Upwelling periodically disturbs the ecological assembly of microbial communities in the Laurentian Great Lakes

Augustus Pendleton[1]*, Mathew Wells[2], & Marian L. Schmidt[1]*

[1]Department of Microbiology, Cornell University, 123 Wing Dr, Ithaca, NY 14850, USA

[2]Department of Physical and Environmental Sciences, University of Toronto Scarborough, 1065 Military Trail, Toronto, Ontario M1C 1A4, Canada

Corresponding Authors: Augustus Pendleton: arp277@cornell.edu; Marian L. Schmidt: marschmi@cornell.edu

Author Contribution Statement: AP contributed to conceptualization, methodology, software, formal analysis, investigation, visualization, data curation, and writing. MW contributed to methodology, data curation, and writing. MLS contributed to the conceptualization, methodology, investigation, visualization, resources, data curation, writing, supervision, project administration, and funding acquisition.

Preprint servers: A version of this article was previously submitted to *bioRxiv* (doi: 10.1101/2025.01.17.633667) under a CC-BY-NC-ND 4.0 International license.

Classification: Major: Biological Sciences Minor: Ecology

Keywords: Microbial Communities - Hydrodynamics - Great Lakes - Metacommunity - Biodiversity - Upwelling

Abstract

The Laurentian Great Lakes hold 21% of the world's surface freshwater and supply drinking water to nearly 40 million people. Here we provide the first evidence that wind-driven upwelling fundamentally restructures microbial communities in Lake Ontario, with its effects extended and redistributed by an internal Kelvin wave propagating along the shoreline. While thermal stratification is known to organize microbial communities by depth and season, we show that this vertical structure arises from contrasting mechanisms: homogenizing selection in surface waters and dispersal limitation with drift in the hypolimnion. Kelvin wave-driven upwelling disrupts this scaffold, displacing rare taxa into the surface and creating novel coastal communities enriched in methane oxidation and sulfur metabolism genes—functional traits absent elsewhere in the lake. We observed a Kelvin wave lasting over two weeks and propagating eastward at ~60 km day ^{1}. Given the ~10–12 day recurrence of wind events (1), at any time, at least one segment of Lake Ontario's coastline is typically experiencing upwelling—producing pulses frequent and sustained enough to remodel microbial communities on ecologically relevant timescales. Recurrent upwellings, sustained and redistributed by Kelvin waves, act as a biological disturbance that overrides stratification, mobilizes rare functional potential, and assembles novel coastal microbial communities. As climate change lengthens stratified periods and reshapes large-lake circulation, understanding how physical forcing governs microbial assembly is essential for forecasting the biogeochemical future of Earth's great lakes—especially in shoreline zones where ecological shifts directly affect human communities.

Significance Statement

The Laurentian Great Lakes hold 21% of the world's surface freshwater and nearly 85% of North America's—supplying drinking water to over 40 million people. Yet the microbial life that underpins their ecological function remains poorly understood. We show that wind-driven upwelling, followed by internal Kelvin waves that redistribute and sustain upwelling, acts as a recurring physical disturbance that overrides thermal stratification, redistributing rare microbes and assembling novel shoreline communities. These shifts bring unexpected functional changes, including enrichment of methane oxidation and sulfur metabolism genes, where people most interact with lake water. As climate change reshapes large lake circulation and intensifies thermal layering, understanding how microbial ecosystems respond is essential for forecasting transformations in water quality, ecosystem function, and biogeochemical resilience.

Data Availability Statement

All raw and processed data for this project are publicly available. The main GitHub repository for this project is available at https://github.com/MarschmiLab/Pendleton_2025_Ontario_Publication_Repo, which includes all processed data (Amplicon Sequence Variants (ASVs), CTD casts, water quality data, all other metadata, etc.), the code for all figures, tables, and summary statistics, including the generation of ASVs using the DADA2 workflow, and a reproducible *renv* environment. All the software versions and citations used for data processing and statistics are described in Table S2. The raw compressed 16S rRNA gene sequencing fastq files are available in the NCBI Sequence Read Archive under the BioProject accession numbers PRJNA1212049. All flow cytometry data are available in the FlowRepository database with the ID number FR-FCM-Z8SJ (2).

Introduction

Microbial communities fuel aquatic ecosystems—transforming energy, cycling nutrients, purifying water, and anchoring aquatic food webs (3–6). In lakes, microbes recycle organic matter and support higher trophic levels via the microbial loop (7). Yet in Earth's largest lakes, we still lack a predictive understanding of how microbial communities are structured—or how they respond to the powerful physical forces of stratification and circulation that define these systems and are shifting under climate change (8–11).

Classic ecological theory holds that community assembly reflects a balance between deterministic forces like environmental selection and stochastic processes such as drift and dispersal limitation (5, 12). In hydrodynamically active systems—like Earth's largest lakes and the global oceans—these ecological processes unfold within a dynamic physical environment shaped by ecosystem-scale drivers like stratification, currents, and wind-driven upwelling (8, 13, 14). Large lakes—upon which hundreds of millions depend—thus offer a tractable but underused model for understanding how physics rewires microbial biogeography and biogeochemical function, with insights that scale to the oceans, where mesoscale circulation shapes the global cycling of carbon, nitrogen, and sulfur (15–17).

Community assembly mechanisms—including selection, dispersal, diversification, and drift—interact dynamically in response to environmental heterogeneity (5, 6, 18–20). Selection may homogenize communities under stable conditions or drive divergence across gradients, while stochasticity introduces probabilistic variation in community structure (12). These processes shift across spatial and temporal scales, and may respond strongly to physical mixing or isolation. Although microbial communities in smaller lakes and along watershed gradients have been widely studied (21–24), studies of large lakes have more often focused on surface waters or treated them as spatially homogeneous systems (25, 26). As a result, the influence of vertical structure and hydrodynamic circulation on microbial community assembly in large lakes remains underexplored (27–30).

One powerful—but largely overlooked—driver of microbial dynamics in large lakes is the wind–upwelling–Kelvin wave cascade. Wind initiates coastal upwelling through

Ekman transport, tilting the thermocline. When the wind relaxes, this displaced density interface releases its stored energy as baroclinic internal waves. In offshore waters, these manifest as near-inertial oscillations (31); along the coast, they generate internal Kelvin waves—coastally trapped gravity waves that roll anticlockwise (in the Northern Hemisphere) within a narrow 3–5 km shoreline band (32–34). Propagating with the wave is a baroclinic coastal jet: a shore-parallel current, strongest near the surface, that redistributes upwelling and downwelling water along the lake's perimeter—a disruption of the thermocline that can last for weeks. This process repeats in Lake Ontario every ~10–12 days (1) and every 5–10 days in Lake Erie (34). While physical limnologists have long studied these features for their role in transporting nutrients, pollutants, and thermal discharges (13, 32, 35), their ecological consequences remain largely unexamined (36).

Here, we leverage Lake Ontario—a 19,000 km^2 freshwater basin with well-characterized seasonal circulation—to test how physical and ecological forcings shape microbial community structure and function. The lake stratifies annually and experiences wind-driven upwelling events that transport cold, nutrient-rich water to surface zones (8, 37–39), creating transient and chemically distinct habitats. Using an interdisciplinary framework that integrates 16S rRNA gene sequencing, flow cytometry, environmental profiling across 47 stations and two seasons, and physical limnology, we reveal that Kelvin wave–driven upwelling fundamentally reshapes microbial biogeography, mobilizes rare functional potential, and challenges existing models of ecological assembly. We show that microbial communities stratify by depth due to homogenizing selection—not dispersal limitation—but, strikingly, when upwelling disrupts this scaffold, it assembles novel shoreline communities enriched in rare taxa and distinctive functional traits. Together, these findings reveal how physical forcing structures microbial ecosystems and challenge assumptions of spatial homogeneity in large lakes.

Results

Upwelling events are frequent drivers of spatial heterogeneity and persist on microbially relevant timescales

We detected upwelling events of cold water along the northern shore in May (stations 29 & 43) and the southern shore in September (stations 35 and 38) (Fig. 1B, Fig. S1). In May, the warm southern shore likely reflected a combination of downwelling and relatively warmer water flowing in from the Niagara River (Fig. S2A-C). In September, strong easterly winds in the days leading up to sampling likely drove the pronounced southern upwelling observed at station 38, as reflected in both temperature (Fig. 1B, Fig. S2D) and nutrient data (Fig. S3). While prevailing winds in Lake Ontario are typically westerly—driving upwelling near Toronto—these easterlies reversed the usual pattern, generating upwelling near Rochester instead. We estimate this upwelling formed near the Niagara outlet on September 23rd and propagated eastward as a Kelvin wave along the southern shore, concluding at the eastern terminus by October 9th (>2 weeks later), at which point winds shifted back from the west-its prevailing direction-causing upwellings along the northern shore (Fig. S2D-E).

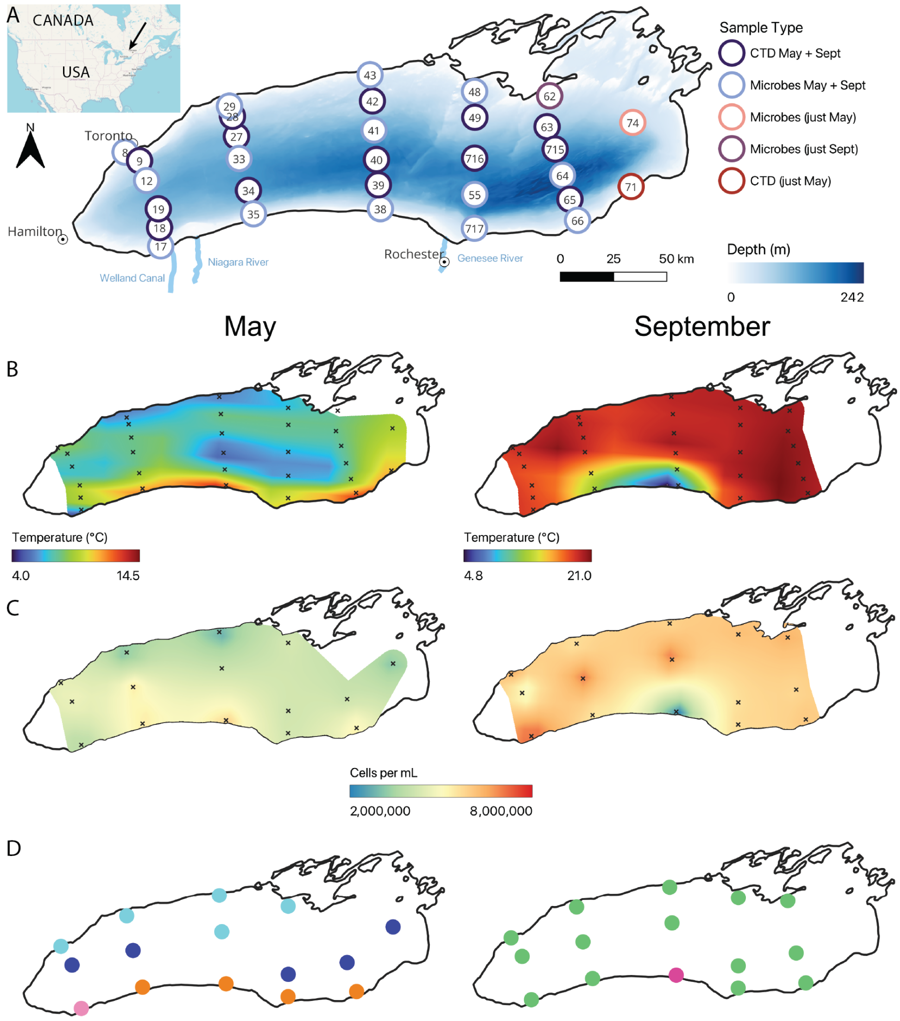

Figure 1. Upwelling reshapes surface microbial abundance and composition. (A) Microbial samples were collected at fifteen stations (light blue) in both months, though station 74 was only sampled in September, and station 62 only in May. All stations with microbial data had associated chemical data collected. Additional stations with only chemical data (dark blue) were collected. Station 71 was only collected in May, and CTD data was unreliable for station 65 in September. (B) Surface (5 m) temperatures in Lake Ontario interpolated spatially across the lake in May and September using multilevel B-splines. (C) Surface cell counts interpolated across May and September using multilevel B-splines. (D) Surface community clusters (shown by color) formed via UPGMA clustering on the community dissimilarity (Fig. S4). Trees were cut at the same height for both months (h = 0.32), resulting in four groups in May and two groups in September.

Upwelling reshapes surface microbial abundance and composition and enriches rare taxa

Spatial patterns of cell abundance varied across the lake surface in both May and September and correlated to surface temperatures (Fig. 1C). In May, cold, upwelling areas along the northern shore had lower cell counts while warmer offshore areas and the southern shore downwelling and Niagara-river impacted regions exhibited higher cell counts. In September, the upwelling zone at station 38 had the lowest cells abundances, whereas station 17, influenced by Welland Canal water, had the highest cell abundances.

Surface communities differentiated spatially in response to patterns in surface temperature cause by upwelling (Fig. 1D). Using UPGMA clustering based on community dissimilarity and cutting the dendrogram at the same height (Fig. S4), we defined 4 clusters in May and two clusters in September, consistent with higher dispersion in May (Fig. S5C, PERMDISP, $p = 0.002$). In May, these clusters corresponded to the northern upwelling, an offshore pelagic group, the southern downwelling, and an isolated station near the Welland canal. In contrast, surface communities across September were comparatively homogeneous, besides in the upwelling zone (Fig. 1D, Fig. S5C). The high dissimilarity was driven by an influx of hypolimnion–associated taxa (Fig. S6A), likely due to mass effects, and many rare taxa not observed elsewhere (Fig. S6B, S7A-D), which collectively increased the potential for unique microbial interactions (Fig. S6C). Rare taxa were also abundant in May near the Welland Canal (Fig. S7B) but not along the northern shore upwelling (Fig. S7B).

Environmental conditions shaped by stratification, not geographic distance, determine community composition

To address what ecological processes structured the spatial patterns shown in Fig. 1, we explored the relative importance of distance- and depth-decay in Lake Ontario's microbial communities. Community similarity (1 - weighted UniFrac dissimilarity) was not correlated with geographic distance but exhibited a strong relationship with depth in both months, particularly when the thermocline intensified in September (Fig. 2A-B). Variance partitioning revealed that depth, month, and other environmental factors together explained 73% of the variation in microbial community composition (weighted UniFrac; Fig. 2C), a signal echoed in the PCA of environmental gradients (Fig. S5B). Geographic location, even when combined with environmental variables, accounted for only 3% of the variance.

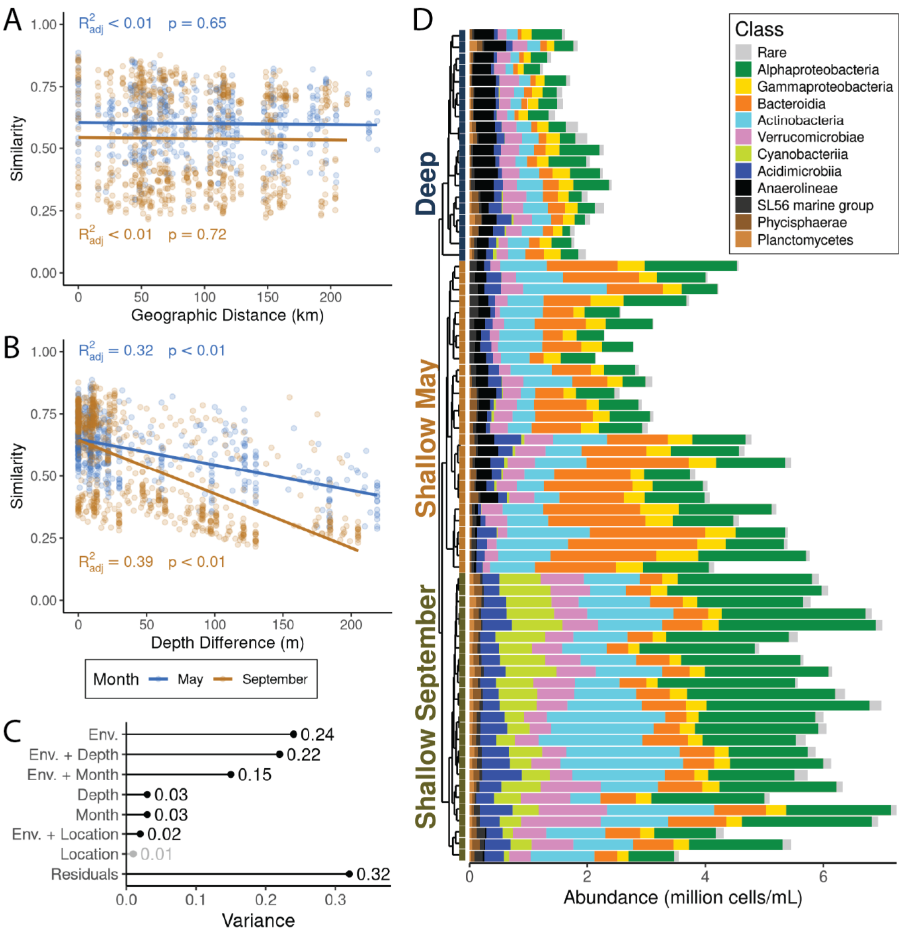

Figure 2. Thermal stratification—not geographic distance—drives microbial community structure in Lake Ontario. Microbial community (A) distance- and (B) depth-decay relationships by month using similarity defined as 1 - community dissimilarity. (C) Variance partitioning of the weighted UniFrac dissimilarity. Environmental variables (Env.) corresponded to scaled physical and chemical data from each sample. Variables which were significant (all $p < 0.001$) when tested using an ANOVA-like permutation test for Constrained Correspondence Analysis post-dbRDA are in black whereas those that were not (i.e., location) are in light gray. (D) Community composition of all samples in absolute abundance, colored at the Class-level. Samples are arranged via hierarchical clustering as shown in Fig. S5A.

As a result of stratification, Lake Ontario bacterial communities were clustered into three distinct groups (Fig. S5A-D). These groups closely hew to sample depth (typically >30 m for deep samples, depending on thermocline depth) and month of collection, and are thus referred to as "depth-month groups." This pattern was consistent across both hierarchical clustering (Fig. S5A) and principal coordinates analysis (Fig. S5C). Deep samples from both months grouped together, while shallow samples from May and September formed distinct clusters ($p < 0.001$, $R^2 = 0.55$, PERMANOVA). Therefore, we categorized bacterial communities into three groups for analysis: Shallow May, Shallow September, and Deep.

Microbial cell abundances across these groups were strongly correlated with temperature (Fig. S8A; Spearman's R =0.92, $p < 2.2e\text{-}16$). Cell abundances were highest in Shallow September, followed by Shallow May, and lowest in Deep samples (Fig. S8B). Unexpectedly, deep September samples had fewer cells than deep May samples ($p = 0.031$, Two-Sample Wilcoxon).

There were a core set of cosmopolitan taxa shared across all three groups, including the Alphaproteobacteria LD12, Actinobacteria acI lineages, Gammaproteobacteria LD28 and PnecB, and Verrucomicrobiae LD19 (Fig. 2D, Fig. S9B). Differentially abundant taxa included an enrichment of Chloroflexi in Deep samples, Bacteroidia in shallow May, and Cyanobacteria in shallow September (Fig. S9C)—including potentially harmful taxa (Fig. S10). A more thorough discussion of abundant microbial groups can be found in the Supporting Information.

Season and depth modulate the strength of ecological selection across the lake

We deepened our analysis with community assembly processes across stratification using iCAMP to quantify selection, dispersal, and drift (Fig. 3). Drift dominated Deep samples (52%), with some influence from dispersal limitation (13%) (Fig. 3A). Dispersal limitation was more pronounced when comparing Deep to Shallow samples (12%), increasing to 23% in September during stronger stratification. In Shallow samples, homogenizing selection was greater, especially in September (67%) (Fig. 2A). Across all depth-month groups, homogenizing selection explained 36%–55% of turnover, indicating consistent selection for a core microbial community throughout the lake (Fig. 2B).

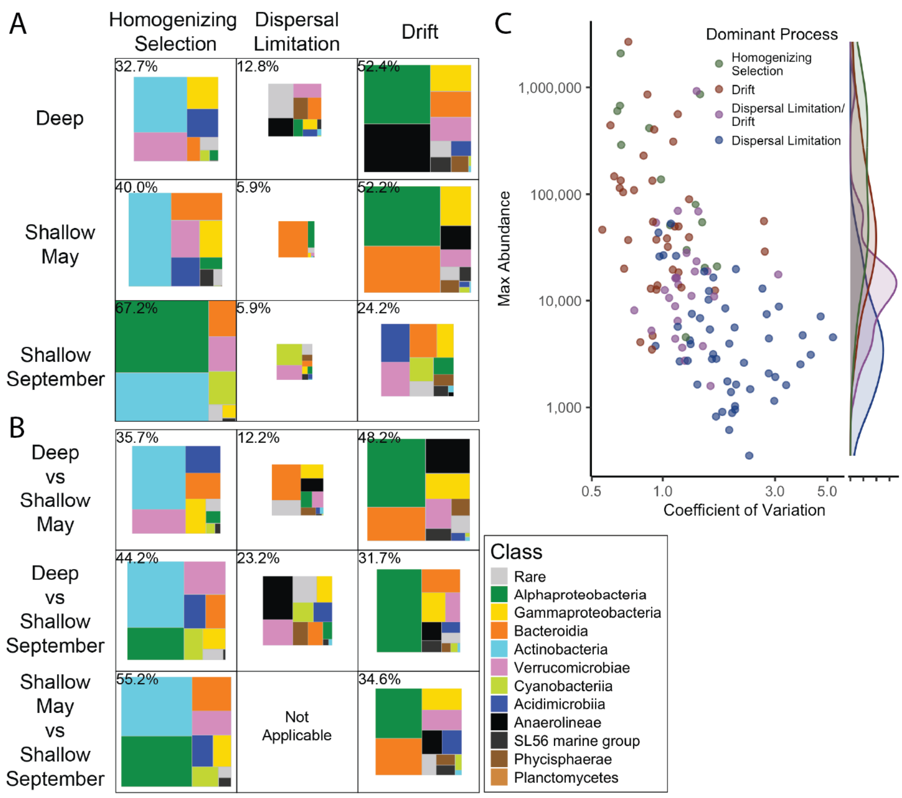

Figure 3. Contrasting forces of selection and drift structure microbial communities across depth. (A-B) Relative contribution of different taxonomic groups to microbial community assembly processes, as determined by iCAMP. The area of each box represents the relative contribution of that process (columns) to turnovers between samples (A) *within* a given depth-month group, or (B) *between* two depth-month groups. The percentage contribution of a given process to community turnover is in the top left corner, and boxes are scaled relative to the most influential process (*i.e.* homogenizing selection in shallow September at 62.7%), and the area represented by each Class is scaled relative to that Class's contribution to that process. (C) The relative importance of community assembly processes within each phylogenetic bin, calculated by iCAMP. Absolute ASV counts were summed per bin per sample while the coefficient of variation was calculated as the standard deviation of abundances across samples by the mean abundance of that bin across all samples. The density plot on the right shows the distribution of each process across the maximum phylogenetic bin abundances. Three bins dominated by heterogeneous selection or a combination of drift and homogenizing selection were excluded.

Distinct microbial taxa were responsible for the relative importance of each assembly process (Fig 3A-B). Homogenizing selection was primarily driven by the Actinobacteria, while drift was dominated by Alphaproteobacteria. Dispersal limitation between deep and shallow groups was primarily influenced by Bacteroidia in May and Anaerolineae in September.

The contribution of taxonomic groups to community assembly was closely tied to their abundance within samples, particularly for within-group comparisons (compare Fig. 2C and Fig. 3A). More abundant bins had lower abundance variation across samples

(Spearman's ranked correlation, ρ = -0.61, p < 0.0001) and were primarily shaped by drift and homogenizing selection (Fig. 3C). In contrast, low-abundance bins with high variance were dominated by dispersal limitation, while intermediate-abundance bins were influenced by both drift and dispersal limitation (Fig. 3C).

Upwellings are an opportunity for novel microbial metabolic potential

Our analyses revealed that stratification structures microbial communities by modulating the relative influence of selection and drift. But what happens to microbial biogeochemical cycling when upwelling disrupts this physical structure? We used FAPROTAX to infer the abundance of genes associated with microbial metabolic functions (Fig. 4). Upwelling zones exhibited a functional signature that blended traits from both the epilimnion (e.g., photoautotrophy; Fig. 4A) and the hypolimnion (e.g., ammonia oxidation; Fig. 4B). However, consistent with the presence of novel microbial taxa and interactions in the September upwelling zone (Fig. S6), we also detected functions that were unique to upwelling, including sulfate respiration (Fig. 4C) and methanotrophy (Fig. 4D). These results suggest that upwelling does not just remix existing microbial functional potential—it creates entirely distinct biogeochemical niches that may alter elemental cycling in the Great Lakes.

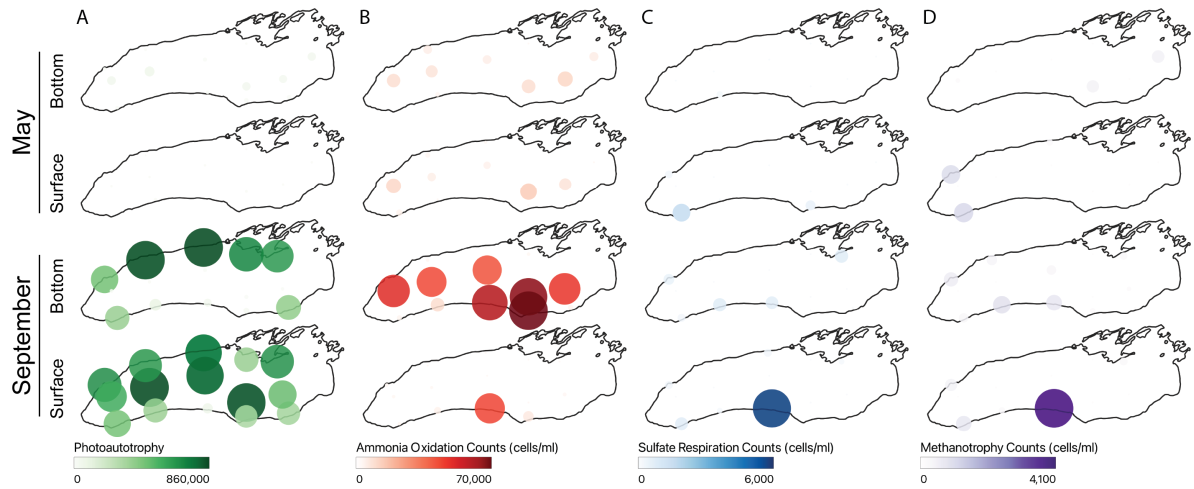

Figure 4. Upwelling creates distinct microbial functional potential. Functional traits were inferred using FAPROTAX (40) using the absolute abundance of each ASV as input, and results are shown for (A) photoautotrophy, (B) aerobic ammonia oxidation, (C) sulfate respiration and (D) methanotrophy. The size and color of each point is scaled by the inferred abundance of each functional trait, in units of predicted cells/ml. Note the range of values varies between functional traits.

Discussion

Physical forcing and ecological processes interact to govern microbial community structure

Microbial community assembly in Lake Ontario is fundamentally shaped by physical forces that structure the water column. Stratification imposed vertical gradients, with depth and season modulating the balance between deterministic and stochastic processes. Yet, when stratification was disrupted—through upwelling, downwelling, or tributary mixing—community composition was rapidly restructured. These transient disturbances reveal how responsive microbial communities are to ephemeral but ecologically significant niches in large lakes.

Upwelling creates transient but consequential microbial niches

Lake Ontario's ocean-like circulation drives microbial turnover at spatial and temporal scales rarely quantified in freshwater systems. In spring, thermal bars trap nutrients and sculpt sharp near-shore gradients (41, 42). In summer, typically westerly wind events force Ekman transport, tilt the thermocline, and inject cold, nutrient-rich hypolimnetic water into the surface (32, 34, 37, 43, 44). As winds relax, energy is released as internal waves: near-inertial oscillations offshore and coastally trapped Kelvin waves (31, 33). Propagating anticlockwise at ~60 km day ^{1} (33), the Kelvin wave–jet complex drags the upwelling signal laterally and, as we show, sustains an easterly microbial disturbance that persisted over two weeks (Fig. S2D–E)—a timescale relevant for microbial growth, succession, and metabolism.

Although physical limnologists have characterized these waves for decades, our data provide the first evidence that they restructure microbial communities, mobilizing rare taxa and their interaction networks and, in turn, reshaping functional potential. Given that upwelling-generating winds occur every ~10–12 days in Lake Ontario (1) and can persist for weeks, some segment of the shoreline is likely experiencing upwelling—and its associated microbial disturbance—at any given point during the stratified season.

Upwelling-favorable winds have already increased by ~45% over the past three decades (39) and are projected to intensify further (45), underscoring the urgency of linking physical forcing to microbial ecological response.

Despite their frequency, the microbial consequences of upwelling remain undercharacterised, especially in freshwater systems (46–48). During the September event we tracked hypolimnetic lineages (e.g., Anaerolineae, Nitrospiria) surfacing at station 38 (Fig. S6A), which we estimate had coupled microbial and thermal signatures persisting for more than two weeks (Fig. S2D–E). Upwelling sites were uniquely enriched in sulfate-respiration and methanotrophic genes (Fig. 4C–D), revealing that upwelling does more than mix water: it creates novel biogeochemical niches capable of rewiring microbial ecosystem function. As wind regimes intensify under climate change, resolving these physical–microbial linkages will be essential for predicting the function and resilience of Earth's largest lakes.

Rare taxa contribute disproportionately to community novelty in upwelling zones

Upwelling and tributary-influenced zones hosted microbial communities distinct from surface or bottom waters, enriched in rare taxa and unique microbial interactions (Fig. S6). In September, low-abundance hypolimnetic microbes dominated at the southern upwelling site. In May, similar enrichment occurred near the Welland Canal but not in the upwelling zone—possibly due to colder temperatures at that station and consequent slower microbial growth rates (49, 50).

These rare taxa are often overlooked, but may serve as a latent reservoir of ecological innovation—a microbial seed bank whose functional potential emerges under transient, physically driven conditions (51). The distribution of these rare taxa is most likely to be structured by dispersal limitation, so their presence may be both spatially and environmentally constrained. While the long-term consequences remain unquantified, our findings point to a cryptic but consequential layer of microbial responsiveness—one that highlights upwellings as a driver of microbial dynamism with the potential to reshape ecosystem function in the world's largest lakes.

Tributary inputs did not generate persistent nearshore–offshore gradients, in contrast to patterns in Lake Michigan (52, 53). Instead, strong offshore selection filtered rare taxa, reinforcing a cosmopolitan microbial core structured by drift and selection (Fig. 3). However, episodic enrichment near tributary inputs combined with downwelling and thermal bar formation (Fig. 1A, Fig. 4) may temporarily boost microbial diversity and function in nearshore zones. These findings suggest that rare taxa, though ephemeral, may play outsized roles in microbial ecosystem response, revealing an under-appreciated layer of adaptability in Great Lakes microbiomes.

Depth and season drive a shift from selection to drift and dispersal limitation

Environmental gradients from stratification and hydrodynamics—not geographic distance—best explain microbial community structure, consistent with metacommunity theory and findings that dispersal is rarely limiting in aquatic systems (6). Surface communities, especially in late summer, were shaped by homogenizing selection, yielding consistent assemblages across sites. In contrast, hypolimnion communities were more stable and governed by drift and dispersal limitation. In May, depth-decay was weaker (Fig. 2B), and surface communities were more spatially variable (Fig. 1D) and similar to those at depth, likely reflecting ongoing niche differentiation after spring mixing. By September, stratification and low turbulence reinforced surface selection. These dynamics mirror trends in the Pacific Ocean, where surface microbes are selected and dispersal limitation intensifies with depth (54). Similar patterns occur in oligotrophic lakes, where nutrient scarcity favors efficient, streamlined taxa (21, 23, 55, 56), while productive systems show more stochasticity (57–59).

In the hypolimnion, stratification constrained dispersal (Figs. 2B & 3A–B), and limited mixing fostered site-specific communities that were consistent across space and time (Fig. S5C), despite environmental stability. These cold, oxygenated deep waters harbor slow-growing microbes, with low cell densities (Fig. S8B) (60), limited gene flow, and dormancy-all weakening selection (61, 62)—explaining the dominance of drift in hypolimnetic bins (Fig. 3C).

While cold, oligotrophic hypolimnia are often thought to be selection-dominated [*see Supplemental Discussion: Selection in the Hypolimion*; (63); (29); (64); (65)], our results align with recent ocean studies showing that drift increases with depth due to sparse biomass and isolation (54). These effects were evident in Lake Ontario's hypolimnion, where distance-decay was weak (Fig. S11E–F) but dispersal limitation was consistent. Rather than forming a continuous biogeographic gradient (25), deep microbial communities appeared as isolated parcels shaped by intermittent mixing—a spatial mosaic of stochastic processes.

Microbial abundance and taxonomy link niche breadth to assembly processes

Abundant taxa were governed by selection and drift (Fig. 3C), consistent with the idea that dominant microbes experience stronger ecological filtering (21, 27). Rare taxa were more dispersal limited, reflecting restricted distributions, though it is unclear if this is a result of narrow niche breadth or a transient introduction from an external source (66). Synthesizing multi-year datasets could help discern which rare taxa are transiently rare versus permanently rare.

Taxonomic identity also influenced assembly. Actinobacteria, a group characterized by diverse metabolic capabilities and small genomes (67), were primarily associated with homogenizing selection. In contrast, Alphaproteobacteria—dominated by the ubiquitous LD12 clade—were governed by drift. This contrast may reflect differences in ecological versatility. Widespread Actinobacteria may exploit more diverse niches via adaptive radiation (67). In contrast, the distribution Alphaproteobacteria like LD12 were structured by both selection and drift (Fig. 3). This highlights the importance of drift in structuring microbial distributions, especially for highly abundant taxa (68).

These patterns underscore that community assembly cannot be explained by abundance alone. Phylogenetic identity, ecological function, and life history traits all shape how taxa respond to environmental gradients and physical forcing. Yet, our interpretation is limited by the temporal resolution of our sampling, which captured only two stratified timepoints. Key transitions—such as thermocline formation, fall mixing, and inverse stratification-remain unresolved (9, 69). Finer habitat delineations unexplored here (such as shorelines, benthic surfaces, and suspended particles) are also likely to influence assembly processes (53, 70, 71). Finally, while iCAMP offers valuable insights, its inferences reflect model assumptions and should be applied with appropriate caution (*see Supplemental Discussion: Biases in iCAMP*).

Upwelling-driven microbial restructuring alters functional potential

Hydrodynamic restructuring of microbial communities carries clear biogeochemical consequences. Upwelling, tributary mixing, and downwelling redistribute microbial taxa and their metabolic potential, creating biogeochemical hot spots that are both spatially and temporally dynamic (37, 72). Our results suggest that upwelling introduces rare, functionally distinct taxa into the surface waters—including lineages associated with sulfur cycling, methane oxidation, and ammonia oxidation—potentially shifting biogeochemical processing. These changes unfolded over weeks, as Kelvin wave–driven upwelling swept laterally along the nearshore, suggesting that microbial metabolism responds rapidly to these common physical forcings.

Functional inference based on metabarcoding, however, has inherent limitations. The 16S rRNA V4 region lacks the resolution to capture fine-scale genomic or functional differences, and taxa with identical ASVs may differ in metabolic capacity (47, 73). Additionally, variable gene copy numbers and high dormancy rates in freshwater microbes can decouple gene presence from metabolic activity (74, 75). As a result, the accuracy of functional inference depends on the trait of interest, the environment, and the size and curation of reference databases (40, 47, 76, 77). For instance, nitrification rates in Lake Erie have been shown to be uncoupled from nitrifier abundance (78, 79).

To stay conservative, we used FAPROTAX, a relatively stringent approach; only 24% of ASVs matched at least one annotated function, likely underestimating the true functional potential of these communities. Future work using metatranscriptomics could more directly quantify metabolic responses to upwelling (74, 80, 81). Ideally, these inferences would be validated by experimental incubations that measure microbial biogeochemical rates in upwelled waters (82–84).

Microbial biogeography emerges from physical-ecological coupling

Microbial communities in Lake Ontario are not static, but dynamically assembled by the interaction of physical forcing and ecological processes. As northern lakes warm, deepening and lengthening stratification will isolate hypolimnetic communities (85–87) while enhancing surface selection, reshaping microbial biogeochemical cycling in ways we are only beginning to quantify. At the same time, upwelling-favorable winds are becoming more frequent and are predicted to intensify, increasing the prevalence and impact of microbial disturbance zones alongshore.

Our data show that such disturbances are not rare—recurring 2-3 times a month-with their influence persisting for weeks. These pulses override stratification and assemble novel shoreline microbial communities, remodeling ecosystems on meaningful ecological timescales.

By pairing high-resolution microbial, ecological, and physical limnology data, we demonstrate that community structure and function in large lakes are governed not just by who arrives or survives, but by how water moves. These systems offer natural laboratories for testing metacommunity theory, bridging freshwater and marine paradigms, and forecasting how microbial life will shape and respond to a changing climate.

Methods

Sample collection

Samples were collected from multiple stations throughout Lake Ontario aboard the *R/V Lake Guardian* in May and September of 2023 as part of the US Environmental Protection Agency's (EPA) Lower Food Web Cooperative Science and Monitoring Initiative (CSMI) survey (Fig. 1A; Table S1). The stations were arranged in 5 north-south transects across the lake, which were sampled in an east-west fashion over May 15th-19th and September 23rd-27th. Water samples were collected using a rosette while depth profiles were taken with a CTD meter equipped with fluorometer, dissolved oxygen, spherical photosynthetic active radiation sensor, and turbidity sensors. "Surface" samples were from 5 m below the surface. "Mid" samples were collected at stations >30 m deep, either at the fluorescence maximum or thermocline in May, or at the station's average depth in September. "Bottom" samples were taken 2 m above the sediment. Samples were pre-filtered consecutively through sterilized 200 μm and 20 μm Nitex mesh (Wildco) into sterile 10L bottles (Nalgene). Microbial samples were filtered onto a polyethersulfone 0.22 μm filter (MilliporeSigma). Filters were flash frozen in liquid N_2 and stored at -80C.

Environmental data

Water chemistry and chlorophyll a data were generated by the U.S. EPA's Great Lakes National Program Office according to their standard protocols (88). See the Supporting Information for details on environmental data, including temperature monitors and meteorological records.

Quantifying cells with flow cytometry

Flow cytometry samples were prepared in the field by adding 1 μL (in May) or 5 μL (in September) of 25% glutaraldehyde to 1 mL of sample water with a 10 minute incubation. Fixed samples were stained with SYBR I green dye (1X concentration) and cell abundance was measured in DNA-positive cells with a BD Accuri C6 Flow Cytometer in triplicate (Fig. S12). Flow data was analyzed using the R package flowCore and ggcyto (89, 90). Additional details are provided in the supporting information.

DNA extraction & Illumina sequencing

All DNA extractions were carried out using the Qiagen DNeasy PowerWater kit, mostly following the manufactuer recommendations (more details in Supporting Information). Half a filter was used for May samples and a quarter filter was used for September samples, translating to a minimum of 540mL extracted, with a median of 1320mL extracted. An extraction negative was produced for each kit using a blank filter.

The V4 region of the 16S rRNA gene was amplified using the Earth Microbiome Project protocol (91–93). Polymerase chain reaction (PCR) was performed, with blanks, in triplicate 25 μL reactions using KAPA HiFi 2X MasterMix (Roche), 0.2 μM 515F (5'-GTGYCAGCMGCCGCGGTAA) and 806R (5'-GGACTACNVGGGTWTCTAAT) primers with Illumina Nextera adapter overhangs, and 5 ng of sample DNA. A ZymoBIOMICS Microbial Community DNA Standard (Zymo) was included to assess the sequencing error rate. Samples and indexing blanks were sequenced using a 2 x 250 bp paired-end strategy on an Illumina MiSeq at the Cornell Biotechnology Resources Center, generating 15.72 million reads and a median sequencing depth of 105,500 reads (excluding blanks).

Microbial bioinformatic processing

Raw Illumina sequences were processed into amplicon sequence variants (ASVs) using the standard workflow in dada2 package in R (94). The phyloseq package was used to organize the ASV count table, metadata, and later taxonomic table and phylogeny (95). Taxonomy was first assigned using the TaxAss freshwater database (96) and then with the Silva 16S rRNA database (97). Mitochondrial, chloroplast, and ASVs with a higher relative abundance in blanks were removed. In the mock community, three spurious ASVs were detected, representing a sequencing error rate of 0.03%. The resulting minimum number of reads in a sample was 7,931 reads, with a maximum of 73,652 reads, and a total of 7,280 unique ASVs.

A phylogenetic tree was constructed using MAFFT for alignment and FastTree under a generalized time-reversible model (98, 99). The tree was rooted at the most-recent-common-ancestor (MRCA) of the Archaea using the ape package (100). Several polyphyletic, anomalous tips had exceptionally long branch lengths and no taxonomic assignment and were removed by filtering for tips with node heights greater than 2 as calculated using the `phytools::nodeheight` function (101).

Ecological Statistics

Between-sample (beta) diversity was calculated using functions from the `vegan` or `GUniFrac` packages (modified to take into account absolute abundances) (102, 103). Unless otherwise noted, community dissimilarity refers to the Generalized weighted Unifrac distance ($\alpha = 0.5$) calculated using absolute abundances. The relative abundance of each ASV in a given sample was normalized by the cell concentration of that sample, providing the absolute abundance of each ASV. Ordinations were constructed using principal coordinates analysis implemented in phyloseq (104). Replicates were highly similar to each other (Fig. S13) and were merged by summing reads, resulting in a minimum of 13,986 reads and a median of 82,765 reads. Depth-month groups were defined with UPGMA clustering based on community dissimilarity, cutting the tree at 3 groups (Fig. S5A). Variance partitioning and post-hoc tests were performed with Vegan functions `varpart`, `dbrda`, and `anova.cca` (102). Permutational multivariate analysis of variance (PERMANOVA) and multivariate homogeneity of groups dispersions tests were run with `adonis2` and `betadisper` from the `vegan` package (102).

Within-sample (alpha) diversity was estimated using the iNEXT package (105, 106). All rarefaction curves were asymptotic, confirming sufficient sequencing depth for representative sample richness (Fig. S14). Differential abundance of microbial ASVs between depth-month groups was calculated at the Class level using the pairwise setting in the ANCOM-BC2 package (107). Significant differences had an FDR-corrected p-value less than 0.05 and passed the sensitivity test across multiple pseudocount values; other parameters were kept at default settings. Microbial community assembly processes were quantified with iCAMP (108). Full details and explanantion of iCAMP are available in the supplemental. We used a maximum bin size of 24 ASVs, a maximum bin distance of 0.4, and a confidence threshold of 0.975 compared to a randomized null distribution with 1000 iterations (108). Microbial functional traits were inferred using FAPROTAX (40), using the absolute abundance of each ASV.

Statistical testing, including Two-Sample Wilcoxon tests following Kruskal-Wallis tests;

correlation tests; and linear models were done in R, using base functions and the ggpubr package (109). If multiple comparisons occurred within a given plot panel, all p-values were corrected with a Bonferroni Correction.

Spatial Analysis

Mapping and spatial interpolation was performed either in R using the packages sf, terra, tmap, tmaptools and gstat, or in QGIS using GRASS (110, 110, 111, 111, 112, 112–116). Data files for Lake Ontario's outline and bathymetry were downloaded from United States Geological Society and National Oceanic and Atmospheric Administration, respectively (117, 118).

Acknowledgments and funding sources

We thank Eric Osantowski, Matthew Pawlowski, Anne Scofield, Ben Alsip, and the science staff in the Great Lakes National Program Office of the US EPA and the captain and crew of the *R/V Lake Guardian* for facilitating sample collection. We thank Evie Brahmstedt, Sophia Aredas, and Sophia Richter for support in field work. We thank Linda Cote and the BRC Genomics Facility (RRID:SCR_021727) at the Cornell Institute of Biotechnology for sequencing experiments. We are grateful to Andrea Giometto and his lab for flow cytometry support. We appreciate the feedback from Joe Atkinson and also the EPA & NOAA Great Lakes Modeling Group led by James Pauer. Funding for this work was made possible via a grant to MLS from the Affinito-Stewart & President's Council of Cornell Women, Cornell University start-up funds to MLS, and a NOAA Margaret A Davidson Fellowship to AP (NA24NOSX420C0016). Thank you to XX, XX, and anonymous reviewers for their valuable feedback.

References

1. Y. R. Rao, C. R. Murthy, Nearshore currents and turbulent exchange processes during upwelling and downwelling events in Lake Ontario. *Journal of Geophysical Research: Oceans* **106**, 2667–2678 (2001).

532 2. J. Spidlen, K. Breuer, C. Rosenberg, N. Kotecha, R. R. Brinkman, FlowRepository: a resource of annotated flow cytometry datasets associated with peer-reviewed publications. *Cytometry A* **81**, 727–731 (2012).

533 3. J. J. Cole, Aquatic Microbiology for Ecosystem Scientists: New and Recycled Paradigms in Ecological Microbiology. *Ecosystems* **2**, 215–225 (1999).

534 4. P. G. Falkowski, T. Fenchel, E. F. Delong, The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science* **320**, 1034–1039 (2008).

535 5. C. A. Hanson, J. A. Fuhrman, M. C. Horner-Devine, J. B. H. Martiny, Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* **10**, 497–506 (2012).

536 6. J. B. H. Martiny, *et al.*, Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**, 102–112 (2006).

537 7. F. Azam, B. C. Cho, D. C. Smith, M. Simon, "Bacterial Cycling of Matter in the Pelagic Zone of Aquatic Ecosystems" in *Large Lakes: Ecological Structure and Function*, M. M. Tilzer, C. Serruya, Eds. (Springer, 1990), pp. 477–488.

538 8. B. Hlevca, T. Howell, M. Madani, N. Benoit, Fine-Scale Hydrodynamic Modeling of Lake Ontario: Has Climate Change Affected Circulation Patterns? *Environ Model Assess* (2024). https://doi.org/10.1007/s10666-024-10003-z.

539 9. T. Ozersky, *et al.*, The Changing Face of Winter: Lessons and Questions From the Laurentian Great Lakes. *Journal of Geophysical Research: Biogeosciences* **126**, e2021JG006247 (2021).

540 10. C. G. Fichot, K. Matsumoto, B. Holt, M. M. Gierach, K. S. Tokos, Assessing change in the overturning behavior of the Laurentian Great Lakes using remotely sensed lake surface water temperatures. *Remote Sensing of Environment* **235**, 111427 (2019).

541 11. E. J. Anderson, B. Tillotson, C. A. Stow, Indications of a changing winter through the lens of lake mixing in Earth's largest freshwater system. *Environ. Res. Lett.* **19**, 124060 (2024).

542 12. M. Vellend, Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology* **85**, 183–206 (2010).

543 13. G. T. Csanady, Intermittent "full" upwelling in Lake Ontario. *Journal of Geophysical Research (1896-1977)* **82**, 397–419 (1977).

544 14. S. Rahmstorf, Ocean circulation and climate during the past 120,000 years. *Nature* **419**, 207–214 (2002).

545 15. S. Sunagawa, *et al.*, Structure and function of the global ocean microbiome. *Science* **348**, 1261359 (2015).

546 16. O. Ulloa, D. E. Canfield, E. F. DeLong, R. M. Letelier, F. J. Stewart, Microbial oceanography of anoxic oxygen minimum zones. *Proceedings of the National Academy of Sciences* **109**, 15996–16003 (2012).

547 17. J. Rigonato, *et al.*, Ocean-wide comparisons of mesopelagic planktonic community structures. *ISME Communications* **3**, 83 (2023).

548 18. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton University Press, 2011).

549 19. M. A. Leibold, *et al.*, The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* **7**, 601–613 (2004).

550 20. G. G. Mittelbach, B. J. McGill, *Community Ecology*, Second Edition, Second Edition (Oxford University Press, 2019).

551 21. H. Li, *et al.*, Contrasting patterns of diversity of abundant and rare bacterioplankton in freshwater lakes along an elevation gradient. *Limnology and Oceanography* **62**, 1570–1585 (2017).

552 22. W. Chen, *et al.*, Stochastic processes shape microeukaryotic community assembly in a subtropical river across wet and dry seasons. *Microbiome* **7**, 138 (2019).

553 23. Z. Gu, *et al.*, Community assembly processes underlying the temporal dynamics of glacial stream and lake bacterial communities. *Science of The Total Environment* **761**, 143178 (2021).

554 24. A. Isabwe, *et al.*, Riverine bacterioplankton and phytoplankton assembly along an environmental gradient induced by urbanization. *Limnology and Oceanography* **67**, 1943–1958 (2022).

555 25. S. E. Jones, T. A. Cadkin, R. J. Newton, K. D. McMahon, Spatial and temporal scales of aquatic bacterial beta diversity. *Front. Microbiol.* **3** (2012).

556 26. A. Shade, S. E. Jones, K. D. McMahon, The influence of habitat heterogeneity on freshwater bacterial community composition and dynamics. *Environmental Microbiology* **10**, 1057–1067 (2008).

557 27. J. P. Niño-García, C. Ruiz-González, P. A. del Giorgio, Landscape-scale spatial abundance distributions discriminate core from random components of boreal lake bacterioplankton. *Ecol Lett* **19**, 1506–1515 (2016).

558 28. A. C. Yannarell, E. W. Triplett, Within- and between-Lake Variability in the Composition of Bacterioplankton Communities: Investigations Using Multiple Spatial Scales. *Applied and Environmental Microbiology* **70**, 214–223 (2004).

559 29. J. P. Niño-García, C. Ruiz-González, P. A. del Giorgio, Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. *ISME J* **10**, 1755–1766 (2016).

560 30. S. F. Paver, R. J. Newton, M. L. Coleman, Microbial communities of the Laurentian Great Lakes reflect connectivity and local biogeochemistry. *Environmental Microbiology* **22**, 433–446 (2020).

561 31. J. Antenucci, "Currents in stratified water bodies 3: effects of rotation" in *Encyclopedia of inland waters*, G. F. Likens, Ed. (Elsevier, 2009), pp. 559–567.

562 32. G. Csanady, Hydrodynamics of large lakes. *Annual Review of Fluid Mechanics* **7**, 357–386 (1975).

563 33. B. Wang, Kelvin waves. *Encyclopedia of atmospheric sciences* **1062** (2002).

564 34. R. Valipour, Y. R. Rao, L. F. León, D. Depew, Nearshore-offshore exchanges in multi-basin coastal waters: Observations and three-dimensional modeling in Lake Erie. *Journal of Great Lakes Research* **45**, 50–60 (2019).

565 35. D. J. Schwab, Internal free oscillations in lake ontario 1. *Limnology and oceanography* **22**, 700–708 (1977).

566 36. Y. R. Rao, T. Howell, S. B. Watson, S. Abernethy, On hypoxia and fish kills along the north shore of Lake Erie. *Journal of Great Lakes Research* **40**, 187–191 (2014).

567 37. G. D. Haffner, M. L. Yallop, P. D. N. Hebert, M. Griffiths, Ecological Significance of Upwelling Events in Lake Ontario. *Journal of Great Lakes Research* **10**, 28–37 (1984).

568 38. Y. Hui, D. Schlea, J. Atkinson, Z. Zhu, T. Redder, Model development in support of the Lake Ontario Cooperative Science and Monitoring Initiative. *Aquatic Ecosystem Health & Management* **25**, 81–96 (2022).

569 39. A. Jabbari, R. Valipour, J. D. Ackerman, Y. R. Rao, Nearshore-offshore exchanges by enhanced turbulent mixing along the north shore of Lake Ontario. *Journal of Great Lakes Research* **49**, 596–607 (2023).

570 40. S. Louca, L. W. Parfrey, M. Doebeli, Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**, 1272–1277 (2016).

571 41. J. C. Makarewicz, T. W. Lewis, G. L. Boyer, Nutrient enrichment and depletion on the shoreside of the spring thermal front. *Journal of Great Lakes Research* **38**, 72–77 (2012).

572 42. X. Bai, *et al.*, Modeling 1993–2008 climatology of seasonal general circulation and thermal structure in the Great Lakes using FVCOM. *Ocean Modelling* **65**, 40–63 (2013).

573 43. D. Beletsky, W. P. O'Connor, D. J. Schwab, D. E. Dietrich, Numerical Simulation of Internal Kelvin Waves and Coastal Upwelling Fronts. (1997).

574 44. M. D. Rowe, *et al.*, Coastal Upwelling Influences Hypoxia Spatial Patterns and Nearshore Dynamics in Lake Erie. *Journal of Geophysical Research: Oceans* **124**, 6154–6175 (2019).

575 45. B. Hlevca, E. T. Howell, R. Valipour, M. Madani, Modeling nearshore-offshore water exchange in Lake Ontario. *PLOS ONE* **20**, e0298702 (2025).

576 46. L. J. Kerkhof, M. A. Voytek, R. M. Sherrell, D. Millie, O. Schofield, Variability in bacterial community structure during upwelling in the coastal ocean. *Hydrobiologia* **401**, 139–148 (1999).

577 47. F. Sun, M. Wu, Y. Wang, C. Sun, Z. Xu, Diversity and potential function of bacterial communities in different upwelling systems. *Estuarine, Coastal and Shelf Science* **237**, 106698 (2020).

578 48. A. De Wever, *et al.*, Bacterial Community Composition in Lake Tanganyika: Vertical and Horizontal Heterogeneity. *Applied and Environmental Microbiology* **71**, 5029–5037 (2005).

579 49. S. W. Wilhelm, *et al.*, Seasonal changes in microbial community structure and activity imply winter production is linked to summer hypoxia in a large lake. *FEMS Microbiology Ecology* **87**, 475–485 (2014).

50. C. R. Lovell, A. Konopka, Thymidine Incorporation by Free-Living and Particle-Bound Bacteria in a Eutrophic Dimictic Lake. *Applied and Environmental Microbiology* **49**, 501–504 (1985).

51. S. M. Neuenschwander, J. Pernthaler, T. Posch, M. M. Salcher, Seasonal growth potential of rare lake water bacteria suggest their disproportional contribution to carbon fluxes. *Environmental Microbiology* **17**, 781–795 (2015).

52. S. R. Mueller-Spitz, G. W. Goetz, S. L. McLellan, Temporal and spatial variability in nearshore bacterioplankton communities of Lake Michigan. *FEMS Microbiology Ecology* **67**, 511–522 (2009).

53. M. Fujimoto, *et al.*, Spatiotemporal distribution of bacterioplankton functional groups along a freshwater estuary to pelagic gradient in Lake Michigan. *Journal of Great Lakes Research* **42**, 1036–1048 (2016).

54. F. Milke, I. Wagner-Doebler, G. Wienhausen, M. Simon, Selection, drift and community interactions shape microbial biogeographic patterns in the Pacific Ocean. *ISME J* 1–13 (2022). https://doi.org/10.1038/s41396-022-01318-4.

55. J. Liao, *et al.*, Similar community assembly mechanisms underlie similar biogeography of rare and abundant bacteria in lakes on Yungui Plateau, China. *Limnology and Oceanography* **62**, 723–735 (2017).

56. P. Aguilar, R. Sommaruga, The balance between deterministic and stochastic processes in structuring lake bacterioplankton community over time. *Molecular Ecology* **29**, 3117–3130 (2020).

587 57. A. Roguet, *et al.*, Neutral community model explains the bacterial community assembly in freshwater lakes. *FEMS Microbiology Ecology* **91**, fiv125 (2015).

588 58. A. C. Yannarell, A. D. Kent, G. H. Lauster, T. K. Kratz, E. W. Triplett, Temporal Patterns in Bacterial Communities in Three Temperate Lakes of Different Trophic Status. *Microb Ecol* **46**, 391–405 (2003).

589 59. S. Drakare, A. Liess, Local factors control the community composition of cyanobacteria in lakes while heterotrophic bacteria follow a neutral model. *Freshwater Biology* **55**, 2447–2457 (2010).

590 60. S. Evans, J. B. H. Martiny, S. D. Allison, Effects of dispersal and selection on stochastic assembly in microbial communities. *The ISME Journal* **11**, 176–185 (2017).

591 61. J. T. Lennon, S. E. Jones, Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* **9**, 119–130 (2011).

592 62. J.-H. Lin, B.-P. Han, M. C. Urban, L. D. Meester, The role of dormant propagule banks in shaping eco-evolutionary dynamics of community assembly. (2024). https://doi.org/10.1101/2024.11.03.621703.

593 63. J. Wang, *et al.*, Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *The ISME Journal* **7**, 1310–1321 (2013).

594 64. P. Huber, *et al.*, Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. *ISME J* **14**, 2951–2966 (2020).

595 65. E. A. Mueller, J. T. Lennon, Residence time structures microbial communities through niche partitioning. Available at: https://www.biorxiv.org/content/10.1101/2024.08.22.609267v1.

596 66. X. Jia, F. Dini-Andreote, J. F. Salles, Community Assembly Processes of the Microbial Rare Biosphere. *Trends in Microbiology* **26**, 738–747 (2018).

597 67. S. M. Neuenschwander, R. Ghai, J. Pernthaler, M. M. Salcher, Microdiversification in genome-streamlined ubiquitous freshwater Actinobacteria. *The ISME Journal* **12**, 185–198 (2018).

598 68. R. J. Newton, S. E. Jones, A. Eiler, K. D. McMahon, S. Bertilsson, A Guide to the Natural History of Freshwater Lake Bacteria. *Microbiology and Molecular Biology Reviews* **75**, 14–49 (2011).

599 69. G. Pu, *et al.*, The Great Lakes Winter Grab: Limnological data from a multi-institutional winter sampling campaign on the Laurentian Great Lakes. *Limnology and Oceanography Letters* **n/a** (2024).

600 70. H. Yuan, *et al.*, Diversity Distribution, Driving Factors and Assembly Mechanisms of Free-Living and Particle-Associated Bacterial Communities at a Subtropical Marginal Sea. *Microorganisms* **9**, 2445 (2021).

601 71. J. Liu, *et al.*, Comparison of assembly process and co-occurrence pattern between planktonic and benthic microbial communities in the Bohai Sea. *Front. Microbiol.* **13** (2022).

602 72. M. A. Neilson, R. J. J. Stevens, Spatial Heterogeneity of Nutrients and Organic Matter in Lake Ontario. *Can. J. Fish. Aquat. Sci.* **44**, 2192–2203 (1987).

603 73. H. B. Hassler, *et al.*, Phylogenies of the 16S rRNA gene and its hypervariable regions lack concordance with core genome phylogenies. *Microbiome* **10**, 104 (2022).

604 74. J. Sadeghi, A. Hashemi Shahraki, S. R. Chaganti, D. Heath, Functional gene transcription variation in bacterial metatranscriptomes in large freshwater Lake Ecosystems: Implications for ecosystem and human health. *Environmental Research* **231**, 116298 (2023).

605 75. G. M. Douglas, *et al.*, PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* **38**, 685–688 (2020).

606 76. R. J. Wright, M. G. I. Langille, PICRUSt2-SC: an update to the reference database used for functional prediction within PICRUSt2. *Bioinformatics* **41**, btaf269 (2025).

607 77. C. Sansupa, *et al.*, Can We Use Functional Annotation of Prokaryotic Taxa (FAPROTAX) to Assign the Ecological Functions of Soil Bacteria? *Applied Sciences* **11**, 688 (2021).

608 78. C. C. Clevinger, R. T. Heath, D. L. Bade, Oxygen use by nitrification in the hypolimnion and sediments of Lake Erie. *Journal of Great Lakes Research* **40**, 202–207 (2014).

609 79. D. K. Hoffman, *et al.*, Nitrification in the water column of Lake Erie: Seasonal patterns, community dynamics, and competition with cyanobacterial harmful algal blooms.

610 80. A. M. Linz, F. O. Aylward, S. Bertilsson, K. D. McMahon, Time-series metatranscriptomes reveal conserved patterns between phototrophic and heterotrophic microbes in diverse freshwater systems. *Limnology and Oceanography* **65**, S101–S112 (2020).

611 81. J. J. Hamilton, *et al.*, Metabolic Network Analysis and Metatranscriptomics Reveal Auxotrophies and Nutrient Sources of the Cosmopolitan Freshwater Microbial Lineage acI. *mSystems* **2**, 10.1128/msystems.00091–17 (2017).

612 82. K. K. Brandt, F. Vester, A. N. Jensen, K. Ingvorsen, Sulfate Reduction Dynamics and Enumeration of Sulfate-Reducing Bacteria in Hypersaline Sediments of the Great Salt Lake (Utah,USA). *Microb Ecol* **41**, 1–11 (2001).

613 83. P. C. J. Reis, S. D. Thottathil, C. Ruiz-González, Y. T. Prairie, Niche separation within aerobic methanotrophic bacteria across lakes and its link to methane oxidation rates. *Environmental Microbiology* **22**, 738–751 (2020).

614 84. W. Li, *et al.*, Microbiome processing of organic nitrogen input supports growth and cyanotoxin production of Microcystis aeruginosa cultures. *The ISME Journal* **18**, wrae082 (2024).

615 85. R. I. Woolway, *et al.*, Phenological shifts in lake stratification under climate change. *Nat Commun* **12**, 2318 (2021).

616 86. J. T. Lehman, Mixing Patterns and Plankton Biomass of the St. Lawrence Great Lakes under Climate Change Scenarios. *Journal of Great Lakes Research* **28**, 583–596 (2002).

87. R. Schwefel, B. Müller, H. Boisgontier, A. Wüest, Global warming affects nutrient upwelling in deep lakes. *Aquat Sci* **81**, 1–11 (2019).

88. U. E. P. Agency, Sampling and analytical procedures for GLNPO's open lake water quality survey of the Great Lakes. [Preprint] (2003).

89. B. Ellis, *et al.*, "flowCore: flowCore: Basic structures for flow cytometry data" (2024).

90. P. Van, W. Jiang, R. Gottardo, G. Finak, ggcyto: Next-generation open-source visualization software for cytometry. *Bioinformatics* (2018).

91. A. Apprill, S. McNally, R. Parsons, L. Weber, Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology* **75**, 129–137 (2015).

92. J. G. Caporaso, *et al.*, Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences* **108**, 4516–4522 (2011).

93. A. E. Parada, D. M. Needham, J. A. Fuhrman, Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* **18**, 1403–1414 (2016).

94. B. J. Callahan, *et al.*, DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**, 581–583 (2016).

625 95. P. J. McMurdie, S. Holmes, with contributions from G. Jordan, S. Chamberlain, "phyloseq: Handling and analysis of high-throughput microbiome census data" (2022).

626 96. R. R. Rohwer, J. J. Hamilton, R. J. Newton, K. D. McMahon, TaxAss: Leveraging a Custom Freshwater Database Achieves Fine-Scale Taxonomic Resolution. *mSphere* **3**, 10.1128/msphere.00327–18 (2018).

627 97. C. Quast, *et al.*, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**, D590–D596 (2013).

628 98. K. Katoh, D. M. Standley, MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, 772–780 (2013).

629 99. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE* **5**, e9490 (2010).

630 100. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

631 101. L. J. Revell, phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ* **12**, e16505 (2024).

632 102. J. Oksanen, *et al.*, "vegan: Community ecology package" (2022).

633 103. J. Chen, *et al.*, Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* **28**, 2106–2113 (2012).

634 104. P. J. McMurdie, S. Holmes, phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).

635 105. A. Chao, *et al.*, Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs* **84**, 45–67 (2014).

636 106. M. Roswell, J. Dushoff, R. Winfree, A conceptual guide to measuring species diversity. *Oikos* **130**, 321–338 (2021).

637 107. H. Lin, M. Eggesbo, S. D. Peddada, Linear and nonlinear correlation estimators unveil undescribed taxa interactions in microbiome data. *Nature communications* **13**, 1–16 (2022).

638 108. D. Ning, *et al.*, A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat Commun* **11**, 4717 (2020).

639 109. A. Kassambara, "ggpubr: ggplot2 based publication ready plots" (2023).

640 110. E. Pebesma, Simple features for R: Standardized support for spatial vector data. *The R Journal* **10**, 439–446 (2018).

641 111. R. J. Hijmans, "terra: Spatial data analysis" (2024).

642 112. M. Tennekes, tmap: Thematic maps in R. *Journal of Statistical Software* **84**, 1–39 (2018).

643 113. M. Tennekes, "tmaptools: Thematic map tools" (2021).

114. E. J. Pebesma, Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* **30**, 683–691 (2004).

115. GRASS Development Team, *Geographic Resources Analysis Support System (GRASS) Software* (Open Source Geospatial Foundation, 2020).

116. QGIS Development Team, *QGIS Geographic Information System* (2021).

117. USGS, Great Lakes and Watersheds Shapefiles - hydro_p_LakeOntario.

118. National Geophysical Data Center, NOAA, Bathymetry of Lake Ontario. https://doi.org/10.7289/V56H4FBH.