Biogeochemistry goes viral: Towards a multifaceted approach to study viruses and biogeochemical cycling

Patricia Q. Tran[1,2], Karthik Anantharaman[1]

[1] Department of Bacteriology, University of Wisconsin-Madison
[2] Department of Integrative Biology, University of Wisconsin-Madison

ABSTRACT

Viruses are ubiquitous on Earth and are keystone components of environments, ecosystems, and human health. Yet, viruses remain poorly studied because most cannot be isolated in a laboratory. In the field of biogeochemistry, which aims to understand the interactions between biology, geology, and chemistry, there is progress to be made in understanding the different roles played by viruses in nutrient cycling, food webs, and elemental transformations. In this commentary, we outline current microbial ecology frameworks for understanding biogeochemical cycling in aquatic ecosystems. Next, we review some existing experimental and computational techniques that are enabling us to study the role of viruses in biogeochemical cycling, using examples from aquatic environments. Finally, we provide a conceptual model that balances limitations of computational tools when combined with biogeochemistry and ecological data. We envision meeting the grand challenge of understanding how viruses impact biogeochemical cycling by using a multifaceted approach to viral ecology.

COMMENTARY

The importance of viruses in aquatic biogeochemistry

Microbial communities are central to biogeochemical cycling, as observed in marine (1), soil (2), and freshwater environments (3). Over the past decades, technological advances have led to the increase of genomic sequencing, resulting in discoveries about the roles of microbes, particularly bacteria and archaea. However, few studies in aquatic microbial ecology transcend the domains of life to the realm of viruses. This lack of understanding of viruses prevents their inclusion in next-generation models that are being used to inform long-term predictions of metabolism, ecosystems, and biogeochemistry.

Most studies either focus on bacteria, archaea, or viruses individually. When combined, studies can explain how sudden shifts in biogeochemical processes in otherwise stable communities are driven by viruses (4). Microorganisms form complex communities that interact with each other through predation mechanisms such as cell lysis, grazing, and competition for resources (Figure 1A). Therefore, studying how all these drivers interact with each other may provide a mechanistic understanding that goes beyond descriptive ecology.

Viral ecology studies have demonstrated that viral roles in ecosystems cannot be ignored. For example, lytic viruses can target microbes, release carbon that fuels the microbial food web (the viral shunt) (5), and have direct effects on the microbial community composition (6). Additionally, viruses encoding auxiliary metabolic genes (AMGs) can manipulate their hosts and

impact microbial metabolism, and processes such as carbon, nitrogen, sulfur, and iron cycling (7). These biogeochemical pathways are often tightly associated with environmental conditions such as oxyclines or chemoclines in aquatic ecosystems (Figure 1B). Most viral genomic studies that specifically address biogeochemical pathways concern marine environments. Other aquatic environments including inland lakes, coastal regions, streams, and rivers, also have dynamic spatio-temporal patterns which are related to microbial (bacterial, archaeal, eukaryotic) roles in biogeochemical cycling but remain understudied in the context of viruses. Evidence points to similarly prominent viral communities in ecosystems such as lakes, where AMG-containing viruses are potentially involved in biogeochemical cycling (4, 8–10). With growing evidence of viral roles in biogeochemical cycling, obtaining a more holistic understanding of functional roles, interactions, and effects of these communities can be achieved by bridging across the bacteria-archaea-eukaryote-virus boundaries.

Techniques to study the role of environmental viruses

Experimental and laboratory techniques exist and provide an initial set of tools to begin integrating different scales of biology (Figure 1). Some methods rely on the ability to culture viruses with their host, whereas others can be performed without. Enumeration of viruses by phage plaque assays show that virus counts vary within an aquatic ecosystem (11). In a global analysis of virus morphology in the oceans, researchers used microscopy to observe that non-tailed viruses dominated surface ocean microbial communities (12). By incorporating ecological context, follow-up studies have showed that non-tailed viruses in marine environments are a major predation mechanism on bacteria (13). Yet, most viruses studied in culture are tailed, thereby showing the importance of both cultivation-based and cultivation-independent lines of evidence for understanding ecological relevance. Dilution-to-extinction, another laboratory method, involves filtering water, followed by enrichment, purification, and isolation to finally obtain a virus-host system (14). Model host-virus systems are useful to explore targeted biogeochemical pathways and host-virus interactions since the controlled environment provides higher reproducibility. For example, carbon regeneration could be addressed by changing the abundance of viruses and measuring the host growth rate and biomass over time. Similarly, a host known to be involved in denitrification can be measurably impaired or improved upon the addition of a virus that targets it, by tracking host, viral, and chemical characteristics over time.

One step towards a more holistic understanding of biogeochemical processes in ecosystems is to move beyond studying model organisms to learn about other components of an ecosystem (biological, chemical, geological). Additionally, biogeochemistry relies on biology, geology, and chemistry, all of which have various techniques that can help understand the overall impact of viral ecology. Whereas there is a generalized recognition of the need to study uncultured microorganisms (archaea, bacteria, eukaryotes) to understand ecosystem processes, this concept is not as common in the field of virology. Since viruses are dependent on a host for cultivation, and most microorganisms in nature cannot be cultivated, few environmentally relevant viruses have been cultured to date.

To circumvent the limitations of culture-dependent viral ecology, the ongoing development of computational techniques that address the interpretative challenges of viral ‘omics’ data will facilitate their analysis in complex environmental ecosystems. In the past years, the field of

microbial metagenomics (mostly bacteria, archaea) has seen a shift from bulk read-based metagenome characterization towards functional understanding at the scale of metagenome-assembled genomes, and even at strain-level understanding of evolutionary processes and ecological patterns. The improved ability to leverage information from metagenomics is in part due to computational advances like high-throughput sequence processing, genome binning, improved algorithmic efficiency, and standardization of data. Such computational advances may be possible in the future for viral omics. Viral genomic tools are being written, tested, compared, and used to gain ecological insights (15–17), and information is becoming standardized (18). In time, these tools will facilitate a better understanding of viruses and their complex biogeochemical interactions.

Transcending laboratory-only and genomics-only boundaries can lead to novel methods for studying viral ecology that take advantage of both strengths. Single-cell viral tagging and sequencing, analogous in some ways to single-cell genome sequencing of bacteria, relies on tagging viruses, using cell sorting and sequencing and was developed for the human-gut (19). Another technique, epicPCR, consists of linking phylogenetic genes to functional genes, and then uses sequencing to obtain high-throughput ecologically relevant information about cells (20). EpicPCR has been adapted to study viral-host interactions without cultivation in estuarine environments (21). All these techniques highlight the future of viral ecology, and the potential for their application across aquatic ecosystems.

Looking forward: Combining multifaceted approaches is important to get a holistic understanding of ecosystem ecology

The amount of genomic data generated has exponentially increased in recent years, and their interpretation benefits from a thorough understanding of the historical and ecological context, and of future challenges that the ecosystem may encounter (Figure 2). We believe that the ability to interpret viral ecology data, particularly omics-based, will be facilitated by collecting metadata and contextualizing the study system. For example, one could study the impact of carbon on bacterial growth at various resolutions ranging from simple studies focused on positive feedbacks at an organismal or community level (Figure 2A), to increasing complexity of interactions (Figure 2B, C). Moving towards more integrative studies, the incorporation of multiple species, multiple realms of life, comprehensive metadata about biogeochemistry and the environment will allow us to determine complex positive and negative feedbacks in the system (Figure 2). Specifically in the case of viral ecology, we suggest that standard virus sampling methods be coupled with detailed metadata collection of biogeochemistry, and microbial communities (bacteria, archaea, eukaryotes), which could greatly increase the ability to interpret and synthesize results.

Figure 2D demonstrates how computational techniques and their results, while offering a glimpse into viral ecology, remain challenging to interpret. In the simplified example, a metagenome generated from a given sample is used as a starting point to computationally identify viruses. Along each step of the pipeline, context is lost because a relatively low percentage of viruses are identified, of which most viral genomes are partial, and even fewer of the identified viruses have an identified ecological function or role. The analysis of viral genomics can be challenging on its own, especially where viral bioinformatics methods remain

in constant development and have their own shortcomings. Given the same genomic dataset and outcome, the ability to interpret ecological functions is significantly increased with the availability of comprehensive metadata and biogeochemical data (Figure 2C, D) compared to without (Figure 2A).

Finally, we envision that full integration of viral ecology into measurable and predictable outcomes would involve its integration into biogeochemical and ecosystem models. Substantial efforts have been made to integrate metagenomic and metatranscriptomic data of microorganisms (bacteria and archaea) in predicting biogeochemical processes such as carbon, nitrogen and sulfur cycling across redox gradients (22, 23). Realistically, it has taken over a decade of work for the field of (bacterial and archaeal) metagenomics to move on from descriptive studies of biodiversity towards mechanistic and predictive models that integrate multiple lines of experimental and genomic evidence. Even so, these integrative studies are not the norm. While challenges and opportunities in viral ecology will involve overcoming resource limitations and cross-disciplinary learning curves, we envision the ability to closely couple viral ecology and biogeochemistry to be made through these multifaceted efforts.

Acknowledgements

We thank Laura Hug and Jack Gilbert for the invitation to contribute to this special series. We thank members of the Anantharaman laboratory at the University of Wisconsin-Madison for helpful feedback and discussions. Figures were created with BioRender.com. P.Q.T is funded by a Natural Science and Engineering Research Council of Canada (NSERC) fellowship. K.A acknowledges support from the National Science Foundation under Grant No. NSF-DBI 2047598.

References

1. Fuhrman JA. 1999. Marine viruses and their biogeochemical and ecological effects. Nature 399:541–548.
2. Jin M, Guo X, Zhang R, Qu W, Gao B, Zeng R. 2019. Diversities and potential biogeochemical impacts of mangrove soil viruses. Microbiome 7:58.
3. Zimmerman AE, Howard-Varona C, Needham DM, John SG, Worden AZ, Sullivan MB, Waldbauer JR, Coleman ML. 2020. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. Nat Rev Microbiol 18:21–34.
4. Okazaki Y, Nishimura Y, Yoshida T, Ogata H, Nakano S-I. 2019. Genome-resolved viral and cellular metagenomes revealed potential key virus-host interactions in a deep freshwater lake. Environ Microbiol 21:4740–4754.
5. Wilhelm SW, Suttle CA. 1999. Viruses and Nutrient Cycles in the SeaViruses play critical roles in the structure and function of aquatic food webs. Bioscience 49:781–788.
6. Weitz JS, Wilhelm SW. 2012. Ocean viruses and their effects on microbial communities and biogeochemical cycles. F1000 Biol Rep 4:17.
7. Mara P, Vik D, Pachiadaki MG, Suter EA, Poulos B, Taylor GT, Sullivan MB, Edgcomb VP. 2020. Viral elements and their potential influence on microbial processes along the permanently stratified Cariaco Basin redoxcline. ISME J 14:3079–3092.
8. Coutinho FH, Cabello-Yeves PJ, Gonzalez-Serrano R, Rosselli R, López-Pérez M, Zemskaya TI, Zakharenko AS, Ivanov VG, Rodriguez-Valera F. 2020. New viral biogeochemical roles revealed through metagenomic analysis of Lake Baikal. Microbiome 8:163.
9. Kieft K, Breister AM, Huss P, Linz AM, Zanetakos E, Zhou Z, Rahlff J, Esser SP, Probst AJ, Raman S, Roux S, Anantharaman K. 2021. Virus-associated organosulfur metabolism in human and environmental systems. Cell Rep 36:109471.
10. Chen L-X, Méheust R, Crits-Christoph A, McMahon KD, Nelson TC, Slater GF, Warren LA, Banfield JF. 2020. Large freshwater phages with the potential to augment aerobic methane oxidation. Nat Microbiol 5:1504–1515.
11. Ortmann AC, Suttle CA. 2005. High abundances of viruses in a deep-sea hydrothermal vent system indicates viral mediated microbial mortality. Deep Sea Res Part I 52:1515–1527.
12. Brum JR, Schenck RO, Sullivan MB. 2013. Global morphological analysis of marine viruses shows minimal regional variation and dominance of non-tailed viruses. ISME J 7:1738–1751.
13. Kauffman KM, Hussain FA, Yang J, Arevalo P, Brown JM, Chang WK, VanInsberghe D, Elsherbini J, Sharma RS, Cutler MB, Kelly L, Polz MF. 2/2018. A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria. Nature 554:118–122.
14. Buchholz HH, Michelsen ML, Bolaños LM, Browne E, Allen MJ, Temperton B. 2021. Efficient dilution-to-extinction isolation of novel virus-host model systems for fastidious heterotrophic bacteria. ISME J 15:1585–1598.
15. Kieft K, Zhou Z, Anantharaman K. 2019. VIBRANT: Automated recovery, annotation and curation of microbial viruses, and evaluation of virome function from genomic sequences. Bioinformatics.

16. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol 39:578–585.
17. Bolduc B, Jang HB, Doulcier G, You Z-Q, Roux S, Sullivan MB. 2017. vConTACT: an iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. PeerJ 5:e3243.
18. Roux S, Adriaenssens EM, Dutilh BE, Koonin EV, Kropinski AM, Krupovic M, Kuhn JH, Lavigne R, Brister JR, Varsani A, Amid C, Aziz RK, Bordenstein SR, Bork P, Breitbart M, Cochrane GR, Daly RA, Desnues C, Duhaime MB, Emerson JB, Enault F, Fuhrman JA, Hingamp P, Hugenholtz P, Hurwitz BL, Ivanova NN, Labonté JM, Lee K-B, Malmstrom RR, Martinez-Garcia M, Mizrachi IK, Ogata H, Páez-Espino D, Petit M-A, Putonti C, Rattei T, Reyes A, Rodriguez-Valera F, Rosario K, Schriml L, Schulz F, Steward GF, Sullivan MB, Sunagawa S, Suttle CA, Temperton B, Tringe SG, Thurber RV, Webster NS, Whiteson KL, Wilhelm SW, Wommack KE, Woyke T, Wrighton KC, Yilmaz P, Yoshida T, Young MJ, Yutin N, Allen LZ, Kyrpides NC, Eloe-Fadrosh EA. 1/2019. Minimum Information about an Uncultivated Virus Genome (MIUViG). Nat Biotechnol 37:29–37.
19. Džunková M, Low SJ, Daly JN, Deng L, Rinke C, Hugenholtz P. 2019. Defining the human gut host-phage network through single-cell viral tagging. Nat Microbiol 4:2192–2203.
20. Spencer SJ, Tamminen MV, Preheim SP, Guo MT, Briggs AW, Brito IL, A Weitz D, Pitkänen LK, Vigneault F, Juhani Virta MP, Alm EJ. 2016. Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. ISME J 10:427–436.
21. Sakowski EG, Arora-Williams K, Tian F, Zayed AA, Zablocki O, Sullivan MB, Preheim SP. 2021. Interaction dynamics and virus–host range for estuarine actinophages captured by epicPCR. Nat Microbiol https://doi.org/10.1038/s41564-021-00873-4.
22. Louca S, Hawley AK, Katsev S, Torres-Beltran M, Bhatia MP, Kheirandish S, Michiels CC, Capelle D, Lavik G, Doebeli M, Crowe SA, Hallam SJ. 2016. Integrating biogeochemistry with multiomic sequence information in a model oxygen minimum zone. Proc Natl Acad Sci U S A 113:E5925–E5933.
23. Reed DC, Algar CK, Huber JA, Dick GJ. 2014. Gene-centric approach to integrating environmental genomics and biogeochemical models. Proc Natl Acad Sci U S A 111:1879–1884.

Figures

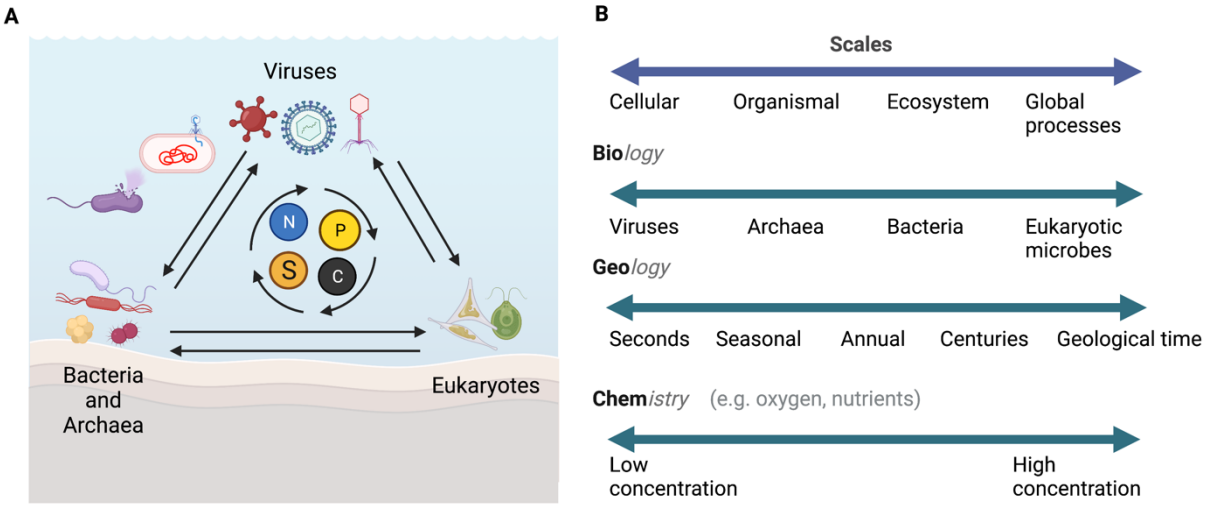

Figure 1. A. Complex microbial communities made up of viruses, bacteria, archaea, and eukaryotes interact with each other and their environment through mechanisms such as predation and competition for resources. B. Different levels of organization contribute to a holistic understanding of ecology, and are associated with challenges of studying viruses. Each of the biology, geology, and chemistry components can be studied across a range of scales, from cellular to global processes.

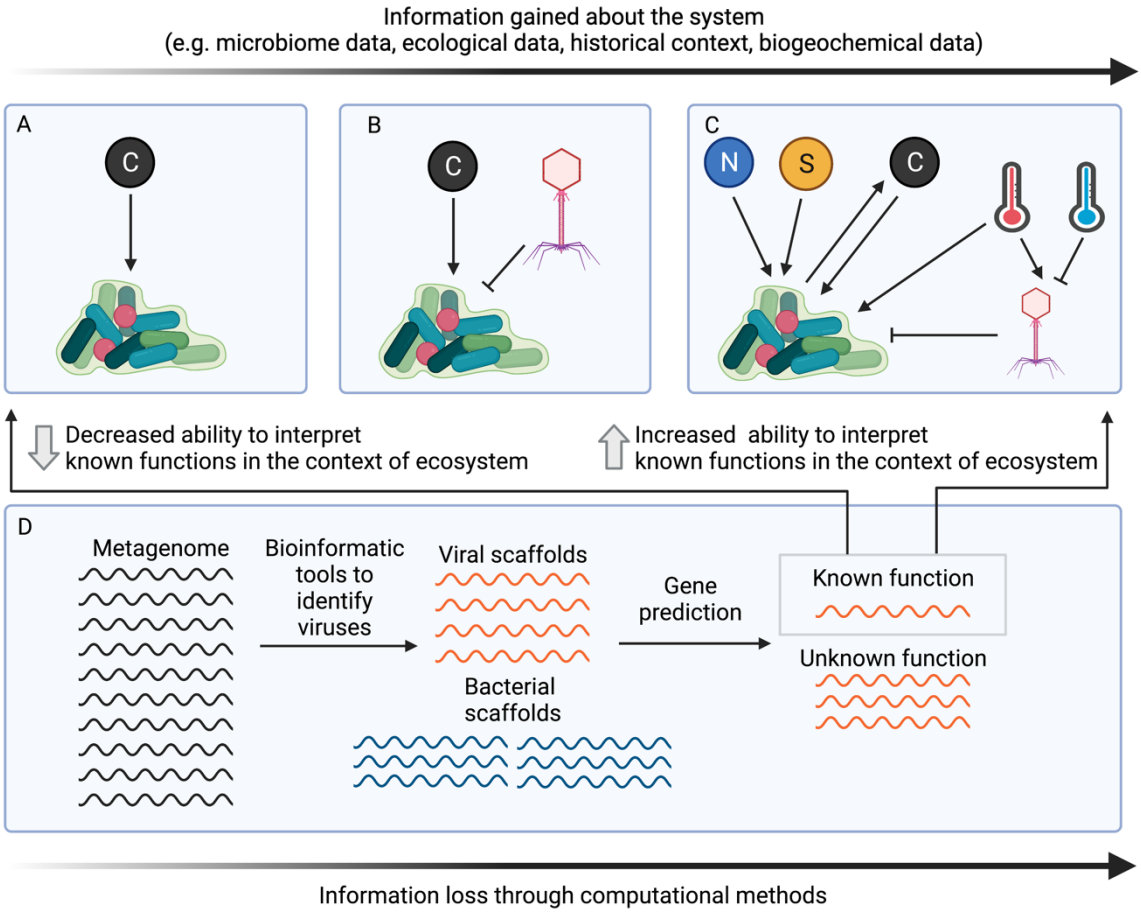

Figure 2. Conceptual framework for maximizing information about viral ecology and biogeochemistry in nature. Along the upper axis are ways to gain more information about a system. A. Example showing the positive feedback of carbon on bacterial growth. B. Addition of viruses increases complexity over A. C. Further adding detailed biogeochemical and environmental metadata such as carbon, nitrogen, sulfur, and temperature can relate complex environmental conditions to ecology but increases complexity over A and B. Examples of positive (arrow tip) and negative (inhibitor tip) interaction shown. D. Loss of information across various steps of computational analyses in viral ecology. The loss of information from computational analysis can be balanced by information gained from biogeochemical and environmental metadata.