The ISME Journal, 2026, 20(1), wrag099
https://doi.org/10.1093/ismejo/wrag099
Published: 23 April 2026
Original Article

Predicting microbial community responses to disturbance using genome-resolved trait-based life-history strategies

Ezequiel Santillan[1], Soheil A. Neshat[1], Stefan Wuertz[1,2,*]

[1]Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore 637551, Singapore, Singapore
[2]School of Civil and Environmental Engineering, Nanyang Technological University, Singapore 639798, Singapore, Singapore
*Corresponding author. Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, 60 Nanyang Drive, SBS-01N-27, 637551, Singapore.
E-mail: swuertz@ntu.edu.sg

Abstract

Understanding how microbial communities respond to disturbance remains a fundamental question in ecology, with broad implications for biodiversity, ecosystem function, and biotechnology. Trait-based approaches offer general rules to predict community responses by linking ecological strategies to measurable traits. Whereas life-history strategy frameworks such as the competitor–ruderal–stress-tolerant (CSR) model are well established in plant and animal ecology, their application to microbial communities has been limited. Here, we experimentally tested how microbial communities shift across a gradient of disturbance frequency in replicated bioreactors treating synthetic wastewater. We applied six conditions by doubling the organic loading rate at different frequencies, from undisturbed to press disturbance, and monitored changes over 42 days using genome-resolved metagenomics, 16S rRNA gene sequencing, biomass quantification, and effluent chemistry. By integrating ordination, network analysis, and machine learning, we identified emergent community-level life-history strategies, with competitor-dominated communities under undisturbed conditions, ruderal-associated strategies at intermediate disturbance frequencies, and stress-tolerant strategies under sustained high-frequency (press) disturbance. These strategies were reflected in functional trade-offs, shifts in community composition, and genomic trait distributions. A simulation-based approach was used to generate a CSR classification of metagenome-assembled genomes, which was consistent with patterns observed in other microbial ecosystems. Our results demonstrate that life-history frameworks can capture predictable microbial dynamics across disturbance regimes. This approach provides a unifying tool for linking microbial structure, function, and traits across scales, helping to reconcile ecological theory with microbial resource management in natural and engineered ecosystems.

Keywords **microbial communities, life-history strategies, community traits, microbiome predictability, disturbance regimes, metagenome-assembled genomes, trait-based ecology**

Introduction

Microbial communities underpin essential functions in both natural and engineered ecosystems, including global biogeochemical cycles [1] and the operation of biotechnological systems [2, 3]. In wastewater treatment reactors, for example, microbial consortia play critical roles in organic matter degradation [4], nutrient removal [5], bioenergy generation [6], and emerging applications such as resource recovery [7] and microbial protein production [8, 9]. However, the stability and performance of these systems depend on how microbial communities respond to environmental disturbances [10], including toxic shocks [11, 12], changes in substrate availability [13–15], temperature fluctuations, or operational disruptions. Predicting such responses remains a major challenge in microbial ecology [16, 17], particularly as many systems function under non-equilibrium and fluctuating conditions [2, 18].

Trait-based approaches, in particular, provide a way to simplify complex dynamics into interpretable axes of ecological strategy [2, 19–21]. Traits, defined as morphologic, physiologic, genomic or phenotypic attributes that affect growth, reproduction, and survival (i.e. fitness) [22], can be combined across taxa into community-level indicators known as community-aggregated traits (CATs) [19, 23]. These traits offer mechanistic insights into how microbial communities reorganize their functional potential in response to environmental change [19].

The competitor–stress-tolerant–ruderal (CSR) life-history theory [24] classifies organisms based on how they allocate resources to growth, resource acquisition, or stress tolerance under different combinations of competition, disturbance, and stress. Competitor (C) strategies are favored under relatively stable, low-disturbance, and productive (resource-rich) conditions, where competitive ability for shared resources determines community structure; ruderal (R) strategies rapidly exploit disturbed niches; and stress-tolerant (S) strategies persist under sustained resource-limited or otherwise

Received: 13 September 2025. **Revised:** 25 January 2026. **Accepted:** 20 April 2026

constraining conditions. Originally conceptualized for plants, the CSR theory has since been adapted to microbes [21, 25, 26] and applied to both laboratory [19] and field studies, especially in soil communities [27–29], using trait-based ordinations and gene-level indicators.

The potential of CSR to explain microbial dynamics in engineered systems was first tested in lab-scale bioreactors [19], which exposed bacterial communities in activated sludge to a gradient of xenobiotic disturbance (3-chloroaniline) and identified life-history strategies based on CATs derived from functional gene profiles. The authors proposed that microbial responses in biotechnological systems could be simplified into ecologically meaningful categories that reflect performance trade-offs. Yet, disturbance is inherently multidimensional [30], and testing other types of disturbance is needed to assess generalizability. Grime's original CSR framework defines disturbance primarily in terms of biomass destruction or mortality at the level of individual plants [24], whereas subsequent ecological theory has broadened this concept to include temporal events that disrupt community structure by altering resource availability or environmental conditions [2, 30]. In microbial systems, such events need not involve direct biomass removal; instead, repeated or sustained shifts in resource supply, oxygen demand, and physicochemical constraints can impose selective pressures functionally analogous to disturbance or stress [13, 19]. In this context, variation in disturbance regimes can be interpreted as a continuum that modulates both the timing and persistence of selective pressures acting on microbial communities.

Advances in metagenomics enable trait inference at a higher resolution. While earlier work relied primarily on short-read functional annotations [19], the analysis of metagenome-assembled genomes (MAGs) can provide a more precise characterization of genotypic traits across microbial populations [31, 32]. MAGs capture features such as genome size and gene content, revealing the presence of genes involved in growth, stress-response, and cellular maintenance. These can be interpreted as microbial analogs of classical ecological traits including growth potential, competitive investment, and stress tolerance. Genomic content reflects functional potential at the level of individual cells or populations [25]. Because organisms face trade-offs in allocating limited resources among growth, maintenance, stress tolerance, and resource acquisition, consistent selective pressures are expected to favor the enrichment of specific gene categories under given disturbance regimes, linking genomic traits to life-history strategies as predicted by CSR theory. MAG-based analyses provide a complementary view to amplicon-based community structure or short-read functional profiles by integrating taxonomy, ecological strategy, and encoded potential functions. This can enhance the applicability of CSR and other trait-based frameworks in complex and dynamic microbial systems [33, 34]. Further, the recently developed MicroEcoTools package [35] facilitates statistical evaluation of trait-based strategies using metagenomic and functional datasets, promoting reproducibility and accessibility.

In this study, we apply and extend the CSR life-history framework to microbial communities in activated sludge bioreactors subjected to a disturbance gradient created by doubling the organic load at varied frequencies. This disturbance regime influences microbial dynamics by altering competition for oxygen, substrate availability, and spatial niches within the biofilm [13, 34]. Here, disturbance is defined as a temporal event that disrupts community structure by altering resource availability or environmental conditions [2, 30]. In this system, increased organic loading elevates oxygen demand and intensifies competition for electron acceptors and space, potentially generating localized stress conditions such as transient anoxia while simultaneously creating resource surpluses. Building on prior work [19], we increased experimental replication (n = 5), applied a different disturbance type, and incorporated both 16S rRNA gene amplicon sequencing and genome-resolved metagenomics to track community structure, functional potential, and trait dynamics. Specifically, to address the multidimensional nature of disturbance [2], we tested whether CSR-based hypotheses previously developed using a chemical stressor (3-chloroaniline) [19] would hold under a distinct, operationally relevant disturbance type based on increased organic loading [13, 34], while increasing replication (n = 5) to account for inherent community variability and strengthen statistical inference [2, 36].

We hypothesized that extremes of the disturbance frequency gradient would favor distinct community-level life-history strategies, with undisturbed conditions favoring competitor strategies, sustained high-frequency organic loading (press disturbance) favoring stress-tolerant strategies, and intermediate disturbance frequencies promoting ruderal-type strategies. These shifts were expected to manifest as differences in taxonomic composition, functional trait distributions, and community-level trade-offs across the disturbance gradient. Using ordination, network, and machine learning analyses of CATs derived from MAGs and 16S rRNA gene amplicon data, we tested whether distinct life-history strategies emerge along this gradient and assessed the extent to which the CSR framework provides predictive insights into microbial community responses under disturbance, with relevance for both applied and natural microbiome systems.

Materials and methods

Experimental design and functional assessments

This study builds upon an experiment described previously [34], which reported detailed experimental procedures and baseline analyses. Here, we focus on additional analyses and datasets not included in that publication. Briefly, we conducted a 42-day experiment using thirty microcosm-scale sequencing batch bioreactors (25 mL working volume), each inoculated with activated sludge from a full-scale municipal wastewater treatment plant in Singapore. Bioreactors were maintained at 30°C and operated in daily cycles. A synthetic medium was supplied following a fixed feeding regime with double organic loading applied at six disturbance frequencies: never (undisturbed), every 8, 6, 4, or 2 days (intermediate disturbance), and daily (sustained high-strength loading; hereafter referred to as press disturbance), each in five independent replicates (n = 5). These were labeled levels 0 to 5, corresponding to increasing disturbance frequency, with calculated frequencies of 0, 1/8, 1/6, 1/4, 1/2, and 1, respectively.

System performance was monitored weekly by measuring soluble chemical oxygen demand (COD) and total Kjeldahl nitrogen (TKN) removal at the end of a cycle using spectrophotometric and ion chromatographic methods [37]. COD removal reflects the microbial community's capacity for organic carbon degradation, a core ecosystem function in wastewater treatment systems, while TKN removal integrates multiple nitrogen transformation processes (ammonification, nitrification, and assimilation), providing an indicator of functional stability and metabolic coordination under disturbance. Settling capacity was assessed via sludge volume index (SVI; mL/g) after 30 min of settling. SVI is a key operational metric linking microbial

community structure to process performance, as poor settleability leads to biomass washout and reduced solids retention, whereas low SVI values indicate well-flocculated, stable communities. For correlation analyses, we used the additive inverse (−SVI) so that higher values reflect better settleability, as lower SVI values correspond to improved settling performance. Feed concentrations in mixed liquor after feeding were ∼306 mg COD/L and ∼46 mg TKN/L under regular conditions and ∼595 mg COD/L and ∼46 mg TKN/L under double loading. These loading conditions defined the disturbance gradient by periodically increasing substrate availability without proportionally increasing nitrogen, thereby altering resource stoichiometry, oxygen demand, and competitive interactions among microbial populations. Across all treatments, soluble substrates were largely consumed within each cycle, as indicated by consistently high COD and TKN removal efficiencies (Table S1), allowing disturbance frequency to be interpreted primarily as modulating the timing and persistence of resource availability rather than causing enduring substrate accumulation. Biomass levels were controlled using a food-to-biomass (F:M) ratio strategy as previously described [13], where biomass (sludge) was measured weekly as total suspended solids (TSS) and wasted as needed to maintain a TSS concentration of 1500 mg/L. Maintaining comparable biomass concentrations across reactors minimized confounding effects of biomass quantity on effluent chemistry and allowed observed differences in community structure, trait composition, and function to be attributed primarily to disturbance frequency rather than differences in standing biomass. The corresponding solids residence times (SRTs) ranged from 30 to 15 days as disturbance frequency increased, remaining well above the minimum doubling times reported for relevant activated sludge bacteria [38]. SRT integrates biomass retention and growth dynamics and can act as an ecological filter on microbial life-history strategies; however, in this study, SRT variation emerged as a secondary consequence of the imposed organic loading disturbance rather than as an independent selective pressure. All SRT values were sufficiently long to retain both fast- and slow-growing activated sludge populations, such that the imposed disturbance regime did not directly constrain population persistence through insufficient biomass retention but instead operated through transient increases in organic loading.

Sludge inoculum and reactor setup

The inoculum was sourced from a nitrifying-denitrifying activated sludge process with ∼200 000 m^3/d flow and ∼80%–90% nutrient removal efficiency. On the day of collection, influent COD and TKN concentrations averaged 221 mg/L and 45 mg/L, respectively. Activated sludge was transferred to the lab in sterile conditions, homogenized, and distributed into 30 sterile 50-mL tubes acting as sequencing batch reactors (SBRs). Each reactor was fed 12.5 mL of either regular or high-strength synthetic wastewater, depending on the disturbance schedule. A daily cycle involved 30 min of settling, decanting 12.5 mL of effluent, and replacing it with fresh medium. This setup yielded a hydraulic residence time (HRT) of 48 h. HRT defines the average time that soluble substrates and metabolites remain in the reactor and therefore constrains nutrient transformation efficiency. By keeping HRT constant across all treatments, we ensured that differences in community structure, functional performance, and trait distributions were driven by disturbance frequency in organic loading, rather than by hydraulic washout or systematic differences in substrate contact time.

Synthetic wastewater composition

The synthetic wastewater composition followed previous formulations [13], including complex organics (yeast extract, peptones), simple carbon sources (acetate, dextrose), inorganic nitrogen sources (urea, ammonium salts), buffering agents, and trace elements. For double organic loading events, COD inputs were doubled primarily through the organic components, whereas inorganic nitrogen was adjusted to maintain TKN levels. Phosphate was added to target an N:P ratio of ∼6, consistent with typical activated sludge operating conditions [38] and chosen to minimize the likelihood of imposing deliberate nitrogen or phosphorus limitation. Under this design, organic carbon availability constituted the primary axis of disturbance, while inorganic nutrients remained non-limiting across treatments. Media were prepared in large batches, sterile-filtered (0.2 μm), and stored at 4°C.

DNA extraction and 16S rRNA gene amplicon sequencing

Sludge samples were collected weekly from all reactors (n = 180), plus inoculum controls (n = 4), and stored at −80°C. Genomic DNA was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals) with previously described protocol modifications to improve yield [19]. DNA was quantified using NanoDrop and Qubit instruments and further purified before sequencing.

The 16S rRNA gene V3–V4 regions were amplified using the 341f/785r primer set in a two-step PCR protocol [13, 39]. Dual-indexed libraries were generated, pooled, and sequenced using the MiSeq System (Illumina) with 300 bp paired-end reads. Amplicon sequence variants (ASVs) were inferred using the DADA2 pipeline [40], with quality filtering, error correction, merging, and chimera removal. Taxonomy was assigned with the SILVA database (v.138) [41]. Rarefaction was performed to 5089 reads per sample to normalize sequencing depth for downstream diversity and community analyses.

Metagenomics bioinformatics

Metagenomic analyses followed a previously described pipeline [32], optimized for recovering and characterizing non-redundant, medium- to high-quality MAGs. Raw reads were trimmed using Trimmomatic [42] to remove adapters and low-quality bases ($Q < 30$), with FastQC [43] used for quality assessment. Reads from replicate and time-series samples were co-assembled using SPAdes [44] in metagenomics mode with multiple k-mer sizes. Coverage profiles were generated by mapping reads to contigs using BBMap [45], and genome binning was performed with MetaBAT2 [46]. Resulting bins were dereplicated using dRep [47], selecting species-level MAGs with ≥50% completeness and ≤10% contamination, assessed using CheckM [48].

Taxonomic classification was carried out with GTDB-Tk [49], and open reading frames (ORFs) were annotated with Prokka [50] and functionally classified via EggNOG-mapper. Ribosomal and transfer RNA genes were annotated using Barrnap [51] and tRNAscan-SE [52], respectively. MAG relative abundances were estimated with CoverM [53], using genome-based quantification with minimum coverage thresholds. This pipeline enabled robust recovery of MAGs and consistent functional annotation, supporting downstream trait-based ecological analyses.

Community structure, genotypic traits, and CSR analysis

Community function, structure, and genotype data were analyzed using the CSR framework as outlined in [19]. Community structure was assessed at the genus level using both metagenomics and 16S rRNA gene amplicon sequencing data, combining ordination methods and multivariate tests in PRIMER (v.7) [54]. Square-root transformed, normalized genus abundances were used to reduce the influence of dominant taxa. To evaluate links between community structure and function across disturbance levels, constrained ordination via canonical analysis of principal coordinates (CAP) was conducted, incorporating Pearson's correlation vectors of normalized functional metrics. Community differences by disturbance level were tested using PERMANOVA on Bray–Curtis dissimilarity matrices [55], with factors treated as fixed. Multivariate dispersion homogeneity was assessed using PERMDISP [56], and all P values were calculated with 9999 permutations.

Relationships between genotype-level traits and community structural patterns were explored using distance-based linear modeling (DistLM) applied to MAG data, and distance-based redundancy analysis (dbRDA) was performed using COG trait-complexes as predictor variables [57]. Pearson's correlation vectors ($r > 0.20$) were overlaid on the dbRDA plots. CSR life-history strategy assignments were derived from both taxonomic and genotypic data using the *CSR_assign()* function from the MicroEcoTools package [35]. Briefly, this function integrates community-level taxonomic or functional trait profiles with replicated experimental designs to statistically evaluate differences across disturbance treatments. The approach identifies traits or trait groups that are significantly enriched or depleted across treatments and maps these patterns onto the CSR life-history framework to infer community-level strategy assignments. Composition-aware network analysis of MAG genotypic profiles was performed using Gephi (v.0.9.2), based on SparCC correlation matrices generated with the *sparcc* function from the SpiecEasi R package (v.1.1.0) [58], following [11]. This analysis was performed to complement ordination and trait-based approaches by examining how disturbance frequency influences patterns of co-variation and community-level organization across the disturbance gradient, with reduced sensitivity to abundance-driven biases. Weak correlations ($r_{adj} < 0.2$) were removed. Node clusters were defined by Louvain modularity class, with clusters colored by their association with undisturbed or disturbed conditions. Networks were visualized using the Fruchterman-Reingold layout; node size was scaled by degree, and edge thickness by correlation strength.

Results

Disturbance shapes bacterial succession, community structure, and functional trade-offs

Bacterial communities differentiated in terms of β-diversity across disturbance levels and with time, as revealed by community analysis through 16S rRNA gene amplicon sequencing. There was a temporal separation of communities at the undisturbed level 0, the intermediately disturbed levels 1–4, and the press-disturbed level 5 as shown by unconstrained ordination (Fig. 1a). This pattern was further supported by constrained ordination at day 42, which emphasized differences across disturbance treatments (Fig. 1b). Multivariate tests yielded significant results for disturbance levels (PERMANOVA $P < .001$), with no significant effect of heteroscedasticity (PERMDISP $P = .86$). Variations in β-diversity corresponded with trade-offs in community-level function, as shown by Pearson's correlation vectors (Fig. 1b), which represent ecosystem function parameters that significantly differed across reactors exposed to different disturbance levels (Table S1). In terms of nitrogen, undisturbed reactors (L0) exhibited the highest effluent nitrate (NO_3^--N) concentrations, while press-disturbed reactors (L5) showed the poorest total Kjeldahl nitrogen (TKN) removal. Intermediately disturbed reactors (L1–L4) achieved the most effective TKN removal. Sludge settleability was also highest at intermediate disturbance levels, as indicated by the additive inverse sludge volume index (-SVI). Carbon removal remained consistently high across all reactors (98%–99%) and did not differ significantly among disturbance levels. The observed patterns in community diversity and function across disturbance levels were consistent with the a priori hypotheses derived from the CSR framework (Fig. 1c).

Genome resolved insights into community clustering and CSR life-history strategies

To gain finer resolution of bacterial dynamics under disturbance, we reconstructed 133 medium- and high-quality MAGs from the day-42 metagenomes (Table S2), following the Minimum Information about Metagenome-Assembled Genomes (MIMAG) framework [59]. Of these, 57 MAGs were classified as having substantial (90%–99.9%) or perfect (100%) completeness, and 105 MAGs had low (5%–9.9%) or no (0%) contamination, ensuring high confidence in downstream trait and taxonomic analyses. Summary data and CSR classifications for all MAGs are available in Supplementary Table S3. Canonical analysis of principal coordinates (CAP), based on Bray-Curtis dissimilarities of MAG relative abundances derived from coverage, revealed distinct clustering patterns corresponding to specific disturbance levels (Fig. 2a), with significant group separation (PERMANOVA $P < .001$) and no dispersion bias (PERMDISP $P = .10$). A SparCC correlation network of MAGs revealed co-occurrence patterns structured by disturbance regime, with node modularity classes aligning with CSR life-history strategy assignments. These modular associations indicate that MAGs sharing similar CSR classifications tended to co-vary across treatments, consistent with the disturbance-driven shifts in community composition and diversity observed at the ASV level (Fig. 1b). Although MAGs represent a defined subset of the microbial community, the observed patterns align with and complement and reinforce the broader trends inferred from 16S rRNA gene data.

We further applied the MicroEcoTools *CSR_assign* function to classify MAGs into CSR life-history strategies (Fig. 3). Of the 133 medium- and high-quality MAGs, 97 received unique assignments: 18 as competitors (C), 28 as ruderals (R), and 29 as stress-tolerants (S), with the remainder showing overlapping or unassigned classifications (Fig. 3a; see Supplementary Table S3 for MAG-level assignments and statistical outputs). Consistent patterns of relative abundance across the disturbance gradient were observed among representative MAGs within each category, with competitor MAGs dominating undisturbed conditions, ruderal MAGs enriched at intermediate disturbance frequencies, and stress-tolerant MAGs prevailing under sustained high-frequency disturbance (Fig. 3b). Genome size distributions generally aligned with CSR expectations: competitors had the largest and least variable genome sizes, ruderals had smaller genomes, and stress-tolerant MAGs spanned a wider range. However, these differences were not statistically significant (Welch's ANOVA $P = .35$; Fig. 3c).

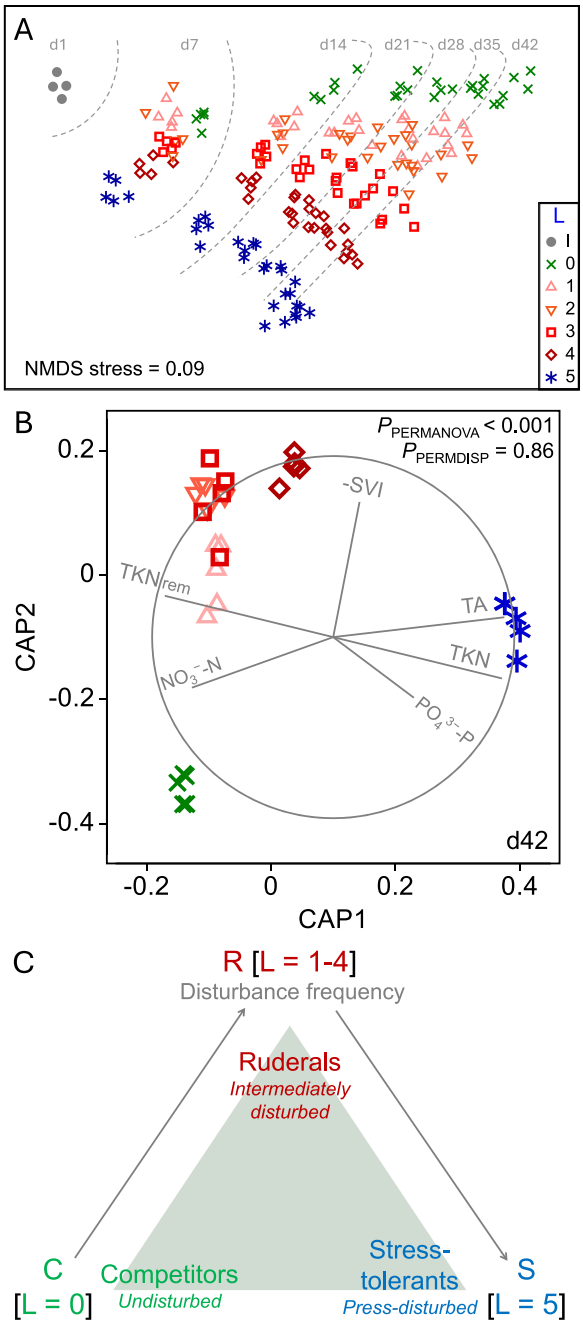

Figure 1 Succession of bacterial communities and functional trade-offs across a disturbance gradient. (a) Non-metric multidimensional scaling (NMDS) of 16S rRNA gene amplicon sequence variant (ASV) data across all time points. Bray-Curtis dissimilarities were calculated from square-root transformed abundances. Symbols represent disturbance levels (L0–L5, n = 5 each), and gray dots represent the full-scale plant inoculum (I, n = 4). Dashed lines trace community trajectories from day 1 to day 42. (b) Canonical analysis of principal coordinates (CAP) at day 42. Same symbols as panel (a). Vectors indicate Pearson correlations for significantly different (Table S1) effluent and performance parameters: Total Kjeldahl nitrogen (TKN), nitrate (NO_3^--N), phosphate (PO_4^{3-}-P), total alkalinity (TA), TKN removal (TKN_{rem}), and additive inverse sludge volume index (–SVI). PERMANOVA and PERMDISP *P*-values indicate significant group separation without dispersion differences. (c) Conceptual CSR triangle for d42 bacterial communities showing the dominance of competitor (C), ruderal (R), and stress-tolerant (S) strategies under undisturbed (L0), intermediately disturbed (L1–L4), and press-disturbed (L5) conditions, respectively.

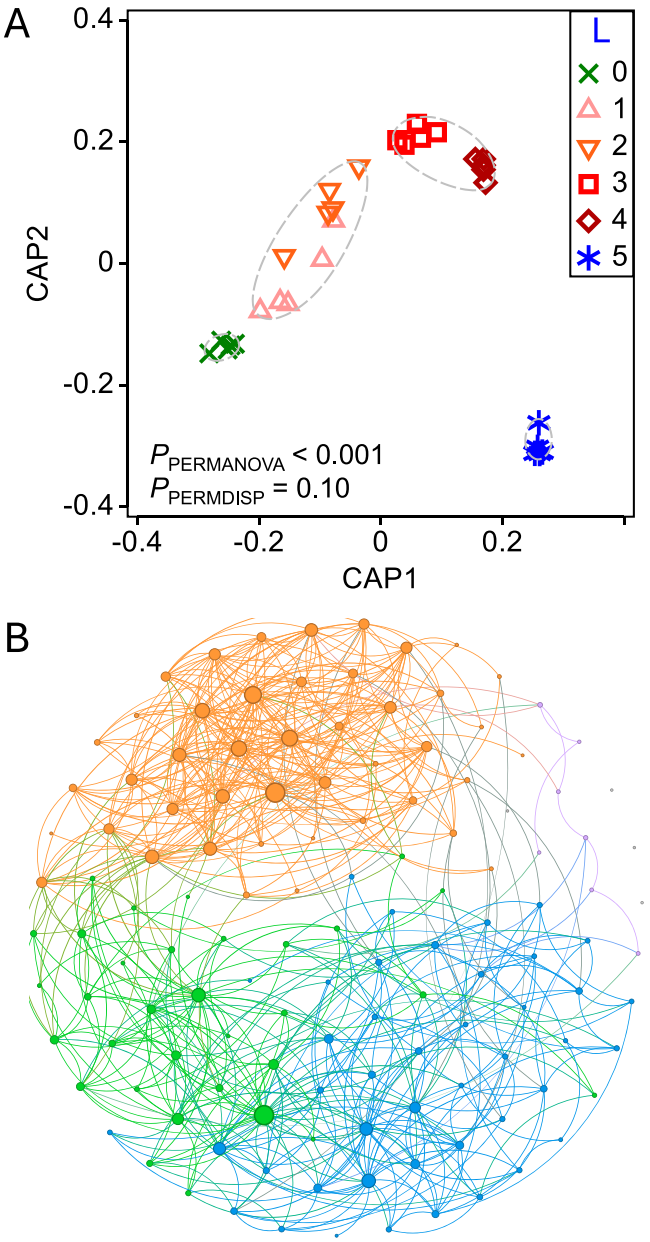

Figure 2 Clustering of bacterial community structure on day 42 based on metagenome-assembled genomes (MAGs). (a) Canonical analysis of principal coordinates (CAP) based on Bray-Curtis dissimilarities of square-root transformed relative abundances from 133 medium- and high-quality MAGs. Symbols represent disturbance levels (L0–L5, n = 5). Dashed ellipses represent 85% similarity thresholds from group-average clustering. PERMANOVA and PERMDISP *P*-values indicate significant group separation without dispersion bias. (b) SparCC-based correlation network showing strong positive associations among MAGs ($r_{adj} \geq 0.20$). Edge thickness indicates correlation strength, and node size reflects degree. Nodes are grouped by modularity class, corresponding to dominance in undisturbed (L0), intermediately disturbed (L1–L4), or press-disturbed (L5, blue) reactors. These groups form distinct clusters in the network layout (bottom left, top, and bottom right, respectively) and are additionally distinguished by colour (green, orange, and blue, respectively). This analysis represents a MAG-defined subset of the microbial community; broader patterns across all taxa are shown in Fig. 1 using 16S rRNA gene ASV data.

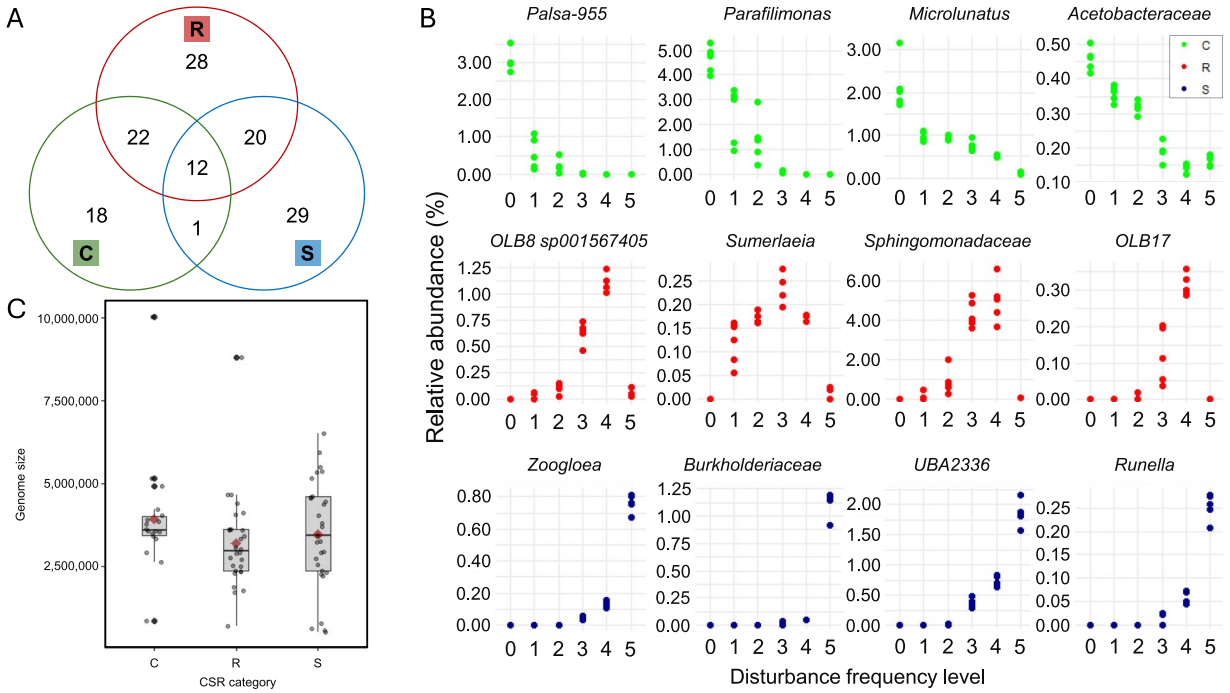

Figure 3 CSR classification of metagenome-assembled genomes (MAGs) and genomic features using the *CSR_assign* function in MicroEcoTools [35]. (a) Venn diagram of CSR assignments among 133 medium- and high-quality MAGs, indicating overlaps and unique classifications. CSR assignments and statistical outputs for all individual MAGs are reported in Supplementary Table S3. (b) Top four MAGs assigned to each life-history strategy, with competitors (C) in the top row, ruderals (R) in the middle row, and stress-tolerants (S) in the bottom row, based on their relative abundance across disturbance frequency levels. These groups are additionally distinguished by colour in the figure (as indicated in the legend). (c) Distribution of estimated genome sizes (in base pairs) for MAGs assigned to CSR categories. Diamond symbols display mean values (Welch's ANOVA $P = .35$). The box bounds the interquartile range (IQR) divided by the median, and Tukey-style whiskers extend to a maximum of 1.5 times the IQR beyond the box.

Linking microbial genotypic traits to CSR strategies through trait-based analysis

Trait-based analysis of MAGs revealed distinct genotypic signatures associated with different life-history strategies along the disturbance gradient (Fig. 4a). By aggregating Clusters of Orthologous Groups (COG) functional categories across the 133 high- and medium-quality MAGs, we observed clear patterns of enrichment and reduction that aligned with competitor (C), ruderal (R), and stress-tolerant (S) strategies. For example, traits linked to biosynthetic and metabolic efficiency such as secondary structure [Q] and lipid metabolism [I] were enriched under undisturbed (L0) conditions, consistent with a competitor strategy favoring stable environments. In contrast, ruderal-associated traits like transcription [K] and inorganic ion transport [P] peaked at intermediate disturbance levels (L1–L4), where faster metabolic activation and turnover would be advantageous. Stress-tolerant traits, including nucleotide metabolism [F], DNA repair [L], and signal transduction [T], were enriched under press-disturbance (L5), reflecting an adaptation to persist in resource-limited or inhibitory environments.

To enable interpretable model-based inference of how trait composition relates to community structure, we applied distance-based linear modeling (DistLM), a machine-learning approach that fits regression models to multivariate resemblance data and performs stepwise model selection and predictor ranking based on the corrected Akaike Information Criterion (AICc). From a total of 19 normalized COG trait categories considered, DistLM identified a 10-variable model that best explained variation in the MAG-based Bray-Curtis community dissimilarities (AICc = 109.15), accounting for 97.2% of the fitted variation. The resulting distance-based redundancy analysis (dbRDA) showed that the first two ordination axes explained 70.8% and 16.6% of the total variation, respectively (Fig. 4b). Predictor vectors such as defense mechanisms [V], transcription [K], and translation, ribosomal structure and biogenesis [J] showed strong correlations with the ordination axes and aligned with separation of communities by disturbance level, indicating that variation in community composition across the disturbance gradient was associated with shifts in trait composition consistent with CSR framework transitions.

Similarly to the MAG-level classification, we applied the MicroEcoTools *CSR_assign* function to assign COG trait categories to CSR strategies based on their enrichment patterns across the disturbance gradient (Table 1). Traits exclusive to each CSR category, such as lipid metabolism [I] for competitors, transcription [K] for ruderals, and DNA repair [L] for stress-tolerants, highlighted distinct ecological trade-offs. Multifunctional traits, including defense mechanisms [V] and cell wall biogenesis [M], appeared across categories (e.g., CS or CR), suggesting context-dependent roles that span resource competition, rapid growth, and stress endurance. These findings support the interpretation that microbial communities can be mapped onto predictable life-history strategies based on their aggregated genotypic potential.

Discussion

This study shows that microbial communities exposed to varied frequencies of organic loading disturbance in bioreactor systems develop distinct configurations of taxonomic composition, functional attributes, and genotypic traits, shaped by both disturbance frequency and temporal dynamics. These patterns align with classical

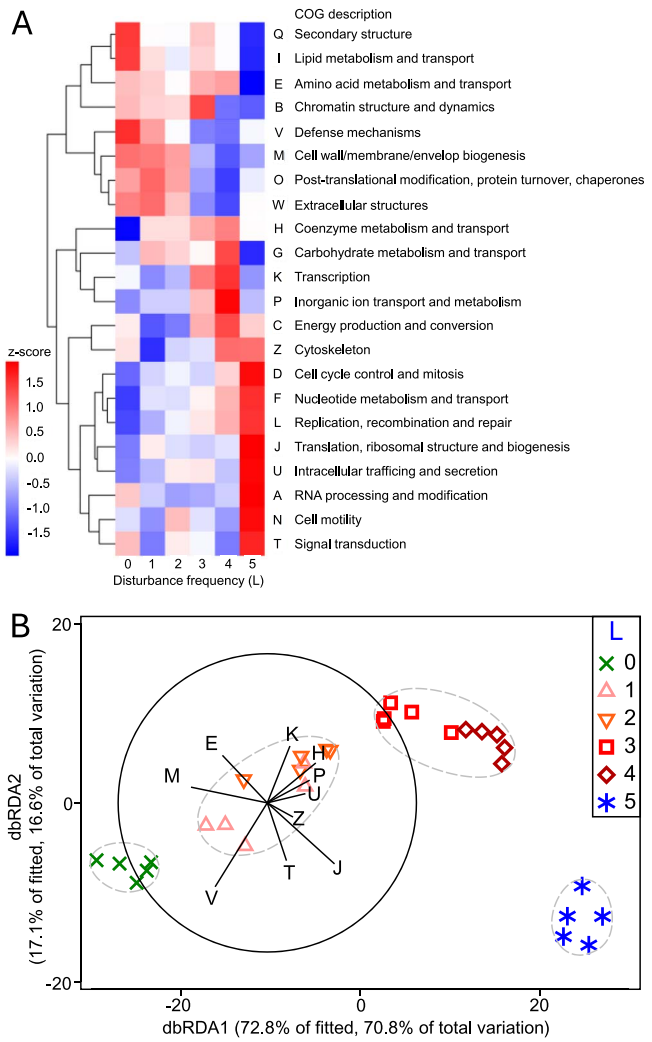

Figure 4 Disturbance frequency shapes bacterial genotypic trait composition across CSR strategies. (a) Z-score normalized abundances of clusters of orthologous groups (COG) trait categories across disturbance levels (L0 to L5), derived from 133 medium- and high-quality metagenome-assembled genomes (MAGs) recovered on day 42. COG categories are hierarchically clustered (average linkage). Values indicate relative enrichment (positive values) and reduction (negative values). Patterns reflect ecological selection of traits associated with competitor (C), ruderal (R), and stress-tolerant (S) strategies. (b) Distance-based redundancy analysis (dbRDA) illustrates how variation in COG trait composition explains community structure across disturbance levels. Bray-Curtis dissimilarities were calculated from square-root transformed MAG abundance data. Predictor selection was performed using distance-based linear modeling (DistLM) with stepwise variable selection and AICc minimization, a machine-learning approach that identifies the most explanatory trait complexes. Predictor vectors represent Pearson correlations of model-selected COG categories. PERMANOVA ($P < .001$) and PERMDISP ($P = .10$) P-values indicate significant group separation without dispersion bias. Symbols denote disturbance levels (L0 to L5; n = 5 per group), with 85% similarity ellipses based on group-average clustering.

Table 1 CSR classification of genotypic traits using the *CSR_assign* function in MicroEcoTools [35], based on 133 medium- and high-quality MAGs from day-42 bioreactor metagenomes.

CSR_assign[a]	Community Aggregated Trait (CAT)[b]	COG group
C	Secondary Structure [Q]	Metabolism
C	Lipid metabolism and transport [I]	Metabolism
CS	Defense mechanisms [V]	Cellular processes and signaling
CR	Amino Acid metabolism and transport [E]	Metabolism
CR	Cell wall/membrane/envelop biogenesis [M]	Cellular processes and signaling
R	Transcription [K]	Information storage and processing
R	Inorganic ion transport and metabolism [P]	Metabolism
S	Nucleotide metabolism and transport [F]	Metabolism
S	Cell cycle control and mitosis [D]	Cellular processes and signaling
S	Replication, recombination and repair [L]	Information storage and processing
S	Signal Transduction [T]	Cellular processes and signaling
S	Translation, ribosomal structure and biogenesis [J]	Information storage and processing

[a]CSR framework: Competitors (C) excel in resource acquisition in stable niches. Ruderals (R) grow rapidly in disturbed environments but are less resource-efficient. Stress-tolerants (S) maintain metabolic function under harsh or resource-poor conditions. [b]Community aggregated traits (CATs) are based on clusters of orthologous genes (COG) trait complexes, with COG category codes shown in bracket.

CSR life-history strategies and build on earlier work demonstrating disturbance-driven organization of microbial communities in activated sludge systems under 3-CA disturbance [19]. Here, we further show that CSR strategies can be resolved under a different disturbance regime and link these dynamics to genome-resolved trait distributions. In contrast to our earlier microcosm study [34], which focused on diversity and assembly patterns across varied disturbance frequencies in the form of doubling organic loading, here we provide a trait-based and genome-resolved ecological interpretation of the same experiment grounded in life-history strategies. Under this framework, undisturbed conditions are associated with competitive dominance, intermediate disturbance frequencies favor ruderal-associated strategies, and sustained high-frequency organic loading selects for stress-tolerant strategies, consistent with a press disturbance that imposes persistent energetic and environmental stress [60]. By integrating taxonomic profiles, ecosystem function metrics, and CATs derived from both amplicon and genome-resolved metagenomic data, we show that disturbance frequency is associated with consistent and statistically supported differentiation of community-level life-history strategies. Permutation-based multivariate analyses of community composition, replicated-design statistical testing of functional traits, and genome-resolved CSR life-history strategy assignments collectively indicate that competitor-associated communities dominate under undisturbed conditions, ruderal-associated communities emerge at intermediate disturbance frequencies, and stress-tolerant strategies prevail under sustained high-frequency disturbance. Together, these results indicate coherent disturbance-driven organization of microbial communities across beta-diversity structure, functional trade-offs, and trait distributions.

Interpreting CSR strategies from MAG-derived CATs requires considering that genomic trait content reflects potential energetic investment and allocation capacity [25]. Under this premise, differential enrichment of functional genes across the disturbance gradient represents

an expected outcome of selection acting on alternative life-history strategies shaped by fundamental trade-offs in resource allocation. Consistent with prior work, we treat genome-resolved community-aggregated traits as proxies for dominant life-history strategies emerging from the combined effects of taxon identity, relative abundance, and trait distributions across populations [19, 27]. Although MAGs do not capture all functional potential present in the community, they provide organism-resolved insights that complement broader taxonomic or gene-centric approaches. Analyses of genes related to biosynthesis, resource use, and stress response supported CSR assignments, reinforcing the ecological interpretation. These results underscore the value of genome-resolved trait inference for predictive microbial ecology [32] and provide a convergent application of CSR life-history strategies in engineered microbiomes, aligning with similar trait-based interpretations in soil ecosystems [27, 28]. Under this interpretation, CSR patterns reflect both the average genomic strategies of dominant taxa and emergent community-level properties arising from trade-offs among taxa, rather than strict within-organism allocation constraints. These interpretations are not based solely on inferred genomic potential, but are supported by directly measured ecosystem functions, including carbon and nitrogen removal performance, which provide empirical, community-level responses aligned with the observed trait and strategy patterns. While genomic proxies cannot directly capture realized allocation without transcriptomic or proteomic validation [29], convergent evidence from theoretical work, soil systems, and engineered microbiomes supports their utility for detecting reproducible life-history trade-offs at the community scale [19–21, 27]. Recognition of these scale-dependent interpretations strengthens the conceptual bridge between classical CSR theory and microbial trait ecology.

Among traits assigned to a single CSR category, patterns generally matched expected ecological trade-offs. Competitor (C) traits such as secondary structure [Q] and lipid metabolism [I] were enriched in undisturbed L0 reactors, where nitrogen removal was highest. These traits likely support efficient resource acquisition and biosynthesis in stable environments, consistent with prior findings from a study that used 3-chloroaniline (3-CA) as a xenobiotic disturbance in lab-scale activated sludge bioreactors [19]. That study, which applied the CSR framework using CATs from read-based functional gene profiles, also observed biosynthetic trait enrichment under undisturbed conditions. However, L0 reactors in the present study showed poorer settling capacity, suggesting that C traits alone do not guarantee optimal system function. At intermediate disturbance levels (L1 to L4), ruderal (R) traits such as transcription [K] and inorganic ion transport [P] were enriched, coinciding with enhanced TKN removal and settling capacity. These traits likely support fast metabolic responses suited to fluctuating conditions, and their presence aligns with trait combinations observed in the intermediate regimes of both the 3-CA study and recent CSR applications in soils [27], reinforcing the cross-system relevance of this ecological framework. Stress-tolerant (S) traits including nucleotide metabolism [F], cell cycle control [D], DNA repair [L], signal transduction [T], and translation [J] were enriched in L5 reactors, where nitrogen removal was lowest, consistent with survival and maintenance under prolonged stress. This aligns with the broader view that S traits emphasize repair, persistence, and stability under constraint [29, 61].

Distance-based Linear Modeling (DistLM), a machine learning approach, added predictive power and interpretability to our trait-based findings. By identifying a minimal yet optimal set of genotypic traits that best explained variation in MAG-defined community composition, the model validated CSR assignments through independent variable selection. The model-selected traits, such as defense mechanisms [V], transcription [K], and translation [J], contributed strongly to ordination axes and mirrored CSR-aligned clustering patterns. Machine-learning approaches such as DistLM enhance trait-based analyses by identifying parsimonious combinations of genotypic traits that explain observed variation in community structure across disturbance regimes. While the resulting models are constrained to the experimental gradient studied here, they provide a quantitative basis for generating testable predictions within this domain and motivate future work aimed at broader predictive generalization using larger and independent datasets. When integrated with hypothesis-driven experimentation and ecological theory, machine learning provides a framework for generating testable predictions and actionable strategies, thereby advancing predictive microbial ecology in both natural and engineered systems [2, 62–64].

Traits assigned to multiple CSR categories showed interpretable but more variable patterns. CR traits like amino acid metabolism [E] and cell wall biogenesis [M] were enriched at intermediate disturbance, suggesting roles in both growth and structural resilience. This is consistent with increased biosynthesis and membrane-related traits reported under moderate disturbance in other systems, including soils [27]. Defense mechanisms [V], classified as CS, were moderately enriched across levels, reflecting functions relevant to both stress buffering and resource competition. Translation [J] appeared in both L0 and L1 to L4 reactors, similar to the 3-CA study where ribosomal traits were shared by competitor and ruderal communities [19]. Some traits diverged from expectations. Lipid metabolism [I], although classified as a C trait, may also contribute to stress tolerance through membrane stabilization, illustrating potential trait multifunctionality. Signal transduction [T], enriched under press disturbance, may mediate rapid environmental sensing, supporting roles in both R and S contexts. Amino acid metabolism [E] also appeared in L0, reinforcing its competitor function noted in prior studies. These results indicate that intermediate disturbance levels do not represent a single ecological state, but rather a continuum in which different disturbance frequencies favor distinct combinations of mixed CSR strategies. These mixed patterns align with findings from soil systems where traits do not always map cleanly to CSR categories and can shift functionally with context [27, 65]. This highlights the importance of refining trait classification empirically, particularly in complex and engineered environments.

Although genome size did not differ significantly across CSR categories, the observed trends align with ecological expectations. MAGs classified as competitors had the largest average genome sizes and the narrowest spread, consistent with metabolic versatility and regulatory capacity suited for stable environments where resource efficiency is favored [65]. Ruderal MAGs showed the smallest genomes with constrained variability, which may reflect selection for rapid replication and low biosynthetic burden under fluctuating disturbance. Stress-tolerant MAGs displayed the broadest genome size distribution, potentially capturing a mix of streamlining for persistence and larger genomes carrying diverse stress-response and repair functions. This heterogeneity matches prior findings that different ecological strategies can be associated with both genome expansion and reduction, depending on selective pressures [61]. Although genome size alone does not determine strategy, its distribution may reflect underlying trade-offs in growth, regulation, and survival, supporting its

complementary role in trait-based CSR classification. However, genome size estimates derived from MAGs have inherent limitations. Even MAGs reported as 100% complete by tools such as CheckM may lack substantial genomic regions, because completeness is assessed based on the presence of marker genes [66]. Assuming the bias in genome size estimation is randomly distributed across all MAGs, comparative trends can still offer meaningful ecological insights.

Even though the focus of this study is on traits, several MAG-assigned taxa align with their CSR classifications based on known biology. Among competitors (C), *Microlunatus* are polyphosphate-accumulating organisms known to store internal phosphorus reserves, supporting stable nutrient removal under undisturbed conditions [67]. Members of *Acetobacteraceae* oxidize low concentrations of simple carbon compounds and can contribute to stable biofilm communities, favoring resource-efficient niches. *Parafilimonas*, part of *Chitinophagaceae*, are associated with degradation of complex polysaccharides like chitin and cellulose, indicating adaptation to slow but sustained substrate availability. Although little is known about *Palsa-955*, its consistent detection in undisturbed conditions suggests traits supporting persistence in stable environments. Among ruderals (R), *Sphingomonadaceae* are well known for degrading a wide range of xenobiotic compounds and for their metabolic flexibility, allowing fast adaptation to fluctuating substrates [68]. *Sumerlaeia* has been associated with nitrogen cycling in Antarctic soil ecosystems where microbiota responded rapidly to augmented nutrient regimes [69]. *OLB8* and *OLB17*, although poorly characterized [70], have been previously enriched in reactors under varied disturbance regimes [19], suggesting rapid growth and flexible substrate uptake as ruderal traits. Stress-tolerant (S) taxa include *Zoogloea*, which form protective extracellular matrices, are motile, and engage in aerobic denitrification, traits that could enhance survival in stressed environments [71]. Members of *Burkholderiaceae* are often capable of surviving in oligotrophic conditions and possess stress-response pathways such as DNA repair and oxidative stress resistance. *UBA2336* and *Runella*, though less studied, belong to lineages that are commonly found in activated sludge and have been linked to biofilm formation and resilience under harsh conditions [72]. These assignments support the use of MAG-level CSR classifications to interpret microbial succession and functional adaptation in bioreactors across disturbance gradients.

Trait-based CSR interpretation is gaining traction in microbial ecology, with studies in soil microbiomes showing trade-offs that parallel Grime's original framework [20, 27, 73, 74]. Several features of microbes support this analogy, including limited mobility, long-term persistence under adverse conditions, and the prevalence of microbial seed banks, which render microbial life strategies in many respects more comparable to plants than animals [25]. Prior work has shown that interpreting microbial functional traits within the CSR framework captures trade-offs among taxa that rapidly exploit resource availability, re-establish following frequent disturbance, or persist under sustained stress through maintenance- and dormancy-associated traits [20]. At the same time, we recognize that microbial evolution, horizontal gene transfer, and metabolic interactions can relax assumptions of fixed lineage-level trade-offs. Accordingly, CSR is applied here as a flexible framework for organizing observed trait-disturbance patterns at the community scale, rather than as a deterministic classification of individual organisms. Together with prior work [19], the present study extends this approach to engineered systems such as activated sludge, demonstrating its relevance for bioreactors where microbial communities are managed for nutrient removal or resource recovery. By reducing microbial complexity into three ecologically meaningful axes, CSR offers a practical framework for integrating structure, function, and traits [2], whereas tools like MicroEcoTools streamline this process using reproducible metrics [35]. Other microbial trait frameworks, such as the Yield, Acquisition, Stress tolerance (Y-A-S) model, emphasize related dimensions of microbial strategy including carbon-use efficiency and substrate uptake and have been applied in soil microbiome studies [75, 76]. Despite conceptual overlaps, these frameworks are best viewed as complementary approaches that highlight different aspects of microbial life-history variation [29].

Further work is needed to evaluate the broader applicability of CSR as a tool for microbial resource management, particularly in systems relevant to biotechnology such as anaerobic digestion [77, 78], and at different operational and spatial scales, including systems with varying degrees of environmental control and bacterial immigration [2]. Controlled comparisons across systems will help assess the relative strengths and limitations of CSR and other frameworks. Our findings support CSR as a practical, interpretable approach for engineered microbial communities, with strong potential for broader integration alongside emerging ecological and predictive modeling tools.

Conclusions

This study reinforces the utility of classical life-history theory as a unifying framework for interpreting microbial community dynamics across environmental contexts. By integrating taxonomic, functional, and genome-resolved trait data across a disturbance gradient, we show that strategy differentiation among competitors, ruderals, and stress-tolerants can be detected and meaningfully interpreted. The use of MAG-based community-aggregated traits improves trait resolution for assembled taxa, offering deeper ecological insights, although some relevant functions within the community may be missed due to assembly limitations. Machine learning approaches, such as DistLM, enhance this framework by identifying trait combinations that predictably explain community structure, contributing to the development of interpretable, data-driven models of microbial responses to environmental change. Open-source tools like MicroEcoTools enable scalable and reproducible classification of microbial communities, supporting wider application of trait-based approaches. The convergence of results from independent studies in activated sludge systems, each employing different disturbance types, further supports the generality of this framework. As microbial ecology increasingly aims to generate predictive, theory-grounded insights, such integrative approaches provide a bridge between classical ecological models and contemporary microbial genomics.

Acknowledgments

This research was supported by the Singapore National Research Foundation and Ministry of Education under the Research Centre of Excellence Program. We thank JYJ Tan, JQ Teo, A Latiff, SS Thi, AFBM Batcha, CK Aw, ABA Aziz, and JHJ Lim for their assistance with laboratory work. We thank LCW Liew for supporting the collection of sludge inoculum.

Author contributions

ES and SW conceived the idea. ES designed and conducted the experiment, and performed data processing and analysis. SAN carried out

the metagenomic bioinformatics. SW secured funding for the study. ES drafted the manuscript, and all authors contributed to its revision and editing.

Supplementary material

Supplementary material is available at *The ISME Journal* online.

Conflicts of interest

The authors declare no competing interests.

Funding

None declared.

Data availability

DNA sequencing data are available in the NCBI BioProject database under accession number PRJNA723443. A set of dereplicated MAGs at the genus level is available via GenBank under BioProject accession number PRJNA108977245. Additional files, including strain-level dereplicated MAGs (FASTA), functional annotations, relative abundances across samples, and phylogenetic tree files, are available through Zenodo (DOI: https://doi.org/10.5281/zenodo.8405311), as described in [32]. Summary statistics, taxonomic classifications, quality metrics, and MicroEcoTools *CSR_assign* classifications for the MAGs are available online in Supplementary Table S3.

References

1. Falkowski PG, Fenchel T, Delong EF. The microbial engines that drive earth's biogeochemical cycles. *Science.* 2008;**320**:1034–9. https://doi.org/10.1126/science.1153213
2. Santillan E, Neshat SA, Wuertz S. Disturbance and stability dynamics in microbial communities for environmental biotechnology applications. *Curr Opin Biotechnol* 2025;**93**:103304. https://doi.org/10.1016/j.copbio.2025.103304
3. Timmis K, Ramos JL, Verstraete W. Microbial biotechnology to assure national security of supplies of essential resources: energy, food and water, medical reagents, waste disposal and a circular economy. *Microb Biotechnol* 2022;**15**:1021–5. https://doi.org/10.1111/1751-7915.14049
4. Zhang X, Fan Y, Hao T., *et al.* Insights into current bio-processes and future perspectives of carbon-neutral treatment of industrial organic wastewater: a critical review. *Environ Res* 2024;**241**:117630. https://doi.org/10.1016/j.envres.2023.117630
5. Verstraete W, Strubbe L, Pikaar I., *et al.* Escaping historical lock-in—redesigning wastewater treatment plants and their microbiomes for the 21st century. *Environ Sci Technol* 2025;**59**:14862–9. https://doi.org/10.1021/acs.est.5c06208
6. Kieft B, Finke N, McLaughlin RJ., *et al.* Genome-resolved correlation mapping links microbial community structure to metabolic interactions driving methane production from wastewater. *Nat Commun* 2023;**14**:5380. https://doi.org/10.1038/s41467-023-40907-5
7. Durkin A, Guo M, Wuertz S., *et al.* Resource recovery from food-processing wastewaters in a circular economy: a methodology for the future. *Curr Opin Biotechnol* 2022;**76**:102735. https://doi.org/10.1016/j.copbio.2022.102735
8. Vethathirri RS, Santillan E, Thi SS., *et al.* Microbial community-based production of single cell protein from soybean-processing wastewater of variable chemical composition. *Sci Total Environ* 2023;**873**:162241. https://doi.org/10.1016/j.scitotenv.2023.162241
9. Vethathirri RS, Santillan E, Woo Y., *et al.* Towards sustainable aquafeeds: safe and consistent microbial protein grown on food-processing wastewater. *Cleaner Circular Bioeconomy* 2025;**10**:100139. https://doi.org/10.1016/j.clcb.2025.100139
10. Shade A, Peter H, Allison SD., *et al.* Fundamentals of microbial community resistance and resilience. *Review Front Microbiol* 2012;**3**:417. https://doi.org/10.3389/fmicb.2012.00417
11. Santillan E, Seshan H, Wuertz S. Press xenobiotic 3-chloroaniline disturbance favors deterministic assembly with a shift in function and structure of bacterial communities in sludge bioreactors. *ACS EST Water* 2021;**1**:1429–37. https://doi.org/10.1021/acsestwater.1c00018
12. Seshan H, Santillan E, Constancias F., *et al.* Metagenomics and metatranscriptomics suggest pathways of 3-chloroaniline degradation in wastewater reactors. *Sci Total Environ* 2023;**903**:166066. https://doi.org/10.1016/j.scitotenv.2023.166066
13. Santillan E, Phua WX, Constancias F., *et al.* Sustained organic loading disturbance favors nitrite accumulation in bioreactors with variable resistance, recovery and resilience of nitrification and nitrifiers. *Sci Rep* 2020;**10**:21388. https://doi.org/10.1038/s41598-020-78314-1
14. Santillan E, Constancias F, Wuertz S. Press disturbance alters community structure and assembly mechanisms of bacterial taxa and functional genes in mesocosm-scale bioreactors. *mSystems.* 2020;**5**:e00471–20. https://doi.org/10.1128/mSystems.00471-20
15. Jia C, Li J, Li Z., *et al.* Influence of high-load shocks on achieving mainstream partial nitrification: microbial community succession. *Water Res X* 2025;**27**:100304. https://doi.org/10.1016/j.wroa.2025.100304
16. van den Berg NI, Machado D, Santos S., *et al.* Ecological modelling approaches for predicting emergent properties in microbial communities. *Nat Ecol Evol* 2022;**6**:855–65. https://doi.org/10.1038/s41559-022-01746-7
17. Cairns J, Jokela R, Becks L., *et al.* Repeatable ecological dynamics govern the response of experimental communities to antibiotic pulse perturbation. *Nat Ecol Evol* 2020;**4**:1385–94. https://doi.org/10.1038/s41559-020-1272-9
18. Widder S, Allen RJ, Pfeiffer T., *et al.* Challenges in microbial ecology: building predictive understanding of community function and dynamics. *Mini Review ISME J* 2016;**10**:2557–68. https://doi.org/10.1038/ismej.2016.45
19. Santillan E, Seshan H, Constancias F., *et al.* Trait-based life-history strategies explain succession scenario for complex bacterial communities under varying disturbance. *Res Article Environ Microbiol* 2019;**21**:3751–64. https://doi.org/10.1111/1462-2920.14725
20. Ho A, Di Lonardo DP, Bodelier PLE. Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiol Ecol* 2017;**93**:fix006. https://doi.org/10.1093/femsec/fix006
21. Krause S, Le Roux X, Niklaus PA., *et al.* Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front Microbiol* 2014;**5**:251. https://doi.org/10.3389/fmicb.2014.00251
22. Violle C, Navas M-L, Vile D., *et al.* Let the concept of trait be functional! *Oikos.* 2007;**116**:882–92. https://doi.org/10.1111/j.0030-1299.2007.15559.x
23. Fierer N, Barberan A, Laughlin DC. Seeing the forest for the genes: using metagenomics to infer the aggregated traits of microbial communities. *Front Microbiol* 2014;**5**:614. https://doi.org/10.3389/fmicb.2014.00614
24. Grime JP. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolution-

ary theory. *Am Nat* 1977;**111**:1169–94. https://doi.org/10.1086/283244
25. Ho A, Kerckhof FM, Luke C., *et al.* Conceptualizing functional traits and ecological characteristics of methane-oxidizing bacteria as life strategies. *Environ Microbiol Rep* 2013;**5**:335–45. https://doi.org/10.1111/j.1758-2229.2012.00370.x
26. Couso LL, Soler-Bistué A, Aptekmann AA., *et al.* Ecology theory disentangles microbial dichotomies. *Environ Microbiol* 2023;**25**:3052–63. https://doi.org/10.1111/1462-2920.16495
27. Piton G, Allison SD, Bahram M., *et al.* Life history strategies of soil bacterial communities across global terrestrial biomes. *Nat Microbiol* 2023;**8**:2093–102. https://doi.org/10.1038/s41564-023-01465-0
28. Sharma U, Rawat D, Mukherjee P., *et al.* Ecological life strategies of microbes in response to antibiotics as a driving factor in soils. *Sci Total Environ* 2023;**854**:158791. https://doi.org/10.1016/j.scitotenv.2022.158791
29. Treseder KK. Ecological strategies of microbes: thinking outside the triangle. *J Ecol* 2023;**111**:1832–43. https://doi.org/10.1111/1365-2745.14115
30. Graham EB, Averill C, Bond-Lamberty B., *et al.* Toward a generalizable framework of disturbance ecology through crowdsourced science. *Rev Front Ecol Evol* 2021;**9**:588940. https://doi.org/10.3389/fevo.2021.588940
31. Singleton CM, Petriglieri F, Kristensen JM., *et al.* Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing. *Nat Commun* 2021;**12**:2009. https://doi.org/10.1038/s41467-021-22203-2
32. Neshat SA, Santillan E, Seshan H., *et al.* Non-redundant metagenome-assembled genomes of activated sludge reactors at different disturbances and scales. *Sci Data* 2024;**11**:855. https://doi.org/10.1038/s41597-024-03601-9
33. Santillan E, Seshan H, Constancias F., *et al.* Frequency of disturbance alters diversity, function, and underlying assembly mechanisms of complex bacterial communities. *NPJ Biofilms Microbiomes* 2019;**5**:8. https://doi.org/10.1038/s41522-019-0079-4
34. Santillan E, Wuertz S. Microbiome assembly predictably shapes diversity across a range of disturbance frequencies in experimental microcosms. *NPJ Biofilms Microbiomes* 2022;**8**:41. https://doi.org/10.1038/s41522-022-00301-3
35. Neshat SA, Santillan E, Wuertz S. MicroEcoTools: an R package for comprehensive theoretical microbial ecology analysis. *Methods Ecol Evol* 2025;**16**:1369–77. https://doi.org/10.1111/2041-210X.70047
36. Prosser JI, Bohannan BJM, Curtis TP., *et al.* The role of ecological theory in microbial ecology. *Nat Rev Microbiol* 2007;**5**:384–92. https://doi.org/10.1038/nrmicro1643
37. Rice EW, Baird RB, Eaton AD. *Standard Methods for the Examination of Water and Wastewater*, 23rd edn. Washington, DC, USA: APHA, AWWA, WEF, 2017.
38. Tchobanoglous G, Stensel HD, Tsuchihashi R., *et al. Wastewater Engineering: Treatment and Resource Recovery*, 5th edn. New York, NY, USA: McGraw Hill Education, 2013.
39. Thijs S, Op De Beeck M, Beckers B., *et al.* Comparative evaluation of four bacteria-specific primer pairs for 16S rRNA gene surveys. *Front Microbiol* 2017;**8**:494. https://doi.org/10.3389/fmicb.2017.00494
40. Callahan BJ, McMurdie PJ, Rosen MJ., *et al.* DADA2: high resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;**13**:581–3. https://doi.org/10.1038/nmeth.3869
41. Glöckner FO, Yilmaz P, Quast C., *et al.* 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J Biotechnol* 2017;**261**:169–76. https://doi.org/10.1016/j.jbiotec.2017.06.1198
42. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;**30**:2114–20. https://doi.org/10.1093/bioinformatics/btu170
43. Andrews S. *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute: Cambridge, UK. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
44. Bankevich A, Nurk S, Antipov D., *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;**19**:455–77. https://doi.org/10.1089/cmb.2012.0021
45. Bushnell B. *BBMap: A fast, accurate, splice-aware aligner*. Lawrence Berkeley National Laboratory, U.S. Department of Energy Joint Genome Institute: Walnut Creek, CA, USA, 2014. Available at: https://escholarship.org/uc/item/1h3515gn.
46. Kang DD, Li F, Kirton E., *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ.* 2019;**7**:e7359. https://doi.org/10.7717/peerj.7359
47. Olm MR, Brown CT, Brooks B., *et al.* dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 2017;**11**:2864–8. https://doi.org/10.1038/ismej.2017.126
48. Parks DH, Imelfort M, Skennerton CT., *et al.* CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;**25**:1043–55. https://doi.org/10.1101/gr.186072.114
49. Chaumeil P-A, Mussig AJ, Hugenholtz P., *et al.* GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics.* 2020;**36**:1925–7. https://doi.org/10.1093/bioinformatics/btz848
50. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;**30**:2068–9. https://doi.org/10.1093/bioinformatics/btu153
51. Seemann T. *Barrnap: Basic rapid ribosomal RNA predictor*. University of Melbourne: Melbourne, VIC, Australia. Available at: https://github.com/tseemann/barrnap.
52. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997;**25**:955–64. https://doi.org/10.1093/nar/25.5.955
53. Woodcroft BJ. *CoverM*. https://github.com/wwood/CoverM.
54. Clarke KR, Gorley RN. PRIMER v7: user manual/tutorial. PRIMER-E Ltd: Plymouth, UK, 2015.
55. Anderson MJ, Walsh DCI. PERMANOVA, ANOSIM, and the mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing? *Ecol Monogr* 2013;**83**:557–74. https://doi.org/10.1890/12-2010.1
56. Anderson MJ. Distance-based tests for homogeneity of multivariate dispersions. *Biometrics.* 2006;**62**:245–53. https://doi.org/10.1111/j.1541-0420.2005.00440.x
57. McArdle BH, Anderson MJ. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology.* 2001;**82**:290–7. https://doi.org/10.1890/0012-9658(2001)082[0290:fmmtcd]2.0.co;2
58. Kurtz ZD, Müller CL, Miraldi ER., *et al.* Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comp Biol* 2015;**11**:e1004226. https://doi.org/10.1371/journal.pcbi.1004226
59. Bowers RM, Kyrpides NC, Stepanauskas R., *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 2017;**35**:725–31. https://doi.org/10.1038/nbt.3893
60. Borics G, Várbíró G, Padisák J. Disturbance and stress: different meanings in ecological dynamics? *Hydrobiologia.* 2013;**711**:1–7. https://doi.org/10.1007/s10750-013-1478-9

61. Westoby M, Nielsen DA, Gillings MR., *et al.* Strategic traits of bacteria and archaea vary widely within substrate-use groups. *FEMS Microbiol Ecol* 2021;**97**:fiab142. https://doi.org/10.1093/femsec/fiab142
62. McElhinney JMWR, Catacutan MK, Mawart A., *et al.* Interfacing machine learning and microbial omics: a promising means to address environmental challenges. *Mini Rev Front Microbiol* 2022;**13**:851450. https://doi.org/10.3389/fmicb.2022.851450
63. Delogu F, Kunath BJ, Queirós PM., *et al.* Forecasting the dynamics of a complex microbial community using integrated meta-omics. *Nat Ecol Evol* 2024;**8**:32–44. https://doi.org/10.1038/s41559-023-02241-3
64. Huang R, Xu H, Santillan E., *et al.* Predicting single-cell protein production from food-processing wastewater in sequencing batch reactors using ensemble learning. *Bioresour Technol* 2025;**430**:132561. https://doi.org/10.1016/j.biortech.2025.132561
65. Beier S, Werner J, Bouvier T., *et al.* Trait-trait relationships and tradeoffs vary with genome size in prokaryotes. Hypothesis and theory. *Front Microbiol* 2022;**13**:985216. https://doi.org/10.3389/fmicb.2022.985216
66. Mirete S, Sánchez-Costa M, Díaz-Rullo J., *et al.* Metagenome-assembled genomes (MAGs): advances, challenges, and ecological insights. *Microorganisms.* 2025;**13**:985. https://doi.org/10.3390/microorganisms13050985
67. Kawakoshi A, Nakazawa H, Fukada J., *et al.* Deciphering the genome of polyphosphate accumulating actinobacterium Microlunatus phosphovorus. *DNA Res* 2012;**19**:383–94. https://doi.org/10.1093/dnares/dss020
68. Stolz A. Molecular characteristics of xenobiotic-degrading sphingomonads. *Appl Microbiol Biotechnol* 2009;**81**:793–811. https://doi.org/10.1007/s00253-008-1752-3
69. Nair GR, Kooverjee BB, de Scally S., *et al.* Changes in nutrient availability substantially alter bacteria and extracellular enzymatic activities in Antarctic soils. *FEMS Microbiol Ecol* 2024;**100**:fiae071. https://doi.org/10.1093/femsec/fiae071
70. Kondrotaite Z, Valk LC, Petriglieri F., *et al.* Diversity and ecophysiology of the genus OLB8 and other abundant uncultured Saprospiraceae genera in global wastewater treatment systems. *Front Microbiol* 2022;**13**:917553. https://doi.org/10.3389/fmicb.2022.917553
71. Dugan PR, Stoner DL, Pickrum HM. The genus zoogloea. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E (eds). *The Prokaryotes*. New York, NY, USA: Springer, 2006, pp 960–70.
72. Trujillo ME, Dedysh S, DeVos P., *et al.* Runella. In: Whitman WB (ed). *Bergey's Manual of Systematics of Archaea and Bacteria.* Hoboken, NJ, USA: Wiley, 2015.
73. Chagnon PL, Bradley RL, Maherali H., *et al.* A trait-based framework to understand life history of mycorrhizal fungi. *Trends Plant Sci* 2013;**18**:484–91. https://doi.org/10.1016/j.tplants.2013.05.001
74. Wood JL, Franks AE. Understanding microbiomes through trait-based ecology. *Microbiology Australia* 2018;**39**:53–6. https://doi.org/10.1071/MA18014
75. Malik AA, Martiny JBH, Brodie EL., *et al.* Defining trait-based microbial strategies with consequences for soil carbon cycling under climate change. *ISME J* 2019;**14**:1–9. https://doi.org/10.1038/s41396-019-0510-0
76. Li C, Liao H, Xu L., *et al.* The adjustment of life history strategies drives the ecological adaptations of soil microbiota to aridity. *Mol Ecol* 2022;**31**:2920–34. https://doi.org/10.1111/mec.16445
77. Krohn C, Khudur L, Dias DA., *et al.* The role of microbial ecology in improving the performance of anaerobic digestion of sewage sludge. *Review Front Microbiol* 2022;**13**:1079136. https://doi.org/10.3389/fmicb.2022.1079136
78. Neshat SA, Santillan E, Wuertz S. Uncovering microbial life-history strategies under disturbance: a trait-based computational analysis of anaerobic systems. *bioRxiv,* 2025, preprint 2025.09.18.676720. https://doi.org/10.1101/2025.09.18.676720

The ISME Journal, 2026, **20(1)**, wrag099
https://doi.org/10.1093/ismejo/wrag099
Original Article