Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multi-method keystone screening reveals distributed ecological resilience in aquaponic microbiomes

Victor Lobanov

victor.lobanov@gu.se

University of Gothenburg

Shima Rezaei

Auburn University

Brendan Higgins

Auburn University

Alyssa Joyce

University of Gothenburg

Research Article

Keywords: aquaponics, keystone species, microbial community, 16S rRNA, PERMANOVA, co-occurrence network, LASSO, distributed resilience

Posted Date: August 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10305887/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Multi-method keystone screening reveals distributed ecological resilience in aquaponic microbiomes

Victor Lobanov[1*], Shima Rezaei[2], Brendan Higgins[2], Alyssa Joyce[1]

[1] Department of Conservation, University of Gothenburg, Gothenburg, 40530, Sweden

[2] Department of Biosystems Engineering, Auburn University, Auburn, AL 36849, USA

[*] Corresponding author: victor.lobanov@gu.se

Abstract

Background: Aquaponic systems couple aquaculture and crop production with microbial communities underpinning both the waste removal needs of the fish and nutrient supply needs of the plants. Keystone taxa, organisms whose removal disproportionately restructures the community, are well-documented in macroecological food webs (1), yet their existence in aquaponic microbiomes and the practical implications for system monitoring remain untested. We applied a multi-method consensus framework for keystone identification to 16S rRNA gene amplicon data from a research-scale aquaponic facility operating both coupled and decoupled hydraulic configurations.

Results: 88 amplicon libraries from three compartments (clarifier, fish tank, grow bed) in four treatment configurations (coupled/decoupled × light/dark conditions in the fish tank) across six sampling dates were processed with DADA2, yielding 825 ASVs after prevalence filtering. Three independent keystone screens were applied: leave-one-out PERMANOVA (67 taxa with FDR < 0.05 and R^2 > 0.02), co-occurrence network hub analysis (four hubs), and LASSO regression on temporal turnover (no individual predictors retained). No taxon was supported by more than one method. *Cetobacterium* sp. (*Fusobacteriota*) ranked highest in the presence-impact screen (R^2 up to 0.11) and was strongly enriched in fish tank samples, consistent with its gut-associated ecology; network hubs were environmental biofilm-associated *Proteobacteria* and *Planctomycetota*.

Conclusions: The absence of cross-method convergence is consistent with a distributed ecological architecture in which community function is spread across habitats and functional guilds rather than concentrated in any single taxon. Aquaponic microbiome monitoring should therefore target compartment-specific indicators and operating conditions rather than universal keystone candidates. *Cetobacterium* nonetheless stood out as a useful fish-tank community marker, but its influence does not extend system-wide. In practice, monitoring compartment-specific assemblages, especially *Cetobacterium* in fish

tanks and *Chitinophagaceae* in biofilm fractions, is more tractable than targeting a single universal keystone organism.

Keywords: aquaponics, keystone species, microbial community, 16S rRNA, PERMANOVA, co-occurrence network, LASSO, distributed resilience

1. Introduction

Aquaponics integrates recirculating aquaculture and hydroponic plant production in a shared water circuit, substantially reducing water consumption and external nutrient discharge relative to conventional agriculture (2,3). The microbial communities within these systems perform nitrification of fish-excreted ammonia, mineralization of organic matter, and facilitation of plant nutrient uptake—the biochemical functions on which system stability depends (4–7). Despite this ecological centrality, identifying which organisms within aquaponic microbiomes are most critical and whether their loss would destabilize community function remains unresolved (5,8,9). Keystone species—taxa whose influence on community structure exceeds what their relative abundance alone would predict—offer one framework for identifying such critical members, and testing whether they occur in aquaponic microbiomes would connect ecological theory directly to system monitoring (3,10,11).

The keystone species concept, originally formulated for macro-ecological food webs (1), has been adapted for microbial systems through statistical (10,12) and model-driven approaches (13–15). Co-occurrence network topology can identify hub taxa with properties consistent with keystones, although such patterns may also reflect shared environmental responses rather than direct ecological interactions (13). Predictive analyses show that small subsets of highly connected taxa—1–5% of community richness—can forecast whole-community compositional change, with greater predictive power in undisturbed systems (14). Crucially, keystone identity can differ by analytical method: top-down influence metrics and network centrality do not necessarily flag the same taxa, making cross-method convergence essential for distinguishing genuine keystones from statistical artifacts (15).

Identifying microbial keystones is further complicated by the multiple mechanisms through which a taxon can appear ecologically important: high relative abundance, central network position, tipping-element dynamics regarding structural robustness, and unique functional capacity each generate keystone-like signals (10,16). Functional redundancy compounds the problem: when several taxa perform equivalent metabolic roles, the function of any one taxon may be wholly or partly compensated by others (17,18). Redundancy is not uniform across functions. For example, nitrogen cycling is generally less redundant than carbon cycling, which may make nitrification-associated taxa disproportionately influential in aquaponic systems (18–20). In contrast to natural aquatic ecosystems such as lake environments, aquaponic systems are inherently compartmentalized: fish tanks, clarifiers, and hydroponic grow beds represent planktonic, attached-biofilm, and rhizosphere-associated habitats, each selecting for distinct microbial assemblages under local

physicochemical conditions (4,21,22). Plant species can exert a stronger selective influence on rhizosphere composition than upstream inoculation from the water column, indicating that community assembly in each compartment is governed largely by local rather than system-wide filters (8,23–25). System design adds a further dimension: coupled configurations, in which water recirculates continuously between fish and plant units, differ in microbial community structure from decoupled designs, in which the hydroponic loop is hydraulically separated (2,22).

Despite the extensive literature on aquaponic microbiomes, no study has yet applied multiple independent keystone detection methods to the same multi-compartment dataset. Here we ask whether convergent candidates emerge when presence-impact analysis, co-occurrence hub structure, and LASSO turnover prediction are applied together, and the extent to which any candidates partition by compartment and hydraulic configuration. To limit circularity, each focal taxon was excluded from the community matrix before Bray-Curtis distances were recomputed and the presence-impact test was applied. Identifying keystone taxa would have direct practical value for aquaponic management: taxa that disproportionately maintain community structure could serve as early-warning biosensors for system dysbiosis or as targets for deliberate microbiome enrichment strategies, potentially also including probiotics (4,11,18,26).

2. Results

2.1 Sequencing and quality control

A total of 88 paired-end amplicon libraries were sequenced from three aquaponic compartments (clarifier, fish tank, and grow bed) across six sampling dates spanning May through November 2024. These libraries were derived from samples collected from twelve aquaponic systems, producing Nile tilapia and cherry tomatoes. Four different aquaponic system configurations were tested with triplicate systems for each (see methods for specific operational details). Primer trimming with cutadapt was critical for accurate chimera detection: an initial analysis conducted without primer removal yielded an 85.7% chimera rate, retaining only approximately 14% of reads; with primer trimming, non-chimeric retention improved substantially. The DADA2 pipeline produced 61,787 unique ASVs across all samples. Prevalence-based filtering (ASVs present in at least 20% of samples) retained 825 ASVs for downstream analysis (Figure 1).

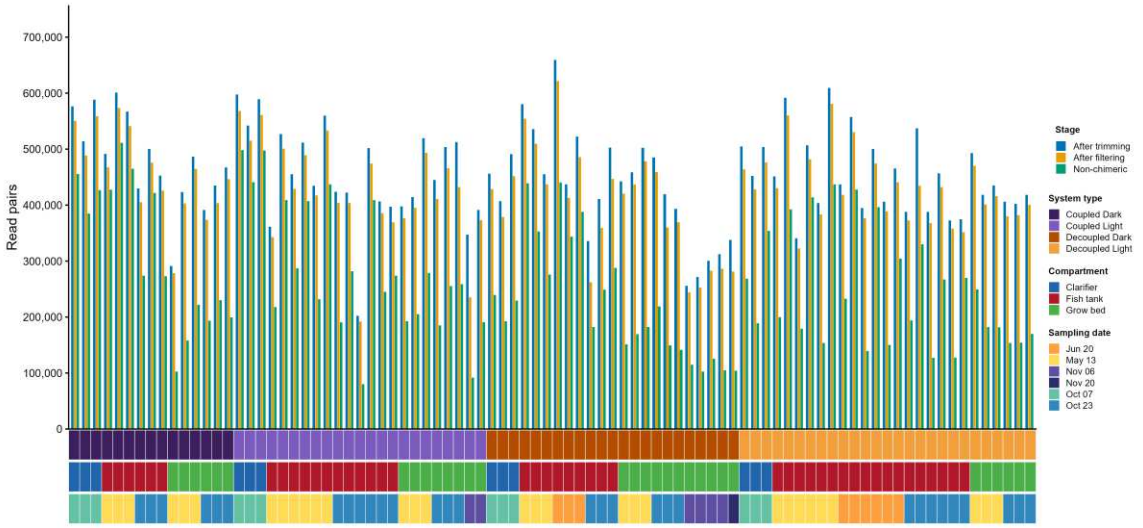


Figure 1. Read retention through the DADA2 pipeline. Per-sample read counts at each processing stage (after primer trimming, after quality filtering, and after chimera removal), demonstrating the importance of primer trimming for accurate chimera detection.

2.2 Microbial community structure across compartments

Alpha diversity, measured as the Shannon index, varied across compartments (Figure 2), and observed ASV richness showed the same compartment ordering (Supplementary Figure S4). Grow bed samples exhibited the highest diversity (Shannon index and observed richness), consistent with the environmental heterogeneity and extensive surface area available for microbial colonization on hydroponic media (pea gravel) and plant root surfaces. Clarifier samples displayed intermediate diversity, while fish tank samples showed the lowest richness, reflecting the comparatively homogeneous planktonic environment. No consistent differences in alpha diversity were observed between coupled and decoupled system configurations. When controlling for sampling date and fish tank illumination, decoupled systems showed significantly higher Shannon diversity (regression coefficient +1.30 Shannon units, $p = 0.002$) and observed ASV richness (+433 ASVs, $p = 0.003$) relative to coupled systems (27), consistent with a broader range of niches in hydraulically independent configurations (9,21,28).

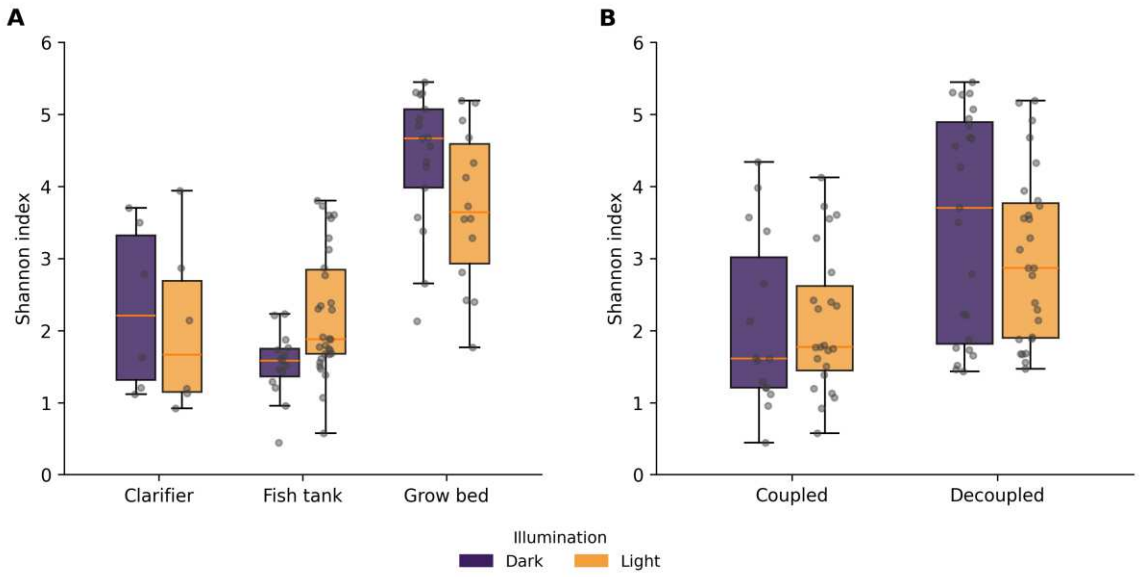


Figure 2. Shannon diversity across compartments and system types. (A) Shannon diversity index by compartment. (B) Shannon diversity by system type (coupled vs. decoupled). Boxes show median and interquartile range; points represent individual samples. Observed ASV richness by compartment is provided as Supplementary Figure S4.

Principal coordinates analysis (PCoA) of Bray-Curtis distances, computed from all 88 samples, revealed clear separation of microbial communities by compartment along the first two ordination axes (Figure 3). Grow bed communities formed a distinct cluster from clarifier and fish tank assemblages, consistent with the fundamentally different ecological niches represented by biofilm-associated (plant root), attached particulate (clarifier), and planktonic (fish tank) habitats. Within each compartment, coupled and decoupled systems showed partially overlapping ordination space, suggesting that hydraulic configuration modulates, but does not fundamentally reorganize community structure. Full PERMANOVA results (compartment $R^2 = 0.256$, coupling R^2, illumination, and season effects) are reported in Supplementary Table S3. Fish tank illumination (light vs. dark) also contributed significantly to community composition (Supplementary Table S3); the illumination effect was most pronounced in the fish tank compartment, consistent with light-driven enrichment of photosynthesis-associated taxa (29).

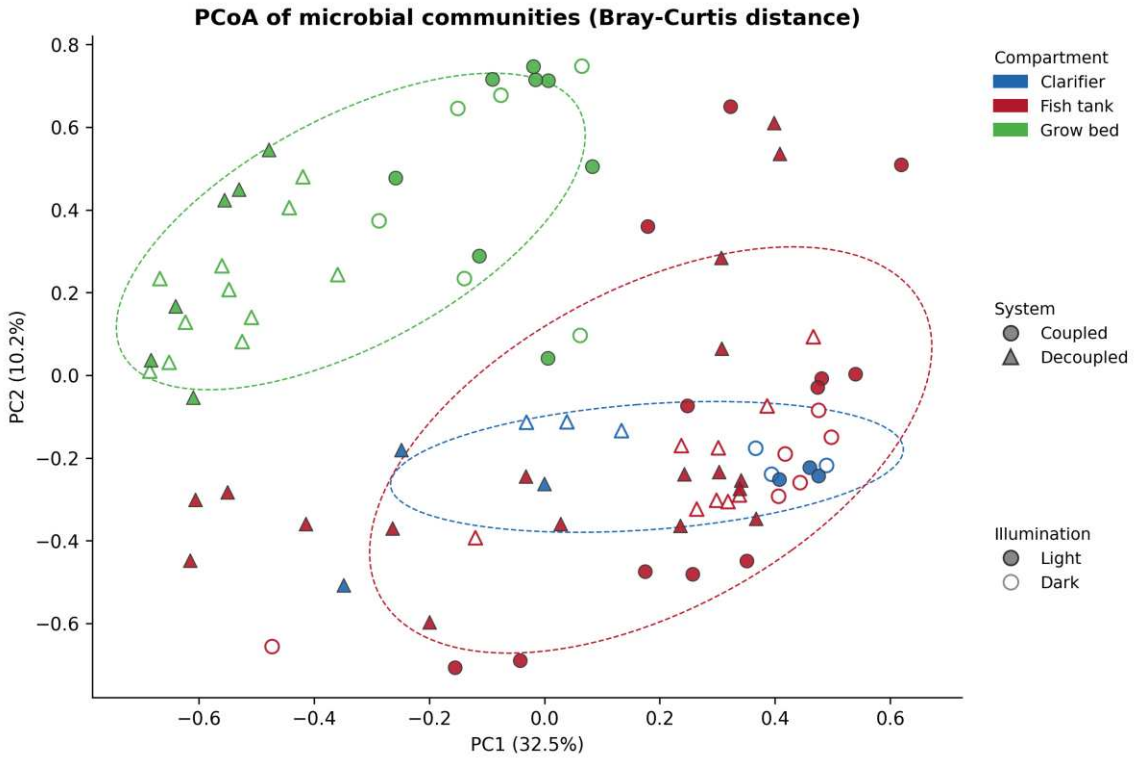


Figure 3. Beta diversity of microbial communities based on Bray-Curtis distances (PCoA, n = 88). Points are colored by compartment and shaped by system type (coupled vs. decoupled). Dashed ellipses represent 68% confidence intervals for each compartment.

2.3 Presence-impact analysis identifies *Cetobacterium* as the primary community-structuring taxon

The leave-one-out PERMANOVA approach tested 150 taxa (top by prevalence, restricted to the 10-90% prevalence range) on all 88 samples, controlling for compartment and system type as covariates. Of these, 75 were statistically significant after FDR correction (p_adj_ < 0.05), and 67 passed both significance and effect-size thresholds (p_adj_ < 0.05 and R^2 > 0.02) (Figure 4). The distribution of R^2 values was right-skewed, with most significant taxa explaining only 2-4% of community variance and a small number of taxa exhibiting substantially larger effects.

Additional high-impact taxa included members of the *Dysgonomonadaceae* (R^2 = 0.067), a family of fermentative *Bacteroidales* with diverse carbohydrate-active enzyme repertoires, associated with organic matter decomposition in fish hindguts and organic-enriched aquatic microenvironments (30); *Thermomonas* sp. (*Proteobacteria*; R^2 = 0.054), a genus isolated from denitrifying biofilm reactors and waste-activated sludge (31,32); *Ferruginibacter* sp. (*Bacteroidota*; R^2 = 0.045), a freshwater sediment genus that also forms a consistent component of activated sludge biofilms (33,34); *Tabrizicola* sp. (*Alphaproteobacteria*; R^2 = 0.045), an aerobic Alphaproteobacterium associated with epilithic biofilm communities in freshwater lakes (35); and members of the *Comamonadaceae* (R^2 = 0.041), a ubiquitous family in aquaponic and recirculating system microbiomes (11,21) (Supplementary Table

S1). The taxonomic diversity of these high-impact taxa, spanning four phyla, suggests that community-structuring influence is not confined to a single lineage but is distributed across functionally distinct groups.

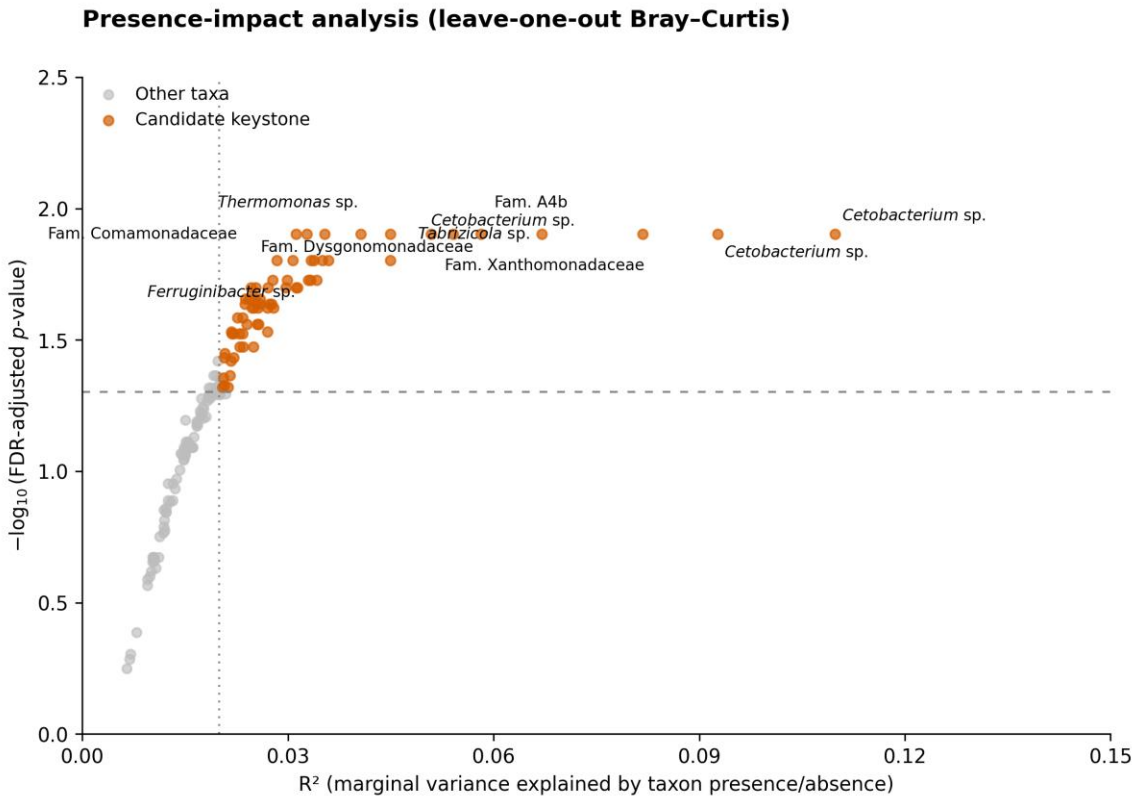


Figure 4. Presence-impact analysis results (leave-one-out PERMANOVA). Each point represents one of 150 tested taxa. The x-axis shows the marginal R^2 (variance explained by focal taxon presence/absence), and the y-axis shows -$\log_{10}$(p_adj). The horizontal dashed line indicates FDR = 5%; the vertical dotted line indicates R^2 = 0.02. Labels identify the top 15 taxa by effect size. *Cetobacterium sp.* ASVs cluster in the upper right, reflecting their disproportionate community-structuring influence.

2.4 Co-occurrence network identifies environmental biofilm hubs

The co-occurrence network, constructed from the 50 most variable taxa in the CLR-transformed 60-sample subset, comprised 50 nodes and 335 edges (density = 0.27) after application of Spearman correlation thresholds (|rho| > 0.6, p_adj_ < 0.001) (Supplementary Figure S3). The Graphical LASSO (GLASSO) was initially attempted across 20 regularization levels (rho = 0.05-2.0), but even at maximal penalization, the minimum achievable density was 30%, indicating that the covariance structure among these taxa does not satisfy the sparsity assumptions underlying GLASSO. This finding itself is informative, suggesting that the most variable taxa in the system are broadly co-responsive to shared environmental drivers rather than organized into discrete, conditionally independent modules(36).

Four taxa were identified as hub organisms based on their position in the top 20% by mean centrality rank across degree, betweenness, and eigenvector centrality: SWB02 sp. (*Proteobacteria*; rank 1 by degree), Fam. SC-I-84 (*Proteobacteria*; rank 2.5), Ord. S-70

(*Planctomycetota*; rank 2.5), and Fam. *Fimbriimonadaceae* (*Armatimonadota*; rank 6). These hub taxa are predominantly environmental biofilm-associated organisms, representing lineages commonly detected in aquatic biofilms and activated sludge systems. Their high connectivity likely reflects co-occurrence driven by shared responses to biofilm-associated environmental conditions (e.g., dissolved oxygen gradients, organic carbon availability) rather than direct metabolic interactions(12).

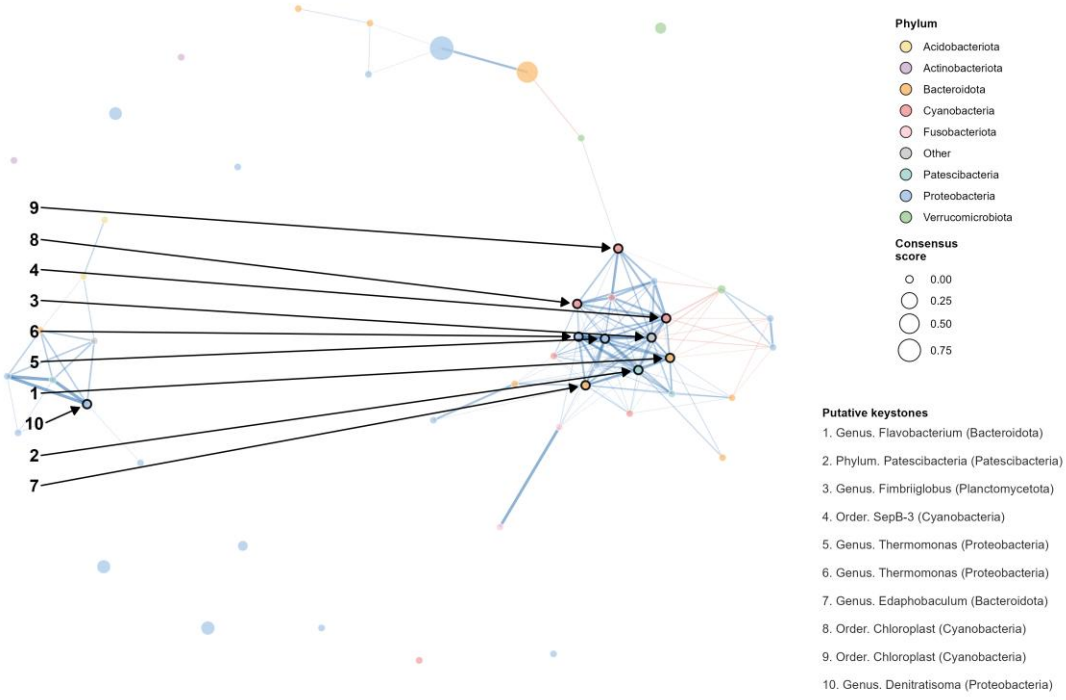


Figure 5. Co-occurrence network among the 50 most variable taxa (Spearman |ρ| > 0.6, p_adj < 0.001). Node size is proportional to degree; node color indicates phylum. Hub taxa (top 10 by degree) are highlighted with black borders and labelled.

2.5 Predictive modeling reveals collective, rather than individual, community turnover

LASSO regression, applied to predict temporal community turnover (Bray-Curtis dissimilarity between consecutive samples at each location), selected zero taxa as individual predictors in both tested configurations: (i) relative abundances across all 88 (37)samples (lambda.1se = 0.093; 85 valid observations, 825 predictors), and (ii) CLR-transformed abundances from the 60-sample subset (lambda.min = 0.065; 57 observations). This null result indicates that temporal community turnover in this system is not attributable to the dynamics of individual taxa but reflects coordinated, community-level shifts, wherein many taxa change concurrently in response to shared environmental drivers such as seasonal transitions, management events, or stochastic processes.

This finding is consistent with the high degree of interconnectedness observed in the co-occurrence network, where even stringent correlation thresholds (|rho| > 0.6) retained 335

edges among 50 taxa, and with the ecological expectation that communities subjected to strong shared environmental forcing (e.g., temperature seasonality in aquaponics) exhibit correlated dynamics across many taxa simultaneously (38).

2.6 Multi-method consensus reveals no classical keystone taxa

The consensus framework identified 71 unique taxa flagged by at least one of the three methods: 67 by presence-impact analysis alone, and four by network centrality alone (Figure 6 and Supplementary Figure S3).

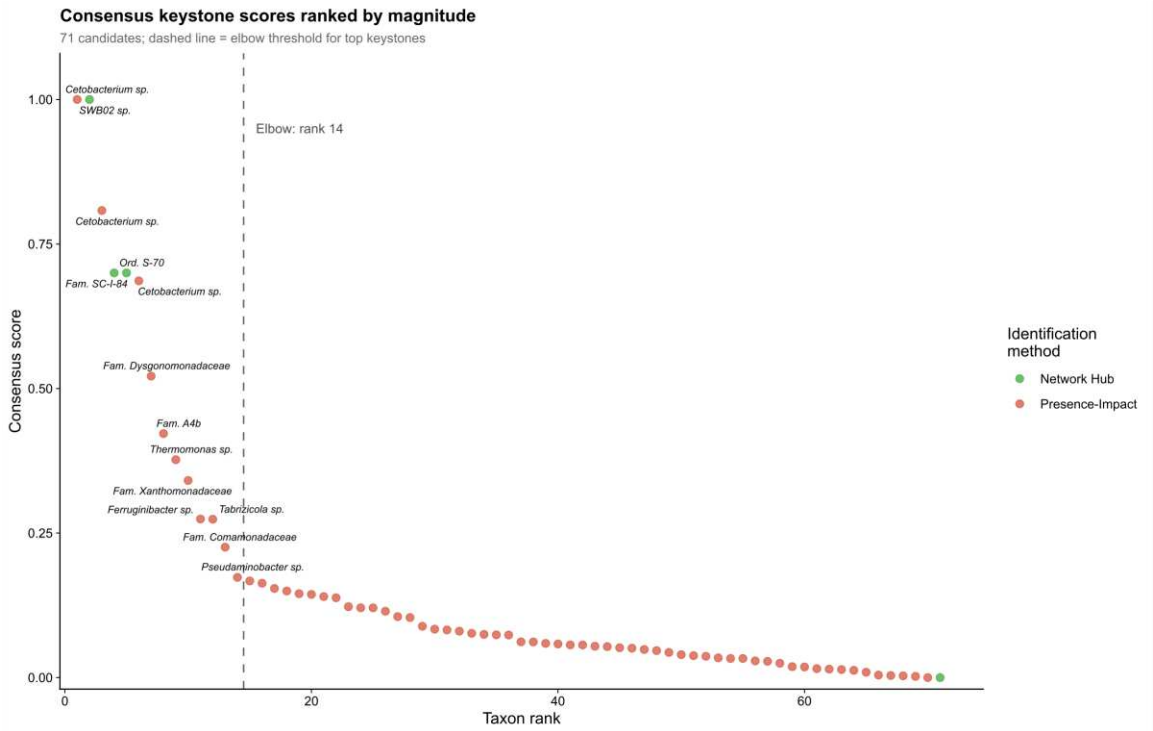


Figure 6. Consensus keystone scores for 71 candidate taxa across all three methods. Points are colored by the number of supporting methods (1 = blue; 2+ = red). Labels indicate taxa identified by multiple criteria (none in this analysis).

This separation reflects the fact that "community-structuring influence" and "network connectivity" represent fundamentally different ecological attributes. A taxon may strongly influence overall community composition through compartment-specific dominance (*i.e.*, by being abundant in one habitat and absent from another, thereby creating large differences in Bray-Curtis distance) without being highly connected to other taxa in a co-occurrence network. Conversely, taxa that are strongly correlated with many other community members, as reflected by network hub status, may be responding to, rather than driving, community dynamics.

The phylum-level composition of the 71 keystone candidates was dominated by *Proteobacteria* and *Bacteroidota*, with additional representation from *Chloroflexi*, *Fusobacteriota*, *Planctomycetota*, and *Armatimonadota* (Figure 7).

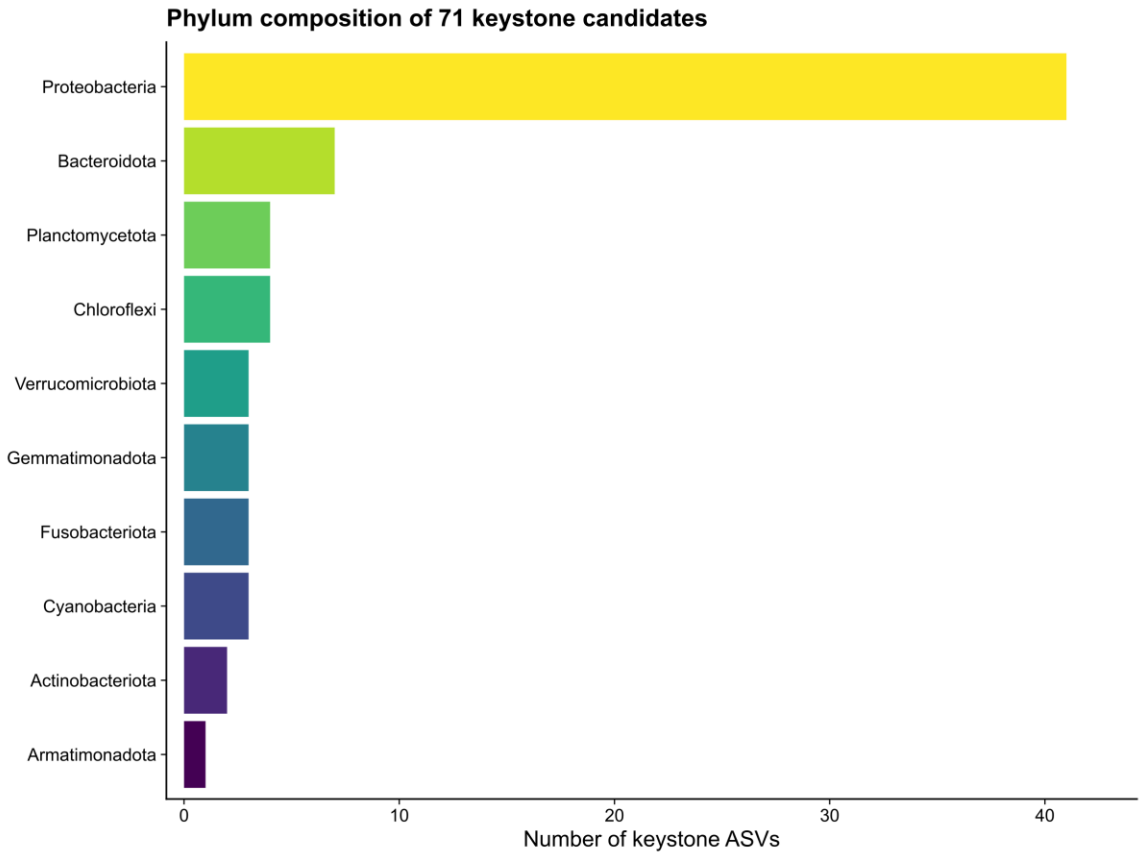


Figure 7. Phylum-level composition of the 71 keystone candidates identified by at least one method. Bar heights indicate the number of candidate ASVs assigned to each phylum.

2.7 Compartment-specific abundance patterns and temporal dynamics

The abundance heatmap of the top 30 keystone candidates revealed distinct clustering by compartment (Figure 8). *Cetobacterium* ASVs were strongly enriched in fish tank samples, consistent with their fish gut-associated ecology, while *Chloroflexi* and *Planctomycetota* representatives were more abundant in grow bed and clarifier biofilm samples. This compartment-specific patterning is consistent with niche specialization and selective assembly within the aquaponic system.

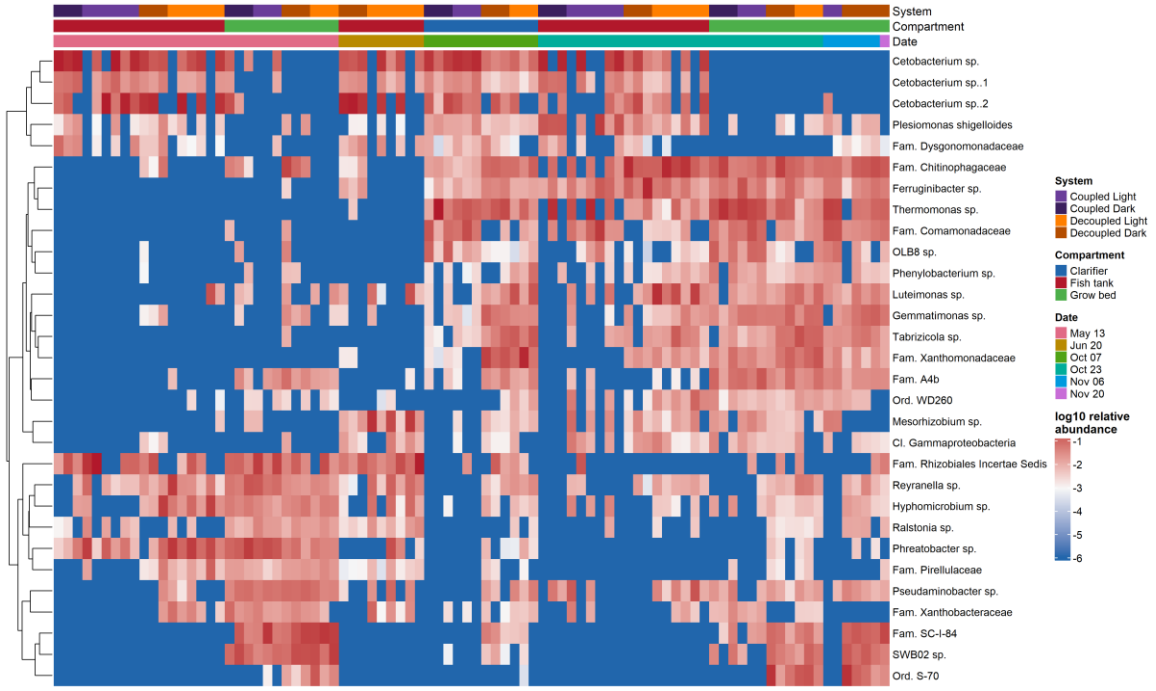


Figure 8. Heatmap of $\log_{10}$-transformed relative abundance for the top 30 keystone candidates. Columns are ordered by sampling date, compartment, system type, and system ID across all 88 samples; annotations identify these groupings, and the complete sample key is provided as Supplementary Table S2.

Temporal trends in candidate keystone taxa are reflected in the date-structured columns of the abundance heatmap (Figure 8). We attribute these temporal patterns to seasonal shifts in water temperature, photoperiod, organic loading, and fish feeding activity, all of which modulate the relative ecological importance of individual taxa over time. Candidate abundance distributions across coupled light, coupled dark, decoupled light, and decoupled dark systems were summarized as boxplots to emphasize central tendencies among hydraulic-light regimes (Figure 9).

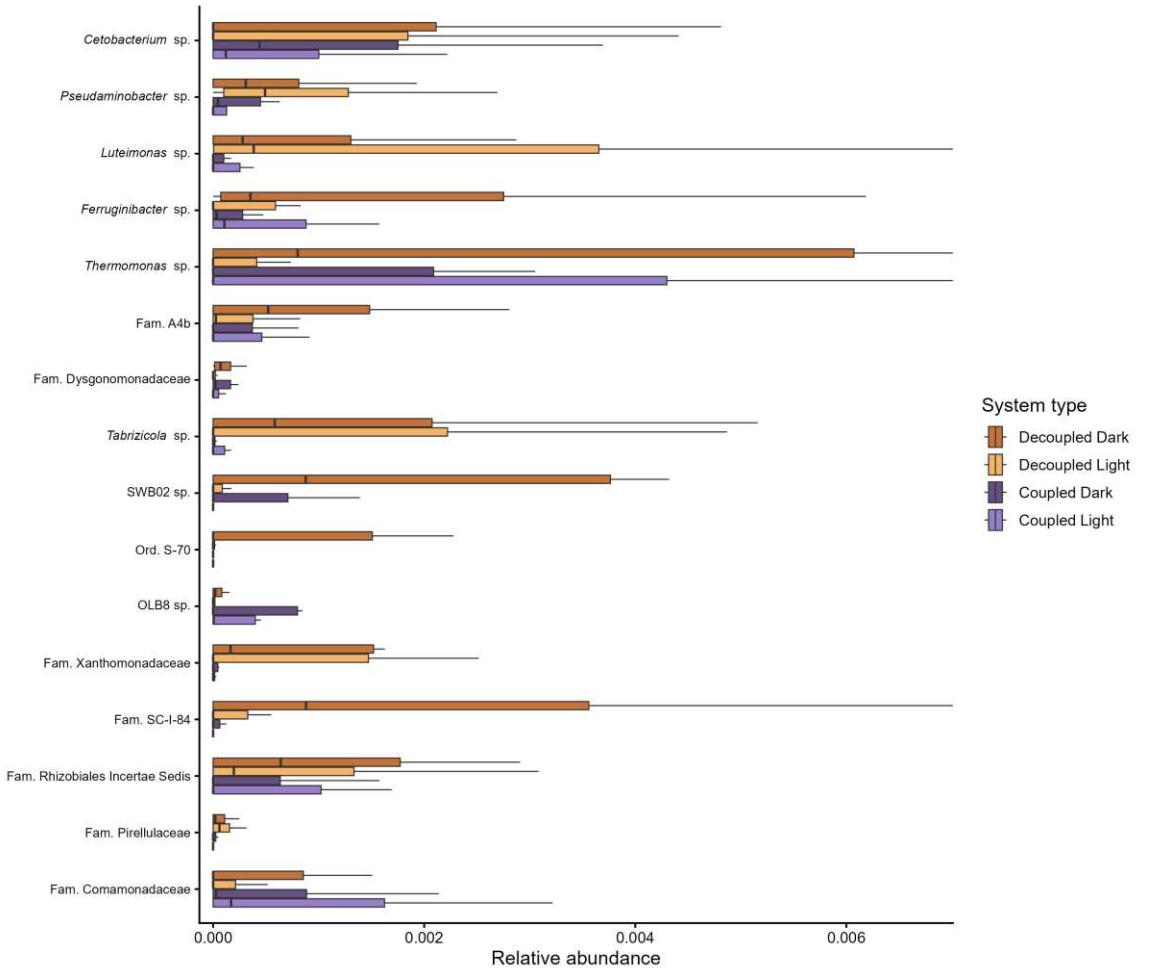


Figure 9. Keystone candidate abundance distributions across system types. Boxplots show central relative-abundance distributions; individual sample points are omitted and the x-axis is zoomed to improve visualization of differences among system types.

Temporal stability, quantified as the coefficient of variation (CV) of relative abundance, varied among keystone candidates (Figure 10). Network hub taxa (environmental biofilm organisms) tended to exhibit lower CV values, suggesting greater temporal persistence, while high-impact presence-impact taxa (e.g., *Cetobacterium*) showed higher variability. This pattern is consistent with a trade-off between ecological influence and temporal stability: the most community-structuring taxa may be dynamically responsive to environmental change, while the most stably connected taxa may serve a buffering or anchoring role within the community.

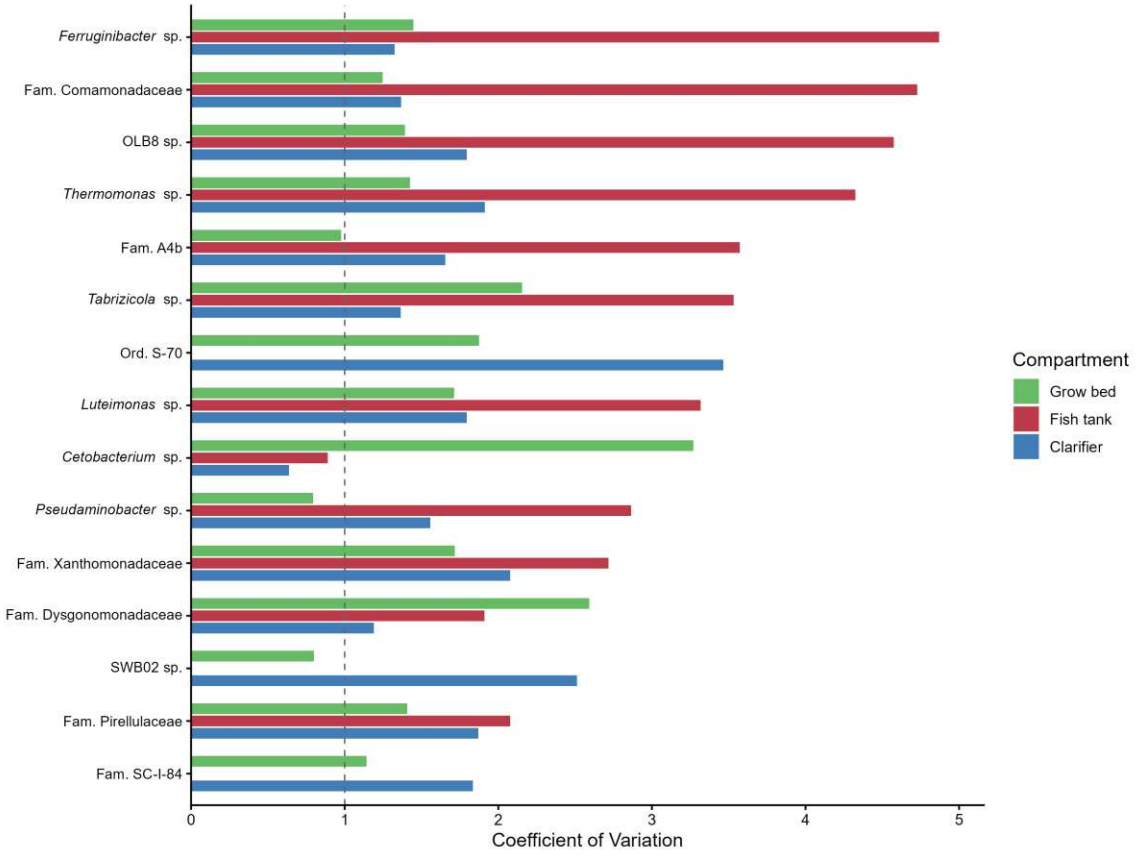


Figure 10. Temporal stability of keystone taxa by coefficient of variation (CV) per compartment. Lower CV values indicate greater temporal persistence. The dashed vertical line marks CV = 1; values below 1 indicate stable taxa, whereas values above 1 indicate more variable taxa.

2.8 Environmental operating context and system-mode contrasts

The operating records are presented as context for interpreting the community patterns, not as primary forcing variables. Water quality parameters (pH, dissolved oxygen, temperature) were actively managed toward target set-points through continuous aeration, pH adjustment, and partial climate control; observed variation in these metrics therefore reflects system responses rather than experimental treatment effects. The primary forcing variables in this experiment were system configuration (coupled vs. decoupled), fish tank illumination (light vs. dark), sampling season, and fish biomass and feed rate, which increased progressively over the grow-out period. Overall, mean pH was 6.6 ± 0.3, mean dissolved oxygen was 7.3 ± 1.7 mg/L, and mean temperature was 23.4 ± 3.3 °C; light-decoupled systems tended toward slightly higher pH and dissolved oxygen than other strata (Figure 11). These differences provide plausible background forcing for the compartment- and method-specific microbial patterns described above.

Trajectories of pH, dissolved oxygen, temperature, feed, and base addition by hydraulic mode and illumination regime across the monitoring period are shown in Figure 11. The photosynthesis-associated proxy is provided as Supplementary Figure S1. Exploratory PERMANOVA using 16S samples with matched environmental data showed that

compartment, coupling, illumination, and season were associated with Bray-Curtis community structure, while sample-matched temperature and dissolved oxygen had smaller effects in the matched subset (Supplementary Table S3; Supplementary Figure S2; Supplementary Table S4).

The PERMANOVA results carry specific mechanistic implications. Compartment structure ($R^2 = 0.256$) dominated because the three aquaponic habitats, fish tank, clarifier, and grow bed, impose qualitatively different selection pressures on microbial communities through gradients in oxygen, organic loading, and surface attachment opportunity. The coupling effect ($R^2 = 0.069$) reflects documented differences between coupled and decoupled system architectures in nutrient delivery and organic-matter dynamics (2). The illumination effect ($R^2 = 0.031$) is consistent with photoautotrophic community contributions in light-exposed systems, as shown by the photosynthesis-associated proxy (Supplementary Figure S1). The season effect ($R^2 = 0.055$) likely reflects temperature-driven and plant phenology-linked shifts in organic inputs and nutrient cycling demand that co-occur with crop rotation. The non-significant effects of sample-matched temperature (adjusted $p = 0.279$) and dissolved oxygen (adjusted $p = 0.372$) are noteworthy: they do not indicate that these variables are biologically unimportant, but that point-in-time values at six sampling dates capture less community variation than the categorical treatment contrasts, which integrate sustained environmental differences across the full monitoring period. Monitoring across more time points or at higher temporal resolution would improve the power to detect continuous environmental drivers.

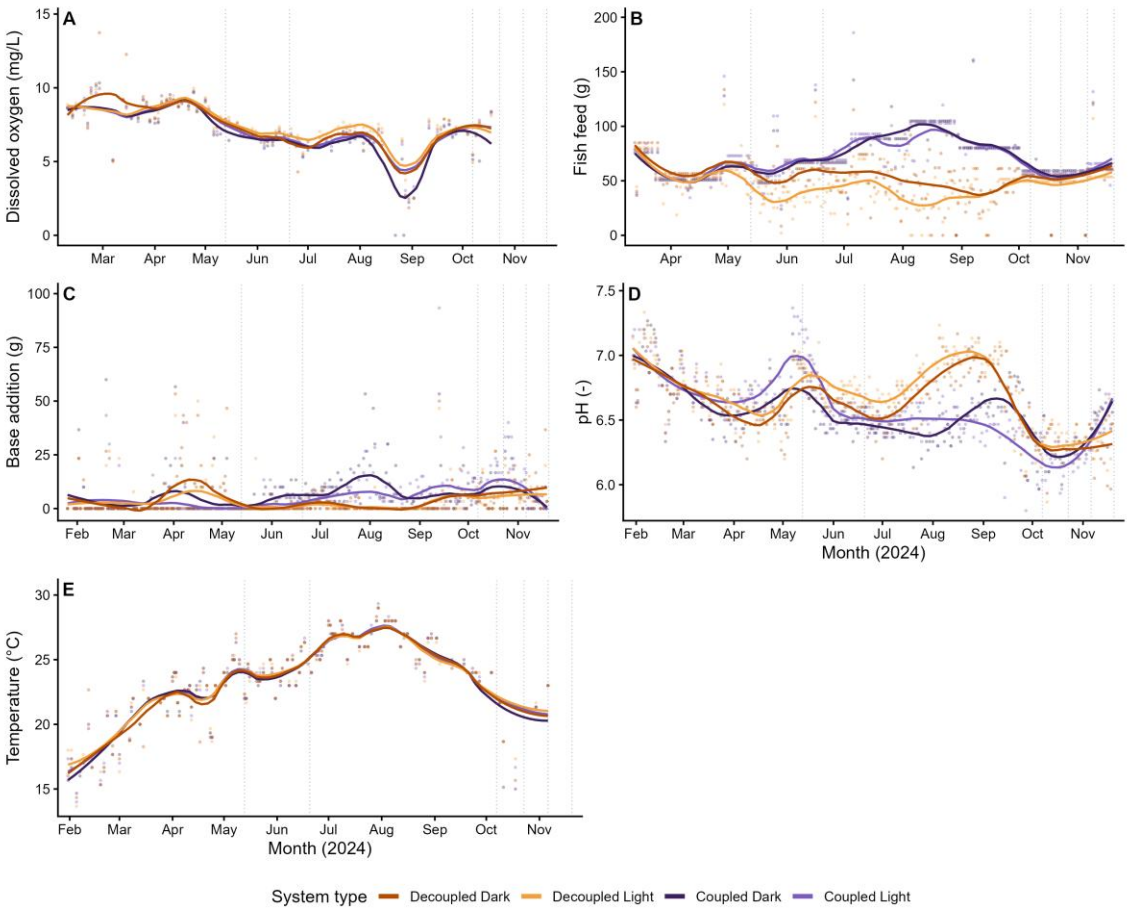


Figure 11. Operational water quality parameters and management inputs across the monitoring period. Colors distinguish hydraulic mode (coupled versus decoupled), line types distinguish illumination (dark versus light), and blue vertical dashed lines indicate 16S sampling dates.

2.9 Treatment-stratified presence-impact analysis

To assess whether community-structuring signals were consistent across hydraulic configurations, the presence-impact screen was repeated separately within coupled (n = 38) and decoupled (n = 50) subsets, with compartment and illumination retained as covariates. In coupled systems, six taxa survived FDR correction ($p < 0.05$): *Chitinophagaceae* ($R^2 = 0.086$), *Micropepsaceae* ($R^2 = 0.082$), *Reyranella* ($R^2 = 0.082$), *Luteolibacter* ($R^2 = 0.072$), *Defluviimonas* ($R^2 = 0.071$), and *Cetobacterium* ($R^2 = 0.057$). In decoupled systems, two taxa survived FDR correction: *Chitinophagaceae* ($R^2 = 0.061$) and *Romboutsia* ($R^2 = 0.032$). *Chitinophagaceae* was the only taxon significant in both subsets, indicating cross-treatment community-structuring activity. The treatment-specific detection of *Cetobacterium* exclusively in coupled systems is consistent with its fish-tank ecology being amplified when fish-tank water circulates directly through all compartments (Supplementary Table S5).

3. Discussion

3.1 Aquaponics as a distributed ecological architecture

No single taxon governed this aquaponic microbiome. Three mechanistically independent screens of the 88-library dataset disagreed almost completely: presence-impact analysis flagged 67 community-structuring taxa, co-occurrence centrality flagged four hubs, and LASSO selected no turnover predictor at all, and no taxon survived more than one screen (Figures 4–6 and Supplementary Figure S3). Such disagreement is signal, not noise. It marks a compartmentalized system in which ecological influence is spread across fish tanks, clarifiers, and grow beds rather than concentrated in any single controlling taxon (4,6,17,21,22,38).

This pattern echoes a growing body of microbial-ecology evidence that the classical keystone concept, formulated for macro-ecological food webs, may not translate directly to complex microbial communities (15,16). Microbial systems often combine high taxonomic diversity, rapid turnover, functional redundancy, and dense co-response networks, all of which can distribute ecological influence across multiple taxa and functions (13,17,18,38). Each screen captured a different dimension of community organization: taxa that most shaped composition, taxa central to the co-occurrence network, and taxa that jointly predicted temporal turnover (Figures 4–6 and Supplementary Figure S3). These method-specific signals reflect complementary ecological processes, not analytical noise (10,17,18,39).

Two recent studies frame this pattern directly. Amit and Bashan showed that top-down influence metrics flag taxa that network topology misses (15), and Herren and McMahon showed that keystone-based community prediction is contingent on context and disturbance (14). The aquaponic data align with this view: Bray-Curtis ordination separated communities primarily by compartment (Figure 3), the heatmap revealed compartment-structured abundance patterns among top candidates (Figure 8), and environmental records documented seasonal and system-type variation in pH, dissolved oxygen, temperature, feed, and base addition inputs (Figure 11). Together, these results link the distributed keystone signal directly to the physical architecture and operating context of the facility.

3.2 Cetobacterium as a compartment-structuring taxon

Cetobacterium sp. was the single clearest compartment-structuring taxon in this dataset. Three *Cetobacterium* ASVs ranked highest in the presence-impact screen, each explaining approximately 8-11% of community variance after the focal taxon was removed from the distance matrix (Figure 4). This signal has a clear biological basis: *Cetobacterium somerae* is an anaerobic bacterium associated with freshwater fish intestines (40). In Nile tilapia specifically, *Cetobacterium* supplementation has been shown to improve carcass ratio and enhance serum amino acid levels (41). Its enrichment in fish-tank samples and near absence from grow beds and clarifiers therefore links the statistical signal to a biologically plausible source habitat within the aquaponic system (Figures 3, 4, and 8).

Read in context, *Cetobacterium* is an indicator of fish-tank-associated community structure, not a universal aquaponic keystone. That distinction matters: aquaponic systems hold planktonic, biofilm-associated, and rhizosphere-associated habitats, each imposing its own environmental filter on the resident community (4,8,21,22). In the present data, *Cetobacterium* was influential in the presence-impact analysis but was not recovered as a network hub or LASSO predictor (Figures 4–6 and Supplementary Figure S3), indicating that its utility is greatest as a compartment-specific marker rather than a system-wide control point. Treatment-stratified analysis confirmed this interpretation: *Cetobacterium* was significant only within coupled systems ($R^2 = 0.057$, $p < 0.05$), whereas *Chitinophagaceae* emerged as the sole cross-treatment community-structuring taxon ($R^2 = 0.086$ coupled; $R^2 = 0.061$ decoupled), consistent with its role in biofilm-associated nutrient cycling across system types. The organism was detected in the fish tank fraction of both coupled and decoupled systems, suggesting that its presence reflects fish-gut seeding into the water column regardless of hydraulic configuration. Of note, anaerobic activity in the Radial Flow Settler (RFS) of decoupled systems has previously been associated with root hypoxia problems in the grow beds (42); *Cetobacterium*'s anaerobic metabolism may contribute to oxygen depletion in compartments where organic matter accumulates, warranting further investigation as a potential system health indicator.

3.3 Presence-impact taxa beyond *Cetobacterium*

Although *Cetobacterium* produced the strongest presence-impact signal, the screen flagged a taxonomically broad set of additional community-structuring taxa whose ecology clarifies why no single organism behaves as a universal keystone. The vitamin B_{12}-producing, network-restructuring role reported for *Cetobacterium somerae* (43) underscores that even the top taxon acts through compartment-specific dominance rather than system-wide control.

Several high-impact taxa are anaerobic or detritus-associated *Bacteroidota* consistent with clarifier and grow-bed niches rather than fish-tank planktonic habitats. The *Dysgonomonadaceae* ($R^2 = 0.067$) are documented fermenters of sugars and amino acids in low-oxygen bioreactors (44,45), while *Ferruginibacter* sp. and related *Chitinophagaceae* ($R^2 = 0.045$) are freshwater sediment and engineered-biofilm *Bacteroidota* (33,46,47). The stronger coupled-system presence-impact of *Chitinophagaceae* ($R^2 = 0.086$; Supplementary Table S5) is consistent with hydraulic integration delivering particulate organic matter to these biofilm compartments.

A second group comprises biofilm-associated *Proteobacteria* and *Chloroflexi* typical of nitrogen-cycling attached-growth systems. The A4b lineage of *Chloroflexi* ($R^2 = 0.058$) matches filamentous *Chloroflexota* described in activated sludge and anammox biofilms (48–50), and *Thermomonas* sp. ($R^2 = 0.054$) is a denitrification-associated genus first isolated from biofilm on plastic carriers and repeatedly linked to nitrification performance in hybrid biofilm/sludge reactors (31,51). Both showed stronger presence-impact in coupled than decoupled systems (Table S5), consistent with integrated biofilter hydraulics sustaining nitrifier-denitrifier niches.

A third group is dominated by light-responsive and low-abundance taxa whose interpretation is more tentative. *Tabrizicola* sp. (R^2 = 0.045) is an aerobic anoxygenic phototroph that produces bacteriochlorophyll a and is enriched in epilithic biofilms rather than open water (35,52,53); its physiology is relevant given the illumination effect on community structure (R^2 = 0.031), although aggregate keystone abundance did not differ between light and dark samples (Wilcoxon p = 0.65). *Pseudaminobacter* sp. (R^2 = 0.036) is an activated-sludge and granule taxon with little aquaponics-specific literature (54,55), and the *Pirellulaceae* (*Planctomycetota*; R^2 = 0.035) align with low-oxygen, biofilm-associated planctomycetes in wastewater and anammox systems (56–58).

No single lineage dominates this set. The candidates span at least four phyla (*Bacteroidota*, *Proteobacteria*, *Chloroflexi*, and *Planctomycetota*) and sort by compartment and hydraulic mode rather than by taxonomy. That breadth is the central point: community-structuring influence here is distributed across functionally distinct, habitat-specific guilds, so each presence-impact taxon is best read as a compartment indicator, not a universal control point.

3.4 Network hubs as environmental biofilm indicators

The network hubs told a different story. SWB02, SC-I-84, S-70, and *Fimbriimonadaceae* surfaced through co-occurrence centrality rather than presence-impact effect size (Figures 5–6 and Supplementary Figure S3), describing how the community is wired rather than which taxa move it. This inference carries a caveat: high connectivity can reflect shared environmental responses as readily as direct interaction. That caveat is relevant here because GLASSO did not yield a sparse conditional-dependence network, while the Spearman network retained many edges among the 50 most variable taxa (Figure 5). Separate from these network hubs, the presence-impact screen flagged one cross-treatment community-structuring taxon: *Chitinophagaceae* was detected in both coupled and decoupled systems across fish tank and grow bed compartments. Members of this family are polysaccharide-degrading *Bacteroidota* associated with organic-matter cycling and biofilm communities in freshwater and engineered aquatic systems (59,60); their cross-treatment presence makes them a candidate indicator taxon for operational monitoring of aquaponic system health(12,13,61).

This biofilm reading fits both the study design and prior work on engineered water systems, where component microenvironments structure microbial distributions and biofilm and suspended communities occupy separate niches (6,62). In our data, the network hubs were not the same taxa that drove fish-tank presence-impact effects (Figures 4–6 and Supplementary Figure S3), which supports a compartment-linked interpretation: network hubs are better viewed as indicators of environmental biofilm organization rather than as direct management targets (5,6).

The individual hubs reinforce that co-response reading. The highest-degree hub, SWB02 (*Hyphomonadaceae*; degree = 27), is enriched in sessile rather than planktonic freshwater communities (63,64). The uncultured lineage SC-I-84 (degree = 26) is reported in

freshwater sediments and activated-sludge taxonomies of uncultured *Proteobacteria* (46,65,66), and the order-level *Planctomycetota* lineage S-70 (degree = 26) is consistent with biofilm planctomycetes in nitrogen-cycling systems (58). *Fimbriimonadaceae* (*Armatimonadota*), the hub with the highest betweenness centrality (44.1), has been detected in anammox and partial-nitritation biofilms (67,68), a profile that fits a connector role between community modules rather than a high presence-impact effect.

3.5 Environmental drivers and operating context

Operating conditions give the distributed structure its mechanistic backdrop. Across the 2024 operating record, mean pH was 6.6 ± 0.3, mean dissolved oxygen was 7.3 ± 1.7 mg/L, and mean temperature was 23.4 ± 3.3 °C (Figure 11; invalid and outlier pH values removed before reporting). These values differed by hydraulic and illumination stratum: light-decoupled systems showed the highest mean pH and dissolved oxygen (pH 6.7; dissolved oxygen 7.5 mg/L), whereas dark-coupled systems had lower corresponding means (pH 6.55 ± 0.28; dissolved oxygen 7.1 ± 1.7 mg/L; Figure 11). Because pH, oxygen availability, temperature, and organic loading collectively shape microbial growth and biogeochemical activity in aquaponic and engineered water systems, these operating gradients offer plausible environmental filters for the compartment-specific candidate profiles observed across the three analyses (61,62).

The sample-matched environmental analysis supported this interpretation while also defining its limits. Across all 88 samples, compartment explained the largest share of Bray-Curtis variation (R^2 = 0.256, p = 0.001), followed by coupling (R^2 = 0.069, p = 0.001), season (R^2 = 0.055, p = 0.001), and illumination (R^2 = 0.031, p = 0.004; Supplementary Table S3). In the subset with matched temperature and dissolved oxygen records (n = 42), compartment, coupling, and illumination remained significant predictors, whereas temperature and dissolved oxygen showed smaller effects that did not survive multiple-testing correction (temperature R^2 = 0.018, adjusted p = 0.279; dissolved oxygen R^2 = 0.014, adjusted p = 0.372; Supplementary Table S3). Exploratory envfit analysis identified a temperature vector associated with the ordination (r^2 = 0.147, p = 0.047) and a weaker vector for dissolved oxygen (r^2 = 0.108, p = 0.109; Supplementary Figure S2; Supplementary Table S4), indicating that both variables contributed to the environmental background without overriding compartment structure (29).

The same environmental framing sharpens the keystone results. *Cetobacterium* was most influential in fish-tank-associated presence-impact analyses (Figure 4), whereas network hubs were environmental biofilm-associated taxa (Figures 5–6 and Supplementary Figure S3); the environmental records show that these biological signals occurred under distinct light/dark, coupled/decoupled, seasonal, and water-chemistry conditions (Figure 11; Supplementary Table S3). Therefore, light regime, pH trajectory, temperature, dissolved oxygen, and management inputs should be interpreted as operating filters that shape which taxa become prominent in a given compartment, rather than as evidence for a single environmental driver of all candidate keystone patterns.

Stratified presence-impact analysis links these environmental gradients to specific candidate taxa. Biofilm-associated taxa including *Thermomonas*, *Tabrizicola*, SC-I-84, and *Chitinophagaceae* showed stronger presence-impact in coupled than decoupled subsets (Supplementary Table S5), indicating that hydraulic integration modulates which taxa structure the biofilm community. Consistent with the whole-community result, an exploratory envfit analysis placed a significant temperature vector on the ordination (r^2 = 0.147, p = 0.047) with a weaker dissolved-oxygen vector (r^2 = 0.108, p = 0.109), while pooled keystone abundance did not differ by illumination (Wilcoxon p = 0.65). These patterns indicate taxon-specific responses to operating conditions rather than a single global environmental driver, in keeping with light-responsive phototrophs and dissolved-oxygen-stratified biofilm hubs.

3.6 Functional redundancy and distributed resilience

The case for distributed resilience rests on both these data and microbial theory. In this study, the absence of cross-method convergence occurred alongside high community diversity after prevalence filtering (825 ASVs), strong compartment structure (Figure 3), and distinct candidate profiles by method (Figures 4–6 and Supplementary Figure S3). Microbial resistance, resilience, and functional redundancy are context dependent, and communities with overlapping functions may maintain process potential even when taxonomic composition changes (17,18,69). These concepts fit the observed pattern: aquaponic function appears more likely to depend on distributed microbial guilds across habitats rather than on one taxon that is universally important across all analytical definitions (17,18).

A limitation of 16S rRNA amplicon data is that they reflect taxonomic composition rather than metabolic function; demonstrating functional keystone activity would require complementary approaches such as metagenomics, metatranscriptomics, or controlled perturbation experiments (16,18). The present results are therefore interpreted in compositional and ecological terms, with gene-level and process-level keystone roles remaining as explicit hypotheses for future investigation. Nitrogen metabolism deserves specific attention in this context. Aquaponic systems depend on microbial nitrification to support fish health and plant nutrition, yet also contain microenvironments where denitrification occurs (4,5,70). The 67 presence-impact candidates span multiple phyla, including *Proteobacteria*, which encompass families known to harbour ammonia-oxidising bacteria, heterotrophic nitrifiers, and denitrifiers, and *Planctomycetota*, which include anaerobic ammonium-oxidising (Anammox) lineages (71). Network hubs SWB02 and SC-I-84 belong to orders commonly found in engineered biofilm environments where nitrogen transformations are active. However, the 16S rRNA amplicon data resolve phylogenetic affiliation rather than gene expression or catalytic activity, so putative nitrogen-cycling roles for any candidate taxon remain inferential. Demonstrating functional keystone status in nitrogen cycling would require quantitative PCR targeting functional genes or metatranscriptomic profiling paired with nutrient flux measurements. The absence of a single dominant nitrogen-cycling candidate in the consensus set is consistent with the

known functional redundancy of nitrification and denitrification guilds observed in recirculating aquaculture and biofilm reactor systems (51,72).

3.7 Methodological considerations

The leave-one-out PERMANOVA correction mattered because it aligned the statistical test with the ecological question. Pre-correction, 802 tested taxa were significant. Removing each focal taxon from the community matrix before recomputing Bray-Curtis distances reduced the candidate set to 67 taxa (Figure 4). This reduction shows that the corrected analysis limited circularity between the focal taxon and the distance matrix, a concern that is especially relevant for compositional microbiome data (73). The corrected candidate set therefore provides a more conservative basis for interpreting taxa that structure the remainder of the community.

The network and LASSO results also hinge on the study design. The network analysis used the 50 most variable taxa from the CLR-transformed 60-sample subset, and GLASSO did not produce a sparse network under the tested regularization range; the final hub analysis therefore used a stringent Spearman co-occurrence network (Figure 5). The predictive analysis evaluated 825 ASVs across 88 samples, a high-dimensional setting in which LASSO regularization is expected to be conservative; machine-learning and model-based keystone approaches may require larger datasets or different experimental designs to recover predictive taxa robustly (16,74). These constraints support cautious interpretation of method-specific candidate lists rather than overextension from any single method(75).

The presence-impact analysis was additionally conducted within coupled and decoupled subsets separately (Section 2.9), confirming that most community-structuring signals are partially treatment-specific. Co-occurrence network construction and LASSO modelling require larger per-group sample sizes than those available here; treatment-stratified network and predictive analyses therefore remain priorities for future work with larger, treatment-balanced datasets.

Finally, topological indices such as degree centrality or betweenness centrality are typically weak predictors of experimentally validated keystone identity. Accurate prediction requires large datasets, yet it is difficult to assess whether a given sample size can be considered sufficiently large (10,16,74). This highlights the need for further studies to corroborate the evidence presented here.

3.8 Implications for aquaponic management

For practice, the message is simple: aquaponic microbiome monitoring should be community- and compartment-aware. The data show that fish-tank structure was strongly associated with *Cetobacterium* in the presence-impact analysis (Figure 4), whereas network hubs reflected a separate set of environmental biofilm-associated taxa (Figures 5–6 and Supplementary Figure S3). Monitoring strategies should therefore track compartment-specific indicators and environmental operating conditions rather than relying on a single universal keystone target.

The operating record shows why. pH, dissolved oxygen, temperature, feed, and base addition records varied across system type and the monitoring period after invalid and robust outlier pH values were removed (Figure 11). Because microbial resistance and resilience depend on disturbance context and environmental history (17,61,69), baseline microbiome expectations in aquaponic systems should be defined by compartment, system configuration, and season. This framing connects the statistical result to practical management: maintaining stable operating conditions that support diverse compartment-specific communities is a more defensible target than manipulating single taxa without functional validation (4,5,16).

4. Methods

4.1 Experimental systems and design

Twelve independent aquaponic systems were operated over a 9-month fish grow-out period at the Auburn University aquaponic research facility. The setup and operation of these systems has been described in detail in previous publications (27,42). Briefly, Nile tilapia (*Oreochromis niloticus*) were grown in 700-L fish tanks and fed with Cargill 3606 pelletized feed at 1.5% of body weight per day. Fish (130 g on average) were placed in tanks in early April and grown to approximately 400 g upon harvest in early November. Fish tanks were intensively aerated to maintain dissolved oxygen at approximately 7.0 $mg \cdot L^{-1}$. Four system configurations were tested in triplicate: coupled-light (LC), coupled-dark (DC), decoupled-light (LD), and decoupled-dark (DD), yielding 12 systems total. In coupled systems, water recirculated continuously between fish tanks and hydroponic grow beds; in decoupled systems, the hydraulic loops were separated. Light and dark designations refer to fish tank illumination: light systems received natural sunlight through the tank walls, while dark systems had opaque coverings preventing light penetration into the fish tank. Cherry tomatoes (Favorita F1, Johnny's Seeds) were cultivated in pea-gravel substrate in 1-m^2 grow beds with five vines per bed. A spring tomato trial ran from March to June and a fall trial from August to November. Water was clarified in a Radial Flow Settler (RFS) before recirculation.

4.2 Study design and sample collection

Water and biofilm samples were collected from a research-scale aquaponic facility at Auburn University encompassing three primary compartments: the clarifier (return-flow solid separator), fish tanks, and grow beds. The facility operated both coupled and decoupled system configurations under light and dark conditions in the fish tanks. Sampling spanned six time points from May through November 2024 (May 13, June 20, October 7, October 23, November 6, and November 20), capturing both the spring and fall growing seasons, for a total of 88 paired-end amplicon libraries. The 88 libraries encompass four distinct sample types: fish tank water (planktonic fraction, FT; n = 30), biofloc pellets collected by centrifugation of fish tank water in decoupled systems (BFF; n = 15), clarifier (Radial Flow Settler) sludge (CL; n = 12), and grow-bed sludge/rhizosphere material (GB; n = 24). An additional seven samples represent specialised November collections (grow-bed

water columns and a single bed-sludge sample from three decoupled systems). All sample types were processed by centrifugation at 4696×g for 10 minutes; the resulting biomass pellets were stored at −20 °C until DNA extraction.

4.3 DNA extraction and 16S rRNA gene amplicon sequencing

Total DNA was extracted from water and biofilm samples using a commercially available extraction kit (MP Biomedicals). The V4 region of the 16S rRNA gene was amplified using the Earth Microbiome Project primer pair 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACNVGGGTWTCTAAT-3'). Libraries were prepared with unique dual indices (UDI) and sequenced on an Illumina platform generating 2 x 251 bp paired-end reads(76,77).

4.4 Bioinformatic processing

Raw demultiplexed FASTQ files were processed using the DADA2 pipeline (v1.34.0) in R (v4.4.1). Residual primer sequences were removed from all reads using cutadapt (v5.1) (78), specifying both forward and reverse primers alongside their reverse complements. Reads lacking identifiable primer sequences were discarded. Trimmed reads were quality-filtered usingfilterAndTrim() with truncation lengths of 220 bp (forward) and 180 bp (reverse), maximum expected errors of 2 per read, and PhiX removal. Error models were learned separately for forward and reverse reads using learnErrors() with 10^8 bases. ASVs were inferred per sample using dada(), paired reads were merged usingmergePairs(), and chimeric sequences were removed usingremoveBimeraDenovo() with the consensus method (79). Taxonomic classifications were assigned usingassignTaxonomy() andaddSpecies() against the SILVA v138.1 SSU NR99 reference database(80).

4.5 Community analysis

Count data were filtered to retain ASVs present in at least 20% of samples across all 88 libraries, yielding 825 ASVs. For compositional transformation, zero-replacement was performed using the count zero multiplicative method (CZM;zCompositions::cmultRepl()) followed by centered log-ratio (CLR) transformation (73,81). Twenty-eight samples with greater than 80% zero-count taxa were excluded from CLR-based analyses, retaining 60 samples for network construction. Bray-Curtis dissimilarity on relative abundances was computed for all 88 samples for presence-impact analysis and ordination.

Alpha diversity (Shannon index, observed ASV richness) was computed using phyloseq (82). Beta diversity was visualized using PCoA on Bray-Curtis distances.

4.6 Keystone species identification

Presence-impact analysis. For each of the top 150 ASVs by prevalence (10-90% range), the focal taxon was removed from the OTU table, relative abundances were recomputed for the remaining taxa, Bray-Curtis distance was calculated, and PERMANOVA (vegan::adonis2(), 999 permutations, by = "margin") was applied with the focal taxon's presence/absence as the predictor and compartment (Location) and system type (Coupling) as covariates. P-values were FDR-corrected using the Benjamini-Hochberg procedure. Taxa were considered

significant at p_adj_ < 0.05 with R^2 > 0.02. To assess treatment specificity, the screen was repeated separately within coupled (n = 38) and decoupled (n = 50) subsets, retaining compartment and illumination as covariates and using 499 permutations (83,84).

Network centrality analysis. A co-occurrence network was constructed from the top 50 ASVs by CLR variance (60-sample subset). The Graphical LASSO (GLASSO) was attempted across 20 regularization levels (rho = 0.05-2.0) but could not achieve density below 30%; a Spearman rank correlation network was therefore constructed, retaining edges where |rho| > 0.6 and p_adj_ < 0.001. Four centrality measures (degree, betweenness, closeness, eigenvector centrality) were computed using igraph; taxa in the top 20% by mean centrality rank were designated hub candidates (36,85).

Predictive modeling. Temporal community turnover (Bray-Curtis dissimilarity between consecutive samples at each location) was modeled using LASSO regression (glmnet::cv.glmnet(), alpha = 1) in two configurations: (i) relative abundances across all 88 samples (lambda at lambda.1se), and (ii) CLR-transformed abundances from the 60-sample subset (lambda at lambda.min). Taxa with non-zero coefficients were designated predictive candidates(37).

Consensus scoring. Qualified taxa from each method were combined into a union set. A consensus score was computed as the sum of scaled PI effect size (0-1), scaled inverse centrality rank (0-1), and a predictive flag (0 or 1). Taxa identified by two or more methods represent the highest-confidence candidates.

4.7 Supplementary environmental community analysis

Exploratory environmental-community analyses were conducted to connect operational records with 16S community structure. Bray-Curtis PERMANOVA was first run on all 88 samples using compartment, coupling, illumination, and tomato season as predictors. A second PERMANOVA was run on the subset of samples with matched temperature and dissolved oxygen data (n = 42), including compartment, coupling, illumination, temperature, and dissolved oxygen. pH, feed, and base addition were summarized as operating-context variables but were not included in sample-matched PERMANOVA because too few 16S samples had matched values at the selected temporal window. For visualization, Bray-Curtis PCoA was calculated for the temperature/dissolved oxygen subset, and envfit() was used to project temperature and dissolved oxygen vectors onto the ordination.

4.8 Statistical environment

All analyses were performed in R v4.4.1 with the following packages: dada2 v1.34.0, phyloseq v1.50.0, vegan v2.6-8, igraph v2.1.1, glasso v1.11, glmnet v4.1-8, zCompositions v1.5.0-4, and ggplot2 v3.5.1. Primer trimming used cutadapt v5.1.

5. Conclusions

We present a systematic multi-method assessment of keystone candidates in aquaponic microbiomes using 88 amplicon libraries from a multi-compartment research facility. No taxon met a stringent cross-method definition of a classical keystone. Presence-impact analysis highlighted *Cetobacterium* sp. and related ASVs as strong compartment-structuring taxa in fish tanks; network analysis identified environmental biofilm-associated hubs; and LASSO selected no individual turnover predictors.

Across all three methods, the evidence points to distributed ecological resilience rather than single-taxon control. The leave-one-out PERMANOVA approach described here limits circularity in presence-impact screening and should be directly transferable to other amplicon studies. Linking community composition to process rates and testing causal claims will require metagenomics, metatranscriptomics, or experimental manipulation. The most actionable conclusion for practice is that compartment-level monitoring and stable operating conditions provide a more robust basis for aquaponic microbiome management than targeting individual taxa whose functional roles remain unverified.

Ethics approval and consent to participate

All experimental protocols involving fish were approved by the Auburn University Institutional Animal Care and Use Committee (IACUC protocol #22-4099). Passive microbial sampling from system water and grow-bed substrate did not require additional ethical review.

Consent for publication

Not applicable.

Availability of data and materials

Raw 16S rRNA gene amplicon sequence reads (n = 88 samples, 176 paired-end fastq.gz files) are available in the NCBI Sequence Read Archive under BioProject PRJNA1333965 (accessions SRR35620316–SRR35620403; (42)). This dataset was deposited by co-authors Rezaei et al. in September 2025; it constitutes the primary sequence archive for the present manuscript. Analysis scripts, processed data tables (including all environmental integration outputs and Supplementary Tables S1–S5), and the final figure files have been deposited at Zenodo under DOI https://doi.org/10.5281/zenodo.20170975. The Zenodo deposit is released under MIT (code) and CC-BY-4.0 (data) licences and contains a README that documents reproduction of the figures from the SRA reads. Sample metadata is provided in MIxS-compliant format as part of the BioSample submission (PRJNA1333965).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Belmont Forum Formas grant 2022-02559 (AquaponicsOpti; PI: Joyce). The Auburn University component of this work was

additionally supported by the National Science Foundation grant #2147832, the USDA Agricultural Research Service grant #6010-32000-028-000-D, and the Alabama Agriculture Experiment Station (ALA0HIGGINS).

Authors' contributions

VL conceived and designed the analysis, processed the amplicon sequencing data, performed the keystone screening (leave-one-out presence-impact, co-occurrence network centrality, and LASSO temporal analyses), produced the figures, and wrote the original draft. SR operated the aquaponic systems, collected the samples, and carried out nucleic acid extraction and sequencing preparation, and contributed to interpretation of the results. BH provided resources and supervision for the experimental work at Auburn University, contributed to funding acquisition, and critically revised the manuscript. AJ conceived the study, acquired funding, supervised the project, and critically revised the manuscript. All authors read and approved the final manuscript.

References

1. Paine RT. A Note on Trophic Complexity and Community Stability. Am Nat. 1969 Jan;103(929):91–3. doi:10.1086/282586

2. Goddek S, Joyce A, Kotzen B, Burnell GM, editors. Aquaponics Food Production Systems: Combined Aquaculture and Hydroponic Production Technologies for the Future [Internet]. Cham: Springer International Publishing; 2019 [cited 2026 Jul 7]. Available from: https://link.springer.com/10.1007/978-3-030-15943-6 doi:10.1007/978-3-030-15943-6

3. Okomoda VT, Oladimeji SA, Solomon SG, Olufeagba SO, Ogah SI, Ikhwanuddin M. Aquaponics production system: A review of historical perspective, opportunities, and challenges of its adoption. Food Sci Nutr. 2023 Mar;11(3):1157–65. doi:10.1002/fsn3.3154

4. Kasozi N, Abraham B, Kaiser H, Wilhelmi B. The complex microbiome in aquaponics: significance of the bacterial ecosystem. Ann Microbiol. 2021 Dec;71(1):1. doi:10.1186/s13213-020-01613-5

5. Bartelme RP, Oyserman BO, Blom JE, Sepulveda-Villet OJ, Newton RJ. Stripping Away the Soil: Plant Growth Promoting Microbiology Opportunities in Aquaponics. Front Microbiol. 2018 Jan 22;9:8. doi:10.3389/fmicb.2018.00008

6. Bartelme RP, Smith MC, Sepulveda-Villet OJ, Newton RJ. Component Microenvironments and System Biogeography Structure Microorganism Distributions in Recirculating Aquaculture and Aquaponic Systems. Tamaki H, editor. mSphere. 2019 Aug 28;4(4):e00143-19. doi:10.1128/mSphere.00143-19

7. Eck M, Sare AR, Massart S, Schmautz Z, Junge R, Smits THM, et al. Exploring Bacterial Communities in Aquaponic Systems. Water. 2019 Feb 2;11(2):260. doi:10.3390/w11020260

8. Lobanov V, Keesman KJ, Joyce A. Plants Dictate Root Microbial Composition in Hydroponics and Aquaponics. Front Microbiol. 2022;13:848057. doi:10.3389/fmicb.2022.848057

9. Schmautz Z, Espinal CA, Bohny AM, Rezzonico F, Junge R, Frossard E, et al. Environmental parameters and microbial community profiles as indication towards microbial activities and diversity in aquaponic system compartments. BMC Microbiol. 2021 Dec;21(1):12. doi:10.1186/s12866-020-02075-0

10. Banerjee S, Schlaeppi K, Van Der Heijden MGA. Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol. 2018 Sep;16(9):567–76. doi:10.1038/s41579-018-0024-1

11. Munguia-Fragozo P, Alatorre-Jacome O, Rico-Garcia E, Torres-Pacheco I, Cruz-Hernandez A, Ocampo-Velazquez RV, et al. Perspective for Aquaponic Systems: " *Omic* " Technologies for Microbial Community Analysis. BioMed Res Int. 2015;2015:1–10. doi:10.1155/2015/480386

12. Faust K, Raes J. Microbial interactions: from networks to models. Nat Rev Microbiol. 2012 Aug;10(8):538–50. doi:10.1038/nrmicro2832

13. Berry D, Widder S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Front Microbiol. 2014 May 20;5. doi:10.3389/fmicb.2014.00219

14. Herren CM, McMahon KD. Keystone taxa predict compositional change in microbial communities. Environ Microbiol. 2018 Jun;20(6):2207–17. doi:10.1111/1462-2920.14257

15. Amit G, Bashan A. Top-down identification of keystone taxa in the microbiome. Nat Commun. 2023 Jul 4;14(1):3951. doi:10.1038/s41467-023-39459-5

16. Garza DR, Gonze D, Faust K. Keystone concept revisited: insights into microbial community dynamics and control. Nat Rev Microbiol. 2026 May;24(5):359–71. doi:10.1038/s41579-025-01266-8

17. Allison SD, Martiny JBH. Resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci. 2008 Aug 12;105(supplement_1):11512–9. doi:10.1073/pnas.0801925105

18. Ramond P, Galand PE, Logares R. Microbial functional diversity and redundancy: moving forward. FEMS Microbiol Rev. 2025 Jan 14;49:fuae031. doi:10.1093/femsre/fuae031

19. Schimel JP, Schaeffer SM. Microbial control over carbon cycling in soil. Front Microbiol. 2012;3:348. doi:10.3389/fmicb.2012.00348

20. Griffiths BS, Philippot L. Insights into the resistance and resilience of the soil microbial community. FEMS Microbiol Rev. 2013 Mar;37(2):112–29. doi:10.1111/j.1574-6976.2012.00343.x

21. Schmautz Z, Graber A, Jaenicke S, Goesmann A, Junge R, Smits THM. Microbial diversity in different compartments of an aquaponics system. Arch Microbiol. 2017 May;199(4):613–20. doi:10.1007/s00203-016-1334-1

22. Ruiz A, Scicchitano D, Palladino G, Nanetti E, Candela M, Furones D, et al. Microbiome study of a coupled aquaponic system: unveiling the independency of bacterial communities and their beneficial influences among different compartments. Sci Rep. 2023 Nov 11;13(1):19704. doi:10.1038/s41598-023-47081-0

23. Schreiter S, Ding GC, Heuer H, Neumann G, Sandmann M, Grosch R, et al. Effect of the soil type on the microbiome in the rhizosphere of field-grown lettuce. Front Microbiol. 2014 Apr 8;5. doi:10.3389/fmicb.2014.00144

24. Zhalnina K, Louie KB, Hao Z, Mansoori N, Da Rocha UN, Shi S, et al. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. Nat Microbiol. 2018 Mar 19;3(4):470–80. doi:10.1038/s41564-018-0129-3

25. Edmonds JW, Sackett JD, Lomprey H, Hudson HL, Moser DP. The aeroponic rhizosphere microbiome: community dynamics in early succession suggest strong selectional forces. Antonie Van Leeuwenhoek. 2020 Jan;113(1):83–99. doi:10.1007/s10482-019-01319-y

26. Stouvenakers G, Massart S, Depireux P, Jijakli MH. Microbial Origin of Aquaponic Water Suppressiveness against Pythium aphanidermatum Lettuce Root Rot Disease. Microorganisms. 2020 Oct 29;8(11):1683. doi:10.3390/microorganisms8111683

27. Rezaei S, Spoor M, Hohn G, Schorer R, Bender G, Cline D, et al. Enhancing yields: The role of fish tank illumination and coupling/decoupling in biofloc aquaponics performance. Aquaculture. 2026 Feb;613:743286. doi:10.1016/j.aquaculture.2025.743286

28. Oliveira V, Martins P, Marques B, Cleary DFR, Lillebø AI, Calado R. Aquaponics using a fish farm effluent shifts bacterial communities profile in halophytes rhizosphere and endosphere. Sci Rep. 2020 Jun 22;10(1):10023. doi:10.1038/s41598-020-66093-8

29. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 2001 Feb;26(1):32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x

30. Takahashi K, Hisatomi A, Mori K, Yuki M, Noda S, Hongoh Y, et al. Reclassification of the genus Dysgonomonas and description of Dysgonomonas reticulitermitis sp. nov. and Viscerimonas tarda gen. nov., sp. nov. from the gut of the subterranean termite Reticulitermes speratus. Int J Syst Evol Microbiol. 2026 Jan 19;76(1). doi:10.1099/ijsem.0.007031

31. Mergaert J, Cnockaert MC, Swings J. Thermomonas fusca sp. nov. and Thermomonas brevis sp. nov., two mesophilic species isolated from a denitrification reactor with poly(epsilon-caprolactone) plastic granules as fixed bed, and emended description of the genus Thermomonas. Int J Syst Evol Microbiol. 2003;53(6):1961–6. doi:10.1099/ijs.0.02489-0

32. Ju JH, Kim JS, Lee DH, Jeon JH, Heo SY, Seo JW, et al. Thermomonas aquatica sp. nov., isolated from an industrial wastewater treatment plant. Int J Syst Evol Microbiol. 2019 Nov 1;69(11):3399–404. doi:10.1099/ijsem.0.003630

33. Lim JH, Baek SH, Lee ST. Ferruginibacter alkalilentus gen. nov., sp. nov. and Ferruginibacter lapsinanis sp. nov., novel members of the family "Chitinophagaceae" in the phylum Bacteroidetes, isolated from freshwater sediment. Int J Syst Evol Microbiol. 2009 Oct 1;59(10):2394–9. doi:10.1099/ijs.0.009480-0

34. Oh S, Choi D, Cha CJ. Ecological processes underpinning microbial community structure during exposure to subinhibitory level of triclosan. Sci Rep. 2019 Mar 14;9(1):4598. doi:10.1038/s41598-019-40936-5

35. Tarhriz V, Thiel V, Nematzadeh G, Hejazi MA, Imhoff JF, Hejazi MS. Tabrizicola aquatica gen. nov. sp. nov., a novel alphaproteobacterium isolated from Qurugöl Lake nearby Tabriz city, Iran. Antonie Van Leeuwenhoek. 2013 Dec;104(6):1205–15. doi:10.1007/s10482-013-0042-y

36. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. Biostatistics. 2008 Jul 1;9(3):432–41. doi:10.1093/biostatistics/kxm045

37. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33(1). doi:10.18637/jss.v033.i01

38. Shade A, Handelsman J. Beyond the Venn diagram: the hunt for a core microbiome. Environ Microbiol. 2012 Jan;14(1):4–12. doi:10.1111/j.1462-2920.2011.02585.x

39. Duan L, Li JL, Yin LZ, Luo XQ, Ahmad M, Fang BZ, et al. Habitat-dependent prokaryotic microbial community, potential keystone species, and network complexity in a subtropical estuary. Environ Res. 2022 Sep;212:113376. doi:10.1016/j.envres.2022.113376

40. Tsuchiya C, Sakata T, Sugita H. Novel ecological niche of Cetobacterium somerae, an anaerobic bacterium in the intestinal tracts of freshwater fish. Lett Appl Microbiol. 2008;46(1):43–8. doi:10.1111/j.1472-765X.2007.02258.x

41. Wang M, Tang C, Zhang Z, Fan Z, Jiang L, Liu Z, et al. Effect of the gut core microbiota Cetobacterium on the growth, physiology, and nutritional metabolism of Nile tilapia (Oreochromis niloticus). Aquac Rep. 2025 Mar;40:102583. doi:10.1016/j.aqrep.2024.102583

42. Rezaei S, Otto M, Hohn G, Wang Q, Arthur W, Hall MB, et al. Coupling, decoupling, and fish tank illumination shape the functional microbiome and nutrient cycling in aquaponics. Bioresour Technol Rep. 2025 Dec;32:102439. doi:10.1016/j.biteb.2025.102439

43. Qi X, Zhang Y, Zhang Y, Luo F, Song K, Wang G, et al. Vitamin B12 produced by Cetobacterium somerae improves host resistance against pathogen infection through strengthening the interactions within gut microbiota. Microbiome. 2023 Jun 15;11(1):135. doi:10.1186/s40168-023-01574-2

44. Maus I, Tubbesing T, Wibberg D, Heyer R, Hassa J, Tomazetto G, et al. The Role of Petrimonas mucosa ING2-E5AT in Mesophilic Biogas Reactor Systems as Deduced from Multiomics Analyses. Microorganisms. 2020 Dec 17;8(12):2024. doi:10.3390/microorganisms8122024

45. Fanfoni E, Sinisgalli E, Fontana A, Soldano M, Garuti M, Morelli L. Microbial Characterisation of a Two-Stage Anaerobic Digestion Process for Conversion of Agri-Based Feedstock in Biogas and Long-Chain Fatty Acids in a Circular Economy Framework. Fermentation. 2024 May 31;10(6):293. doi:10.3390/fermentation10060293

46. Roberto AA, Van Gray JB, Leff LG. Sediment bacteria in an urban stream: Spatiotemporal patterns in community composition. Water Res. 2018 May;134:353–69. doi:10.1016/j.watres.2018.01.045

47. Zhou X, Liu H, Fan X, Wang X, Bi X, Cheng L, et al. Comparative Analysis of Bacterial Information of Biofilms and Activated Sludge in Full-Scale MBBR-IFAS Systems. Microorganisms. 2024 May 31;12(6):1121. doi:10.3390/microorganisms12061121

48. Bovio-Winkler P, Guerrero LD, Erijman L, Oyarzúa P, Suárez-Ojeda ME, Cabezas A, et al. Genome-centric metagenomic insights into the role of Chloroflexi in anammox, activated sludge and methanogenic reactors. BMC Microbiol. 2023 Feb 21;23(1):45. doi:10.1186/s12866-023-02765-5

49. Petriglieri F, Kondrotaite Z, Singleton C, Nierychlo M, Dueholm MKD, Nielsen PH. A comprehensive overview of the Chloroflexota community in wastewater treatment plants worldwide. McGrath J, editor. mSystems. 2023 Dec 21;8(6):e00667-23. doi:10.1128/msystems.00667-23

50. Conroy K, Poelstra J, Mancl K. Impact of salinity and time on structure and functional potential of wastewater treatment biofilms in intermittent sand bioreactors. J Gen Appl Microbiol. 2024;70(1):n/a. doi:10.2323/jgam.2023.12.003

51. Godzieba M, Zubrowska-Sudol M, Walczak J, Ciesielski S. Development of microbial communities in biofilm and activated sludge in a hybrid reactor. Sci Rep. 2022 Jul 22;12(1):12558. doi:10.1038/s41598-022-16570-z

52. Tarhriz V, Hirose S, Fukushima S ichi, Hejazi MA, Imhoff JF, Thiel V, et al. Emended description of the genus Tabrizicola and the species Tabrizicola aquatica as aerobic anoxygenic phototrophic bacteria. Antonie Van Leeuwenhoek. 2019 Aug;112(8):1169–75. doi:10.1007/s10482-019-01249-9

53. Galachyants AD, Krasnopeev AY, Podlesnaya GV, Potapov SA, Sukhanova EV, Tikhonova IV, et al. Diversity of Aerobic Anoxygenic Phototrophs and Rhodopsin-Containing Bacteria in the Surface Microlayer, Water Column and Epilithic Biofilms of Lake Baikal. Microorganisms. 2021 Apr 14;9(4):842. doi:10.3390/microorganisms9040842

54. Kämpfer P, Müller C, Mau M, Neef A, Auling G, Busse HJ, et al. Description of Pseudaminobacter gen. nov. with two new species, Pseudaminobacter salicylatoxidans

sp. nov. and Pseudaminobacter defluvii sp. nov. Int J Syst Evol Microbiol. 1999 Apr 1;49(2):887–97. doi:10.1099/00207713-49-2-887

55. Hahn YK, Kim MS, Im WT. Pseudaminobacter granuli sp. nov., isolated from granules used in a wastewater treatment plant. J Microbiol. 2017 Aug;55(8):607–11. doi:10.1007/s12275-017-7257-y

56. Chouari R, Le Paslier D, Daegelen P, Ginestet P, Weissenbach J, Sghir A. Molecular Evidence for Novel Planctomycete Diversityin a Municipal Wastewater TreatmentPlant. Appl Environ Microbiol. 2003 Dec;69(12):7354–63. doi:10.1128/AEM.69.12.7354-7363.2003

57. Dedysh SN, Kulichevskaya IS, Beletsky AV, Ivanova AA, Rijpstra WIC, Damsté JSS, et al. Lacipirellula parvula gen. nov., sp. nov., representing a lineage of planctomycetes widespread in low-oxygen habitats, description of the family Lacipirellulaceae fam. nov. and proposal of the orders Pirellulales ord. nov., Gemmatales ord. nov. and Isosphaerales ord. nov. Syst Appl Microbiol. 2020 Jan;43(1):126050. doi:10.1016/j.syapm.2019.126050

58. Speth DR, In 'T Zandt MH, Guerrero-Cruz S, Dutilh BE, Jetten MSM. Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. Nat Commun. 2016 Mar 31;7(1):11172. doi:10.1038/ncomms11172

59. Thomas F, Hehemann JH, Rebuffet E, Czjzek M, Michel G. Environmental and Gut Bacteroidetes: The Food Connection. Front Microbiol. 2011;2. doi:10.3389/fmicb.2011.00093

60. McKee LS, La Rosa SL, Westereng B, Eijsink VG, Pope PB, Larsbrink J. Polysaccharide degradation by the Bacteroidetes: mechanisms and nomenclature. Environ Microbiol Rep. 2021 Oct;13(5):559–81. doi:10.1111/1758-2229.12980

61. Choi H, Roy P, Park H, Yang J, Oh SJ, Choi H, et al. Microbial community dynamics and functional shifts in aquaponic systems: Implications for sustainable aquaculture. Aquac Rep. 2025 Dec;45:103248. doi:10.1016/j.aqrep.2025.103248

62. Deng Y, Jiang YH, Yang Y, He Z, Luo F, Zhou J. Molecular ecological network analyses. BMC Bioinformatics. 2012 Dec;13(1):113. doi:10.1186/1471-2105-13-113

63. Lhoste E, Comte F, Brown K, Delisle A, Jaclin D, Ponsin V, et al. Bacterial, Archaeal, and Eukaryote Diversity in Planktonic and Sessile Communities Inside an Abandoned and Flooded Iron Mine (Quebec, Canada). Appl Microbiol. 2023 Jan 12;3(1):45–63. doi:10.3390/applmicrobiol3010004

64. Sjöberg S, Stairs CW, Allard B, Homa F, Martin T, Sjöberg V, et al. Microbiomes in a manganese oxide producing ecosystem in the Ytterby mine, Sweden: impact on metal mobility. FEMS Microbiol Ecol. 2020 Oct 28;96(11):fiaa169. doi:10.1093/femsec/fiaa169

65. Nierychlo M, Andersen KS, Xu Y, Green N, Jiang C, Albertsen M, et al. MiDAS 3: An ecosystem-specific reference database, taxonomy and knowledge platform for activated

sludge and anaerobic digesters reveals species-level microbiome composition of activated sludge. Water Res. 2020 Sep;182:115955. doi:10.1016/j.watres.2020.115955

66. Wasmund K, Singleton C, Dahl Dueholm MK, Wagner M, Nielsen PH. The predicted secreted proteome of activated sludge microorganisms indicates distinct nutrient niches. Newton RJ, editor. mSystems. 2024 Oct 22;9(10):e00301-24. doi:10.1128/msystems.00301-24

67. Shao YH, Wu YW, Naufal M, Wu JH. Genome-centered metagenomics illuminates adaptations of core members to a partial Nitritation–Anammox bioreactor under periodic microaeration. Front Microbiol. 2023 Jan 26;14:1046769. doi:10.3389/fmicb.2023.1046769

68. Zhou S, Zhu W, He Y, Zhang T, Jiang Z, Zeng M, et al. A comprehensive analysis of microbial community differences in four morphologies of mainstream anaerobic ammonia oxidation systems using big-data mining and machine learning. Front Mar Sci. 2024 Sep 4;11:1458853. doi:10.3389/fmars.2024.1458853

69. Shade A, Peter H, Allison SD, Baho DL, Berga M, Bürgmann H, et al. Fundamentals of Microbial Community Resistance and Resilience. Front Microbiol. 2012;3. doi:10.3389/fmicb.2012.00417

70. Mulder A, Graaf AA, Robertson LA, Kuenen JG. Anaerobic ammonium oxidation discovered in a denitrifying fluidized bed reactor. FEMS Microbiol Ecol. 1995 Mar;16(3):177–84. doi:10.1111/j.1574-6941.1995.tb00281.x

71. Kuypers MMM, Marchant HK, Kartal B. The microbial nitrogen-cycling network. Nat Rev Microbiol. 2018;16(5):263–76. doi:10.1038/nrmicro.2018.9

72. Deng Y, Ruan Y, Ma B, Timmons MB, Lu H, Xu X, et al. Multi-omics analysis reveals niche and fitness differences in typical denitrification microbial aggregations. Environ Int. 2019 Nov;132:105085. doi:10.1016/j.envint.2019.105085

73. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. Front Microbiol. 2017 Nov 15;8:2224. doi:10.3389/fmicb.2017.02224

74. Wang XW, Sun Z, Jia H, Michel-Mata S, Angulo MT, Dai L, et al. Identifying keystone species in microbial communities using deep learning. Nat Ecol Evol. 2023 Nov 16;8(1):22–31. doi:10.1038/s41559-023-02250-2

75. Bauchinger F, Seki D, Berry D. Characteristics of putative keystones in the healthy adult human gut microbiota as determined by correlation network analysis. Front Microbiol. 2024 Nov 20;15:1454634. doi:10.3389/fmicb.2024.1454634

76. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol. 2016 May;18(5):1403–14. doi:10.1111/1462-2920.13023

77. Apprill A, McNally S, Parsons R, Weber L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquat Microb Ecol. 2015 Jun 4;75(2):129–37. doi:10.3354/ame01753

78. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011 May 2;17(1):10. doi:10.14806/ej.17.1.200

79. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016 Jul;13(7):581–3. doi:10.1038/nmeth.3869

80. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2012 Nov 27;41(D1):D590–6. doi:10.1093/nar/gks1219

81. Palarea-Albaladejo J, Martín-Fernández JA. zCompositions — R package for multivariate imputation of left-censored data under a compositional approach. Chemom Intell Lab Syst. 2015 Apr;143:85–96. doi:10.1016/j.chemolab.2015.02.019

82. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. Watson M, editor. PLoS ONE. 2013 Apr 22;8(4):e61217. doi:10.1371/journal.pone.0061217

83. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Stat Methodol. 1995 Jan 1;57(1):289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

84. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. vegan: Community Ecology Package [Internet]. 2001 [cited 2026 Jul 7]. p. 2.7-5. Available from: https://CRAN.R-project.org/package=vegan doi:10.32614/CRAN.package.vegan

85. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal. 2006.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Keystonesupplementaryfigures.docx
- SupplementaryTables.xlsx
- Keystonesupplementaryfigures.docx
- SupplementaryTables.xlsx
- FigS1phototrophlightcoupling.png
- FigS2environmentalpcoatempdo.png
- FigS3upsetmethods.png
- FigS4observedasvsrichness.png