Microbiome

RESEARCH Open Access

Viral activity in lake analogs of anoxic early Earth oceans

Natascha S. Varona[1*], Bailey A. Wallace[1], Alice Bosco-Santos[2], Julianna Mullinax[1], Alexandra K. Stiffler[1], Molly D. O'Beirne[3], Josh Ford[4], James M. Fulton[4], Josef P. Werne[3], William P. Gilhooly III[5] and Cynthia B. Silveira[1,6*]

Abstract

Background Meromictic lakes, with their stratified water columns, are modern analogs for ancient euxinic (anoxic and sulfidic) oceans, where anaerobic sulfur-oxidizing purple and green sulfur bacteria (PSB and GSB) dominated as primary producers. Recent studies suggest a potential role of viruses in the metabolisms and biosignatures of these bacteria, but conclusive evidence of viral replication and activity in such lakes is still lacking.

Results Here, we investigate viral activity in the upper mixed layer (mixolimnion), the anoxic bottom (monimolimnion), and the microbial plate (a dense layer of phototrophic sulfur bacteria forming at the boundary between the oxygenated mixolimnion and the anoxic monimolimnion) of three meromictic lakes: Poison and Lime Blue Lakes (WA, USA) and Mahoney Lake (BC, CA). Geochemical profiles of two lakes, Mahoney and Poison, which are dominated by PSB, show a sharp chemocline, whereas Lime Blue displays a less steep chemical gradient and hosts a mixture of PSB and GSB. Viral gene transcription and epifluorescence microscopy revealed depth-dependent patterns in viral activity. The two strongly stratified, PSB-dominated lakes showed a significant decrease in the virus-to-microbe ratio (VMR) in their microbial plates, suggesting reduced viral particle production via lysis. Metatranscriptome data corroborated this trend by showing lower levels of viral gene expression in these microbial plates, higher expression of CRISPR defense and lysogeny-related genes, and relatively high expression of photosynthesis-related viral genes. Conversely, the third lake, which harbors a mix of PSB and GSB, exhibited low microbial density, high VMR, and high viral transcriptional activity. Viral transcription levels significantly correlated with VMR in the microbial plates and bottom layers, but this relationship was absent in low-density, oxic surface samples.

Conclusions Here, two independent lines of evidence, abundances and gene expression, show reduced viral lytic production in microbial plates dominated by PSB in stratified lakes. This suggests that viral lysis may contribute less to bacterial community structuring in these high-density microbial plates. Rather, other viral-mediated mechanisms, such as lysogeny and the expression of auxiliary metabolic genes, may represent a more significant viral influence on bacterial physiology and geochemistry. These patterns in virus-bacteria interactions may be consequential for the interpretations of biosignatures left by these bacterial groups in the geologic record.

Keywords Bacteriophage, Viral activity, Meromictic lake, Viral metatranscriptomics

*Correspondence:
Natascha S. Varona
natascha.varona@miami.edu
Cynthia B. Silveira
cynthiasilveira@miami.edu
Full list of author information is available at the end of the article

Introduction

Meromictic lakes consist of highly stratified water columns considered analogs of ancient oceans [1]. This stratification typically comprises an oxygenated surface layer (mixolimnion), a microbial plate associated with sharp chemical gradients (chemocline), and an anoxic bottom layer (monimolimnion). The chemocline is marked by a rapid decline in oxygen and a corresponding rise in hydrogen sulfide concentrations, creating euxinic conditions. In these euxinic waters, light penetration supports the development of a unique microbial community dominated by anaerobic sulfur-oxidizing purple and green sulfur bacteria (PSB and GSB, respectively). These bacteria use sulfide for anoxygenic photosynthesis and can form syntrophism with sulfate reducers [2, 3]. Therefore, the microbial plate supports a microbial community that is distinct from the community in the mixolimnion and the monimolimnion [4–8]. Modern euxinic lakes offer valuable insights into the ecological conditions during periods of ocean anoxia in Earth's history, where anoxygenic photosynthetic bacteria, such as PSB and GSB, are believed to have played important roles as primary producers [9, 10]. This hypothesis is supported mainly by the detection of biomarkers produced by PSBs and GSBs, such as carotenoids, in ancient sedimentary rocks [11, 12]. These biomarkers, along with distinct carbon and sulfur isotopic ratios resulting from their metabolic processes, serve as biosignatures of past or present microbial life with specific activities and environmental conditions, such as light availability, sulfide requirements, and oxygen sensitivity.

Recently, viruses have been proposed to significantly affect PSB and GSB community dynamics [13, 14]. Specifically, viruses could affect biosignatures by modulating community composition via lysis and interfering with the rates at which these bacteria make their metabolic products [13]. Virus-mediated changes in metabolic pathways could decouple microbial biosignatures from environmental factors such as light and sulfide concentration, and confound the interpretation of these markers. This may help explain inconsistencies observed between the abundance and activity of extant bacterial phototrophs and their geochemical signatures in meromictic lakes. For instance, a growing body of evidence suggests that the abundance of PSB does not directly explain the concentrations of okenone, a carotenoid pigment unique to PSB, nor the isotopic composition of carbon and sulfur that is expected to reflect their metabolic activities [15–17]. One such case is found in euxinic Fayetteville Green Lake (NY, USA), where even though GSB was the dominant group in the water column, okenone was the predominant biomarker observed [18]. In Lake Cadagno (Switzerland), one species of PSB, *Chromatium okenii*, made up 0.3% of the bacterial community and yet was responsible for 70% of the carbon uptake [16]. Recent work showing that viruses in euxinic lakes encode auxiliary metabolic genes (AMGs) related to carbon, sulfur, and pigment metabolisms supports the idea that these biogeochemical inconsistencies may be caused by viral infection [13]. Viruses commonly encode genes that can alter the metabolisms and behavior of the bacteria they infect [19–24]. In Lime Blue Lake (WA, USA), a meromictic lake with a chemocline harboring both PSB and GSB, viruses encode AMGs associated with the light reactions of photosynthesis (*psbD, psbA*), carbon fixation and recycling (*CP12*, *rpe*, *rbsK*, *G6PD/zwf*, *talA/B*, *PGD/gnd/gntz*), sulfur relay system (*this*, *moeB*, *moaA*, *mec*, *iscS/NFS1*, *phsA/psrA*, *nrnA*, *cysH*, *cyst*), and pigment production (*crtF*). However, it remains unclear whether these viruses are actively replicating or expressing metabolic genes in situ. Determining the rate at which viruses can manipulate PSB and GSB, thereby influencing their biosignatures, could have implications for interpreting Earth's redox evolution and improve the accuracy of proxies for euxinic water columns. Additionally, viral predation and gene transfer may accelerate the rate of cellular evolution by facilitating adaptation to euxinic environments.

Here, we investigate viral activity in three euxinic lakes by comparing 33 metatranscriptomes and 43 epifluorescence microscopy counts across three water column layers (the mixolimnion, the microbial plate within the chemocline, and the monimolimnion) informed by the lakes' geochemical profiles. We show that viral activity in these lakes significantly differs across depths and between PSB-dominated and GSB-dominated lakes.

Methods

Sample collection

The three study lakes were sampled from July 19 to July 25, 2023 (see Table S1 for exact sampling locations and dates). Water column temperature, dissolved oxygen (DO), specific conductivity, and pH were measured using a multiparameter digital water quality meter (ProDSS; YSI Incorporated, Yellow Springs, OH, USA). Photosynthetically active radiation (PAR) was measured using a spherical underwater quantum sensor and data logger (LI-193 and LI-1500; LI-COR Environmental, Lincoln, NE, USA). Both sensors were cast from the side of an inflatable boat. Measurements were recorded at every meter to detect the location of the mixolimnion, chemocline, and monimolimnion and to create a vertical profile of lake water chemistry. We then increased vertical measurements to every 10 cm when noticing a decrease in oxygen to obtain a higher resolution of chemical profiles across the chemocline. We also pumped small volumes

of water and passed them through a 0.45-µm filter to observe color changes in the filter as an indication of the position of the microbial plate. Samples were collected by pumping water using a peristaltic pump (Electra V2.0; Proactive Environmental Products, Lakewood Ranch, FL, USA) with a custom-assembled acrylic plate to ensure water intake laterally from the measured depths. Once the position of the microbial plate was identified, sampling depths for geochemical and microbiological analyses were determined (mixolimnion, microbial plate, and monimolimnion, respectively): 3 m, 8.8 m, and 13 m for Mahoney Lake; 3 m, 6.7 m, and 7.5 m for Poison Lake; and 3 m, 12 m, and 13.7 m for Lime Blue Lake. For viral activity estimates, two approaches were employed: epifluorescence microscopy and metatranscriptomics. Samples were collected from each lake and depth in five replicates.

Metal extraction

After collection, water samples were filtered through a 0.45-µm PCTE filter (Cytiva, Marlborough, MA, USA) and immediately acidified with 2% HNO_3 and stored in 15 mL acid-cleaned tubes (cleaning solution with 8% HNO_3 and 2% HCl; Merck, Germany). Metal concentrations were determined using an ICP–MS (Thermo Fisher Scientific, XSERIES 2, Dreieich, Germany), equipped with Collision Cell Technology (CCT), a conical chamber, and a Meinhard® concentric nebulizer. The instrument was optimized following the manufacturer's guidelines and calibrated with 10 mg/L multi-elemental solutions (High Purity Standards, Charleston, SC, USA). Certified reference material (SPS-SW2 from LGC Standards) was used for analytical quality control. Calculated recoveries ranged from 92 to 110%, and for uncertified elements, the variation among recoveries ($N=6$) was below 10%, demonstrating the robustness and reliability of the adopted procedures.

Pigment analysis

Lipid extracts were obtained from freeze-dried POM filters (one entire 142 mm GFF per depth) using a modified Bligh and Dyer ultrasonic extraction [25, 26] with methanol (MeOH)/dichloromethane (DCM)/phosphate buffer (50 mM, pH 7.4), (2:1:0.8, v/v/v), respectively [27]. The total lipid extracts were reduced to dryness under a gentle stream of nitrogen, flushed with argon, sealed, and stored in the freezer (−20 ℃) until analysis. Pigments were resolved by reverse-phase high-performance liquid chromatography (HPLC)—tandem atmospheric pressure chemical ionization (APCI) ion trap mass spectrometry in positive ion mode. A Dionex Ultimate 3000 HPLC system with a vacuum degasser, ternary gradient pump, refrigerated autosampler, and diode array detector connected to a Thermo Scientific LTQ XL ion trap mass spectrometer controlled by *Xcalibur 4.1* software. Two coupled ES Industries Spherisep ODS2 (3 µm, 80 Å, 100 mm × 2.1 mm) columns protected by a guard column were used for chromatographic separation. A 32-min ternary gradient method was adapted from a previously published method [28] (for details, see Supplementary Data). Pigments were identified by retention time, UV–vis absorbance spectra, molecular ion m/z ratios, and MS^n fragmentation patterns. Pigments were quantified based on UV–visible light absorption peaks integrated at the maximum absorbance wavelength for carotenoids and at specific Q_y band maxima for bacteriochlorins [29]. Response factors (f_i) for each identified pigment were calculated in reference to the response factors for pure standards (f_{std}), using chlorophyll *a* (Chl *a*; 7.03 absorbance units (AU)/nmol, $R^2=0.999$) for tetrapyrroles and lutein (13.6 AU/nmol, $R^2=0.999$) and beta-carotene (8.66 AU/nmol, $R^2=0.999$) for carotenoids. Differences in molar extinction coefficients (ε) were corrected using Eq. 1. Molar extinction coefficients in acetone for the identified pigments and standards were obtained from literature values [30–32].

$$\mathrm{f_i} = \frac{\varepsilon_i}{\varepsilon_{std}} \times \mathrm{f_{std}} \quad (1)$$

Epifluorescence microscopy

Samples of 1 mL were collected in cryovials and fixed with paraformaldehyde (2% final concentration) for 20 min on site and kept refrigerated for 3 to 5 h until further processing [33]. Depending on densities estimated from filter color and volume taken to clog filters, 20 to 1000 µL of the filtered sample were then filtered through a 0.02-µm Anodisc (Cytiva, Marlborough, MA, USA) using a vacuum hand pump, dried, and stored on dry ice until staining in the laboratory at the University of Miami. For samples where less than 1000 µL was used, molecular grade water was added to the sample for a final volume of 1000 µL before filtration onto the Anodisc (Table S1 for volumes). Nucleic acid-containing particles on the filters were stained with SYBR Gold Nucleic Acid Gel Stain (10X final concentration; Invitrogen, Waltham, MA, USA). Excess stain was removed by wetting the Anodisc from below with two drops of 100 µL molecular grade water and gently blotting the bottom of the filter with a lint-free wipe. Filters were air-dried for 30 min without exposure to light and mounted between a glass microscope slide and coverslip with 15 µL of mounting solution (0.02-µm-filtered 1X PBS, 0.1% ascorbic acid, and 50% glycerol). Slides were visualized under oil immersion at 630 × magnification using a ZEISS Axio Imager.A2 equipped with an Axiocam 506

mono camera and the X-Cite Mini light source (Excelitas Technologies, Pittsburgh, PA, USA) and stored at − 20 ℃. Ten images were captured per slide using the Zeiss software Zen. Cells and virus-like particles were size-delimited (> or < 0.22 µm) using Zen and counted manually to obtain bacterial cells and viral particle abundances per mL of sample after correcting for the volume of sample filtered onto the Anodisc.

Transcriptomics

Water samples were directly filtered through a 0.22-µm Sterivex filter (Millipore-Sigma, Burlington, MA, USA) coupled to the peristaltic pump output. Due to high cell densities in microbial plates and monimolimnion, sampling volumes varied between layers (see Table S1). One mL of RNAlater (Thermo Fisher Scientific, Waltham MA, USA) was immediately added to metatranscriptome samples, and filters were flash-frozen in liquid nitrogen on site. RNA extraction followed a Sterivex extraction protocol [34]. This protocol modifies a DNeasy PowerWater Sterivex kit (Qiagen, Germantown, MD, USA) by adding B-Marcaptoethanol and Invitrogen Turbo DNase. Twelve samples yielded too little RNA for sequencing. However, we had backup raw lake samples (frozen at -80 ℃) for some of these and were able to extract sufficient RNA from two of them using a manual Chloroform/TRIzol Reagent (Invitrogen, Waltham, MA, USA) extraction, leading to a total of 33 RNA samples. RNA was quantified using a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA), and RNA integrity was measured using the RNA Screen Tape on an Agilent 2200 TapeStation (Agilent Technologies, Palo Alto, CA, USA). To remove DNA contaminants, samples were treated with TURBO DNase (Thermo Fisher Scientific, Waltham, MA, USA). rRNA depletion was performed using QIAseq® FastSelect™ – rRNA 5S/16S/23S kit (Qiagen, Germantown, MD, USA) following the manufacturer's protocol. RNA sequencing libraries were constructed with the NEBNext Ultra II RNA Library Preparation Kit for Illumina following the manufacturer's recommendations. Briefly, enriched RNAs are fragmented for 15 min at 94 ℃. First-strand and second-strand cDNA were subsequently synthesized. cDNA fragments were end-repaired and adenylated at 3'ends, and universal adapters were ligated to cDNA fragments, followed by index addition and library enrichment with limited cycle PCR. Sequencing libraries were validated using the Agilent Tapestation 4200 (Agilent Technologies, Palo Alto, CA, USA) and quantified using Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) as well as by quantitative PCR (KAPA Biosystems, Wilmington, MA, USA).

The sequencing libraries were multiplexed and clustered onto a flow cell on the Illumina NovaSeq instrument according to the manufacturer's instructions. The samples were sequenced using a 2 × 150 bp paired-end (PE) configuration at a target depth of 20 M reads per sample. The NovaSeq Control Software (NCS) conducted image analysis and base calling. Raw sequence data (.bcl files) generated from Illumina NovaSeq was converted into fastq files and de-multiplexed using Illumina bcl2fastq 2.20 software. One mismatch was allowed for index sequence identification.

Gene expression analyses

Gene expression analyses were adapted from Coclet et al. [35]. Raw reads were quality-filtered using BBDuk v.38 [36]. The adapters were trimmed, and reads annotated as ribosomal via *k*-mer matching to the RNA Silva 132_99 database were removed [35]. Reads containing four or more *N* bases with an average quality score of less than ten or a length of 51 bp or less were removed [35]. Reads were assembled by sample using Spades v3.15.5 at default with the –meta flag option, resulting in 9,999,946 contigs with an average length of 303.5 bp (min 55 bp, max 41,336 bp) [37]. Reads per kilobase million (rpkm) values and coverage statistics were calculated using BBMap v.38 [36]. Contigs with a minimum of 200 bp were annotated using MetaCerberus using all available HMMs [38]. Lysogenic and CRISPR-associated genes were identified from MetaCerberus annotations if they contained the keywords "CRISPR-Cas" or "CRISPR-associated protein" for CRISPR genes and "transposase," "integrase," "recombinase," and "excisionase," for lysogenic genes. If a gene contained both terms, CRISPR keywords took precedence. The rpkm values were normalized by rpkm per sample to allow comparisons of CRISPR and lysogeny-related genes across samples. Viral contigs were identified with geNomad v1.8.0 using end-to-end configuration [39]. Viral contigs were only considered active if reads mapped to at least 10% of the viral contig [35]. Viral contigs were dereplicated using CheckV anicalc.py and aniclust.py scripts at MIUViG of 95% ANI and a minimum of 85% coverage [40, 41]. Only viruses with no host gene contamination, as defined by CheckV, were used for downstream processes. Statistical analysis was carried out in RStudio version 2022.12.0 + 353. Rpkm values were clustered using pheatmap v1.0.12 with default parameters, which uses an agglomerative hierarchical clustering method with Euclidean distances. Viral diversity at the class level was calculated using the Shannon Index from the R package vegan v.2.6–4 based on the rpkm values at the class level. For statistical comparisons, the R package rstatix v 0.7.2 was used to test the normality of the distribution of data using Shapiro–Wilk

test, and statical differences using one-way ANOVA and Tukey's HSD post hoc test were performed using the R package stats v4.2.2.

Results

Water column stratification

Three meromictic lakes, Lime Blue and Poison Lakes in Washington, USA, and Mahoney Lake in British Columbia, Canada, were sampled to determine viral activity (Fig. 1a). The mixolimnion, chemocline, and monimolimnion were detected through geochemical vertical profiling of dissolved oxygen (DO), salinity, pH, turbidity, temperature, and light. A sharp decrease in DO was observed in Mahoney Lake from the maximum at 4 m (248.0%) to 1% at 9.5 m and at Poison Lake (from 205.5% at 5 m to 1.4% at 7.1 m), whereas DO in Lime Blue Lake decreased more slowly from approximately 5 m (135.8%) to 12.5 m (1.3%) (Fig. 1b). Mahoney Lake had the highest difference in DO, with a 248-fold decrease in the chemocline (Fig. 1b, Table S2), followed by Poison Lake with a 146-fold reduction, albeit over a shorter distance (3.5 m vs. 2.1 m, respectively). The lake

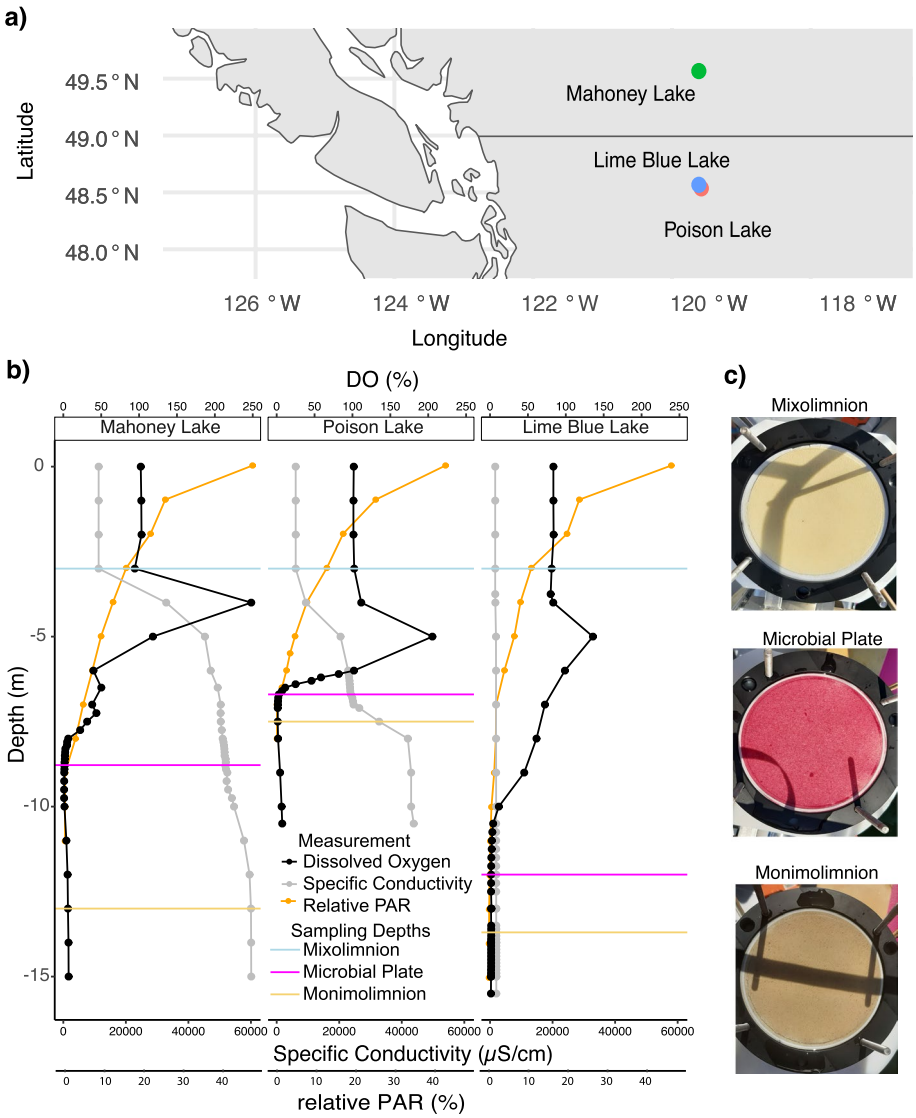

Fig. 1 Sampling sites and euxinic lake profiles. **a** Geographical location of lakes sampled in this study: Mahoney Lake (49°17′19″ N, 119° 34′58″ W), Lime Blue Lake (48°34′01″ N, 119° 36′45″ W), and Poison Lake (48°32′05″ N, 119° 34′36″ W). **b** A vertical profile of the three lakes, including dissolved oxygen (DO), specific conductivity, and Photosynthetically Active Radiation (PAR). Horizontal lines indicate the sampling depth for the mixolimnion, microbial plate, and monimolimnion at each lake. **c** Visual coloration of filters from the mixolimnion, microbial plate, and monimolimnion samples from Mahoney Lake

with less sharp stratification, Lime Blue Lake, displayed a 103-fold decrease in DO over 7.50 m. Salinity varied widely in Mahoney and Poison Lake with a total increase from 11,298 µS/cm on the surface to 60,090 µS/cm at the bottom in Mahoney and from 6077 to 43,823 µS/cm in Poison Lake (Fig. 1b). Along the depth profiles, all lakes showed a drop in pH, the most pronounced in Poison Lake where pH dropped from 8.9 at the surface (3.00 m) to 7.2 at 9 m (Fig. S1a). Lime Blue Lake was mainly stratified by temperature, ranging from 25.5 to 12.0 ℃ (Fig. S1b), while Mahoney and Poison Lake exhibited strong salinity stratification. The microbial plate in each lake was detected through changes in 0.45-µm filter coloration. For Mahoney and Poison Lakes, the location of the microbial plate, just at the end or below the chemocline, was indicated by a bright pink coloration of the filter due to the dominance of PSBs (at 8.8 m and 7.5 m depth, respectively; Fig. 1c). The mixolimnion was generally light in color, the microbial plate pink to purple, and the monimolimnion brown. In Lime Blue Lake, no obvious microbial plate could be identified, except for a slight increase in the green–brown coloration of the filters at a depth of 12 m, concurrent with an increase in turbidity, where the sample identified as the microbial plate was chosen for subsequent analyses.

Stratification of photosynthetic pigments

Mahoney Lake and Poison Lake both had maximum pigment concentrations, including mostly okenone and bacteriochlorophyll *a* (Bchl *a*) derivatives, in the microbial plate (Table 1). This confirms that PSBs are the predominant photosynthetic organisms in this sample. In Mahoney Lake, Bchl *a* was the main tetrapyrrole pigment, and there was a lower concentration of its oxidized derivative hydroxy-Bchl *a*. In contrast, hydroxy-Bchl *a* was dominant in Poison Lake, and Bchl *a* was below the level of quantification in the microbial plate sample. The Lime Blue Lake microbial plate sample also included okenone, but isorenieratene from GSB was more abundant. Its occurrence extended downward below the chemocline and was more concentrated in the monimolimnion. This apparent stratification of PSB above GSB in meromictic lakes is common, including in Fayetteville Green Lake [29]. The GSB tetrapyrrole pigments Bchl *d* and Bchl *e*, as well as bacteriopheophytins *c*, *d*, and *e*, were only detected in Lime Blue Lake, and their concentrations were greater in the monimolimnion compared with the microbial plate. Our reported concentrations consist of the sum of all structural homologs for each pigment type. The mixolimnion was omitted from Table 1, as none of the pigments was detected in these samples.

Low virus-to-microbe ratio in microbial plates of strongly stratified lakes

In Mahoney and Poison Lakes, both cell (Table S3 and Fig. 2a) and viral abundances (Fig. 2b) varied significantly between each of the three layers ($p < 0.01$, Tukey HSD test), increasing with depth. In contrast, cell abundance in Lime Blue Lake was lower at the microbial plate relative to the other lakes and did not vary significantly with depth ($p > 0.05$, Tukey HSD test). Viral abundance in Lime Blue Lake's mixolimnion was significantly lower than abundances at and below the microbial plate ($p < 0.01$, Tukey HSD test).

Virus-to-microbe ratios (VMR) were calculated by dividing the viral abundance by the number of microbial cells. A high VMR is a proxy for lytic viral production,

Table 1 Pigment concentrations (µg/L). Absent values indicate that the pigment concentration was below detection thresholds

		Microbial Plate			Monimolimnion		
	Pigment (µg/L)	Mahoney	Poison	Lime Blue	Mahoney	Poison	Lime Blue
Purple sulfur bacteria	Bacteriochlorophyll a	1.648			0.185	0.417	
	OH-bacteriochlorophyll a	0.497	3.626				
	Bacteriopheophytin a	0.269	0.364	0.397	0.460	0.212	0.478
	Okenone	2.542	5.275	0.172	0.279	2.557	
Green sulfur bacteria	Bacteriochlorophyll d						2.111
	Bacteriopheophytin d			0.350			
	Bacteriopheophytin c			0.260			1.588
	Chlorobactene						0.886
	Bacteriochlorophyll e			1.256			7.448
	Bacteriopheophytin e			1.855			14.376
	Renieratene			0.111			0.182
	Isorenieratene			0.219			0.486

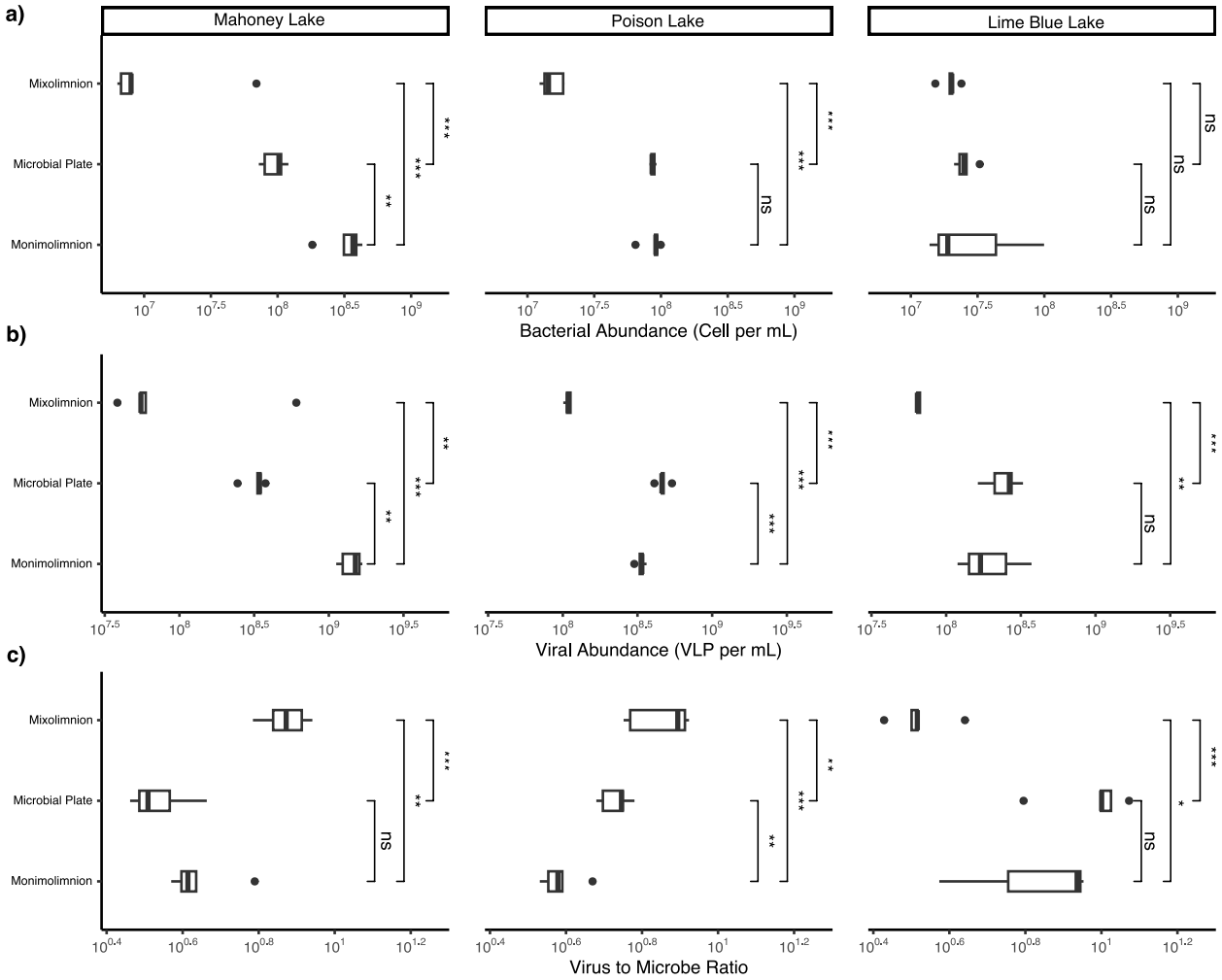

Fig. 2 Microbial densities across depths at each lake. **a** Cell abundances. **b** Viral-like particle (VLP) abundances. **c** Virus-to-microbe ratios (VMR). Comparisons were made using the Tukey HSD test, where * indicates $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and ns indicates not significant

whereas a low VMR could be associated with lysogeny [42–45]. Mahoney and Poison Lakes showed a significant decrease in VMR in the microbial plate, indicating decreased viral lytic activity relative to the mixolimnion (Fig. 2c). Monimolimnion samples also had significantly lower VMR than the surface water samples. This pattern was the opposite for Lime Blue Lake, where VMR was lowest at the mixolimnion and significantly higher at and below the microbial plate.

Viral gene transcription

A total of 19,152 assembled contigs were identified as viral genomes or genome fragments (Table S4), with 11,953 representative sequences after dereplication with MIUViG standards (95% sequence identity and 85% coverage). Of these viral contigs, 83.8% were predicted to belong to the viral realm *Duplodnaviria*, which was dominated (99.8%) by the viral class *Caudoviricetes*, commonly known as tailed bacteriophages. Other abundant contigs belonged to *Riboviria* (6.52%), followed by *Varidnaviria* (1.93%), and *Monodnaviria* (0.05%), while 8.15% of contigs could not be classified at the realm level. Across all lakes and depths, *Caudoviricetes* was the viral group with the highest normalized transcript abundance (reads per kilobase million, rpkm), ranging from 68.2% of transcript reads (SE = 0.64) in the mixolimnion of Lime Blue Lake to 87.3% (SE = 1.90) in the Mahoney Lake microbial plate (Fig. 3a, Fig. S2, Table S4). The transcript reads mapping to the class *Megaviricites* were enriched in mixolimnion samples, with the highest representation in Lime Blue Lake (7.9%, SE = 0.38). Compared to other microbial plates, Mahoney Lake's microbial plate was enriched in *Chrymotiviricetes*, making it the second most active viral group (1.8%, SE = 0.28). In Lime Blue Lake's chemocline, the second most abundant viral class was *Pisonvirecetes* (1.93%, SE = 0.93), whereas in Poison Lake,

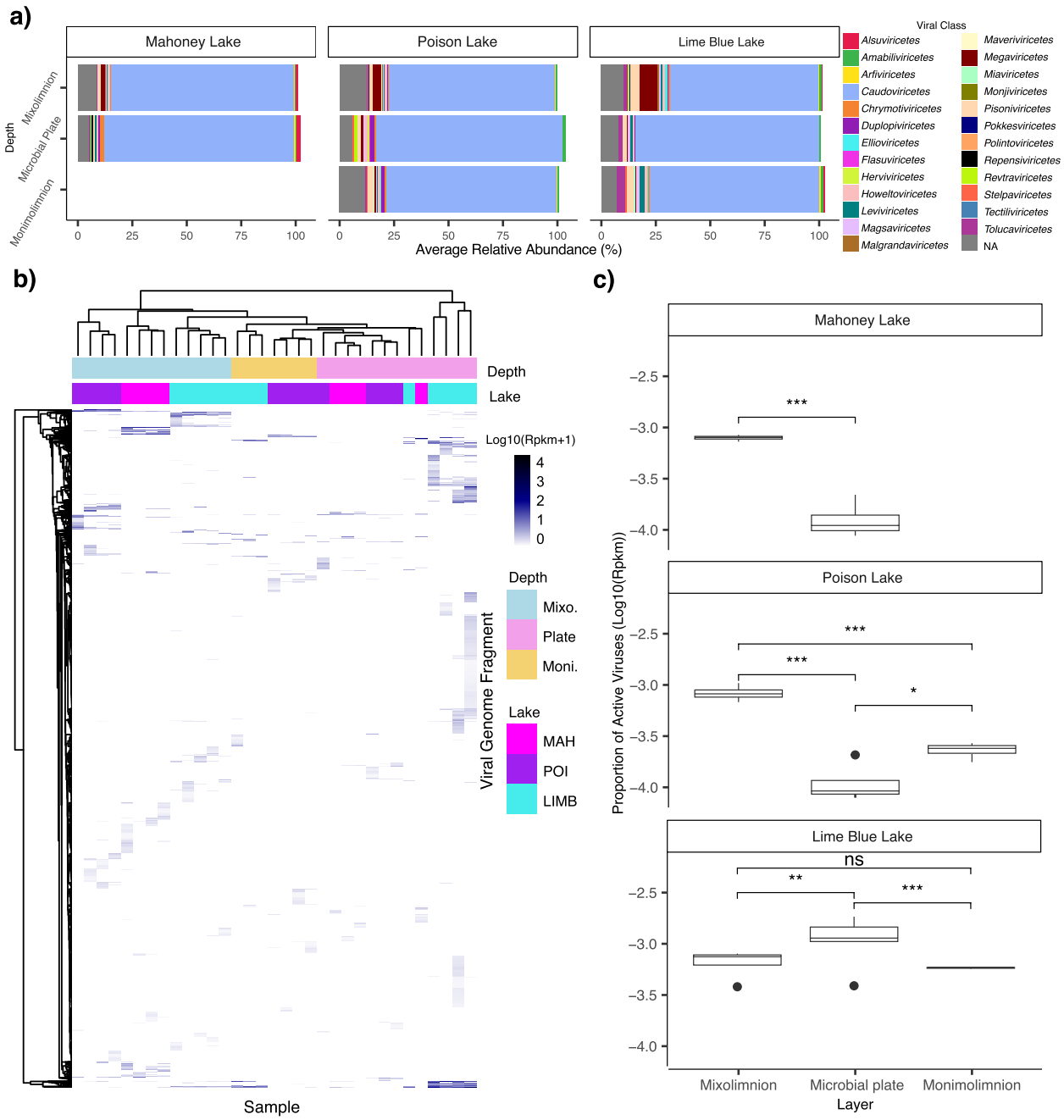

Fig. 3 Viral activity levels by depth and lake. **a** Taxonomic annotation of viral genomes and genome fragments across three sampling depths and lakes. The relative abundances of viral genomes were averaged across biological replicates. **b** Hierarchical clustering of viral genomes by layer and lake. In blue, viral activity is shown in $\log_{10}(x+1)$ transformed reads per kilobase million (rpkm). The upper color bar indicates the lake depth, where Mixo. = mixolimnion, Plate = microbial plate, and Moni. = monimolimnion. The second color bar indicates the lake, abbreviated as MAH = Mahoney Lake, POI = Poison Lake, and LIMB = Lime Blue Lake. The dendrogram is based on hierarchical clustering of $\log_{10}(x+1)$ rpkm values. **c** Number of active viruses based on transcriptional activity in each lake across layers. Significant differences from Tukey's HSD test are displayed with stars, where * indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

it was *Duplopiviricetes* (1.94%, SE = 0.44). In all lakes, the mixolimnion had the highest average viral diversity, indicated by their Shannon Index, which was calculated using base e (Fig. S3, Table S5). Poison Lake's mixolimnion had the highest diversity (H' = 1.47, SD = 0.08), followed by the monimolimnion (H' = 0.881, SD = 0.21, ANOVA,

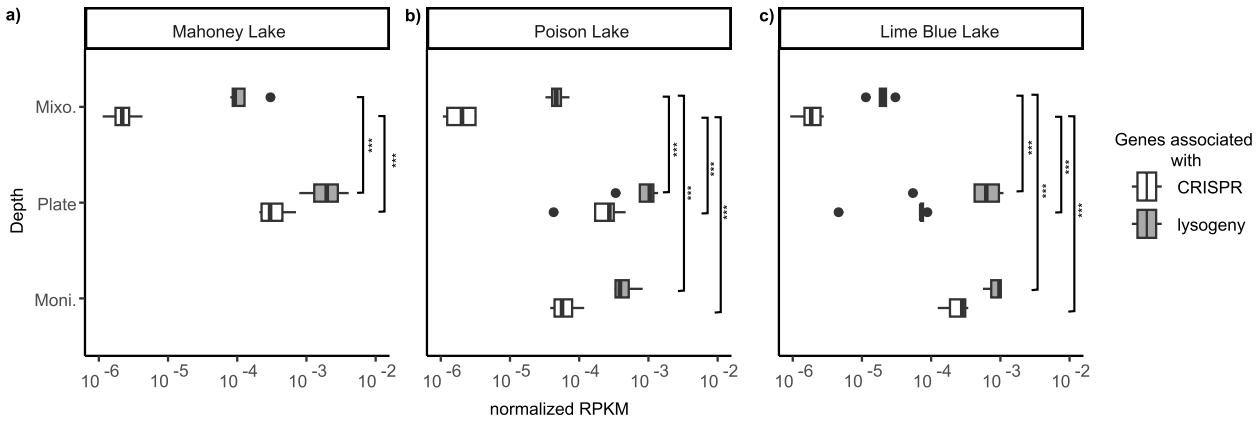

Fig. 4 Expression of CRISPR- and lysogeny-related genes across depths. **a** The relative rpkm of genes associated with CRISPR defense and lysogeny with depth in Mahoney Lake, panel **b** in Poison Lake, and **c** Lime Blue Lake. *** indicates significant differences ($p < 0.001$) based on Tukeys HSD test

$p = 0.01$) and the microbial plate ($H' = 0.749$, SD = 0.45, ANOVA, $p = 0.001$). Viral diversity was lower in the microbial plate and monimolimnion for all lakes.

Hierarchical clustering of total viral transcript abundance using Euclidean distance measures of rpkm values by sample revealed that viral activity clustered primarily by the depth and then by the lake (Fig. 3b). Mahoney and Poison Lakes' microbial plates clustered together, while four of five Lime Blue Lake microbial plate samples formed a separate cluster from all other samples. We quantified the number of active viruses in each sample to compare how viral activity varied across depths within lakes. An active virus was defined as those recruiting transcripts covering at least 10% of the viral contig length [35]. To obtain the proportion of active viruses, the number of active viral contigs was normalized by the total number of contigs per sample and compared with Tukey's HSD test (Fig. 3c, Table S7). Lime Blue Lake's microbial plate viruses were significantly more active than those in the mix- or monimolimnion ($p < 0.01$ and $p < 0.001$, respectively). Conversely, Mahoney and Poison Lake showed significant decreases in viral activity at the microbial plate ($p < 0.001$). There was also a slight increase in the proportion of active viruses in Poison Lake's monimolimnion relative to its microbial plate ($p < 0.05$).

CRISPR defense systems and lysogeny signatures

To investigate whether the decrease in viral activity corresponded with lysogeny or CRISPR defense systems, we investigated the relative abundance (sample-normalized rpkm) of CRISPR-associated or lysogeny-associated genes. Expression of these genes displayed significant differences across the three depths in all lakes examined (ANOVA, $p < 0.001$; Fig. 4). In the Mahoney Lake microbial plate, CRISPR defense gene expression increased on average by 156-fold (Tukey's, $p < 0.001$), and genes associated with lysogeny increased by 15.7-fold (Tukey's, $p < 0.001$) compared to the mixolimnion (Fig. 4a). Similarly, the Poison Lake microbial plate exhibited a 121-fold increase in CRISPR-associated gene expression (Tukey's, $p < 0.001$) and a 19-fold increase in expression of genes associated with lysogeny (Tukey's, $p < 0.001$; Fig. 4b). Between the mixolimnion and monimolimnion in both lakes, CRISPR and lysogeny genes also showed significant increases in expression, although at lower magnitudes (30-fold for CRISPR and 9.77-fold for lysogeny; Tukey's, $p < 0.001$). Lime Blue Lake followed the same trend, with CRISPR gene expression increasing by 32.7-fold and lysogeny genes by 31.02-fold (Tukey's, $p < 0.001$) at the microbial plate compared to the mixolimnion (Fig. 4c). However, the CRISPR defense system expression in Lime Blue Lake's microbial plate was significantly lower compared to Poison Lake (Tukey's, $p = 0.021$) and Mahoney Lake (Tukey's, $p = 0.004$). Additionally, the average proportion of lysogeny-associated genes in Lime Blue Lake was lower than in Mahoney Lake and Poison Lake, although these differences were not statistically significant (Tukey's, $p = 0.072$, $p = 0.505$, respectively).

Relationship between gene expression, virus-to-microbe ratios, and chemical and physical variables

To investigate the relationship between VMR and viral activity, we compared the epifluorescence microscopy counts to the proportion of active viruses in each sample (Fig. 5). Viral-like particle (VLP) abundance displayed a weak negative relationship with viral activity in the microbial plate (Linear regression $R^2 = 0.34$, $p = 0.045$, slope = −0.158; Fig. 5a), while cell abundance

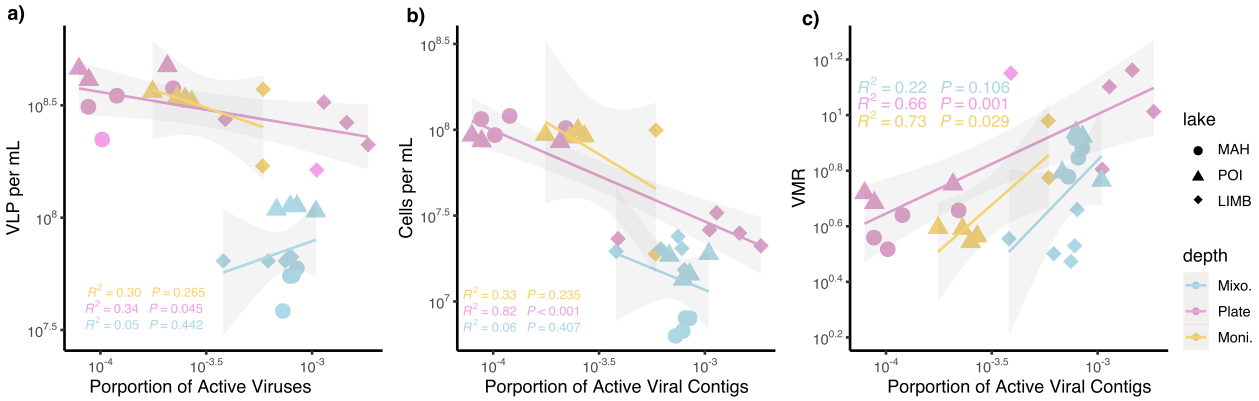

Fig. 5 Relationship between densities and viral activity from transcriptomics. **a** Linear regression between viral-like particle (VLP) abundance and the proportion of active viruses (mixolimnion $R^2 = 0.30$, $p = 0.265$; microbial plate $R^2 = 0.34$, $p = 0.045$; monimolimnion $R^2 = 0.05$, $p = 0.442$). **b** Linear regression between cell abundance and the proportion of active viruses (mixolimnion $R^2 = 0.33$, $p = 0.235$; microbial plate $R^2 = 0.82$, $p < 0.001$; monimolimnion $R^2 = 0.06$, $p = 0.407$). **c** Linear regression between virus-to-microbe ratio (VMR) and the proportion of active viruses (mixolimnion $R^2 = 0.22$, $p = 0.106$; microbial plate $R^2 = 0.66$, $p = 0.001$; monimolimnion $R^2 = 0.73$, $p = 0.029$). MAH = Mahoney Lake, POI = Poison Lake, LIMB = Lime Blue Lake, Mixo = mixolimnion, Plate = microbial plate, Moni = monimolimnion

displayed a strong and significant negative relationship with viral activity (Linear regression $R^2 = 0.82$, $p < 0.001$, slope = −0.532, Fig. 5a). VMR was a strong predictor of viral activity (Linear regression $R^2 = 0.66$, $p = 0.001$, slope = 0.358; $R^2 = 0.73$, $p = 0.029$, slope = 0.666, respectively) in both the microbial plate and monimolimnion (Fig. 5c). These trends were driven by the large density differences between the depth layers in the highly stratified lakes (Mahoney and Poison Lakes) and the lesser stratified Lime Blue Lake.

We examined the relationship between viral activity and environmental variables using linear regression for normally distributed data (Fig. S4). Variables that met the normality threshold (Shapiro–Wilk test, $p > 0.05$) included molybdenum concentration, specific conductivity, turbidity, PAR, bacteriopheophytin *a*, and pH. Among these, specific conductivity and turbidity showed significant negative correlations with viral activity (specific conductivity: $R^2 = 0.45$, $p = 0.040$; turbidity: $R^2 = 0.55$, $p = 0.020$). Turbidity also displayed a significant negative linear relationship with viral Shannon diversity ($R^2 = 0.44$, $p = 0.043$). For non-normally distributed variables with at least five data entries, temperature (℃), okenone (μg/L) dissolved oxygen (%), nickel (ppb), iron (ppb), and manganese (ppb), we utilized Spearman's rank correlation; however, no significant correlations with viral activity were observed (temperature: $p = 0.793$, okenone: $p = 0.075$, dissolved oxygen: $p = 0.755$, nickel: $p = 0.882$, iron: $p = 0.556$, manganese: $p = 0.793$). None of the above variables correlated significantly with VMR.

Highly expressed viral genes

In Lime Blue Lake and Mahoney Lake, major capsid (head) proteins and other structural genes were the most expressed viral genes (320.47 to 1029.14 rpkm) in most samples across all depths, with many serving as strong indicators for specific sample types (Fig. 6, Table S8, Table S9). Poison Lake was unique because viruses in the anaerobic layers (microbial plate and monimolimnion) primarily expressed genes associated with host metabolism. In the microbial plate of Poison Lake, the most active viral contig carried a gene annotated as a D1-like protein of the photosystem II reaction center. This gene was a significant indicator of Poison Lake's microbial plate (indicator analysis, $p = 0.048$, Table S9). The most expressed viral gene in the monimolimnion encoded a gas vesicle structural protein used by prokaryotes to control buoyancy [46]. The second most expressed gene, a nitric oxide reductase NorQ protein, is a key enzyme for the assembly of catalyst cNOR for denitrification [47].

A principal coordinate analysis (PCoA) based on rpkm abundances using Bray–Curtis dissimilarity explained 34.75% of the variance between genes expressed across sample types (Fig. S4). A PERMANOVA with 9999 permutations revealed significant differences in viral gene composition among sample types ($p < 0.001$). The PCoA shows Poison Lake's and Mahoney Lake's microbial plates clustering with Poison Lake's monimolimnion, while all mixolimnion samples formed a cluster, with some overlap with Lime Blue Lake's monimolimnion and microbial plate. A pairwise PERMANOVA revealed that most gene profiles of sample types were significantly different from each other ($p < 0.05$); however, some gene profiles were similar: Mahoney Lake and Poison Lake microbial

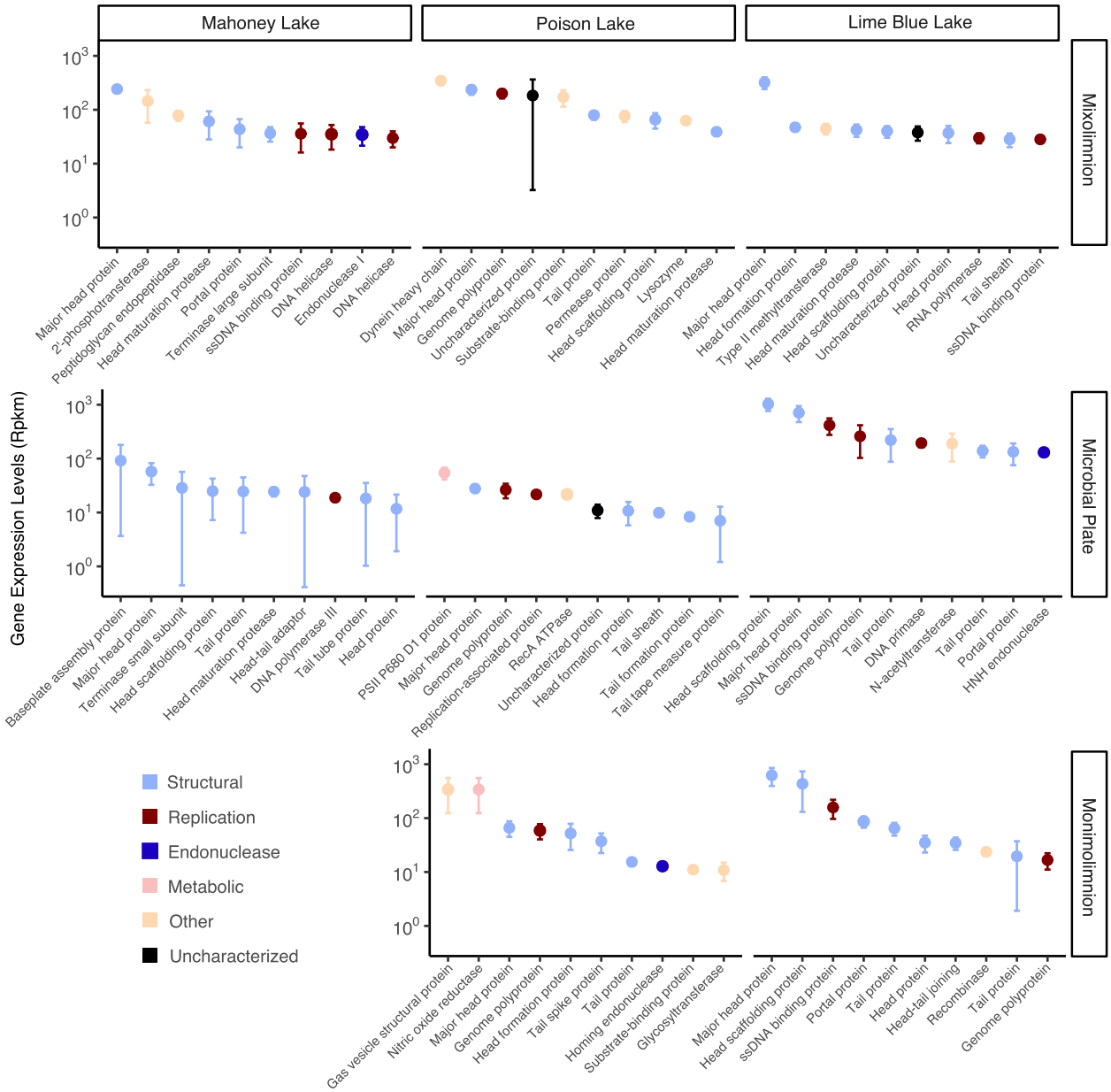

Fig. 6 Highly expressed viral genes at each lake and layer. The ten genes of viral contigs with the highest expression levels in each lake and layer (sorted by the mean across the replicates) are displayed in decreasing order, with mixolimnion samples at the top, chemocline in the center, and monimolimnion samples at the bottom. Each row displays one of three lakes, except for Mahoney Lake's monimolimnion, which was lost during sample processing. Rpkm values are based on the contig where the gene was identified

plates ($p=0.056$), Lime Blue Lake monimolimnion vs. its microbial plate ($p=0.38$), and Mahoney Lake microbial plate vs. Lime Blue Lake monimolimnion ($p=0.063$).

The ten genes with the highest expression levels (Fig. 6) in each sample group also contributed to the largest average dissimilarity between clusters, i.e., major head proteins differentiated the Lime Blue Lake microbial plate from the other two lakes (SIMPER, 22.9% of the dissimilarity with Poison Lake and 17.7 with Mahoney Lake, $p=0.002$ and 0.043, respectively). Although contributing less to the dissimilarity, expression of an 8.4 kDa Cro protein was significantly higher in the microbial plate of Lime Blue Lake than in Mahoney Lake (SIMPER, $p=0.027$, average dissimilarity = 0.071%) and Poison Lake (SIMPER, $p=0.013$, average dissimilarity = 0.080%) microbial plates. This may indicate that

there are more temperate viruses in the microbial plate of Lime Blue Lake undergoing the lytic cycle, as the expression of Cro is responsible for inducing prophages in the model phage lambda [48].

Discussion

Dense microbial plates display low virus-to-microbe ratios

The two highly stratified, PSB-dominated lakes studied here, Mahoney and Poison Lakes, exhibited lower VMR at and below the microbial plate compared to the mixolimnion (Fig. 2). This was also observed in a study investigating VMR using flow cytometry in Lake Cadagno, a highly stratified meromictic lake in Switzerland [14]. These results contrast with the pattern observed in Lime Blue Lake, where VMR was higher at the microbial plate and monimolimnion compared to the mixed layer. Though difficult to fully decouple from other environmental factors, this pattern could be partially attributed to variations in cell density among these lakes. Lime Blue Lake had the lowest cell densities among the three study lakes, and the cell densities did not differ significantly across the water column (Fig. 2). In contrast, the cell densities increased by one order of magnitude between the surface and microbial plate of Mahoney and Poison Lakes. The decrease in VMR in these lakes indicates a decrease in the frequency of lytic viral infections. One possible mechanism explaining the reduced lysis is the increase in lysogenic phage infections by temperate phages, causing an overall decrease in VMR at the community level [49–51]. During lysogeny, the viral genome integrates into the host genome and is replicated alongside the host as it grows, as opposed to the viral particle replication and cell death caused by lytic infections. Temperate phages have been detected in meromictic lakes, and in one case, GSB populations were associated with two prophages consistently infecting over 10 years [52]. The Piggyback-the-Winner framework has proposed that the prevalence of lysogeny increases in high-density microbial communities, driven by an increase in phage coinfections in these conditions of high encounter rates [49–51, 53]. Higher coinfections increase the likelihood of lysogeny by increasing the concentration of repressors of the lytic cycle and activating integrase expression [51, 54]. This explanation would fit the observations of the present study, where higher densities were observed in Mahoney Lake and Poison Lake, concurrent with the lowest VMRs. A non-mutually exclusive hypothesis to explain the decrease in VMR is that defense mechanisms such as CRISPR systems arise in these dense microbial plates. Previous research has shown that CRISPR systems are prevalent in PSB [55], which dominate in both Poison and Mahoney Lakes [6, 8, 13]. These defense systems could contribute to a reduction in detectable viral activity by inhibiting viruses' lytic cycles. CRISPR-Cas systems are abundant in environments with high viral abundance and low viral diversity [56], which is the case in the microbial plates of this study (Fig. S3).

Four lines of evidence indicate that both CRISPR defense and lysogeny contribute to the difference in observed VMR and viral activity. (1) High transcription levels of lysogenic and CRISPR-associated genes in the microbial plates compared to the mixolimnion (Fig. 4) and higher in Poison Lake and Mahoney Lake relative to Lime Blue Lake. (2) Higher expression of the prophage-inducing *cro* in Lime Blue Lake's microbial plate compared to the other two lake's microbial plates. (3) High expression of structural genes as a driver of the dissimilarity between Lime Blue Lake samples compared to other lakes (Fig. 6). (4) The negative relationship between specific conductivity, a proxy for salinity, and viral activity (Fig. S4). Salinity has been positively associated with lysogeny in high-altitude lakes [57] and estuaries [58]. Although the mechanism behind this remains to be explored, salinity can impact phage viability and particle stability [59, 60], which could lead to a preferential lifestyle of lysogeny. This agrees with our finding, where the two saline lakes, Mahoney and Poison Lake, also have the highest estimated frequency of lysogeny.

Transcribed viral genes show evidence for high viral activity in Lime Blue Lake and lower activity in PSB-dominated microbial plates

The most transcriptionally active viral contigs encoded genes related to viral particle production during the lytic cycle, specifically major capsid and tail proteins. Lime Blue Lake's microbial plate sample had the highest viral transcription levels (Fig. 6), with SIMPER analysis identifying the major head protein as a significant contributor to the dissimilarity between Lime Blue Lake and the other lakes, which corroborates the inference of higher lytic activity inferred from high VMR (Fig. 2c) and the proportion of active viruses (Fig. 3c). Conversely, the viral transcripts with the highest expression levels in Poison Lake's microbial plate and monimolimnion were associated with host metabolism, followed by structural genes. This result indicates that these viruses may be primarily influencing the metabolic processes of the host organisms through either lytic or lysogenic infections. The ability of phages to manipulate host metabolism has been observed in marine environments [18, 19, 44, 61]. The most active virus in the microbial plate of Poison Lake encoded a gene with significant similarity to the D1 protein of the Photosystem II in *Cyanobacteria* (Fig. 6). The *L* and *M* subunits of the PSB photosystem reaction center belong to the type II reaction center proteins and are homologs of the D proteins in the reaction centers

of *Cyanobacteria* [62]. Therefore, the highly expressed viral gene observed here in the PSB-dominated Poison Lake likely encodes for the *L* subunit of the PSB reaction center. The gene observed here has only 63.73% amino acid sequence identity with the closest relative Photosystem II D1 protein from NCBI's Refseq protein database (WP_224340432.1), indicating significant divergence. The expression of the *L* protein of the PSB reaction center observed here is likely analogous to the expression of phage-encoded *psbA* genes for the D1 protein during infection of *Cyanobacteria*. This expression keeps the reaction centers running during infection, powering the light reactions of photosynthesis and providing energy in the form of ATP for viral particle production [63–65].

The Poison Lake microbial plate displayed a higher concentration of hydroxy-Bchl *a*, a product of Bchl *a* oxidation, compared to the microbial plate in Mahoney Lake (Table 1), although the microbial plates from both lakes are dominated by PSB. This observation coincides with the high expression of viral-encoded metabolic genes, including photosystem genes, in Poison Lake compared to Mahoney Lake, as discussed above (Fig. 6), and leads us to speculate where viral gene expression is associated with the change in pigment profile. Cyanophages have been shown to affect pigment production. For instance, they can downregulate host genes involved in chlorophyll biosynthesis and other light-harvesting molecules [66]. Inversely, cyanophages can catalyze the production of phycocyanobilin by skipping an intermediary step in the biosynthesis pathway by encoding genes such as *pebS* [67]. A similar mechanism using a different gene has also been observed in phages infecting *Alphaproteobacteria* [68]. Changes in pigment profiles have also been reported in *Serratia marcescens*, where lytic infection increases the pigment prodigiosin [69]. A eukaryotic virus infecting the marine prymnesiophyte *Emiliania huxleyi* causes structural changes to the photosynthetic apparatus and increased hydroxylation of Chl *a*, culminating in the decreased quantum efficiency of Photosystem II and reduced chlorophyll fluorescence [70, 71]. An analogous process could explain the higher concentration of hydroxy-Bchl *a* in the Poison Lake microbial plate observed here. If that process is occurring, light and dark reactions of photosynthesis could be decoupled, resulting in a mismatch between the sulfur oxidation and carbon fixation rates, ultimately observed as mismatched sulfur and carbon isotopic signatures of PSB [13]. Because the light reactions of photosynthesis in PSB are also dependent on carotenoid pigments such as okenone, an increase in the light reactions of photosynthesis could also indirectly increase the demand for okenone synthesis, further decoupling the organic and inorganic geochemical signatures of PSB.

The metabolic effect of viruses on their hosts would have important implications for interpreting geological evidence used to infer the biogeochemical dynamics of ancient oceans. The stratified meromictic lakes examined in this study are modern analogs for ancient euxinic oceans. Viral activity could be decisive in shaping microbial metabolisms, particularly through lysogenic infection and the expression of auxiliary metabolic genes (AMGs), and significantly influence the production and preservation of metabolic outputs, which are used as biosignatures to reconstruct these ecosystems in the geological record. The observed decoupling in biosignatures may help explain inconsistencies in the different biosignatures of PSB in the sediments of meromictic lakes [17, 18] and improve the interpretation of the geologic record of anoxic water columns. Physiological studies quantifying the coupled rates of sulfur oxidation, carbon fixation, and pigment production in infected and uninfected PSB are necessary to test this hypothesis.

Viral-encoded metabolic genes were absent in the ten most abundant viral contigs in Mahoney Lake, which may be due to the difficulty in assembling long viral genome sequences solely through metatranscriptomic data. Highly transcribed AMGs from integrated prophages are more likely to be missed with our method, especially if they have a high sequence identity with the host version of the gene, leading to misidentification as host (not viral) gene transcription. Viruses in these lakes contain genes that could disrupt key metabolic pathways such as sulfur oxidation, carbon fixation, and pigment production [13]. Because these biosignatures are used to interpret the geological record, incorporating the role of viruses into models of ancient ocean biogeochemistry is crucial for refining our understanding of Earth's redox evolution.

Conclusions

Here, we show that the viral life strategies and activity are strongly stratified in lakes dominated by photosynthetic sulfur-oxidizing bacteria. In all lakes, the viral community composition differed across depths, and viral diversity decreased in anoxic layers compared to the mixolimnion. The highly stratified lakes with high densities of PSB, Mahoney and Poison Lakes, exhibited lower VMRs at the microbial plate, due to decreased lytic replication, likely caused by lysogeny and defense systems. In contrast, Lime Blue Lake, which displays a lower-density microbial plate with a mix of GSB and PSB, showed higher VMR at the microbial plate and the highest viral transcriptional activity, including genes for viral particle production, and indicating higher lytic activity than in the microbial

plates of the other two lakes. Transcriptional analysis of viral genes indicates active viral influence on host metabolism, particularly in Poison Lake, although the detection of AMGs remains challenging. These findings underscore the need for further metagenomic approaches to fully understand viral dynamics in stratified lakes and potential impacts on early Earth oceans, highlighting the complex interactions between viruses, their hosts, and environmental factors.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40168-025-02085-y.

Supplementary Material 1: Fig. S1. pH and Temperature gradients of three meromictic lakes. a) The change of pH across depth in meters. b) In dark blue, the temperature change across depth. Fig. S2. Taxonomy of viral contigs identified in each sample sorted by lake and depth. Each bar represents one replicate, grouped by lake and depth, showing their relative abundance in (%). Relative abundance was calculated based on the proportion of viral contigs identified of a viral class. Fig. S3. Viral diversity across lakes and depth. Bar plots show the average Shannon diversity index for viral taxonomy at the class level based on each lake, with error bars representing the standard deviation. Shannon diversity index was calculated based on rpkm values of viral contigs. Fig. S4. Linear relationships between viral activity and environmental variables. All normally distributed environmental variables were log-normalized and used as response variables against log normalized average viral activity (proportion of active viruses per total contigs) via linear regression. Fig. S5. Similarity of gene expression profiles of sample types. Principal Coordinates Analysis (PCoA) of viral gene abundance (rpkm) grouped by sample type. Lakes are abbreviated as Poison Lake (POI), Mahoney Lake (MAH), and Lime Blue Lake (LIMB), while depths are categorized as microbial plate (C), mixolimnion (S), and monimolimnion (B). PERMANOVA based on Bray-Curtis dissimilarity indicates significant clustering among sample types ($p < 0.001$).

Supplementary Material 2: Supplementary Table 1. Sampling information of three lakes, including coordinates, depth, number of replicates and sampling Volumes. Supplementary Table 2. Vertical profile measurements of three lakes for pH, Temperature, Light, salinity, and Dissolved Oxygen. Supplementary Table 3. Epifluorescence microscopy counts of viral like particles, bacteria and Virus-to-microbe ratio. Supplementary Table 4. Bioinformatic information of the number of sequence reads before and after quality filtration steps and assembly. Supplementary Table 5. Viral taxonomy at the class level and their relative abundance in each sample. Abundance is given in percent and was based on the counts of viral transcript per class. Supplementary Table 6. The calculated average Shannon index for viral taxonomy at the class level, including standard deviation and number of samples per sample type. Shannon indexes were calculated based on viral transcript's rpkm values. Supplementary Table 7. Calculation of active viruses per sample and the proportion of active viruses by sample. Supplementary Table 8. Gene annotation for the top 10 most abundant genes in each sample, including rpkm values and which rank the gene is (1-10). Supplementary Table 9. Viral genes identified as significant ($p<0.05$) indicators through indicator species analysis for each lake and depth.

Acknowledgements

We thank the Frost Institute for Data Science and Computing (IDSC) for providing access to the University of Miami's high-performance computing system. We thank Liar's Cove camping resort for offering power supply and sample refrigeration. We also thank Trinity Hamilton for loaning her water quality sonde and PAR meter. We thank the residents of the Alkali Lake community for providing us with road access to Poison Lake. Mahoney Lake sampling was approved by British Columbia Parks (Permit No. 111692) and facilitated by Wendy Pope, area supervisor of South Okanagan. We thank the Indigenous Peoples of Okanagan for permitting sampling, specifically the SnPink'tn, Osoyoos, and Lower Similkameen Indian Band.

Authors' contributions

CBS, ABS, WPG, JPW, and NSV designed the study. CBS directed the field, laboratory, and analytical operations. CBS, WPG, JPW, MDOB, NSV, and BAW collected samples. NSV, BAW, and JM processed the metatranscriptome and microscopy samples. MDOB processed and JMF and JF analyzed the pigment samples. NSV and AKS prepared figures. NSV analyzed the results and wrote the first version of the manuscript. All authors contributed to the manuscript.

Funding

This work was primarily funded by the NASA Exobiology Program (80NSSC23K0676 to CBS, WPG, JPW, and ABS). NSV was supported by the College of Arts and Sciences Fellowship (914000001910). AKS and BAW were supported by the NSF GRFP (2023349872 to AKS and 2023353157 to BAW). ABS acknowledges the Swiss National Science Foundation Spark grant (CRSK-2–220721). NSV, AKS, and BAW were supported by UM's Frost Institute for Data Science and Computing (IDSC) early career research award for computing resources.

Data availability

Sequence Data have been deposited to the National Center for Biotechnology Sequence Read Archive (NCBI SRA) under accession code PRJNA1161448.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Biology, University of Miami, Coral Gables, FL 33146, USA. [2]Institute of Earth Surface Dynamics, University of Lausanne, Vaud, Switzerland. [3]Department of Geology & Environmental Science, University of Pittsburgh, Pittsburgh, PA 15260, USA. [4]Geosciences Department, Baylor University, One Bear Place #97354, Waco, TX 76798-7354, USA. [5]Department of Earth & Environmental Sciences, Indiana University Indianapolis, Indianapolis, IN 46202, USA. [6]Department of Marine Biology and Ecology, Rosenstiel School of Marine, Atmospheric, and Earth Science, University of Miami, Miami, FL 33149, USA.

Received: 18 September 2024 Accepted: 9 March 2025
Published online: 26 April 2025

References

1. Xiong Y, Guilbaud R, Peacock CL, Cox RP, Canfield DE, Krom MD, et al. Phosphorus cycling in Lake Cadagno, Switzerland: a low sulfate euxinic ocean analogue. Geochimica et Cosmochimica Acta. 2019;251:116–35.
2. Bhatnagar S, Cowley ES, Kopf SH, Pérez Castro S, Kearney S, Dawson SC, et al. Microbial community dynamics and coexistence in a sulfide-driven phototrophic bloom. Environmental Microbiome. 2020;15(1):3. https://doi.org/10.1186/s40793-019-0348-0.
3. Wilbanks EG, Jaekel U, Salman V, Humphrey PT, Eisen JA, Facciotti MT, et al. Microscale sulfur cycling in the phototrophic pink berry consortia of the Sippewissett Salt Marsh. Environ Microbiol. 2014;16(11):3398–415.
4. Overmann J, Beatty JT, Krouse HR, Hall KJ. The sulfur cycle in the chemocline of a meromictic salt lake. Limnol Oceanogr. 1996;41(1):147–56.
5. Overmann J, Beatty JT, Hall KJ, Pfennig N, Northcote TG. Characterization of a dense, purple sulfur bacterial layer in a meromictic salt lake. Limnol Oceanogr. 1991;36(5):846–59.

6. Hamilton TL, Bovee RJ, Thiel V, Sattin SR, Mohr W, Schaperdoth I, et al. Coupled reductive and oxidative sulfur cycling in the phototrophic plate of a meromictic lake. Geobiology. 2014;12(5):451–68.
7. Hamilton TL, Bovee RJ, Sattin SR, Mohr W, Gilhooly WP, Lyons TW, et al. Carbon and sulfur cycling below the chemocline in a meromictic lake and the identification of a novel taxonomic lineage in the FCB superphylum, Candidatus Aegiribacteria. Front Microbiol. 2016;7:598.
8. Klepac-Ceraj V, Hayes CA, Gilhooly WP, Lyons TW, Kolter R, Pearson A. Microbial diversity under extreme euxinia: Mahoney Lake, Canada. Geobiology. 2012;10(3):223–35.
9. Johnston DT, Wolfe-Simon F, Pearson A, Knoll AH. Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age. Proc Natl Acad Sci. 2009;106(40):6925–9.
10. Lyons TW, Diamond CW, Planavsky NJ, Reinhard CT, Li C. Oxygenation, life, and the planetary system during Earth's middle history: an overview. Astrobiology. 2021;21(8):906–23.
11. Brocks JJ, Love GD, Summons RE, Knoll AH, Logan GA, Bowden SA. Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea. Nature. 2005;437(7060):866–70.
12. Brocks JJ, Schaeffer P. Okenane, a biomarker for purple sulfur bacteria (Chromatiaceae), and other new carotenoid derivatives from the 1640 Ma Barney Creek Formation. Geochimica et Cosmochimica Acta. 2008;72(5):1396–414.
13. Hesketh-Best PJ, Bosco-Santos A, Garcia SL, O'Beirne MD, Werne JP, Gilhooly WP, et al. Viruses of sulfur oxidizing phototrophs encode genes for pigment, carbon, and sulfur metabolisms. Commun Earth Environ. 2023;4(1):126.
14. Saini JS, Hassler C, Cable R, Fourquez M, Danza F, Roman S, et al. Bacterial, phytoplankton, and viral distributions and their biogeochemical contexts in Meromictic Lake Cadagno offer insights into the proterozoic ocean microbial loop. MBio. 2022;13(4):e00052-22.
15. Storelli N, Peduzzi S, Saad MM, Frigaard NU, Perret X, Tonolla M. CO2 assimilation in the chemocline of Lake Cadagno is dominated by a few types of phototrophic purple sulfur bacteria. FEMS microbiology ecolog. 2013;84(2):421–32.
16. Musat N, Halm H, Winterholler B, Hoppe P, Peduzzi S, Hillion F, et al. A single-cell view on the ecophysiology of anaerobic phototrophic bacteria. Proc Natl Acad Sci. 2008;105(46):17861–6.
17. Posth NR, Bristow LA, Cox RP, Habicht KS, Danza F, Tonolla M, et al. Carbon isotope fractionation by anoxygenic phototrophic bacteria in euxinic Lake Cadagno. Geobiology. 2017;15:798–816.
18. Meyer KM, Macalady JL, Fulton JM, Kump LR, Schaperdoth I, Freeman KH. Carotenoid biomarkers as an imperfect reflection of the anoxygenic phototrophic community in meromictic Fayetteville Green Lake. Geobiology. 2011;9:321–9.
19. Howard-Varona C, Lindback MM, Bastien GE, Solonenko N, Zayed AA, Jang H, et al. Phage-specific metabolic reprogramming of virocells. ISME J. 2020;14:881–95.
20. Ankrah NYD, May AL, Middleton JL, Jones DR, Hadden MK, Gooding JR, et al. Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. ISME J. 2014;8(5):1089–100.
21. Brüssow H, Canchaya C, Hardt W-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol Mol Biol Rev. 2004;68:560–602.
22. Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science. 1996;272:1910–4.
23. Weynberg KD, Voolstra CR, Neave MJ, Buerger P, Van Oppen MJH. From cholera to corals: viruses as drivers of virulence in a major coral bacterial pathogen. Sci Rep. 2015;5(1):17889.
24. Manzer HS, Brunetti T, Doran KS. Identification of a DNA-cytosine methyltransferase that impacts global transcription to promote group B streptococcal vaginal colonization. MBio. 2023;14(6):e02306-23.
25. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. Can J Biochem Physiol. 1959;37(8):911–7. https://doi.org/10.1139/o59-099/M13671378.
26. Pitcher A, Schouten S, Sinninghe Damste JS. In situ production of crenarchaeol in two California hot springs. Appl Environ Microbiol. 2009;75(13):4443–51. https://doi.org/10.1128/AEM.02591-08.
27. O'Beirne MD, Sparkes R, Hamilton TL, van Dongen BE, Gilhooly WP, Werne JP. Characterization of diverse bacteriohopanepolyols in a permanently stratified, hyper-euxinic lake. Org Geochem. 2022;168. https://doi.org/10.1016/j.orggeochem.2022.104431.
28. Airs RL, Atkinson JE, Keely BJ. Development and application of a high resolution liquid chromatographic method for the analysis of complex pigment distributions. J Chromatogr A. 2001;917(1):167–77. https://doi.org/10.1016/S0021-9673(01)00663-X.
29. Fulton JM, Arthur MA, Thomas B, Freeman KH. Pigment carbon and nitrogen isotopic signatures in euxinic basins. Geobiology. 2018;16(4):429–45. https://doi.org/10.1111/gbi.12285.
30. Korthals HJ, Steenbergen CLM. Separation and quantification of pigments from natural phototrophic microbial populations. FEMS Microbiol Ecol. 1985;1(3):177–85. https://doi.org/10.1111/j.1574-6968.1985.tb01146.x.
31. Britton G. UV/visible spectrometry. Carotenoids. 1995;1:13–62.
32. Taniguchi M, Lindsey JS. Absorption and fluorescence spectral database of chlorophylls and analogues. Photochem Photobiol. 2021;97(1):136–65. https://doi.org/10.1111/php.13319.
33. Stiffler AK, Hesketh-Best PJ, Varona NS, Zagame A, Wallace BA, Lapointe BE, et al. Genomic and induction evidence for bacteriophage contributions to sargassum-bacteria symbioses. Microbiome. 2024;12(1):143.
34. Krausfeldt L. RNA Extraction from Sterivex filters. protocols.io. 2017. https://doi.org/10.17504/protocols.io.gmkbu4w.
35. Coclet C, Sorensen PO, Karaoz U, Wang S, Brodie EL, Eloe-Fadrosh EA, et al. Virus diversity and activity is driven by snowmelt and host dynamics in a high-altitude watershed soil ecosystem. Microbiome. 2023;11(1):237.
36. Bushnell B. BB Tools software package. DOE Joint Genome Institute; 2014. Available from: https://sourceforge.net/projects/bbmap/.
37. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.
38. Figueroa JL, Dhungel E, Bellanger M, Brouwer CR, White RA. MetaCerberus: distributed highly parallelized HMM-based processing for robust functional annotation across the tree of life. Bioinformatics. 2024;40(3):btae119.
39. Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, et al. Identification of mobile genetic elements with geNomad. Nat Biotechnol. 2024;42:1303–12.
40. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol. 2020;39(5):578–85.
41. Roux S, Adriaenssens EM, Dutilh BE, Koonin EV, Kropinski AM, Krupovic M, et al. Minimum information about an uncultivated virus genome (MIUViG). Nat Biotechnol. 2018;37(1):29–37.
42. Marbouty M, Thierry A, Millot GA, Koszul R. Metahic phage-bacteria infection network reveals active cycling phages of the healthy human gut. eLife. 2021;10:e60608.
43. Kieft K, Anantharaman K. Deciphering active prophages from metagenomes. MSystems. 2022;7(2):e00084.
44. Engelhardt T, Kallmeyer J, Cypionka H, Engelen B. High virus-to-cell ratios indicate ongoing production of viruses in deep subsurface sediments. ISME J. 2014;8(7):1503–9.
45. Varona NS, Hesketh-Best PJ, Coutinho FH, Stiffler AK, Wallace BA, Garcia SL, et al. Host-specific viral predation network on coral reefs. The ISME Journal. 2024;18(1):wrae240.
46. Walsby AE. Gas vesicles. Microbiol rev. 1994;58(1):94–144.
47. De Boer APN, Van Der Oost J, Reijnders WNM, Westerhoff HV, Stouthamer AH, Van Spanning RJM. Mutational analysis of the nor gene cluster which encodes nitric-oxide reductase from *Paracoccus denitrificans*. Eur J Biochem. 1996;242(3):592–600.
48. Svenningsen SL, Costantino N, Court DL, Adhya S. On the role of Cro in λ prophage induction. Proc Natl Acad Sci. 2005;102(12):4465–9. https://doi.org/10.1073/pnas.0409839102.

49. 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(7595):466–70.
50. Luque A, Silveira CB. Quantification of lysogeny caused by phage coinfections in microbial communities from biophysical principles. MSystems. 2020;5(5):10–128.
51. Silveira CB, Luque A, Rohwer F. The landscape of lysogeny across microbial community density, diversity and energetics. Environ Microbiol. 2021;23(8):4098–111.
52. Berg M, Goudeau D, Olmsted C, McMahon KD, Yitbarek S, Thweatt JL, et al. Host population diversity as a driver of viral infection cycle in wild populations of green sulfur bacteria with long standing virus-host interactions. ISME J. 2021;15(6):1569–84.
53. Silveira CB, Rohwer FL. Piggyback-the-Winner in host-associated microbial communities. npj Biofilms Microbiomes. 2016;2:1–5. https://doi.org/10.1038/npjbiofilms.2016.10.
54. Cheng HH, Muhlrad PJ, Hoyt MA, Echols H. Cleavage of the cII protein of phage lambda by purified HflA protease: control of the switch between lysis and lysogeny. Proc Natl Acad Sci. 1988;85(21):7882–6. https://doi.org/10.1073/pnas.85.21.7882.
55. Kosmopoulos JC, Campbell DE, Whitaker RJ, Wilbanks EG. Horizontal gene transfer and CRISPR targeting drive phage-bacterial host interactions and coevolution in "Pink Berry" marine microbial aggregates. Appl Environ Microbiol. 2023;89:e00177–23. https://doi.org/10.1128/aem.00177-23.
56. Meaden S, Biswas A, Arkhipova K, Morales SE, Dutilh BE, Westra ER, et al. High viral abundance and low diversity are associated with increased CRISPR-Cas prevalence across microbial ecosystems. Curr Biol. 2022;(1)31:220-227.e5. https://doi.org/10.1016/j.cub.2021.10.038.
57. Zang L, Liu Y, Jiao N, Zhong KX, Song X, Yang Y, et al. Salinity as a key factor affecting viral activity and life strategies in alpine lakes. Limnol Oceanogr. 2024;69:961–75.
58. Bettarel Y, Bouvier T, Agis M, Bouvier C, Van Chu T, Combe M, et al. Viral distribution and life strategies in the Bach Dang Estuary. Vietnam Microbial Ecology. 2011;62(1):143–54. https://doi.org/10.1007/s00248-011-9835-6.
59. Ranveer SA, Dasriya V, Ahmad MF, Dhillon HS, Samtiya M, Shama E, et al. Positive and negative aspects of bacteriophages and their immense role in the food chain. npj Science of Food. 2024;8(1):1; https://doi.org/10.1038/s41538-023-00245-8.
60. Jończyk E, Kłak M, Międzybrodzki R, Górski A. The influence of external factors on bacteriophages—review. Folia Microbiol. 2011;56(3):191–200. https://doi.org/10.1007/s12223-011-0039-8.
61. Breitbart M, Bonnain C, Malki K, Sawaya NA. Phage puppet masters of the marine microbial realm. Nat Microbiol. 2018;3:754–66. https://doi.org/10.1038/s41564-018-0166-y.
62. Cardona T, Sanchez-Baracaldo P, Rutherford AW, Larkum AW. Early Archean origin of Photosystem II. Geobiology. 2019;17:127–50. https://doi.org/10.1111/gbi.12322.
63. Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J, et al. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. Proc Natl Acad Sci U S A. 2011;108(39):E757–64. https://doi.org/10.1073/pnas.1102164108.
64. 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. https://doi.org/10.1038/nature04111.
65. Sullivan MB, Lindell D, Lee JA, Thompson LR, Bielawski JP, Chisholm SW. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. PLOS Biology. 2006;4:e234. https://doi.org/10.1371/journal.pbio.0040234.
66. Puxty RJ, Evans DJ, Millard AD, Scanlan DJ. Energy limitation of cyanophage development: implications for marine carbon cycling. ISME J. 2018;12(5):1273–86. https://doi.org/10.1038/s41396-017-0043-3.
67. Dammeyer T, Bagby SC, Sullivan MB, Chisholm SW, Frankenberg-Dinkel N. Efficient phage-mediated pigment biosynthesis in oceanic cyanobacteria. Curr Biol. 2008;18(6):442–8. https://doi.org/10.1016/j.cub.2008.02.067.
68. Ledermann B, Beja O, Frankenberg-Dinkel N. New biosynthetic pathway for pink pigments from uncultured oceanic viruses. Environ Microbiol. 2016;18(12):4337–47. https://doi.org/10.1111/1462-2920.13290.
69. Esteves NC, Scharf BE. Serratia marcescens ATCC 274 increases production of the red pigment prodigiosin in response to Chi phage infection. Sci Rep. 2024;14(1):17750. https://doi.org/10.1038/s41598-024-68747-3.
70. Kimmance SA, Allen MJ, Pagarete A, Martínez Martínez J, Wilson WH. Reduction in photosystem II efficiency during a virus-controlled Emiliania huxleyi bloom. Mar Ecol Prog Ser. 2014;495:65–76. https://doi.org/10.3354/meps10527 .
71. Bale NJ, Airs RL, Kimmance SA, Llewellyn CA. Transformation of chlorophyll a during viral infection of Emiliania huxleyi. Aquat Microb Ecol. 2013;69(3):205–10. https://doi.org/10.3354/ame01640 .

Publisher's Note