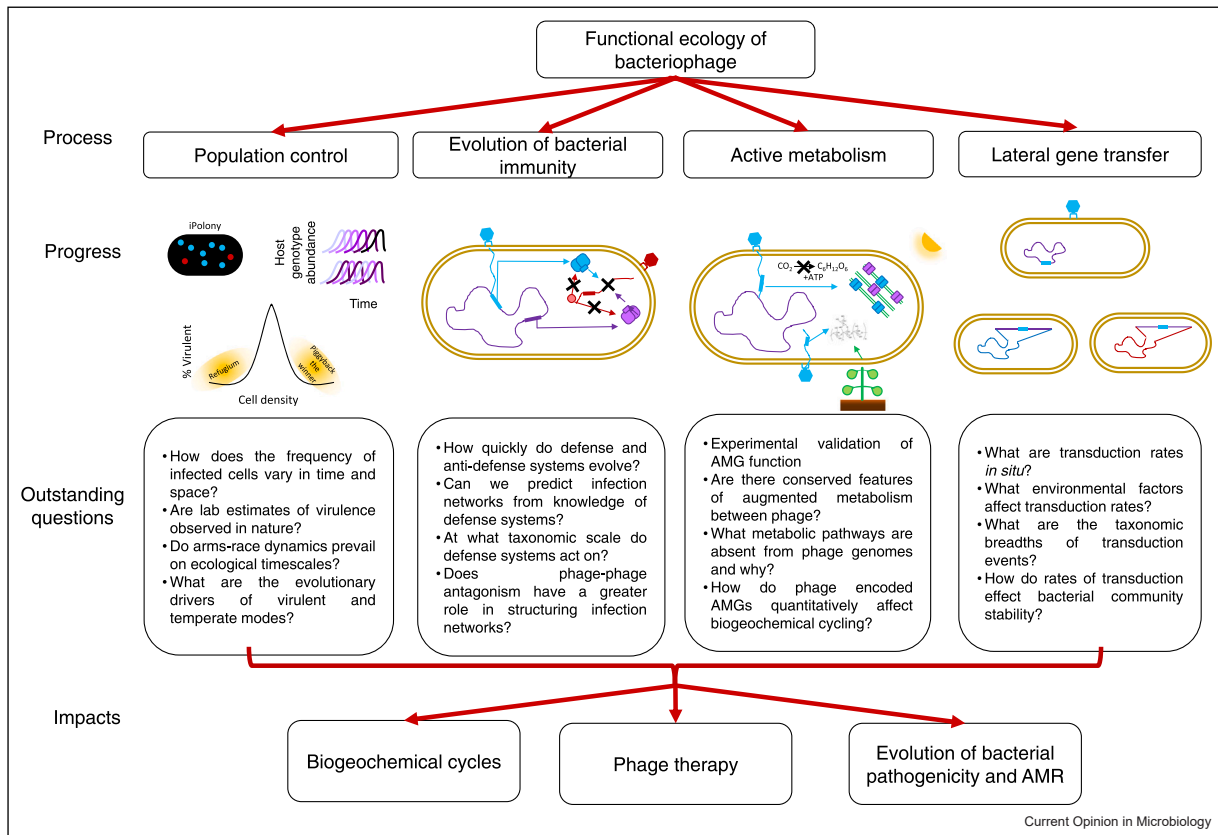
Lateral gene transfer

Phages play an important role in the horizontal movement of genes between bacterial taxa. Through transduction, they are thought to alter host cell function and therefore evolutionary trajectories. In addition to traditional modes of transduction (generalised and specialised), a recent distinct mechanism has been identified in phages of *Staphylococcus* and *Salmonella* [55,56]. The key difference here is the timing of events post induction with replication occurring before excision. Importantly, this leads to the packaging of large (hundreds of kilobases) portions of the host chromosome. Notably, rates of lateral transduction are thought to exceed those by other classical mechanisms (generalised transduction and conjugation) [57]. Thus, this mechanism has been implicated in the formation of bacterial genomic islands [56]. These distinct portions of bacterial chromosomes have disproportionate impacts on acclimation to environmental conditions [58,59], and resistance to phages themselves [60]. Regardless of the mechanism of transduction, understanding rates in the environment has been virtually impossible to measure.

Conclusions

Technological advancements continue to enhance our understanding of the functional ecology of phages. With the iPolony approach allowing for quantification of the number of infected cells [15], such approaches are required in a greater number of studies to gain a broader understanding of how phages contribute to cell lysis. Developments in sequencing technologies and methods have further increased our understanding of phage diversity, in particular RNA phages, and the range of AMG phages may encode. More laboratory model systems are required to confirm the function and understand the mechanism of these AMGs. Current laboratory models continue to produce novel insights into the molecular complexity of phage-host interactions. The next frontier is to integrate sequence data with laboratory studies to provide an improved understanding of processes important for phage ecology. To this end, we provide several outstanding questions that need to be critically addressed by phage ecologists (Figure 1). We believe

Figure 1

Overview of the functional ecology of bacteriophages and current limitations. Bacteriophages can drive the four processes of **(1)** population control, **(2)** evolution of bacterial immunity, **(3)** active metabolism and **(4)** lateral gene transfer.

answering these questions will have broad importance for understanding the role phages play in controlling Earth's biogeochemical cycles, for the spread of bacterial pathogenicity and for the therapeutic use of phages in controlling disease.

Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data were used for the research described in the article.

References and recommended reading

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

- of special interest
- •• of outstanding interest

1. Dion MB, Oechslin F, Moineau S: **Phage diversity, genomics and phylogeny**. *Nat Rev Microbiol* 2020, **18**:125-138.

2. Suttle CA: **Viruses in the sea**. *Nature* 2005, **437**:356-361.

3. Bodor A, Bounedjoum N, Vincze GE, Kis ÁE, Laczi K, Bende G, et al.: **Challenges of unculturable bacteria: environmental perspectives**. *Rev Environ Sci Bio/Technol* 2020, **19**:1-22, https://doi.org/10.1007/s11157-020-09522-4

4. Zaragoza-Solas A, Haro-Moreno JM, Rodriguez-Valera F, López-Pérez M: **Long-read metagenomics improves the recovery of viral diversity from complex natural marine samples**. *MSystems* 2022, **7**:e0019222.

5. Cook R, Hooton S, Trivedi U, King L, Dodd C, Hobman J, et al.: **Hybrid assembly of an agricultural slurry virome reveals a diverse and stable community with the potential to alter the metabolism and virulence of veterinary pathogens**. *Microbiome* (65) 2021, **9**.

6. Zablocki O., Michelsen M., Burris M., Solonenko N., Warwick-Dugdale J., Ghosh R., et al.**: VirION2: a short- and long-read sequencing and informatics workflow to study the genomic diversity of viruses in nature**. *PeerJ*. n.d. ⟨https://doi.org/10.1101/2020.10.28.359364⟩.

7. Warwick-Dugdale J, Solonenko N, Moore K, Chittick L, Gregory
 •   AC, Allen MJ, et al.: **Long-read viral metagenomics captures abundant and microdiverse viral populations and their niche-defining genomic islands**. *PeerJ* 2019, https://doi.org/10.7717/peerj.6800.
The authors demonstrate how the use of long read sequencing technology improves the assembly of phage genomes in combination with short reads. Allowing the assembly of a greater number of phage genomes.

8. Michniewski S, Rihtman B, Cook R, Jones MA, Wilson WH, Scanlan DJ, et al.: **A new family of "megaphages" abundant in the marine environment**. *ISME Commun* 2021, **1**:58, https://doi.org/10.1038/s43705-021-00064-6

9. Muscatt G, Hilton S, Raguideau S, Teakle G, Lidbury IDEA, Wellington EMH, et al.: **Crop management shapes the diversity and activity of DNA and RNA viruses in the rhizosphere**. *Microbiome* 2022, **10**:181.

10. Callanan J, Stockdale SR, Shkoporov A, Draper LA, Ross RP, Hill
 •• C: **Expansion of known ssRNA phage genomes: from tens to over a thousand**. *Sci Adv* 2020, **6**:5981, https://doi.org/10.1126/sciadv.aay5981.
One of the first studies to use HMMs to identify ssRNA phages in assembled metatranscriptomics. Resulting in a 60 fold increase in the number of ssRNA phages.

11. Zayed AA, Wainaina JM, Dominguez-Huerta G, Pelletier E, Guo J, Mohssen M, et al.: **Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome**. *Science* 2022, **376**:156-162.

12. Wolf YI, Silas S, Wang Y, Wu S, Bocek M, Kazlauskas D, et al.: **Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome**. *Nat Microbiol* 2020, **5**:1262-1270.

13. Hillary LS, Adriaenssens EM, Jones DL, McDonald JE: **RNA-viromics reveals diverse communities of soil RNA viruses with the potential to affect grassland ecosystems across multiple trophic levels**. *ISME Commun* 2022, **2**:1-10.

14. Suttle CA: **The significance of viruses to mortality in aquatic microbial communities**. *Microb Ecol* 1994, **28**:237-243.

15. Mruwat N, Carlson MCG, Goldin S, Ribalet F, Kirzner S, Hulata Y,
 •• et al.: **A single-cell polony method reveals low levels of infected *Prochlorococcus* in oligotrophic waters despite high cyanophage abundances**. *ISME J* 2021, **15**:41-54.
The authors demonstrate using the iPolony method direct quantification of the number of infected cyanobaterial cells and contribution of different phage lineages to cell lysis.

16. Brockhurst MA, Koskella B, Zhang Q-G: **Bacteria-phage antagonistic coevolution and the implications for phage therapy**. In *Bacteriophages: Biology, Technology, Therapy*. Edited by Harper DR, Abedon ST, Burrowes BH, McConville ML. Springer International Publishing; 2021:231-251.

17. Gómez P, Buckling A: **Bacteria-phage antagonistic coevolution in soil**. *Science* 2011, **332**:106-109.

18. Laanto E, Hoikkala V, Ravantti J, Sundberg L-R: **Long-term genomic coevolution of host-parasite interaction in the natural environment**. *Nat Commun* 2017, **8**:111.

19. Piel D, Bruto M, Labreuche Y, Blanquart F, Goudenège D, Barcia-Cruz R, et al.: **Phage-host coevolution in natural populations**. *Nat Microbiol* 2022, **7**:1075-1086.

20. Tuttle MJ, Buchan A: **Lysogeny in the oceans: lessons from cultivated model systems and a reanalysis of its prevalence**. *Environ Microbiol* 2020, **22**:4919-4933.

21. Sirén K, Millard A, Petersen B, Gilbert MTP, Clokie MRJ, Sicheritz-Pontén T: **Rapid discovery of novel prophages using biological feature engineering and machine learning**. *NAR Genom Bioinform* 2021, **3**:lqaa109.

22. Kieft K, Zhou Z, Anantharaman K: **VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences**. *Microbiome* 2020, **8**:90.

23. Wendling CC, Refardt D, Hall AR: **Fitness benefits to bacteria of carrying prophages and prophage-encoded antibiotic-resistance genes peak in different environments**. *Evolution* 2021, **75**:515-528.

24. Knowles B, Silveira CB, Bailey BA, Barott K, Cantu VA, Cobian-Guëmes AG, et al.: **Lytic to temperate switching of viral communities**. *Nature* 2016, **531**:466-470.

25. Weitz JS, Beckett SJ, Brum JR, Cael BB, Dushoff J: **Lysis, lysogeny and virus-microbe ratios**. *Nature* 2017, **549**:E1-E3.

26. Anthenelli M, Jasien E, Edwards R, Bailey B, Felts B, Katira P, et al.: **Phage and bacteria diversification through a prophage acquisition ratchet**. *BioRxiv* 2020, https://doi.org/10.1101/2020.04.08.028340

27. Silveira CB, Luque A, Rohwer F: **The landscape of lysogeny across microbial community density, diversity and energetics**. *Environ Microbiol* 2021, **23**:4098-4111.

28. Bernheim A, Sorek R: **The pan-immune system of bacteria: antiviral defence as a community resource**. *Nat Rev Microbiol* 2020, **18**:113-119.

29. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, et al.: **Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants**. *Nat Rev Microbiol* 2020, **18**:67-83.

30. Dion MB, Plante P-L, Zufferey E, Shah SA, Corbeil J, Moineau S: **Streamlining CRISPR spacer-based bacterial host predictions to decipher the viral dark matter**. *Nucleic Acids Res* 2021, **49**:3127-3138.

31. Martínez Arbas S, Narayanasamy S, Herold M, Lebrun LA, Hoopmann MR, Li S, et al.: **Roles of bacteriophages, plasmids and CRISPR immunity in microbial community dynamics revealed using time-series integrated meta-omics**. *Nat Microbiol* 2021, **6**:123-135.

32. Coclet C, Roux S: **Global overview and major challenges of host prediction methods for uncultivated phages**. *Curr Opin Virol* 2021, **49**:117-126.

33. Athukoralage JS, McMahon SA, Zhang C, Grüschow S, Graham S, Krupovic M, et al.: **An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity**. *Nature* 2020, **577**:572-575.

34. S. Shehreen N. Birkholz, P.C. Fineran and C.M. Brown, Widespread repression of anti-CRISPR production by anti-CRISPR-associated proteins, *Nucleic Acids Res*, 50 , 2022, 8615–8625, https://doi.org/10.1093/nar/gkac674.

35. Gussow AB, Park AE, Borges AL, Shmakov SA, Makarova KS, Wolf YI, et al.: **Machine-learning approach expands the repertoire of anti-CRISPR protein families**. *Nat Commun* 2020, **11**:3784.

36. Meeske AJ, Jia N, Cassel AK, Kozlova A, Liao J, Wiedmann M, et al.: **A phage-encoded anti-CRISPR enables complete evasion of type VI-A CRISPR-Cas immunity**. *Science* 2020, **369**:54-59.

37. Mahendra C, Christie KA, Osuna BA, Pinilla-Redondo R, Kleinstiver BP, Bondy-Denomy J: **Broad-spectrum anti-CRISPR proteins facilitate horizontal gene transfer**. *Nat Microbiol* 2020, **5**:620-629.

38. León LM, Park AE, Borges AL, Zhang JY, Bondy-Denomy J: **Mobile element warfare via CRISPR and anti-CRISPR in *Pseudomonas aeruginosa***. *Nucleic Acids Res* 2021, **49**:2114-2125.

39. Vassallo CN, Doering CR, Littlehale ML, Teodoro GIC, Laub MT: **A functional selection reveals previously undetected anti-phage defence systems in the *E. coli* pangenome**. *Nat Microbiol* 2022, **7**:1568-1579.

40. Srikant S, Guegler CK, Laub MT: **The evolution of a counter-defense mechanism in a virus constrains its host range**. *Elife* 2022, **11**:e79549, https://doi.org/10.7554/eLife.79549

41. Pilosof S, Alcalá-Corona SA, Wang T, Kim T, Maslov S, Whitaker R, et al.: **The network structure and eco-evolutionary dynamics of CRISPR-induced immune diversification**. *Nat Ecol Evol* 2020, **4**:1650-1660.

42. Breitbart M, Thompson LR, Suttle CA, Sullivan MB: **Exploring the vast diversity of marine viruses**. *Oceanography* 2007, **20**:135-139.

43. Braga LPP, Spor A, Kot W, Breuil M-C, Hansen LH, Setubal JC, et al.: **Impact of phages on soil bacterial communities and nitrogen availability under different assembly scenarios**. *Microbiome* 2020, **8**:52.

44. Gazitúa MC, Vik DR, Roux S, Gregory AC, Bolduc B, Widner B, et al.: **Potential virus-mediated nitrogen cycling in oxygen-depleted oceanic waters**. *ISME J* 2021 **15**:981-998, https://doi.org/10.1038/s41396-020-00825-6

45. Zhao F, Lin X, Cai K, Jiang Y, Ni T, Chen Y, et al.: **Biochemical and structural characterization of the cyanophage-encoded phosphate-binding protein: implications for enhanced**

phosphate uptake of infected cyanobacteria. *Environ Microbiol* 2022, **7**,:3037-3050, https://doi.org/10.1111/1462-2920.16043

46. Kieft K, Zhou Z, Anderson RE, Buchan A, Campbell BJ, Hallam SJ, •• et al.: **Ecology of inorganic sulfur auxiliary metabolism in widespread bacteriophages**. *Nat Commun* 2021, **12**:3503.
Demonstrated the widespread abundance of genes phage carry that alter the biogeochemical cycling of sulfur.

47. Emerson JB, Roux S, Brum JR, Bolduc B, Woodcroft BJ, Jang HB, • et al.: **Host-linked soil viral ecology along a permafrost thaw gradient**. *Nat Microbiol* 2018, **3**:870-880.
The first study to demonstrate the function of a phage encoded CAZyme and their abundance in soil viromes.

48. Jin M, Guo X, Zhang R, Qu W, Gao B, Zeng R: **Diversities and potential biogeochemical impacts of mangrove soil viruses**. *Microbiome* 2019, **7**:58.

49. Bi L, Yu D-T, Du S, Zhang L-M, Zhang L-Y, Wu C-F, et al.: **Diversity and potential biogeochemical impacts of viruses in bulk and rhizosphere soils**. *Environ Microbiol* 2021, **23**:588-599.

50. Wang Y, Liu Y, Wu Y, Wu N, Liu W, Wang X: **Heterogeneity of soil bacterial and bacteriophage communities in three rice agroecosystems and potential impacts of bacteriophage on nutrient cycling**. *Environ Microbiome* 2022, **17**:17.

51. N.H. Mann, A. Cook, A. Millard, S. Bailey and M. Clokie, Bacterial photosynthesis genes in a virus, *Nature*, **424**, 2003, 741.

52. Lindell D, Jaffe JD, Johnson ZI, Church GM, Chisholm SW: **Photosynthesis genes in marine viruses yield proteins during host infection**. *Nature* 2005, **438**:86-89.

53. R.J. Puxty, D.J. Evans, A.D. Millard and D.J. Scanlan, Energy limitation of cyanophage development: implications for marine carbon cycling, *ISME J*, **12**, 2018,1273–1286.

54. Puxty RJ, Millard AD, Evans DJ, Scanlan DJ: **Viruses inhibit CO_2** • **fixation in the most abundant phototrophs on earth**. *Curr Biol* 2016, **26**:1585-1589.
Demonstrated cyanophages decouple photosynthetic electron transport from CO2 fixation and cyanophages have a direct role in the biogeochemical cycling of C.

55. Chen J, Quiles-Puchalt N, Chiang YN, Bacigalupe R, Fillol-Salom A, Chee MSJ, et al.: **Genome hypermobility by lateral transduction**. *Science* 2018, **362**:207-212, https://doi.org/10.1126/science.aat5867

56. Fillol-Salom A, Bacigalupe R, Humphrey S, Chiang YN, Chen J, Penadés JR: **Lateral transduction is inherent to the life cycle of the archetypical Salmonella phage P22**. *Nat Commun* 2021, **12**:6510.

57. Humphrey S, Fillol-Salom A, Quiles-Puchalt N, Ibarra-Chávez R, •• Haag AF, Chen J, et al.: **Bacterial chromosomal mobility via lateral transduction exceeds that of classical mobile genetic elements**. *Nat Commun* 2021, **12**:6509.
The first study to quantify the rate of gene transfer by plasmids, transposons and generalised transduction in comparison to the recently discovered method of lateral transduction.

58. Li W, Wang A: **Genomic islands mediate environmental adaptation and the spread of antibiotic resistance in multiresistant Enterococci - evidence from genomic sequences**. *BMC Microbiol* 2021, **21**:55.

59. Dufresne A, Ostrowski M, Scanlan DJ, Garczarek L, Mazard S, Palenik BP, et al.: **Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria**. *Genome Biol* 2008, **9**:R90.

60. Avrani S, Wurtzel O, Sharon I, Sorek R, Lindell D: **Genomic island variability facilitates Prochlorococcus-virus coexistence**. *Nature* 2011, **474**:604-608.