International Journal of Marine Science, 2024, Vol.14, No.6, 377-384
http://www.aquapublisher.com/index.php/ijms

Review and Progress **Open Access**

Advancements in Metagenomic and Metatranscriptomic Techniques for Studying Marine Viruses

Linhua Zhang
Institute of Life Science, Jiyang College of Zhejiang A&F University, Zhuji, 311800, Zhejiang, China
Corresponding email: linhua.zhang@jicat.org

International Journal of Marine Science, 2024, Vol.14, No.6, doi: 10.5376/ijms.2024.14.0042
Received: 25 Oct., 2024
Accepted: 12 Dec., 2024
Published: 26 Dec., 2024

Preferred citation for this article:
Zhang L.H., 2024, Advancements in metagenomic and metatranscriptomic techniques for studying marine viruses, International Journal of Marine Science, 14(6): 377-384 (doi: 10.5376/ijms.2024.14.0042)

Abstract Marine viruses play a critical role in shaping marine ecosystems, influencing microbial diversity, biogeochemical cycles, and global climate. However, their immense diversity and complex interactions with hosts pose significant challenges to traditional research methods. This study focuses on recent advancements in metagenomic and metatranscriptomic techniques, highlighting their principles and workflows, as well as their applications in identifying novel viral species, active viral populations, and host-virus interactions. Through case studies, the practical applications of these techniques are demonstrated. The study underscores the importance of metagenomic and metatranscriptomic technologies in advancing marine virology research and envisions future directions for exploring marine viral ecosystems. It provides a scientific basis for developing strategies to address climate change and protect marine ecosystems.

Keywords Marine viruses; Metagenomics; Metatranscriptomics; Host-virus interactions; Ecological roles

1 Introduction

The matter of Marine viruses is quite interesting. We know that these little things in the sea have a significant impact on the ecosystem (Brum and Sullivan, 2015). They not only can control the number of microorganisms, but also help nutrients circulate around, keeping the entire system balanced. However, to clarify this matter, we first need to talk about the progress in research methods. In recent years, omics technology has been of great help, allowing us to understand the relationship between bacteria, archaea and their viruses (Coutinho et al., 2018; Wang, 2024.

But here comes the problem. Traditional methods such as cell culture and immunoassay are particularly laborious to use (Munang'andu et al., 2017). For instance, when it comes to new viruses or those that cannot be sustained, these methods are at a loss. What's more troublesome is that the genetic differences within the viral population are particularly large. Piecing together the genome using short-read sequencing is like playing a jigsaw puzzle (Warwick-Dugdale et al., 2019). Also, the research on archaeal viruses lags far behind that on bacterial viruses, mainly because there are simply too few samples (Zhou et al., 2022).

This time, we mainly want to see what breakthroughs these new technologies such as metagenomics and metatranscriptomics can bring. In simple terms, these methods can not only help us identify viruses, but also figure out how viruses interact with their hosts, and even discover new functions of viruses. To put it bluntly, it's about using new methods to solve old problems and taking the research on Marine viruses a step forward.

2 Metagenomics in Marine Virus Research

2.1 How metagenomics works for viruses

Instead of growing viruses in the lab (which often fails), scientists now directly study genetic material from seawater samples. Here's how it usually goes: First, they collect water samples, then pull out all the DNA/RNA. Machines like NGS sequencers read these genetic codes super fast, producing tons of data. Later, computer programs piece together these fragments like a puzzle, helping spot virus genes and see how diverse they are (Osunmakinde et al., 2018). Fun fact - this method catches viruses that old-school culturing would totally miss.

2.2 Better tech, better results

The tools keep improving - take PacBio and Nanopore sequencers. Unlike older methods, these can read much longer DNA strands, which really helps when dealing with tricky viruses that have complex or varied genes (Warwick-Dugdale et al., 2019). Some labs now mix both short and long-read tech for best results. Of course, all this data needs special software like metaSPAdes to make sense of it. These programs sort through the genetic "noise" to find viral needles in the haystack (Roux et al., 2017).

2.3 Finding hidden viral worlds

Here's the cool part: metagenomics keeps uncovering viruses we never knew existed. From ones infecting seafloor microbes to shellfish pathogens (Zhou et al., 2022), the ocean's full of surprises. Sometimes these viruses have unique DNA "islands" that show how they adapt to specific hosts. Researchers even find useful genes, like those coding for virus-killing enzymes. It's not just about counting species - this work shows how viruses secretly shape ocean life.

3 Metatranscriptomic Technology in Marine Virus Research

3.1 Basic principles and operating procedures of metatranscriptomes

In simple terms, metatranscriptomics involve extracting RNA from environmental samples such as seawater and analyzing it through sequencing to observe the genes that microorganisms (including viruses) are actively expressing. In actual operation, scientists will first collect seawater samples to extract RNA. However, there is a problem here - most of the sample is ribosomal RNA (rRNA), so special methods (such as grabbing RNA with polyA tails or removing rRNA) are needed to enrich the viral transcript. This step will directly affect the final analysis result (Gann et al., 2021). The processed RNA will be reverse transcribed into cDNA, and the sequence will be read by a high-throughput sequencer. Finally, the viral genes will be identified through bioinformatics tools to see which viruses are actively infecting the host.

3.2 New discoveries in viral gene expression and host interaction

This technology has enabled us to observe many interesting phenomena. For instance, during algal blooms, giant viruses like the Bacteroideviridae (AaV) become particularly active. They not only infect algae but also seek out other hosts (Moniruzzaman et al., 2016; 2017). Interestingly, some viruses carry genes that break down host cells, directly interfering with the host's metabolism - which may change the way the entire ecosystem operates (Voorhies et al., 2016). Through metatranscriptomics, we have not only discovered known virus-host combinations but also found many pairs that have never been seen before, indicating that the actual situation is much more complex than what is observed in the laboratory.

3.3 Application in the dynamic research of marine viruses

Macrotranscriptomics are particularly effective for tracking viruses. For instance, in the research conducted in Narragansett Bay, it was found that the infection patterns of giant viruses are completely different: some break out suddenly and then disappear rapidly, while others maintain a long-term stable infection state (Labonte et al., 2015; Terron-Camero et al., 2022). This method can not only catch the viruses that are committing crimes, but also follow the clues to find their new hosts. Moreover, by observing how viral infections alter the composition of microbial communities, we have begun to understand the role viruses play in material circulation and ecosystem balance - they are not merely destroyers, but more like invisible regulators in the ocean.

4 Combine the Study of Metagenomics and Metatranutomics

4.1 How to have a more comprehensive understanding of the marine virome

4.1.1 Simultaneously detect DNA and RNA in environmental samples

Nowadays, scientists often use metagenomic and metatranutomic technologies together, so that both DNA and RNA can be measured from the same environmental sample. Interestingly, although DNA data can tell us what genetic potential viruses have, it is RNA data that truly reveals which viruses are "working". For instance, a study last year found that during the season when algae multiply in large numbers, the number of RNA viruses (including single-stranded and double-stranded ones) that infect single-celled eukaryotes might be even greater than that of bacteriophages (Figure 1). However, it should be noted that the results obtained at different sampling times may vary significantly.

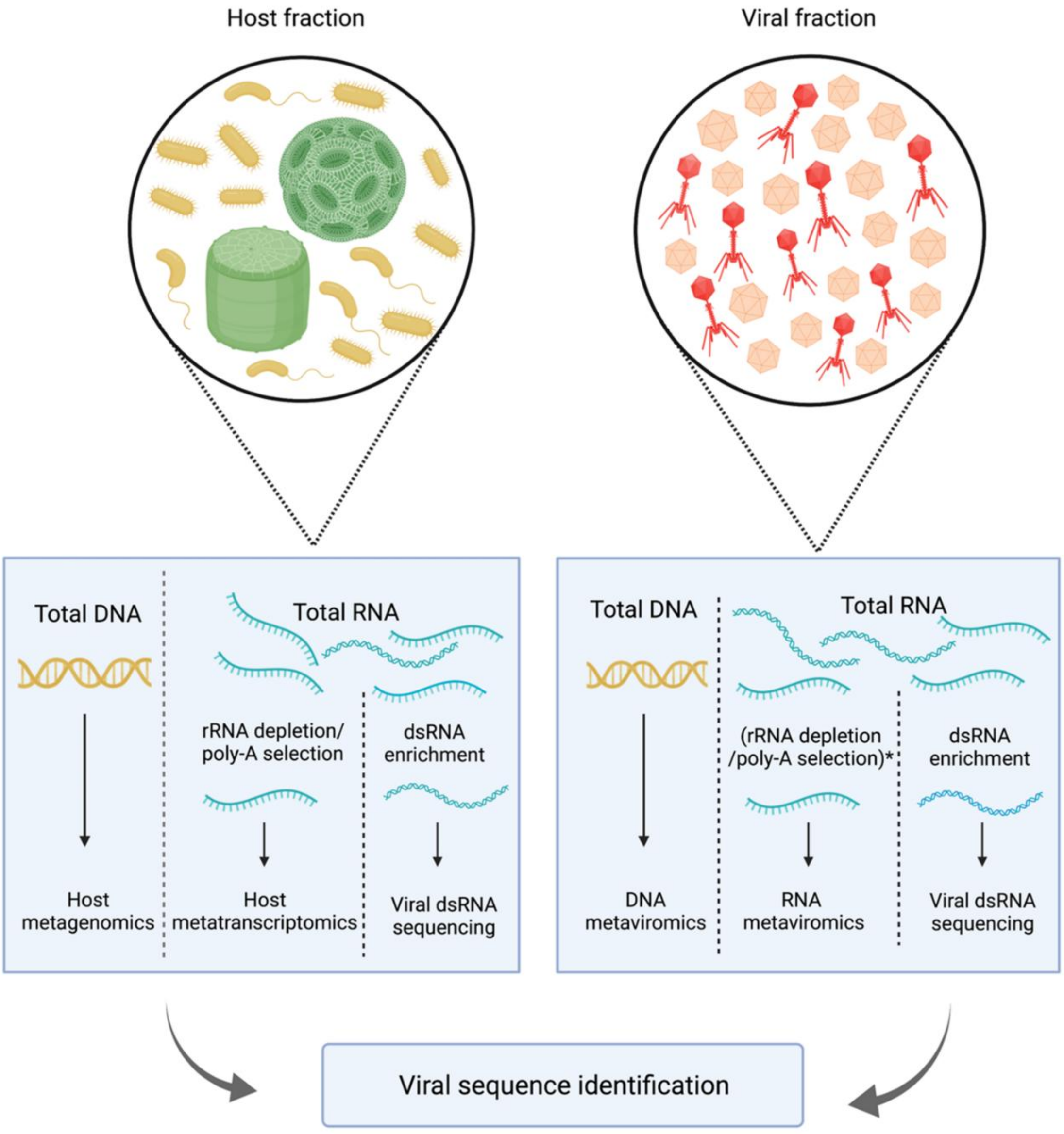

Figure 1 Most commonly used meta-omic sequencing approaches in marine virology (Adopted from Kolundžija et al., 2022)

4.1.2 Identify active and dormant viruses

Another advantage of using these two technologies in combination is that they can distinguish active viruses from dormant ones. In simple terms, metagenomics is like taking a photo, recording all viruses. Metatranscriptomes, on the other hand, are more like videos, specifically designed to capture active viruses. Studies have found that the number of RNA viruses changes particularly rapidly, sometimes being different within just a few hours. This indicates that their impact on Marine microorganisms is more flexible than we imagined. Of course, it is not easy to accurately distinguish between these two types of viruses in actual operation.

4.1.3 What troubles may be encountered in data integration

Although this method works well, it is indeed quite a headache to integrate the data obtained from the two techniques. First of all, handling massive amounts of data requires particularly professional analysis software. Secondly, the sequencing methods used by different laboratories may vary - for instance, the two commonly used techniques, MDA and MTT, each have their own preferences, which can lead to deviations when counting the types and quantities of viruses. Recently, a team has attempted to use mixed long and short read sequencing, and the results seem to be quite good. However, the standardized operation process still needs to be further improved.

4.2 Match the viral genes with their functions

4.2.1 How are gene sequences and expression data matched

The key is to match the viral gene sequences measured by metagenomics with the gene expression data detected by metatranscriptomics. Doing so not only enables us to know which viral genes are actively expressed, but also to discover how they interact with the host. For instance, some current research has begun to incorporate long and short read sequencing, which not only enables the acquisition of a more complete viral genome but also helps to more accurately identify the correspondence between the virus and its host.

4.2.2 Discover the key genes for the interaction between the virus and the host

Through this integrated analysis, scientists have identified many functional genes involved in virus-host interaction. Some of these genes are responsible for viral replication, some manipulate host cells, and others may affect the entire ecological environment. For instance, some viral genes have been found in biofilms, which might be related to polysaccharide metabolism and could potentially affect the formation of biofilms. In addition, there are some discoveries like the endolysin gene, which might be applied in biotechnology in the future.

4.3 Clarify the role of marine viruses in the ecosystem

The greatest value of the combination of these two technologies might lie in helping us understand exactly what role Marine viruses play in the ecological environment. Because they can not only tell us how many types of viruses there are, but also show which viruses are at work. It has now been discovered that viruses can affect nutrient cycling, changes in microbial communities, and even the global carbon cycle - for instance, those special viral communities found in archaea may have a significant impact on the Marine environment. However, to clarify exactly how they work, more research is needed.

5 Tech Breakthroughs and New Research Methods

5.1 High-throughput sequencing gets an upgrade

Studying marine viruses used to be tricky because we had to grow them in labs first. But now, tools like Next Generation Sequencing (NGS) changed the game—they can read DNA/RNA super fast without needing to culture viruses (Aerle and Santos, 2017). Sure, older short-read methods missed some details, but newer stuff like MinION’s long-read tech catches tiny genetic differences in viruses, even spotting special gene clusters that define their habitats. This means we’re finally piecing together fuller pictures of how diverse these viruses really are (Eren et al., 2015; Kumar et al., 2017).

5.2 Zooming in on single viruses

Here’s the thing: bulk sequencing can’t show what individual viruses are up to. That’s where single-virus genomics/transcriptomics steps in—it’s like giving each virus a microscope session. Fosmid-based tricks and other tools help, sure, but they’re pricey and finicky. Still, when combined with high-throughput sequencing, they’ve uncovered how specific viruses evolve and mess with their hosts. Not perfect, but way better for studying common virus types (Cammen et al., 2016; Warwick-Dugdale et al., 2019).

5.3 When AI meets virus hunting

Funny enough, traditional methods often miss weird or short viral sequences in metagenomic data. That’s why tools like DeepVirFinder are a big deal—they use AI to spot viruses way more accurately, even unknown ones. Unlike old-school approaches that rely on reference databases, these machine learning tricks find patterns we’d overlook. They’re not just labeling viruses; they’re helping untangle how these tiny players interact in ocean ecosystems (Ren et al., 2020; Jiang and Wang, 2024).

6 Case Studies

6.1 Metagenomic analysis of viral diversity in the Tara oceans dataset

Scientists have obtained a large amount of Marine data through the Tara Oceans expedition, which is particularly useful for studying the diversity of Marine viruses. They discovered many virus sequences that had never been seen before using metagenomic technology. For instance, using PacBio CCS, a high-precision long-read sequencing technology, more than 30,000 unique viral sequences were identified from complex seawater samples, which have never appeared in other databases (Zaragoza-Solas et al., 2022). In addition, when analyzing the Tara Oceans data, many different viral genomes were discovered, including those that infect Marine archaea. This indicates that these viruses are widely distributed in the global oceans and have a significant impact on the ecosystem (Tominaga et al., 2020).

6.2 Metatranomic study on the interaction between algae and viruses during Phytoplankton blooms

Nowadays, metatranscriptomics has become an important method for studying the interaction between viruses and algal hosts. For instance, in the study of algal blooms by Aureococcus anophagefferens, scientists discovered that giant viruses (such as NCLDVs) and RNA viruses were highly active during algal blooms. Through time series analysis, they observed that viruses of the Mimiviridae family (such as AaV) were particularly active during the peak and decline periods of algal blooms. After these viruses infect host cells, they cause cell lysis, alter gene expression, directly reduce the number of algae, and accelerate the end of algal blooms (Figure 2) (Moniruzzaman et al., 2016; 2017). Interestingly, the activity levels of different viruses vary at different stages of algal blooms, which indicates that viruses have a significant impact on the changes in algal populations. This method can not only detect active viral infections but also help determine the correspondence between viruses and hosts, allowing us to have a clearer understanding of the role of viruses in Marine ecosystems.

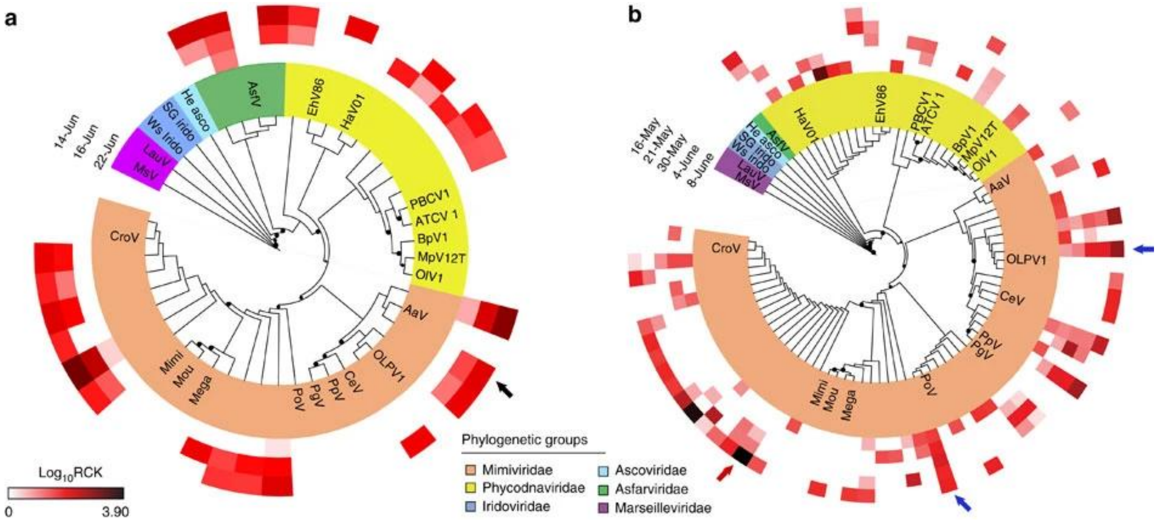

Figure 2 Phylogenetic reconstruction of NCLDV major capsid protein sequences (Adopted from Moniruzzaman et al., 2017)

Image caption: Phylogenetic placement of MCP contigs from (a) Quantuck Bay and (b) Narragansett Bay on a reference tree of NCLDVs with icosahedral capsids (Adopted from Moniruzzaman et al., 2017)

6.3 New bacteriophage-host dynamics discovered in the Arctic Ocean

The environment of the Arctic Ocean is very special, and the way phages interact with their hosts here is also very unique. Recent studies, through metagenomic analysis, have discovered some new phage-host relationships, especially those involving Marine archaea and bacteria. For instance, scientists have identified viruses within individual bacterial and archaeal cells using single-cell genomics techniques and found that some viruses may exist in a lysogenic state (Labonte et al., 2015; Alexander et al., 2021). Some studies have also used computers to predict the interaction between bacteriophages and hosts, helping us understand the role of these viruses in the Arctic Ocean ecosystem (Edwards et al., 2015). These findings indicate that metagenomic and metatranutomic technologies are particularly useful for studying viral diversity in extreme environments.

7 Concluding Remarks

Recently, there have been significant breakthroughs in the methods for studying Marine viruses, especially in metagenomic and metatranscriptomic technologies. Scientists can now analyze viral communities more accurately and discover their roles in microbial death, gene transfer and biogeochemical cycles. For instance, by using long-read sequencing technology (like VirION's method), it is possible to better piece together a complete viral genome, even capturing those viral populations with large numbers but small differences, and identifying their unique gene fragments. Interestingly, during phytoplankton outbreaks, the number of RNA viruses sometimes exceeds that of bacteriophages, and they change particularly rapidly, being different within just a few hours. In addition, incorporating viral metagenomic data into ecological models has also helped us clarify the diversity and distribution patterns of viruses, as well as how they interact with their hosts and the environment.

These technologies are now key tools for studying Marine viruses. Viruses that were previously undetectable in ordinary seawater samples, such as those related to biofilms and archaea, have now all been discovered. They also helped us identify some new genes. For instance, viruses can carry metabolic genes to alter the activities of microorganisms. Through these technologies, scientists have observed the complex interaction between viruses and their hosts and have also clarified their impact on the global elemental cycle. Even more impressive is that now, with the existing RNA sequencing data, a nearly complete viral genome can be pieced together. This is helpful for understanding the relationship between viruses and their hosts, and even for controlling viral diseases in aquaculture.

In the future, research on Marine viruses will definitely still rely on the continuous development of these technologies. As sequencing becomes increasingly cheaper and more convenient, we can analyze viral communities in more detail and discover more new viruses and their functions. If the ecological model can be built more perfectly, it will be clearer how viruses and hosts interact with each other and the effect of virus activities on the entire ecosystem. Moreover, if these technologies are applied to less studied areas, such as the deep sea and the polar regions, more diversity of viruses will surely be discovered. However, new bioinformatics tools also need to be developed; otherwise, so much data simply cannot be analyzed. In conclusion, only by understanding all these can we better understand how important Marine viruses are to the global ecosystem.

Acknowledgments

Thanks Cai Z.T. for his assistance in references collection and discussion for this work completion.

Conflict of Interest Disclosure

The author affirms that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Aerle R., and Santos E., 2017, Advances in the application of high-throughput sequencing in invertebrate virology, Journal of Invertebrate Pathology, 147: 145-156.
https://doi.org/10.1016/j.jip.2017.02.006

Alexander H., Hu S.K., Krinos A.I., Pachiadaki M., Tully B., Neely C., and Reiter T., 2021, Eukaryotic genomes from a global metagenomic data set illuminate trophic modes and biogeography of ocean plankton, MBio, 14(6): e01676-23.
https://doi.org/10.1128/mbio.01676-23

Bikel S., Valdez-Lara A., Cornejo-Granados F., Rico K., Canizales-Quinteros S., Soberón X., Del Pozo-Yauner L., and Ochoa-Leyva A., 2015, Combining metagenomics metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiome, Computational and Structural Biotechnology Journal, 13: 390-401.
https://doi.org/10.1016/j.csbj.2015.06.001

Brum J., and Sullivan M., 2015, Rising to the challenge: accelerated pace of discovery transforms marine virology, Nature Reviews Microbiology, 13: 147-159.
https://doi.org/10.1038/nrmicro3404

Cammen K., Andrews K., Carroll E., Foote A., Humble E., Khudyakov J., Louis M., McGowen M., Olsen M., and Van Cise A., 2016, Genomic methods take the plunge: recent advances in high-throughput sequencing of marine mammals, The Journal of Heredity, 107(6): 481-495.
https://doi.org/10.1093/jhered/esw044

Coutinho F., Gregoracci G., Walter J., Thompson C., and Thompson F., 2018, Metagenomics sheds light on the ecology of marine microbes and their viruses, Trends in Microbiology, 26(11): 955-965.
https://doi.org/10.1016/j.tim.2018.05.015

Ding W., Wang R., Liang Z., Zhang R., Qian P., and Zhang W., 2021, Expanding our understanding of marine viral diversity through metagenomic analyses of biofilms, Marine Life Science and Technology, 3: 395-404.
https://doi.org/10.1007/s42995-020-00078-4

Edwards R., McNair K., Faust K., Raes J., and Dutilh B., 2015, Computational approaches to predict bacteriophage–host relationships, FEMS Microbiology Reviews, 40: 258-272.
https://doi.org/10.1093/femsre/fuv048

Eren A.M., Esen Ö.C., Quince C., Vineis J.H., Morrison H.G., Sogin M., and Delmont T., 2015, Anvi'o: an advanced analysis and visualization platform for 'omics data, PeerJ, 3: e1319.
https://doi.org/10.7717/peerj.1319

Gann E., Kang Y., Dyhrman S.T., Gobler C.J., and Wilhelm S., 2021, Metatranscriptome library preparation influences analyses of viral community activity during a brown tide bloom, Frontiers in Microbiology, 12: 664189.
https://doi.org/10.3389/fmicb.2021.664189

Garretto A., Hatzopoulos T., and Putonti C., 2019, VirMine: automated detection of viral sequences from complex metagenomic samples, PeerJ, 7: e6695.
https://doi.org/10.7717/peerj.6695

Jiang X.L., and Wang W.F., 2024, Climate change and aquatic ecosystem health: impacts adaptation strategies and future challenges, International Journal of Aquaculture, 14(4): 221-231.
https://doi.org/10.5376/ija.2024.14.0022

Kolundžija S., Cheng D.Q., and Lauro F.M., 2022, RNA viruses in aquatic ecosystems through the lens of ecological genomics and transcriptomics, Viruses, 14(4): 702.
https://doi.org/10.3390/v14040702

Kumar A., Murthy S., and Kapoor A., 2017, Evolution of selective-sequencing approaches for virus discovery and virome analysis, Virus Research, 239: 172-179.
https://doi.org/10.1016/j.virusres.2017.06.005

Labonté J., Swan B., Poulos B., Luo H., Koren S., Hallam S., Sullivan M., Woyke T., Wommack E., and Stepanauskas R., 2015, Single-cell genomics-based analysis of virus–host interactions in marine surface bacterioplankton, The ISME Journal, 9: 2386-2399.
https://doi.org/10.1038/ismej.2015.48

Moniruzzaman M., Wurch L.L., Alexander H., Dyhrman S.T., Gobler C.J., and Wilhelm S.W., 2016, Virus-host infection dynamics of marine single-celled eukaryotes resolved from metatranscriptomics, bioRxiv, 2016: 093716.
https://doi.org/10.1101/093716

Moniruzzaman M., Wurch L.L., Alexander H., Dyhrman S.T., Gobler C.J., and Wilhelm S.W., 2017, Virus-host relationships of marine single-celled eukaryotes resolved from metatranscriptomics, Nature Communications, 8(1): 16054.
https://doi.org/10.1038/ncomms16054

Munang'andu H.M., Mugimba K.K., Byarugaba D.K., Mutoloki S., and Evensen Ø., 2017, Current advances on virus discovery and diagnostic role of viral metagenomics in aquatic organisms, Frontiers in Microbiology, 8: 406.
https://doi.org/10.3389/fmicb.2017.00406

Osunmakinde C.O., Selvarajan R., Sibanda T., Mamba B.B., and Msagati T.A.M., 2018, Overview of trends in the application of metagenomic techniques in the analysis of human enteric viral diversity in Africa's environmental regimes, Viruses, 10(8): 429.
https://doi.org/10.3390/v10080429

Ren J., Song K., Deng C., Ahlgren N., Fuhrman J., Li Y., Xie X., Poplin R., and Sun F., 2020, Identifying viruses from metagenomic data using deep learning, Quantitative Biology, 8: 64-77.
https://doi.org/10.1007/s40484-019-0187-4

Rosani U., and Gerdol M., 2017, A bioinformatics approach reveals seven nearly-complete RNA-virus genomes in bivalve RNA-seq data, Virus Research, 239: 33-42.
https://doi.org/10.1016/j.virusres.2016.10.009

Roux S., Emerson J.B., Eloe-Fadrosh E.A., and Sullivan M.B., 2017, Benchmarking viromics: an in silico evaluation of metagenome-enabled estimates of viral community composition and diversity, PeerJ, 5: e3817.
https://doi.org/10.7717/peerj.3817

Sun Y., Qu Y., Yan X., Yan G., Chen J., Wang G., Zhao Z., Liu Y., Tu C., and He B., 2022, Comprehensive evaluation of RNA and DNA viromic methods based on species richness and abundance analyses using marmot rectal samples, mSystems, 7(4): e00430-22.
https://doi.org/10.1128/msystems.00430-22

Terrón-Camero L.C., Gordillo-González F., Salas-Espejo E., and Andrés-León E., 2022, Comparison of metagenomics and metatranscriptomics tools: a guide to making the right choice, Genes, 13(12): 2280.
https://doi.org/10.3390/genes13122280

Tominaga K., Morimoto D., Nishimura Y., Ogata H., and Yoshida T., 2020, In silico prediction of virus-host interactions for marine bacteroidetes with the use of metagenome-assembled genomes, Frontiers in Microbiology, 11: 738.
https://doi.org/10.3389/fmicb.2020.00738

Voorhies A., Eisenlord S., Marcus D., Duhaime M., Biddanda B., Cavalcoli J., and Dick G., 2016, Ecological and genetic interactions between cyanobacteria and viruses in a low-oxygen mat community inferred through metagenomics and metatranscriptomics, Environmental Microbiology, 18(2): 358-371.
https://doi.org/10.1111/1462-2920.12756

Wang L.T., 2024, Advances in monitoring and managing aquatic ecosystem health: integrating technology and policy, International Journal of Aquaculture, 14(2): 101-111.
https://doi.org/10.5376/ija.2024.14.0012

Warwick-Dugdale J., Solonenko N., Moore K., Chittick L., Gregory A.C., Allen M.J., Sullivan M., and Temperton B., 2019, Long-read viral metagenomics captures abundant and microdiverse viral populations and their niche-defining genomic islands. PeerJ, 7: e6800.
https://doi.org/10.7717/PEERJ.6800

Zaragoza-Solas A., Haro-Moreno J.M., Rodríguez-Valera F., and López-Pérez M., 2022, Long-read metagenomics improves the recovery of viral diversity from complex natural marine samples, mSystems, 7(3): e00192-22.
https://doi.org/10.1128/msystems.00192-22

Zhou Y., Zhou L., Yan S., Chen L., Krupovic M., and Wang Y., 2022, Diverse viruses of marine archaea discovered using metagenomics, Environmental Microbiology, 25(2): 367-382.
https://doi.org/10.1111/1462-2920.16287