



## OPEN ACCESS

## EDITED BY

Ajaya Kumar Rout,  
Rani Lakshmi Bai Central Agricultural  
University, India

## REVIEWED BY

Maggie Williams,  
Central the transmission of Michigan  
University, United States  
Weimin Gao,  
Barrow Neurological Institute (BNI),  
United States  
Weiyan Wang,  
Northwest A&F University, China  
Satya Narayan Parida,  
Rani Lakshmi Bai Central Agricultural  
University, India  
Sushree Swati Rout,  
Fakir Mohan University, India  
Sujata Dey,  
Central Inland Fisheries Research Institute  
(ICAR), India

## \*CORRESPONDENCE

Ryan A. Blaustein  
✉ rblauste@umd.edu

RECEIVED 26 November 2024

ACCEPTED 12 May 2025

PUBLISHED 04 June 2025

## CITATION

Blaustein RA, Smith JE, Toro M,  
Pachepsky Y and Stocker MD (2025) Water  
metagenomes reflect physicochemical water  
quality throughout a model agricultural pond.  
*Front. Microbiol.* 16:1535096.  
doi: 10.3389/fmicb.2025.1535096

## COPYRIGHT

© 2025 Blaustein, Smith, Toro, Pachepsky and  
Stocker. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Water metagenomes reflect physicochemical water quality throughout a model agricultural pond

Ryan A. Blaustein<sup>1\*</sup>, Jaclyn E. Smith<sup>2,3</sup>, Magaly Toro<sup>4</sup>,  
Yakov Pachepsky<sup>3</sup> and Matthew D. Stocker<sup>3</sup>

<sup>1</sup>Department of Nutrition and Food Science, University of Maryland, College Park, MD, United States, <sup>2</sup>Oak Ridge Institutes for Science and Education, Oak Ridge, TN, United States, <sup>3</sup>United States Department of Agriculture, Agricultural Research Service, Environmental Microbial and Food Safety Laboratory, Beltsville, MD, United States, <sup>4</sup>Joint Institute for Food Safety and Applied Nutrition, University of Maryland, College Park, MD, United States

Agricultural ponds are essential irrigation resources, though may also serve as reservoirs for pathogens and antimicrobial resistance (AMR) genes. While monitoring microbiological water quality is critical for food safety, the influence of sampling factors (e.g., when and where to collect samples) in making risk assessments and potential applications for using environmental covariates as indicators remain unclear. Here, we explored the hypothesis that metagenomes of agricultural waters change with spatiotemporal shifts in physicochemical water quality, i.e., across water depths over time. Water samples and underlying sediments were collected at a model pond at the surface and within the water column (0, 1, 2 m depths) throughout one day (i.e., 9:00, 12:00, 15:00). All samples were processed for shotgun metagenomic sequencing analysis and enumeration of various water quality parameters (e.g., temperature, nutrient concentrations, turbidity, pH, culturable *Escherichia coli*). At the pond surface, *Microcystis aeruginosa* and members of Cyanobacteria, along with genes encoding pathways related to photosynthesis and nucleotide biosynthesis, were enriched throughout the day. In contrast, within the water column (1–2 m depths) and sediments, diverse members of Proteobacteria and Actinobacteria were more dominant, along with encoded pathways related to respiration and amino acid biosynthesis. Various aspects of water quality (i.e., chlorophyll dissolved organic matter, ammonia, *E. coli* concentrations) correlated with water metagenome diversity, albeit not with any specific AMR genes or virulence factors. Nevertheless, *de novo* assembly of sequenced reads uncovered 22 unique strains encoding several AMR, virulence, or stress response genetic elements, thus linking metagenome functional potential to key taxa. Overall, our findings highlight distinctions in agricultural pond water metagenomes at the surface and in the water column and demonstrate the potential for metagenomic surveillance in water quality monitoring to support food safety.

## KEYWORDS

water microbiome, agriculture, water quality, metagenomics, antimicrobial resistance

## 1 Introduction

Aquatic microbiomes contain diverse planktonic and sediment-resuspended bacteria, viruses, algae, and fungi that collectively play important roles in nutrient biotransformation and cycling, decomposition of organic matter, and transmission and degradation of biological and chemical contaminants, among other ecosystem functions (Andreote et al., 2018; Escalas

et al., 2019; Rout et al., 2024a). Abiotic factors (e.g., temperature, nutrient availability, salinity) vary spatially and temporally in surface waters, including gradients along different depths or flow channels (Neneng et al., 2020). The water quality properties and the intrinsic microbiota both fluctuate at broad and fine spatial and temporal scales, such as between seasons or even within a given day (Newton et al., 2011; Mayr et al., 2020; Stocker et al., 2024; Zhao et al., 2024). Changes in temperature, dissolved oxygen, and nutrient influx in salt marshes have been attributed to restructuring of the aquatic microbial community (Kearns et al., 2016). Likewise, populations of cyanobacteria among the broader water microbiome have been reported to respond to diurnal changes in UV light intensity at different depths within lakes (Cameron et al., 2024). As such, key members of water microbiomes have been suggested as novel indicators for predicting pollution and other aspects of water quality (Wang et al., 2020; Stocker et al., 2024).

To date, most research on water microbiomes has applied 16S rRNA gene sequencing for taxonomic characterization or used more traditional culture-dependent approaches (Andreote et al., 2018; Wang et al., 2020; Cameron et al., 2024; Stocker et al., 2024; Zhao et al., 2024). Understanding relationships between water quality and microbiome functional potential remains largely unexplored. Shotgun metagenomic sequencing is a powerful tool that is increasingly used for microbiome taxonomic and functional profiling. Using deep metagenomic sequencing, He et al. (2024) discovered novel lineages of SAR202 bacteria and their encoded functions, which play important roles in global carbon cycling at different ocean depths. Another recent report on metagenomes in surface water samples in Chile and Brazil ( $n = 404$  metagenomes) identified a variety of microbial contaminants (i.e., *Escherichia*, *Listeria*, *Salmonella*) and diverse antimicrobial resistance (AMR) genes spanning 25 antimicrobial classes (Huang et al., 2024). Advancements in molecular surveillance tools to monitor changes in water quality and for genetic elements involved in AMR, virulence, or even microbial production of toxins (e.g., cyanotoxins) in water resources warrant further investigation for applications to support public health (Escalas et al., 2019; Liguori et al., 2022; Linz et al., 2023).

The microbial quality of agricultural irrigation water is a critical factor for food safety. The U.S. Food and Drug Administration (FDA) recently released the final rule on pre-harvest water as part of the Food Safety Modernization Act Produce Safety rule, requiring farms to perform thorough water quality evaluations, as least on an annual basis, for risk management decision-making purposes (U.S. Food and Drug Administration, 2024). Nevertheless, the influence of water sampling factors (e.g., numbers and volume of samples, water depth, timing of collection) and analytical frameworks on predictive outcomes is not well understood. Moreover, agricultural ponds are an important component in the “one-health” narrative as a sink and source of foodborne pathogens and AMR genes from the environment to fresh produce for human consumption (Franklin et al., 2024). While monitoring pond metagenomes has demonstrated seasonal or regional variations in microbial diversity, AMR, and functional profiles (Chopyk et al., 2020; Malayil et al., 2022), understanding the relationships between microbiome functional potential and environmental covariates associated with water quality (e.g., nutrient concentrations, turbidity) hold potential as novel indicators and improving monitoring recommendations.

Our previous work that utilized 16S rRNA gene sequencing demonstrated that bacterial community taxonomic diversity exhibits variation throughout a model pond at different depths over the course of the day (i.e., 9:00, 12:00, 15:00), with implications for improving water quality monitoring strategies (Stocker et al., 2024). Expanding upon these findings, the present study aimed to explore fine-scale spatial and temporal variations in the pond metagenomes to (1) determine differences at the surface and at depths in the water column over time to establish the impacts of sampling design on the observed microbiome diversity and (2) identify associations between metagenome functional potential (i.e., encoded pathways, AMR genes, virulence factors) and an array of water quality properties.

## 2 Methods

### 2.1 Sample collection, DNA extraction, and sequencing

Sampling was conducted at an agricultural pond at the University of Maryland’s Wye Research and Education Center on 09/15/2022, as reported previously (Stocker et al., 2024). The pond that was sampled (38.916 N 76.141 W) had an approximate surface area of 4,000 m<sup>2</sup> with average depth of 3 m. Here, we extended the original study by processing a subset of the water and sediment samples, which were collected at one site nearshore and two sites at the interior transect of the pond (i.e., L4, L5, and L11 in Stocker et al., 2024; Supplementary Figure S1), for shotgun metagenomics analysis. The sample selection enabled investigation of differences in pond water and sediment metagenomes as a factor of sampling location ( $n = 3$  within-pond sites) and water metagenomes further as a factor of sampling depth ( $n = 3$  gradient depths at L5 and L11) and time ( $n = 3$  time points at all locations). Specifically, surface water samples were collected at all locations from boat at 9:00, 12:00, and 15:00 using sterile 500 mL bottles with aseptic technique ( $n = 9$  water samples taken at the surface, i.e., 3 locations  $\times$  3 times). Gradient depth samples in the water column were also collected at these times at L5 and L11 at 1 m and 2 m depths ( $n = 12$  water samples taken at depth, i.e., 2 locations  $\times$  2 depths  $\times$  3 times). We used a weighted strainer attached to depth-marked vinyl tubing connected to a peristaltic pump to fill the bottles (SigmaMAX 900, Loveland, CO, United States) following an initial 30 s flush. Sediment cores were then collected after the final timepoint at 15:00 at each location ( $n = 3$  sediment samples taken, one at each location), i.e., to avoid microbial resuspensions during the water sampling events. The sediment was sampled at the bank site (L4) using sterile 50 mL conical tubes and at the internal sites (L5, L11) using an Eckman dredge.

DNA was extracted from the water and sediment samples using the DNeasy PowerWater Kit (Qiagen, Hilden, Germany) (Stocker et al., 2024). Water samples (50 mL) and mixed sediment samples (i.e., slurries prepared from 3 g sediment vortexed in 27 mL sterile DI water; 1:10 dilution) were passed through 0.45  $\mu$ m filters using a manifold vacuum. DNA was extracted from the filters and quantified with a Qubit 4 Fluorometer (Invitrogen, Waltham, MA) ( $n = 24$  samples,  $20.0 \pm 2.4$  ng/ $\mu$ l, mean  $\pm$  standard error). Negative controls included DNA extractions from DI water passed through

the same filters ( $n = 1$ ) and reagent blanks ( $n = 3$ ) that were processed alongside water/sediment samples, all of which yielded DNA below detection range. Metagenome libraries were prepared with the Illumina DNA Prep kit with UD Indexes (Illumina, San Diego, CA). Paired-end sequencing (300 cycles) targeting >15 million reads per sample was performed on an Illumina NextSeq1000 at the Joint Institute for Food Safety and Applied Nutrition.

## 2.2 Metagenome analysis

Raw sequenced reads were processed to remove potential sequencing contaminants and low-quality reads using Kneaddata v.0.6.1 with default parameters.<sup>1</sup> Microbiome taxonomic classification was performed with Kraken2 v2.1.2 (Wood et al., 2019) with the option “-confidence 0.1” (Saheb Kashaf et al., 2022; Blaustein et al., 2023), followed by species-level read correction with Braken (Lu et al., 2017). Metagenome functional profiling was performed using HUMAnN2 v.0.11.1 (Franzosa et al., 2018) with default parameters, and MetaCyc pathway abundances were approximated based on copies per million (CPM) of mapped reads. AMR genetic element and virulence factor (VF) reads per kilobase per million reads (RPKM) were determined with ShortBRED v0.9.5 (Kaminski et al., 2015) by running *shortbred\_quantify* with *shortbred\_identify* markers (length >15 amino acids) constructed from protein sequences of the Comprehensive Antibiotic Resistance Database (CARD) v3.2.7 (Alcock et al., 2020) (accessed 08/11/2023) and the Virulence Factor Database (Liu et al., 2022) (VFDB; accessed 08/27/2023), each with reference to UniRef50 (Suzek et al., 2015).

We further employed *de novo* assembly to recover metagenome-assembled genomes (MAGs) as previously described (Blaustein et al., 2023). In brief, metaSPAdes v.3.15.0 (Bankevich et al., 2012), MetaWRAP v.1.2.2 (Uritskiy et al., 2018), and GUNC v1.0.5 (Orakov et al., 2021) were used, each with default parameters, for assembly, binning (with *concoct*, *maxbin2*, and *metabat2*), and prediction and removal of chimeric MAGs, respectively. For the latter, MAGs with contamination greater than 0.05, clade separation greater than 0.45 and a reference representation score greater than 0.5 were considered chimeric and removed (Saheb Kashaf et al., 2022). We then used *dRep* v2.6.2 (Olm et al., 2017) to dereplicate refined bins with predicted completeness greater than 50% and contamination less than 5% based on an average nucleotide identity (ANI) of 99% and minimum overlap of 30% (i.e., to approximate the strain level). The same parameters with 95% ANI were used to refine dereplicated MAGs to the species level. MAGs were given taxonomic assignment using GTDB-Tk v1.3.0 (Chaumeil et al., 2022) and novelty was established based on relative evolutionary divergence (RED) score. The generated protein sequence alignments were used to construct a phylogenetic tree via *FastTree* v.2.1.10 (Price et al., 2010) that was visualized with *ape* (Paradis and Schliep, 2019). MAGs were then queried for genes encoding proteins with homology to AMR and virulence-associated features using *AMRFinderPlus* v3.12.8

(Feldgarden et al., 2021) (database 07-22-2024) with settings for 50% amino acid alignment identity (i.e., since the MAGs are non-model organisms) and the flag “-plus” to return the comprehensive set of AMR, virulence, and stress response genes.

## 2.3 Water quality

Water quality properties associated with each sample were measured *in situ* and via laboratory analysis. In the field, a water sample collection strainer was attached to the sensor guard of a YSI EXO-2 multiparameter water quality sonde that was calibrated with manufacturer standards prior to field use (Yellow Springs Instruments, Yellow Springs, Ohio, United States). The sonde provided *in situ* measurements of temperature (°C), dissolved oxygen (DO) (mg L<sup>-1</sup>), specific conductivity (SPC) (μS cm<sup>-1</sup>), turbidity (NTU), phycocyanin (YSI-PC) (relative fluorescent unit; RFU), chlorophyll-a (YSI-CHL) (RFU), and fluorescent dissolved organic matter (FDOM) (μg L<sup>-1</sup>). At the lab, aliquots of the water samples were processed for *Escherichia coli* concentrations (CFU 100 mL<sup>-1</sup>) via enumeration using membrane filtration and modified mTEC agar (ECCFU) (BD Difco, Sparks, Maryland). The water samples were also processed to measure phycocyanin (PHYC) (μg L<sup>-1</sup>), chlorophyll-a (INV) (RFU), colored dissolved organic matter (CDOM) (μg L<sup>-1</sup>), and nutrient concentrations of orthophosphate (PO<sub>4</sub>) (mg L<sup>-1</sup>), nitrate (NO<sub>3</sub>) (mg L<sup>-1</sup>), dissolved ammonia (NH<sub>3</sub>) (mg L<sup>-1</sup>), total organic carbon (TOC) (mg L<sup>-1</sup>), total inorganic carbon (TIC) (mg L<sup>-1</sup>), and total nitrogen (TN) (mg L<sup>-1</sup>) using methods described in our previous study (Stocker et al., 2024).

## 2.4 Statistical analysis

All statistics and data visualizations were completed in R v.4.3.1. *Vegan*<sup>2</sup> was used to compute  $\alpha$ -diversity (Shannon index) and  $\beta$ -diversity (Bray–Curtis dissimilarity) for microbiome taxonomic profiles as relative abundances at the species level and functional profiles as relative abundances of MetaCyc pathway CPMs. The effects of sampling factors (i.e., sampling location, water column depth, time of day) on microbiome  $\alpha$ -diversity and  $\beta$ -diversity were determined with ANOVA and PERMANOVA, respectively. Correlations with the measured water quality properties were evaluated with the Spearman’s rank test ( $\alpha$ -diversity) and PERMANOVA ( $\beta$ -diversity) as well. A linear mixed-effects (LME) model was employed to determine differential abundances of microbial species (i.e., those that were >0.1% abundant on averages across all samples), MetaCyc pathways, and AMR and VF genetic elements based on fixed effects for the sampling depth (i.e., water surface, water column, sediments) and random effects for sampling location (i.e., L4, L5, L11) and time of day (i.e., 9:00, 12:00, 15:00). *p*-values were adjusted with Benjamini–Hochberg correction and  $q < 0.05$  was considered significant. The total number of genes per MAG with homology to AMR and VF genetic elements was compared for MAGs belonging to different bacterial phyla using the Dunn’s test with Benjamini–Hochberg correction.

<sup>1</sup> <http://huttenhower.sph.harvard.edu/kneaddata>

<sup>2</sup> <https://github.com/vegandevs/vegan>

### 3 Results

#### 3.1 Pond water and sediment metagenomic profiles correlate with sampling factors and water quality

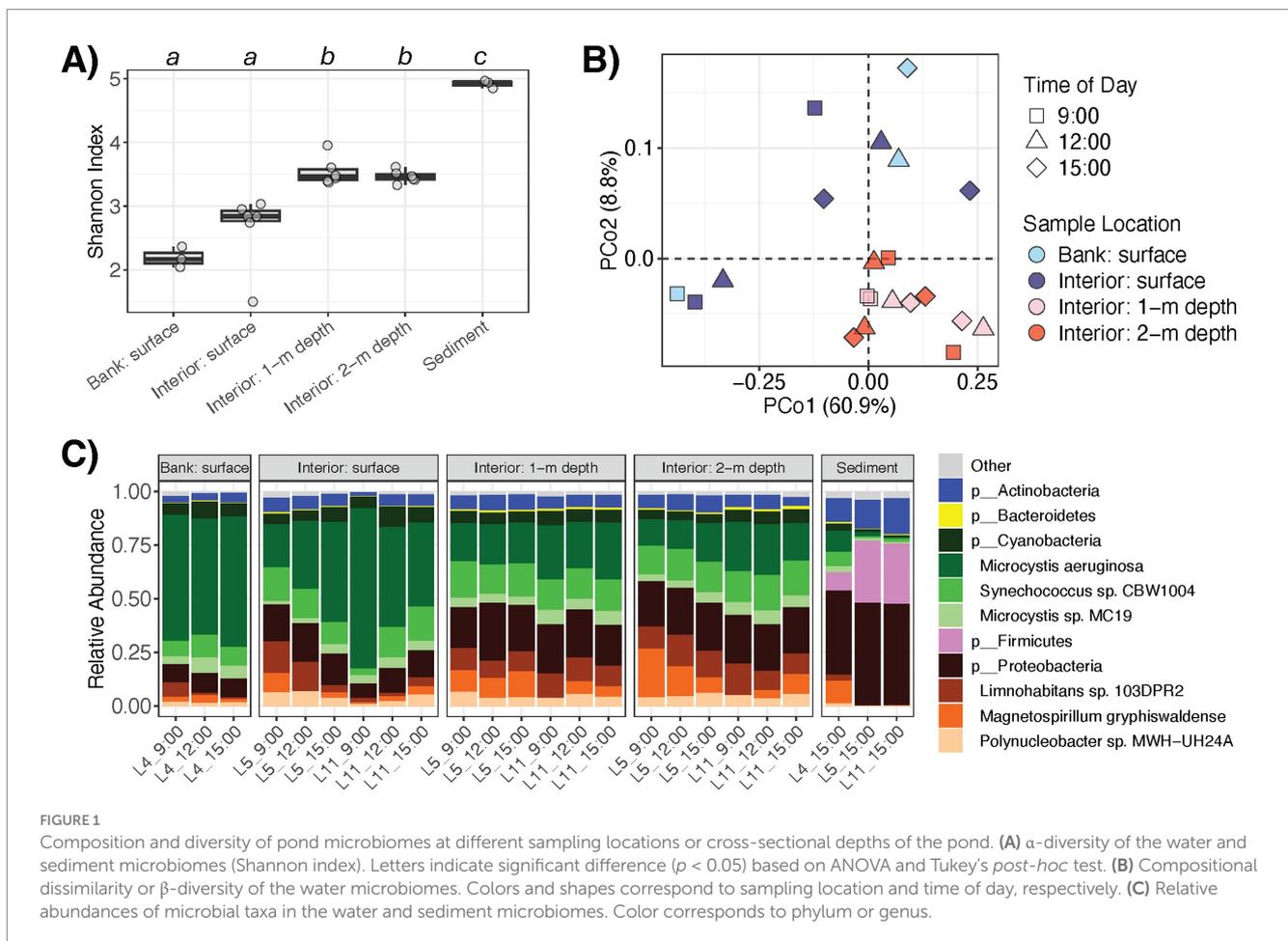
Metagenomic sequencing of the water ( $n = 21$ ) and sediment ( $n = 3$ ) samples generated 302.8 M sequenced reads (or 200 Gb data), with  $12.6 \pm 2.6$  M paired end reads per sample (mean  $\pm$  standard deviation). After pre-processing, there were  $8.8 \pm 2.0$  M high quality non-human paired reads per sample from which 1,303 microbial species representing 613 genera, 238 families, 134 orders, 70 classes, and 33 phyla were detected.

Pond microbiome  $\alpha$ -diversity was lowest at the water surface, especially near shore (L4) compared to the interior sites (L5 and L11), and it was significantly higher within the water column and highest in the pond sediments (Figure 1A). At L5 and L11, the water microbiomes at the surface exhibited significantly lower  $\alpha$ -diversity as compared to those at the 1 m ( $q = 0.002$ ) and 2 m ( $q = 0.004$ ) depths throughout the day, though there were no differences detected between the 1 m and 2 m depths ( $q = 0.977$ ). Accordingly, microbial community profiles at the water surface (all sites) and those in the water column (i.e., L5 and L11 at depths of 1 m and 2 m) clustered apart (PERMANOVA  $R^2 = 0.206$ ,  $p = 0.008$ ) (Figure 1B). On the contrary, the time of day (i.e., 9:00, 12:00, 15:00) did not appear to

significantly associate with  $\alpha$ -diversity ( $p = 0.669$ ) or  $\beta$ -diversity (PERMANOVA  $R^2 = 0.132$ ,  $p = 0.228$ ) of the water microbiota. Thus, the pond water microbiomes were substantially different between the surface and depths in the water column with variation that appeared to be consistent throughout the day.

The most abundant phyla within the water microbiomes were Cyanobacteria, Proteobacteria, Actinobacteria, and Bacteroidetes, which comprised on average 55.4, 36.5, 5.6, and 0.7% of the microbial communities, respectively (Figure 1C). Members of Cyanobacteria such as dominant *Microcystis aeruginosa* were significantly more abundant at the water surface, while members of Proteobacteria, Actinobacteria, and Bacteroidetes were more abundant within the water column ( $q < 0.05$  for each) (Figure 2A). In total, there were three microbial species (*M. aeruginosa*, *Microcystis viridis*, *Halomonas* sp. JS92-SW72) that were more prominent at the pond water surface and 36 microbial species (e.g., *Synechococcus* sp. CBW1004, *Limnohabitans* sp. 103DPR2, *Magnetospirillum gryphiswaldense*, *Microcystis* sp. MC19, *Phenylobacterium parvum*, *Acidovorax* sp. T1) that had higher relative abundances in the water column ( $q < 0.05$ ; Supplementary Table S1).

The pond sediments harbored microbial communities with the highest  $\alpha$ -diversity with comparable abundances of Proteobacteria to the water column microbiome, along with higher abundances of Actinobacteria and Firmicutes and lower abundances of Cyanobacteria, suggesting gradient ecological dynamics with microbial settling and resuspension (Figure 1C). There were 25 species



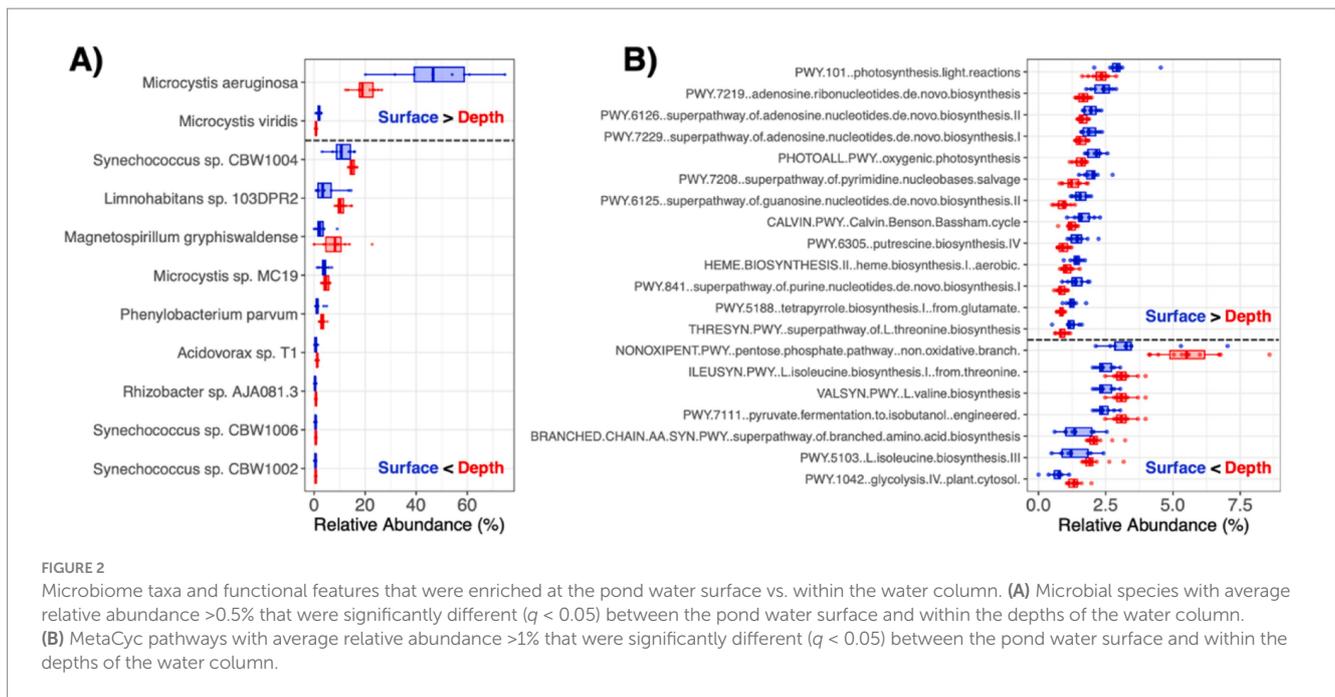


FIGURE 2

Microbiome taxa and functional features that were enriched at the pond water surface vs. within the water column. (A) Microbial species with average relative abundance >0.5% that were significantly different ( $q < 0.05$ ) between the pond water surface and within the depths of the water column. (B) MetaCyc pathways with average relative abundance >1% that were significantly different ( $q < 0.05$ ) between the pond water surface and within the depths of the water column.

that were significantly more abundant in the water column than the in the sediments, such as *M. aeruginosa* and *M. viridis*, which were further dominant at the water surface. Alternatively, there were 11 microbial species enriched in the sediment microbiome, including *Serratia marcescens*, *Rhizobacter* sp. AJA081-3, *Shinella* sp. XGS7, *Methylocystis parvus*, *Rubrivivax gelatinosus*, among others ( $q < 0.05$ ; Supplementary Table S1).

As described in our previous report (Stocker et al., 2024), the set of studied water quality properties also exhibited spatial variation throughout the pond. Here, most of these measured water variables correlated with taxonomic profile  $\alpha$ -diversity (Table 1), which may be attributed to confounded differences in both water quality and microbial diversity at the surface and within the water column (Figure 1A). For example, water temperature was lower at increasing depths and inversely correlated with water microbiota  $\alpha$ -diversity ( $\rho = -0.686, p = 0.001$ ), perhaps reflecting microbial resuspensions around the underlying sediments. In addition, *E. coli* (CFU 100 mL<sup>-1</sup>) was more abundant in the water column than at the surface ( $p = 0.001$ ) and correlated with  $\alpha$ -diversity ( $\rho = 0.519, p = 0.016$ ) (Figure 3). Moreover, water properties linked to proliferation of algae or Cyanobacteria (e.g., chlorophyll and phycocyanin contents, colored dissolved organic matter), which was dominant at the surface (Figure 1C), were significantly associated  $\beta$ -diversity of the water microbiomes as well (Table 1).

### 3.2 Pond microbiome functional potential

In line with trends for the microbial taxa, the encoded metabolic profiles significantly differed between the surface and within the water column (PERMANOVA  $R^2 = 0.196, p = 0.013$ ; Supplementary Figure S2), though not at the different sampling times of day (PERMANOVA  $R^2 = 0.101, p = 0.400$ ). As water quality also differed across scale, particularly between the surface and deeper into the water column, the water metagenome functional profiles

TABLE 1 Metagenome correlations with pond water quality properties.

Water property	Microbial taxa $\alpha$ -diversity		Microbial taxa profile		Functional (MetaCyc) profile	
	$\rho$	$P$	$R^2$	$P$	$R^2$	$P$
Temp, °C	<b>-0.686</b>	<b>0.001</b>	0.063	0.240	0.064	0.254
DO	<b>-0.757</b>	<b>&lt;0.001</b>	0.062	0.234	0.085	0.176
SPC	0.344	0.138	0.046	0.407	0.073	0.181
pH	<b>-0.760</b>	<b>&lt;0.001</b>	0.056	0.330	0.048	0.362
NTU	<b>-0.589</b>	<b>0.007</b>	0.088	0.164	<b>0.192</b>	<b>0.015</b>
YSI-PC	<b>-0.749</b>	<b>&lt;0.001</b>	<b>0.398</b>	<b>0.002</b>	<b>0.377</b>	<b>0.001</b>
YSI-CHL	<b>-0.698</b>	<b>0.001</b>	<b>0.413</b>	<b>0.001</b>	<b>0.368</b>	<b>0.001</b>
$f$ DOM	<b>0.674</b>	<b>0.001</b>	<b>0.193</b>	<b>0.015</b>	<b>0.213</b>	<b>0.012</b>
PO4	0.205	0.371	0.054	0.293	0.026	0.693
NH <sub>3</sub>	<b>0.745</b>	<b>&lt;0.001</b>	0.102	0.098	<b>0.145</b>	<b>0.031</b>
NO <sub>3</sub>	0.149	0.517	0.053	0.304	0.033	0.579
ECCFU	<b>0.519</b>	<b>0.016</b>	0.094	0.116	<b>0.149</b>	<b>0.030</b>
TC	0.243	0.287	0.030	0.627	0.028	0.628
TIC	<b>0.469</b>	<b>0.033</b>	0.046	0.382	0.060	0.259
TN	<b>0.608</b>	<b>0.004</b>	0.096	0.102	0.082	0.153
TOC	-0.042	0.858	0.023	0.765	0.031	0.622
CDOM	<b>-0.679</b>	<b>0.001</b>	<b>0.291</b>	<b>0.003</b>	<b>0.259</b>	<b>0.002</b>
INV	<b>-0.584</b>	<b>0.006</b>	<b>0.306</b>	<b>0.001</b>	<b>0.252</b>	<b>0.003</b>
PHYC	<b>-0.808</b>	<b>&lt;0.001</b>	<b>0.288</b>	<b>0.002</b>	<b>0.248</b>	<b>0.005</b>

Significant associations ( $p < 0.05$ ) are displayed in bold font.

correlated with many of the measured variables (Table 1). In addition to significantly correlating with the same water properties as microbial taxonomic profiles, functional profiles further exhibited significant

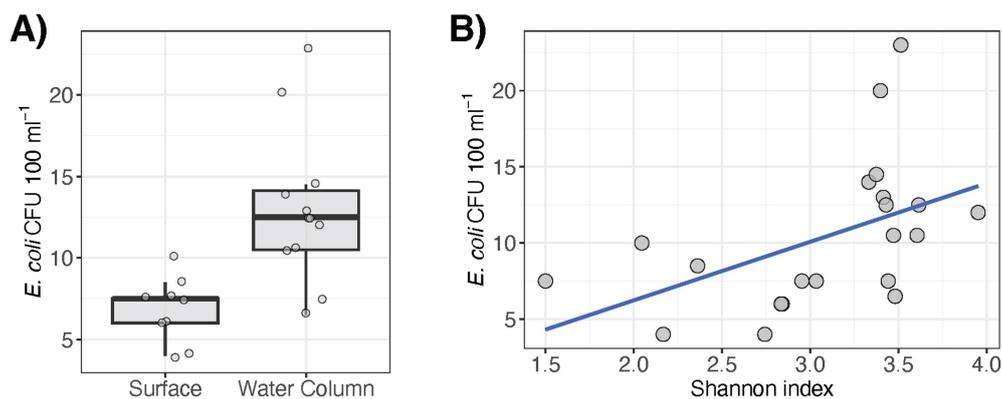


FIGURE 3

*E. coli* variation in pond water correlates with (A) sampling depth ( $p = 0.001$ ) and (B) microbiome  $\alpha$ -diversity ( $\rho = 0.519$ ,  $p = 0.016$ ).

associations with NTU (PERMANOVA  $R^2 = 0.192$ ,  $p = 0.015$ ), NH<sub>3</sub> (PERMANOVA  $R^2 = 0.145$ ,  $p = 0.031$ ), and ECCFU (PERMANOVA  $R^2 = 0.149$ ,  $p = 0.030$ ), suggesting that microbiome functional diversity may be more tightly linked to water quality than microbial taxa composition and structure.

There were 41 and 40 total metabolic pathways that were enriched ( $q < 0.05$ ) at the surface vs. within the water column ( $q < 0.05$ ; Supplementary Table S2). At the pond surface, pathways involved in photosynthesis (e.g., photosynthesis light reactions, oxygenic photosynthesis, Calvin-Benson-Bassham cycle) and nucleotide biosynthesis (e.g., synthesis of pyrimidines, purines, adenosine and guanosine) were significantly more abundant (Figure 2B). Alternatively, pathways involved in respiration (e.g., pentose-phosphate pathway, pyruvate fermentation, glycolysis) and amino acid biosynthesis (e.g., isoleucine, valine, branched amino acids) were more abundant in metagenomes in the water column (Figure 2B). These trends extended into water column-sediment dynamics. There were 46 pathways significantly more abundant in the water column (e.g., pentose-phosphate pathway, oxygenic photosynthesis, Calvin-Benson-Bassham cycle, hydrogen production), while 27 were enriched in the underlying sediments (e.g., pyruvate fermentation, synthesis of pyrimidines, purines, adenosine and guanosine) ( $q < 0.05$  for each; Supplementary Table S2).

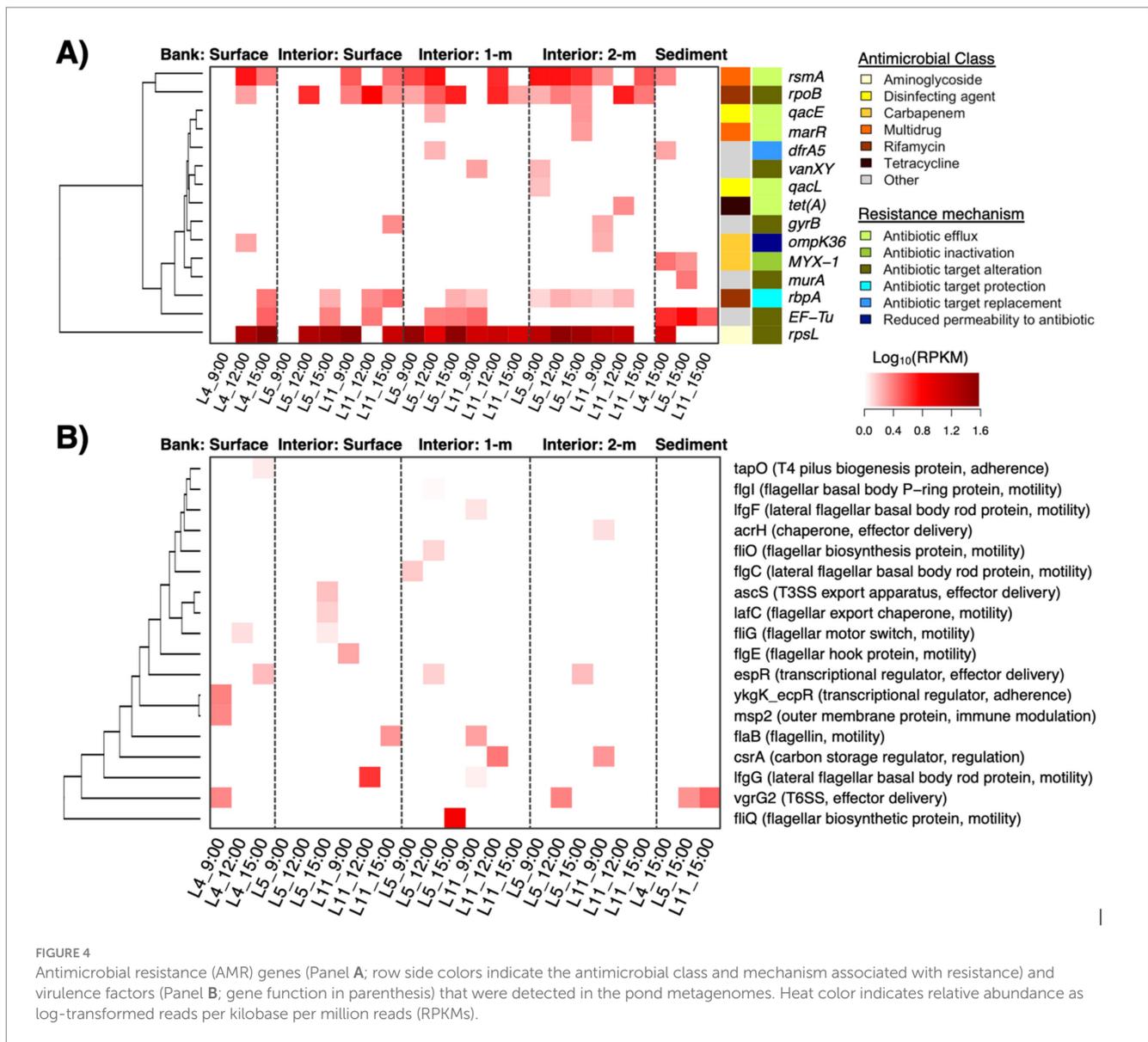
We further identified 15 AMR and 18 VF genetic elements within the pond metagenomes. The AMR elements encoded resistance to aminoglycosides, carbapenems, rifamycin, tetracyclines, and multidrug efflux. Spatial distributions of certain genes were generally sporadic and there were several instances of detection in only one sample type (e.g., *marR*, *qacL*, or *tet(A)* at the 2 m water depth) (Figure 4A). Only four AMR genes were detected in water samples across all pond depths (i.e., *rsmA*, *rpoB*, *rbpA*, and *rpsL*) and most of these commonly occurring genes were largely housekeeping features that may be involved in AMR via antibiotic target alteration under specific mutations. Moreover, notable AMR genes that were detected in pond water and not sediments included *qacE* that is involved in resistance to disinfecting agents via efflux, *ompK36* that can result in reduced permeability to carbapenems, and *gyrB* and *vanXY*, both of which are conferred via target alteration. On the other hand, AMR genes detected in sediment samples and not the water column were with

*MYX-1* and *murA*, which encode enzymatic inactivation (carbapenemase) and potential antibiotic target alteration, respectively. Moreover, the VFs primarily consisted of flagellar machinery for motility, effector systems (e.g., T3SS, T6SS), and proteins involved in adherence (Figure 4B). In contrast to the metabolic pathways that were differentially abundant at the surface and within the water column, the relative abundances of AMR and VF genes detected were low and did not correlate with the sampling factors or the water quality properties measured in this study ( $q > 0.05$  for each genetic element).

### 3.3 MAGs reflect dominant taxa carrying AMR and virulence traits

Processing the metagenomes with *de novo* assembly yielded 67 MAGs with at least medium-quality (i.e., predicted completeness >50% and contamination <5%) that were recovered from the water microbiomes, with at least one MAG recovered from each sample. Dereplication at 99% ANI yielded 22 distinct strains spanning Actinobacteriota ( $n = 1$ ; referred to as Actinobacteria in the built database for the Kraken2 analysis), Bacteroidota ( $n = 3$ ; referred to as Bacteroidetes in the built database for the Kraken2 analysis), Cyanobacteria ( $n = 12$ ), Proteobacteria ( $n = 5$ ), and Verrucomicrobiota ( $n = 1$ ) (Figure 5). These taxa comprised 12 non-redundant species, all of which were considered novel taxa based on ANI and RED values, i.e., the strain for *Microcystis wesenbergii* in Figure 5 clustered with *Microcystis* sp. at the species-level (>95% ANI). The frequency of Cyanobacteria and Proteobacteria strains among the recovered MAG was consistent with their dominance among the pond microbiomes (i.e., 54.6 and 22.7% of MAGs and 55.4 and 36.5% in Kraken2/Bracken, respectively).

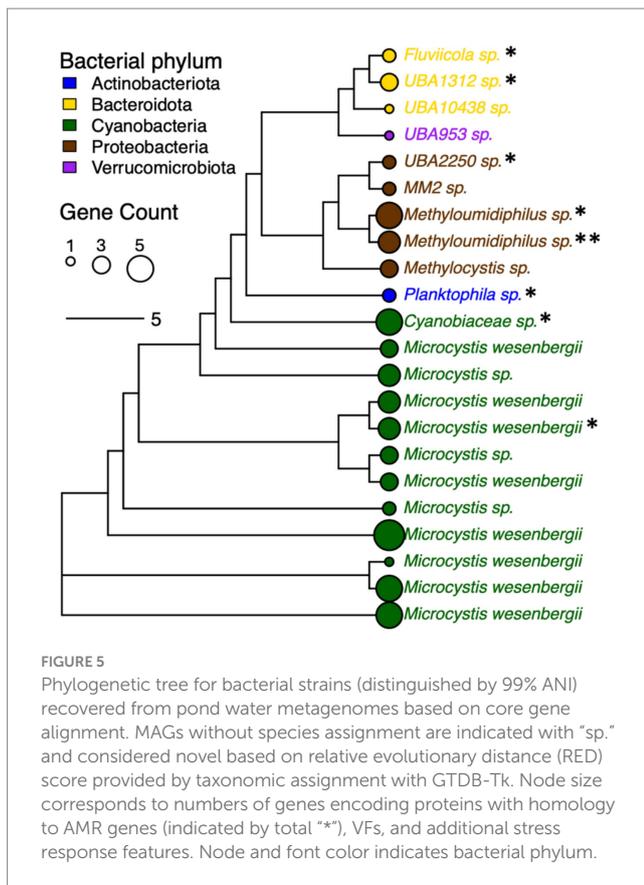
The various strains encoded several genes with homology to genes associated with AMR ( $n = 6$ ), VFs ( $n = 2$ ), and stress response ( $n = 18$ ) (Supplementary Table S3). The AMR features identified among MAGs were involved in resistance to beta-lactams (*blaOXA*, *blaPAU*), tetracycline (*tetA(58)*), trimethoprim (*dfrA3*), pleuromutilin (*taeA*), or multidrug efflux (*ranA*). The VFs included *bpsD* and *katP*. The stress response features included biocide resistance via multidrug



efflux (e.g., *ssmE*, *smr*, *qacE*, *qacF*), heat shock (*clpK*), and a variety of genes involved in tolerance to metals such as nickel and copper (Supplementary Table S3). Thus, there were consistencies in key features identified with the more general read-based characterization of metagenomes with ShortBRED described earlier, such as *dfrA5*, *acrH*, *tetA*, *qacE*, and two genes associated with resistance to beta-lactams or carbapenems (Figure 4A). At least one AMR genetic element was found to be associated with 66.7 and 60% of strains representing Bacteroidota ( $n = 2/3$  MAGs) and Proteobacteria ( $n = 3/5$  MAGs), while for Cyanobacteria this was only 16.7% of strains ( $n = 2/12$  MAGs) (Figure 5). Although specific gene RPKMs were not correlated with sampling factors, perhaps reflecting sequencing depth limitations as described earlier, the observed trends in MAGs suggested that AMR may be disproportionately associated with the Proteobacteria and other water column/sediment-enriched taxa. Nevertheless, we were able to link AMR, virulence, and stress response genes to the range of dominant strains throughout the agricultural pond.

## 4 Discussion

The physicochemical properties of agricultural waters fluctuate spatiotemporally, concurrent with shifts in microbial community dynamics (Stocker et al., 2024). While the US FDA has implemented recommendations for farm-scale monitoring of pre-harvest water quality to enhance applications for food safety (U.S. Food and Drug Administration, 2024), understanding how specific sampling strategy and design may impact observations is important for practical risk management decision-making frameworks. The present study aimed to characterize metagenomes of agricultural pond water as a factor of sampling location, water depth, and time of day. Microbial  $\alpha$ -diversity was lowest at the water surface, where members of Cyanobacteria were most abundant, and increased in the water column toward the sediments. In line with the surface-enriched phytoplankton, we observed correlations between microbiome  $\beta$ -diversity and phycocyanin, chlorophyll-a, and concentrations of *f*DOM and *c*DOM. Previous studies on lake microbiota have also demonstrated that DOM plays an important



role in shaping microbial community composition (Crump et al., 2003; Kent et al., 2006; Amaral et al., 2016; Hengy et al., 2017). Moreover, shifts toward a more diverse microbiome within the water column, thereby being inversely proportional to DO, were consistent with previous reports for lakes (Bernard et al., 2019). Overall, the pond water microbiota correlated with water quality properties that differed between the pond surface and within the water column.

Similarly, the pond water microbiome functional potential was depth-dependent and correlated with many of the measured water quality variables as well. This was somewhat expected, as a stratification within water bodies gives rise to distinct changes in light, temperature, DO, and pH (i.e., epilimnion at the surface and hypolimnion with depth) (Fenchel and Finlay, 2008). Metabolic pathways at the surface prominently featured photosynthesis and nucleotide biosynthesis, while those in the water column and sediments transitioned to increases in respiration, the non-oxidative branch of the pentose phosphate pathway (PPP), and amino acid biosynthesis. Cyanobacteria and other dominant water microflora can shift metabolic pathways in response to predation and environmental cues, which may follow diurnal patterns (Pattanayak et al., 2020; Selim et al., 2021; Grzesiuk et al., 2022). When oxygen becomes limited, bacterioplankton can perform anaerobic glycolysis via the PPP to produce NADPH and other metabolic intermediates (Mckinley, 1977; Bjerkas and Bjornstad, 1999). Phytoplankton also activate PPP under dark conditions (e.g., in the water column with increasing depth) when Photosystem II cannot be utilized (Stal and Moezelaar, 2006). The role of microbiome metabolism in mediating water quality changes warrants further investigation.

Microbial diversity increased with pond water depth and was greatest in the underlying sediments, which are a sink for settling organic matter. Consistent with our observations, Firmicutes have

been reported to often comprise a major fraction of lake and pond sediment microbiota (Wang et al., 2019; Merino et al., 2023). These taxa degrade cellulose and lignin of decomposing plant matter and are involved in denitrification (Zhang et al., 2019). Accordingly, levels of ammonia (NH<sub>3</sub>) in the water column were higher than at the surface, which likely reflected flux from bacterial activity in the sediments (Leoni et al., 2018). As a sink for nutrients and microbiota diversity, sediments can also become a source due to diffusion and resuspension (Hassard et al., 2016; Dadi et al., 2017) that may play an important role in the water quality dynamics throughout the pond.

Developing an effective water quality monitoring strategy requires consideration of variation across potential sampling locations. While the proportions of bacteria at the water surface were relatively stable throughout the day (i.e., 9:00, 12:00, 15:00) at the pond bank (L4), larger shifts at the surface were noted for the interior sampling sites (L5 and L11). This may be related to *Microcystis* changing its vertical position in the water column via gas vacuoles in response to solar radiation, which was greater at surface-interior than surface-bank sites due to partial shading from vegetation at the perimeter of the pond (Smith et al., 2022). Limited water depths at the nearshore sites (e.g., approximately 0.3–0.5 m maximum) also inherently reduced the capacity for vertical translocation. Thus, microbial water quality toward the interior of agricultural ponds may have been more influenced by diurnal changes in light than that at the shoreline due to both broader depths and relatively limited cover. Future research at additional ponds or along more sampling time points is needed to better understand within-site spatiotemporal variation in microbiome interactions with water quality.

Water metagenome diversity correlated with turbidity and *E. coli* concentrations, both of which were greater in the water column than at the surface. As *E. coli* is a routinely monitored indicator organism, the moderate, yet significant positive association with microbiome  $\alpha$ -diversity may have implications for water quality monitoring for food safety risk analysis. The emergence of AMR and virulence features among environmental microbiota is also an emerging concern for public health. A variety of associated genes were identified in the pond microbiomes. For example, variants in *rpsL*, which can confer resistance to streptomycin and other aminoglycosides via key mutations, were detected at the highest abundance and frequency in the pond. These results are consistent with the study by Chopyk et al. (2020) that described *rpsL* to be the most prevalent resistance element among 33 detected AMR genes across 9 different irrigation water sources in the United States. In addition, *rpoB* and *Ef-Tu* (i.e., genes also involved in resistance via potential mutations) were also widely detected in the water and sediments, respectively, which is consistent with previous reports on rivers, lakes, and subaqueous soils (Mittal et al., 2019; Zhao et al., 2019; Su et al., 2020). Whether these housekeeping genes that are implicated in potential AMR exhibited variants of clinical significance was outside of the scope of this work. While efflux pump *rsmA* was found in more than half of the samples collected in this work, which was also seen in other studies (Rout et al., 2024b), additional AMR genes associated with multidrug efflux and carbapenem or tetracycline resistance via enzymatic resistance mechanisms were detected sporadically in the pond and may have implications for food safety as well. Like most of the AMR features, the VFs identified (e.g., encoding adherence, flagellar motility, secretion system) were detected in a limited

number of samples. Although the sources for these genetic elements remain unclear, emerging chemical contaminants (e.g., heavy metals, antibiotics and other antimicrobials) have gained attention for possible adverse impacts on environmental microbiomes, and even more in wastewater systems (Bueno et al., 2021; Guruge et al., 2021; Sharma et al., 2025). Exploring how chemical drivers may associate with AMR in agricultural resources, especially under changing environmental conditions, is needed to inform strategies to mitigate potential risks.

While differences in AMR genetic elements across sampling locations and depths were not elucidated in the scope of this work, perhaps reflective of limited sample size and sequencing depth, we were able to link key features to dominant microbial strains within Cyanobacteria, Proteobacteria, and other phyla. There were a variety of AMR, virulence, and/or general stress response genes found in all MAGs recovered, which is consistent with other reports (Huang et al., 2024; Jiang et al., 2024). Notably, AMR genes were less frequently carried by strains of the surface-associated Cyanobacteria compared to other phyla. The greater microbial diversity within the sediments and water column (e.g., where Proteobacteria and Bacteroidetes were more dominant) suggests these sampling locations as more of a hotspot for genes encoding AMR and other stress responses. As bottom sediments of streams, lakes, rivers, and wastewaters have been reported as key reservoirs for AMR genes and the mobile genetic elements that mediate their transmission (Calero-Cáceres et al., 2017; Hess et al., 2018; Chen et al., 2020; Eramo et al., 2020), similar dynamics may occur in agricultural water systems. Employing long read metagenomics in future studies may help to better understand whether mobile elements are co-localized with AMR or virulence genes, as well as key metabolic pathways (e.g., PPP), and potentially enriched among taxa that exhibit pond depth variation.

We recognize several limitations to this study. While we focused on discerning the influence of sampling factors on observed water metagenome diversity, future studies collecting samples from various agricultural water sources known to differ in intrinsic properties would be essential to better understand the relationship between water metagenomes and water quality. Moreover, although our metagenomic approach offered opportunity for non-targeted molecular surveillance, the sensitivity of AMR gene detection was limited, likely due to low abundances in the natural microbiomes. Deeper metagenomic sequencing is essential in future studies to determine whether AMR and virulence in agricultural waters may exhibit patterns across space and time. For example, a recent study that characterized broad diversity of 1,582 AMR genetic elements (representing 25 antibiotic classes) in 404 surface water samples in South America targeted >20 million metagenomic read pairs per sample (Huang et al., 2024), which was about two times greater than the reads obtained in this work. In addition, employing “quasi-metagenomics” that involves some extent of culture enrichments prior to sequencing (Commichaux et al., 2021; Kocurek et al., 2023) may be necessary to enhance strategies that will help provide the systems-level understanding between foodborne pathogens and water quality.

## 5 Conclusion

The microbiological and physicochemical quality of agricultural water has significant implications for food safety, including emerging

concerns for transmission of genetic elements involved in AMR. However, correlations between specific biological and environmental parameters, which may concurrently exhibit spatial and temporal variation (e.g., changes with water depth and time of day), are not clear. We addressed this gap by incorporating metagenomics into a comprehensive water quality assessment of an agricultural pond. Microbial taxonomic profiles and metabolic pathways largely varied between the water surface and at one and two meter depths in the water column, perhaps reflecting microbial exchanges with the underlying sediments. A variety of genetic elements encoding AMR, virulence, and other forms of bacterial stress response were identified throughout the pond, some of which were linked to specific bacterial strains within the water microbiome. Overall, our findings highlight key differences in water metagenomes throughout the pond that significantly correlated with dynamic water properties (e.g., concentrations of chlorophyll, DOM, ammonia, culturable *E. coli*). Spatiotemporal variation in water quality must be considered when developing applications for molecular surveillance in food safety risk assessment frameworks.

## Data availability statement

Raw sequencing data for shotgun metagenomes of pond water and sediment samples, along with the MAG genome assemblies, are available under NCBI BioProject accession number PRJNA1062331. Scripts and code that may be used to reproduce our analyses are available at [https://github.com/rblaustein/2025\\_Wye\\_pond\\_metagenomes](https://github.com/rblaustein/2025_Wye_pond_metagenomes).

## Author contributions

RB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. JS: Investigation, Writing – review & editing. MT: Methodology, Writing – review & editing. YP: Conceptualization, Writing – review & editing. MS: Conceptualization, Data curation, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported in part by start up funds to Blaustein from the University of Maryland.

## Acknowledgments

We thank Erin Harrelson for assistance with running AMRFinderPlus on the MAGs and for providing helpful feedback on the manuscript.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2025.1535096/full#supplementary-material>

## References

- Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525. doi: 10.1093/nar/gkz935
- Amaral, V., Graeber, D., Calliari, D., and Alonso, C. (2016). Strong linkages between DOM optical properties and main clades of aquatic bacteria. *Limnol. Oceanogr.* 61, 906–918. doi: 10.1002/lno.10258
- Andreote, A. P. D., Dini-Andreote, F., Rigonato, J., Machineski, G. S., Souza, B. C. E., Barbiero, L., et al. (2018). Contrasting the genetic patterns of microbial communities in soda lakes with and without cyanobacterial bloom. *Front. Microbiol.* 9:244. doi: 10.3389/fmicb.2018.00244
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Bernard, C., Escalas, A., Villeriot, N., Agogue, H., Hugoni, M., Duval, C., et al. (2019). Very low phytoplankton diversity in a tropical saline-alkaline Lake, with co-dominance of *Arthrospira fusiformis* (Cyanobacteria) and *Picocystis salinarum* (Chlorophyta). *Microb. Ecol.* 78, 603–617. doi: 10.1007/s00248-019-01332-8
- Bjerkas, E., and Bjornstad, E. (1999). Is there a connection between rapid fluctuation in water temperature and Catatract development in Atlantic Salmon (*Salmo Salar* L.). *Bull. Eur. Assoc. Fish Pathol.* 4:166.
- Blaustein, R. A., Shen, Z., Kashaf, S. S., Lee-Lin, S. Q., Conlan, S., Thomas, J., et al. (2023). Expanded microbiome niches of RAG-deficient patients. *Cell Rep. Med.* 4:101205. doi: 10.1016/j.xcr.2023.101205
- Bueno, I., Beaudoin, A., Arnold, W. A., Kim, T., Frankson, L. E., LaPara, T. M., et al. (2021). Quantifying and predicting antimicrobials and antimicrobial resistance genes in waterbodies through a holistic approach: a study in Minnesota, United States. *Sci. Rep.* 11:18747. doi: 10.1038/s41598-021-98300-5
- Calero-Cáceres, W., Méndez, J., Martín-Díaz, J., and Muniesa, M. (2017). The occurrence of antibiotic resistance genes in a Mediterranean river and their persistence in the riverbed sediment. *Environ. Pollut.* 223, 384–394. doi: 10.1016/j.envpol.2017.01.035
- Cameron, E. S., Krishna, A., Emelko, M. B., and Müller, K. M. (2024). Sporadic diurnal fluctuations of cyanobacterial populations in oligotrophic temperate systems can prevent accurate characterization of change and risk in aquatic systems. *Water Res.* 252:121199. doi: 10.1016/j.watres.2024.121199
- Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2022). GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 38, 5315–5316. doi: 10.1093/bioinformatics/btac672
- Chen, H., Li, Y., Sun, W., Song, L., Zuo, R., and Teng, Y. (2020). Characterization and source identification of antibiotic resistance genes in the sediments of an interconnected river-lake system. *Environ. Int.* 137:105538. doi: 10.1016/j.envint.2020.105538
- Chopyk, J., Nasko, D. J., Allard, S., Bui, A., Treangen, T., Pop, M., et al. (2020). Comparative metagenomic analysis of microbial taxonomic and functional variations in untreated surface and reclaimed waters used in irrigation applications. *Water Res.* 169:115250. doi: 10.1016/j.watres.2019.115250
- Commichaux, S., Javkar, K., Ramachandran, P., Nagarajan, N., Bertrand, D., Chen, Y., et al. (2021). Evaluating the accuracy of *Listeria monocytogenes* assemblies from quasimetagenomic samples using long and short reads. *BMC Genomics* 22:389. doi: 10.1186/s12864-021-07702-2
- Crump, B. C., Kling, G. W., Bahr, M., and Hobbie, J. E. (2003). Bacterioplankton community shifts in an Arctic lake correlate with seasonal changes in organic matter source. *Appl. Environ. Microbiol.* 69, 2253–2268. doi: 10.1128/AEM.69.4.2253-2268.2003
- Dadi, T., Wendt-Potthoff, K., and Koschorreck, M. (2017). Sediment resuspension effects on dissolved organic carbon fluxes and microbial metabolic potentials in reservoirs. *Aquat. Sci.* 79, 749–764. doi: 10.1007/s00027-017-0533-4
- Eramo, A., Morales Medina, W. R., and Fahrenfeld, N. L. (2020). Factors associated with elevated levels of antibiotic resistance genes in sewer sediments and wastewater. *Environ. Sci. (Camb.)* 6, 1697–1710. doi: 10.1039/d0ew00230e
- Escalas, A., Hale, L., Voordeckers, J. W., Yang, Y., Firestone, M. K., Alvarez-Cohen, L., et al. (2019). Microbial functional diversity: from concepts to applications. *Ecol. Evol.* 9, 12000–12016. doi: 10.1002/ece3.5670
- Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., et al. (2021). AMRFinderPlus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* 11:12728. doi: 10.1038/s41598-021-91456-0
- Fenchel, T., and Finlay, B. (2008). Oxygen and the spatial structure of microbial communities. *Biol. Rev.* 83, 553–569. doi: 10.1111/j.1469-185X.2008.00054.x
- Franklin, A. M., Weller, D. L., Durso, L. M., Bagley, M., Davis, B. C., Frye, J. G., et al. (2024). A one health approach for monitoring antimicrobial resistance: developing a national freshwater pilot effort. *Front. Water* 6:10.3389/frwa.2024.1359109. doi: 10.3389/frwa.2024.1359109
- Franzosa, E. A., McIver, L. J., Rahnnavard, G., Thompson, L. R., Schirmer, M., Weingart, G., et al. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* 15, 962–968. doi: 10.1038/s41598-018-0176-y
- Grzesiuk, M., Pietrzak, B., Wacker, A., and Pijanowska, J. (2022). Photosynthetic activity in both algae and cyanobacteria changes in response to cues of predation. *Front. Plant Sci.* 13:907174. doi: 10.3389/fpls.2022.907174
- Guruge, K. S., Tamamura, Y. A., Goswami, P., Tanoue, R., Jinadasa, K. B. S. N., Nomiya, K., et al. (2021). The association between antimicrobials and the antimicrobial-resistant phenotypes and resistance genes of *Escherichia coli* isolated from hospital wastewaters and adjacent surface waters in Sri Lanka. *Chemosphere* 279:130591. doi: 10.1016/j.chemosphere.2021.130591
- Hassard, E., Gwyther, C. L., Farkas, K., Andrews, A., Jones, V., Cox, B., et al. (2016). Abundance and distribution of enteric bacteria and viruses in coastal and estuarine sediments—a review. *Front. Microbiol.* 7:1692. doi: 10.3389/fmicb.2016.01692
- He, C., Fuchil, D., Sosa, A., Wang, H., Kan, J., Liu, J., et al. (2024). Deep metagenomic sequencing unveils novel SAR202 lineages and their vertical adaptation in the ocean. *Commun. Biol.* 7:853. doi: 10.1038/s42003-024-06535-5
- Hengy, M. H., Horton, D. J., Uzarski, D. G., and Learman, D. R. (2017). Microbial community diversity patterns are related to physical and chemical differences among temperate lakes near Beaver Island, MI. *PeerJ* 5:e3937. doi: 10.7717/peerj.3937

### SUPPLEMENTARY FIGURE S1

Map of sampling locations in Stocker et al. (2024). Water and sediment samples that were collected at locations 4, 5, and 11 (orange arrows) were processed for metagenomics analysis in the present study.

### SUPPLEMENTARY FIGURE S2

Compositional dissimilarity or  $\beta$ -diversity of the relative abundances of MetaCyc pathways within water metagenomes. Colors represent sampling location/depth and shapes represent time of day.

### SUPPLEMENTARY TABLE S1

Relative abundances of microbial species with an average abundance >0.1% across all samples at the water surface, in the water column, and in sediments of the agricultural pond. LME model statistics for differences between values in microbiomes at the water surface vs. in the water column and the latter vs. in sediments are displayed. Bold font indicates significance ( $q < 0.05$ ).

### SUPPLEMENTARY TABLE S2

Relative abundances of MetaCyc pathways encoded in the metagenomes at the water surface, in the water column, and in sediments of the agricultural pond. LME model statistics for differences between values in microbiomes at the water surface vs. in the water column and the latter vs. in sediments are displayed. Bold font indicates significance ( $q < 0.05$ ).

### SUPPLEMENTARY TABLE S3

MAG amino acid sequences with homology (>50% identity) to AMR, virulence, and stress response genetic elements.

- Hess, S., Berendonk, T. U., and Kneis, D. (2018). Antibiotic resistant bacteria and resistance genes in the bottom sediment of a small stream and the potential impact of remobilization. *FEMS Microbiol. Ecol.* 94:fy128. doi: 10.1093/femsec/fy128
- Huang, X., Toro, M., Reyes-Jara, A., Moreno-Switt, A. I., Adell, A. D., Oliveira, C. J., et al. (2024). Integrative genome-centric metagenomics for surface water surveillance: elucidating microbiomes, antimicrobial resistance, and their associations. *Water Res.* 264:122208. doi: 10.1016/j.watres.2024.122208
- Jiang, C., Zhao, Z., Grossart, H. P., Ju, F., Zhao, Y., Gadd, G. M., et al. (2024). Health risk ranking of antibiotic resistance genes in the Yangtze River. *Environ. Sci. Ecotechnol.* 21:100388. doi: 10.1016/j.ese.2024.100388
- Kaminski, J., Gibson, M. K., Franzosa, E. A., Segata, N., Dantas, G., and Huttenhower, C. (2015). High-specificity targeted functional profiling in microbial communities with ShortBRED. *PLoS Comput. Biol.* 11:e1004557. doi: 10.1371/journal.pcbi.1004557
- Kearns, P. J., Angell, J. H., Howard, E. M., Deegan, L. A., Stanley, R. H. R., and Bowen, J. L. (2016). Nutrient enrichment induces dormancy and decreases diversity of active bacteria in salt marsh sediments. *Nat. Commun.* 7:12881. doi: 10.1038/ncomms12881
- Kent, A. D., Jones, S. E., Lauster, G. H., Graham, J. M., Newton, R. J., and McMahon, K. D. (2006). Experimental manipulations of microbial food web interactions in a humic lake: shifting biological drivers of bacterial community structure. *Environ. Microbiol.* 8, 1448–1459. doi: 10.1111/j.1462-2920.2006.01039.x
- Kocurek, B., Ramachandran, P., Grim, C. J., Morin, P., Howard, L., Ottesen, A., et al. (2023). Application of quasimetagenomics methods to define microbial diversity and subtype *Listeria monocytogenes* in dairy and seafood production facilities. *Microbiol. Spectr.* 11:e0148223. doi: 10.1128/spectrum.01482-23
- Leoni, B., Patelli, M., Soler, V., and Nava, V. (2018). Ammonium transformation in 14 lakes along a trophic gradient. *Water (Switzerland)* 10:265. doi: 10.3390/w10030265
- Liguori, K., Keenum, I., Davis, B. C., Calarco, J., Milligan, E., Harwood, V. J., et al. (2022). Antimicrobial resistance monitoring of water environments: a framework for standardized methods and quality control. *Environ. Sci. Technol.* 56, 9149–9160. doi: 10.1021/acs.est.1c08918
- Linz, D. M., Sienkiewicz, N., Struewing, I., Stelzer, E. A., Graham, J. L., and Lu, J. (2023). Metagenomic mapping of cyanobacteria and potential cyanotoxin producing taxa in large rivers of the United States. *Sci. Rep.* 13:2806. doi: 10.1038/s41598-023-29037-6
- Liu, B., Zheng, D., Zhou, S., Chen, L., and Yang, J. (2022). VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res.* 50, D912–D917. doi: 10.1093/nar/gkab1107
- Lu, J., Breitwieser, F. P., Thielen, P., and Salzberg, S. L. (2017). Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* 3:e104. doi: 10.7717/peerj-cs.104
- Malayil, L., Ramachandran, P., Chattopadhyay, S., Allard, S. M., Bui, A., Butron, J., et al. (2022). Variations in bacterial communities and antibiotic resistance genes across diverse recycled and surface water irrigation sources in the mid-Atlantic and Southwest United States: a CONSERVE two-year field study. *Environ. Sci. Technol.* 56, 15019–15033. doi: 10.1021/acs.est.2c02281
- Mayr, M. J., Besemer, K., Sieczko, A., Demeter, K., and Peduzzi, P. (2020). Bacterial community composition and function along spatiotemporal connectivity gradients in the Danube floodplain (Vienna, Austria). *Aquat. Sci.* 82:28. doi: 10.1007/s00027-020-0700-x
- Mckinley, K. R. (1977). Light-mediated uptake of H-glucose in a small hard-water. *Ecology* 58, 1356–1365. doi: 10.2307/1935087
- Merino, N., Wasserman, N. L., Coutelot, F., Kaplan, D. I., Powell, B. A., Jiao, Y., et al. (2023). Microbial community dynamics and cycling of plutonium and iron in a seasonally stratified and radiologically contaminated pond. *Sci. Rep.* 13:19697. doi: 10.1038/s41598-023-45182-4
- Mittal, P., Prasoodanan Pk, V., Dhakan, D. B., Kumar, S., and Sharma, V. K. (2019). Metagenome of a polluted river reveals a reservoir of metabolic and antibiotic resistance genes. *Environ. Microbiome* 14:5. doi: 10.1186/s40793-019-0345-3
- Neneng, L., Nugroho, R. A., Komai, Y., Takayama, N., and Kawamura, K. (2020). Water quality measurements with a simple molecular analysis (PCR-RFLP) of the microbiome in a Metropolitan River system in Japan. Available online at: <http://wjst.wu.ac.th>.
- Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., and Bertilsson, S. (2011). A guide to the natural history of freshwater Lake Bacteria. *Microbiol. Mol. Biol. Rev.* 75, 14–49. doi: 10.1128/mmb.00028-10
- Olm, M. R., Brown, C. T., Brooks, B., and Banfield, J. F. (2017). DRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* 11, 2864–2868. doi: 10.1038/ismej.2017.126
- Orakov, A., Fullam, A., Coelho, L. P., Khedkar, S., Szklarczyk, D., Mende, D. R., et al. (2021). GUNC: detection of chimerism and contamination in prokaryotic genomes. *Genome Biol.* 22:178. doi: 10.1186/s13059-021-02393-0
- Paradis, E., and Schliep, K. (2019). Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. doi: 10.1093/bioinformatics/bty633
- Pattanayak, G. K., Liao, Y., Wallace, E. W. J., Budnik, B., Drummond, D. A., and Rust, M. J. (2020). Daily cycles of reversible protein condensation in Cyanobacteria. *Cell Rep.* 32:108032. doi: 10.1016/j.celrep.2020.108032
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490
- Rout, A. K., Dixit, S., Tripathy, P. S., Rout, S. S., Parida, S. N., Parida, P. K., et al. (2024a). Metagenomic landscape of sediments of river Ganga reveals microbial diversity, potential plastic and xenobiotic degradation enzymes. *J. Hazard. Mater.* 471:134377. doi: 10.1016/j.jhazmat.2024.134377
- Rout, A. K., Tripathy, P. S., Dixit, S., Behera, D. U., Behera, B., Das, B. K., et al. (2024b). Metagenomics analysis of sediments of river Ganga, India for bacterial diversity, functional genomics, antibiotic resistant genes and virulence factors. *Curr. Res. Biotechnol.* 7:100187. doi: 10.1016/j.crbiot.2024.100187
- Saheb Kashaf, S., Proctor, D. M., Deming, C., Saary, P., Hölzer, M., Mullikin, J., et al. (2022). Integrating cultivation and metagenomics for a multi-kingdom view of skin microbiome diversity and functions. *Nat. Microbiol.* 7, 169–179. doi: 10.1038/s41564-021-01011-w
- Selim, K. A., Haffner, M., Burkhardt, M., Mantovani, O., Neumann, N., Albrecht, R., et al. (2021). Diurnal metabolic control in cyanobacteria requires perception of second messenger signaling molecule c-di-AMP by the carbon control protein SbtB. *Sci. Adv.* 7:eabk0568. doi: 10.1126/sciadv.abk0568
- Sharma, P., Pal, N., Kumawat, M., Singh, S., Das, D., Tilwari, A., et al. (2025). Investigating the antibiotic resistance genes and mobile genetic elements in water systems impacted with anthropogenic pollutants. *Environ. Res.* 269:120814. doi: 10.1016/j.envres.2025.120814
- Smith, D. J., Berry, M. A., Cory, R. M., Johengen, T. H., Kling, G. W., Davis, T. W., et al. (2022). Heterotrophic Bacteria dominate catalase expression during *Microcystis* blooms. *Appl. Environ. Microbiol.* 88:e0254421. doi: 10.1128/aem.02544-21
- Stal, L. J., and Moezelaar, R. (2006). Fermentation in cyanobacteria. *FEMS Microbiol. Rev.* 21, 179–211. doi: 10.1111/j.1574-6976.1997.tb00350.x
- Stocker, M. D., Smith, J. E., Pachepsky, Y. A., and Blaustein, R. A. (2024). Fine-scale spatiotemporal variations in bacterial community diversity in agricultural pond water. *Sci. Total Environ.* 915:170143. doi: 10.1016/j.scitotenv.2024.170143
- Su, S., Li, C., Yang, J., Xu, Q., Qiu, Z., Xue, B., et al. (2020). Distribution of antibiotic resistance genes in three different natural water bodies: a lake, river and sea. *Int. J. Environ. Res. Public Health* 17:552. doi: 10.3390/ijerph17020552
- Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B., and Wu, C. H. (2015). UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* 31, 926–932. doi: 10.1093/bioinformatics/btu739
- U.S. Food and Drug Administration (2024). FSMA Final Rule on Pre-Harvest Agricultural Water. Available online at: <https://www.fda.gov/food/food-safety-modernization-act-fsma/fsma-final-rule-pre-harvest-agricultural-water> (Accessed November 5, 2024).
- Uritskiy, G. V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP - a flexible pipeline for genome-resolved metagenomic data analysis 08 information and computing sciences 0803 computer software 08 information and computing sciences 0806 information systems. *Microbiome* 6:158. doi: 10.1186/s40168-018-0541-1
- Wang, W., Yi, Y., Yang, Y., Zhou, Y., Jia, W., Zhang, S., et al. (2019). Response mechanisms of sediment microbial communities in different habitat types in a shallow lake. *Ecosphere* 10:e02948. doi: 10.1002/ecs2.2948
- Wang, B., Zheng, X., Zhang, H., Xiao, F., He, Z., and Yan, Q. (2020). Keystone taxa of water microbiome respond to environmental quality and predict water contamination. *Environ. Res.* 187:109666. doi: 10.1016/j.envres.2020.109666
- Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with kraken 2. *Genome Biol.* 20:257. doi: 10.1186/s13059-019-1891-0
- Zhang, L., Zhao, T., Shen, T., and Gao, G. (2019). Seasonal and spatial variation in the sediment bacterial community and diversity of Lake Bosten, China. *J. Basic Microbiol.* 59, 224–233. doi: 10.1002/jobm.201800452
- Zhao, Y., Wang, C., Wang, X., Wang, W., Zhang, T., He, J., et al. (2024). Insights into the plankton community seasonal variations in a finer scale of the Bohai Sea: biodiversity, trophic linkage, and biotic-abiotic interplay. *Front. Mar. Sci.* 11:1498869. doi: 10.3389/fmars.2024.1498869
- Zhao, H., Yan, B., Mo, X., Li, P., Li, B., Li, Q., et al. (2019). Prevalence and proliferation of antibiotic resistance genes in the subtropical mangrove wetland ecosystem of South China Sea. *Microbiology* 8:e871. doi: 10.1002/mbo3.871