Life Sciences and Environmental Research

LSER
LIFE SCIENCES AND ENVIRONMENTAL RESEARCH

Site: https://lserjournal.com
Email: editor@lserjournal.com
ISSN:
3104-3968 (Print)
3104-3976 (Online)

MICROBIAL ECOLOGY AND GENOMICS IN BIOGEOCHEMICAL CYCLING

Muhammad Asadullah Usman[1*], **Zia Ur Rehman**[2]

[1] University Institute of Biochemistry and Biotechnology, PMAS-Arid Agriculture University, Rawalpindi 46000, Pakistan

[2] Institute of Biological Sciences, Gomal University, Dera Ismail Khan 29050, Khyber Pakhtunkhwa, Pakistan

*Corresponding Author E-mail: asadsarfraz420@gmail.com

Received: September 08, 2023 --- Revised: October 10, 2023 Accepted: November 15, 2023

Abstract: Microbial communities are key drivers of biogeochemical cycling since they affect the transformation and movements of such essential elements like carbon, nitrogen, and sulfur. The role of microbes in a range of ecosystem of study ranged with wetlands, forest soils, marine sediments, and permafrost terrains were explored in research studies with view of a mixed- methods experimental strategy that incorporated metagenomics, isotopic labeling, as well as environmental chemistry in the study. A total of nine datasets were generated and these contained distinct functional and geographical signatures of microbial activity. Though the highest production of methane was observed in marine and wetland samples, revealing the existence of active methanogenic niches, the rates of nitrate reduction differed distinctly according to the environments studied, with the highest rates observed in forest soils and estuary-based sampling sites. The number of active genes such as dsrA and mcrA was positively related to the sulfate respiration that had a moderate consistency among samples. Isotopic analyses that involved the use of ^13C and ^15N tracers allowed accurate quantification of rates as they helped in validation of active elemental conversions. Log10-transformed abundance data on functional genes were normalized with respect to process rates that were measured, and they revealed that the abundance was significantly correlated with those measured processes and that it was a strong predictor of metabolic potentials. Visual analysis including hybrid plots or multidimensional scatter projections depicts that coupled or decoupled elemental pathways mediated by microbial guilds depend on the type of the ecosystem and environmental aspects. The reproducibility of such patterns was found in terms of temporal replication, which further confirmed the reliability of this methodology. Importantly, the microbial responses to anthropogenic stress varied, which indicates that nutrient loading and land use altered the structure and the functionality of communities. The paper reaches a conclusion that along with biogeochemical observations, microbial genomics provides a solid premise to understand as well as model the active participation of microorganisms in the saga of global nutrient cycles. This integrative approach can be very helpful in monitoring the microbial resources, in controlling the ecosystems, and in climate modeling.

Keywords: Microbial Ecology, Biogeochemical Cycling, Metagenomics, Nitrate Reduction, Methane Production, Functional Genes

INTRODUCTION

Metagenomics has transformed our understanding of the roles that the microbial communities serve in various ecosystems, and it has been possible because of the study of these microbes in their natural environment (Reynoso-García et al., 2022). Such studies lead to positive insights into the diversity and functions of microbes, which ends up providing new applications in various fields, such as agriculture, human health, and the food industry (Nam et al., 2023). Metagenomic studies specifically have been critical to study the complex microbial community, specifically in the marine environment, because it enables the study of non-culturable microorganisms (Santiago et al., 2023). These culture-independent methods, including bulk metagenomics and single amplified genomes, can be used to examine the microbial diversity and functions in more detail by bypassing the disadvantages of cultivation-based approaches (Sysoev et al., 2021). These methods can also be employed in testing microbiological safety as well as authenticity of food through identification of the bacteria and tracking of antibiotic resistance genes (Sabater et al., 2021). High-throughput sequencing technology, which rapidly develops, has exponentially increased the balance of microbial genomic data (Devika et al., 2023). Consequently, advanced technologies and strategies of studying microbial genomes and their contribution to the biogeochemical variation at one-organism and community scale have become simpler to construct (Zhou et al., 2022). The advancements allow scientists to study molecular basis of complex microbiological phenomena and the interactions that exist between organisms and the environment in vivo (Abdullah et al., 2023). Genome-resolved metagenomics has become very popular in different fields to study the diversity, metabolic potential, and spatiotemporal dynamics of microbes (McDaniel et al., 2021). Functional metagenomics and sequence-based metagenomics are two culture-independent methods that are grown in popularity (Sysoev et al., 2021). The whole-metagenome shotgun sequencing type of sequencing has enabled a species-level resolution of microbial functions (Mi et al., 2024). Microbiomes are masses of bacteria that interact with both biotic and abiotic factors of their environment in accordance with the conditions that enable the community to thrive (Lee et al., 2020). It is possible to get a comprehensive understanding of how and what microorganisms do in the ecosystems through the incorporation of multi-omics methods such as transcriptomics, proteomics, metabolomics, and genomics (Krassowski et al., 2020). These methods may address many modern clinical and diagnostic challenges, such as identifying species, species phylogeny, species strain typing, antibiotic resistance statistics, and epidemiological studies (Salem-Bango et al., 2023). The development of multi-omics technologies and the reduced price of DNA sequencing have resulted in biological breakthroughs to date and enhanced our knowledge of living things (Bodein et al., 2021). Omics tools provide an in-depth insight into molecular, physiological, and biochemical aspects of plant-microbe interaction in different conditions (Jain et al., 2024). The gains have transformed the study of biological mechanisms entirely as it becomes feasible to yield high-throughput biomolecular data (Procopio & Bonicelli, 2024). The integration of multiple types of data in omics is superior in comparison to single omics data, predicting the results more accurately due to the holistic picture of a biological process (Wang et al., 2021). With the help of the so-called omics technologies, including metabolomics, researchers can gain a detailed understanding of the metabolome in a biological sample, including the general composition of metabolites present in any microbial community (Sarkar et al., 2023). Any multi-omics method, including integrated omics, pan-omics, and trans-omics, according to Krassowski et al. (2020), integrates two or more datasets to facilitate the processing, visualization, and interpretation of data with an aim of determining the process involved in the biological system. Snyder and Babu, 2023. Multi-omics technique may lead to a better understanding of how biological processes work, how molecules interact, and how cells can die or live as the techniques have the ability to screen molecular markers and identify regulatory units on numerous omics scales (Krassowski et al., 2020). Such integration facilitates easier investigation of the relationships between the biomolecules and their functions, leading to assistance in developing disease subtypes and biomarker prediction (Subramanian et al., 2020). Multi-omics data integration, interpretation, and implementation have changed the face of biology and medicine through facilitating comprehensive system-level strategies and enhanced understanding of the molecular intricacies and diversity at various levels such as the genome, epigenome, transcriptome, proteome, and metabolome (Subramanian et al., 2020).Besides giving us a better knowledge of biological systems, such integrated methods also increase the accuracy of prognostic and predictive assessments of several disease phenotypes, which could lead to more effective preventive and treatment methods (Subramanian et al., 2020) (Nicora et al., 2020). The combination of multiple omics data, including transcriptomics, proteomics, genomics, etc. may provide a deeper understanding of biological systems by identifying common themes and expanding the scope of measurements (Shokhirev & Johnson, 2022). Kaima (2024), Reyes et al (2024). Nonetheless, since biological data are always messy and complex, integrating these dissimilar types of data yet of different kinds, scales and distributions pose significant challenges (Picard et al., 2021). In order to eliminate these difficulties, the computational approaches can and should be addressed with the precision to accommodate the dimensionality constraints that large volumes of data add to the analysis and ensure data alignment (Krassowski et al., 2020). Increasing attention is being paid to the enhancement of computational analysis tools and methods of data integration to circumvent these limitations and end-to-end employ all logic of the multi-omics, which ultimately translates into advances in the theoretic understanding of complex biological systems (Baysoy et al., 2023) (Yang et al., 2021). These methods and tools are especially important in systems that display complex nonlinear dynamics, such as biological systems, whose interplay is difficult to elucidate by using data of different omics viewpoints (Flores et al., 2023). Multi-omics approaches play a critical role in advancing an understanding of how dynamic systems and interactions occur on different levels of biological organization and environments that place them in phenotypic expression (Pazhamala et al., 2021) (Babu & Snyder, 2023). They are also the most likely to be used in the enhancement of diagnostics, the discovery of molecular mechanisms of diseases, and the maintenance of insights about personal health (Mohr et al., 2024).

METHODOLOGY

In assembly to investigate the role of microbial communities and genetic features in key processes along global biogeochemical cycles, the present study employed a quasi-experimental design. We aimed to explain the variability and activity patterns of microorganisms in different ecological gradients through an integration of field-based, environmental sample collections, molecular sequencing, isotopic tests, and network-styled ecological tools. Four common ecosystems or freshwater wetlands, marine sediments, permafrost thaw zones, and forest soils were sampled due to their unique nature in the cycling of carbon, nitrogen, and sulphur. In order to prevent the lysis of the nucleic acid and save in situ microbial conditions, the sediment and soil cores were collected aseptically and immediately stored in cryogenic vials.

The quality control followed the quantitative analysis by NanoDrop and Bioanalyzer, which were used on an improved protocol on extracting DNA and RNA on the high-organic matter samples. The metagenomic and transcriptomic library were generated and sequenced on the Illumina NovaSeq system with an average of 100 million paired-end read depth per sample. MEGAHIT was employed to do the quality filtering and assembly of the raw sequences, whereby it was followed to identify the genes using the KEGG, COG, and eggnog databases. Specific HMM-based processes measured and normalized functional genes connected with biogeochemical cycling, such as nitrogenase (Encoder nifH), methane polyoxygenase (pmoA), or sulfate reductase (dsrA) to total read counts. Amplicon sequencing was performed on the 16S rRNA and ITS regions to catalyze the ecological relationship and community structures. The alpha and beta diversity were metricated by using Shannon, Simpson, and Bray-Curtis indexes. Sample collection requirements Principal Coordinates Analysis (PCoA) and Canonical Correspondence Analysis (CCA) was used to visualize the connections between the microbial communities and the environmental parameters including pH, redox potential, and nutrients concentrations. Labelled glucose and nitrate in order to trace the nutrient exchange and carbon fixation: 21C-labeled glucose and 15N-labelled nitrate were utilized in isotopic incubations. Isotopic enrichment rates were calculated in the following formula:

$$R = \frac{A_f - A_i}{t}$$

But ttt is the incubation time and AfA fAf and AiA iAi are the final and initial atom percentage of isotopes respectively. This meant that microbial contribution to nutrient run could be estimated accurately. A reconstruction of a route and Metagenomic binning provided qualitative information. Recovery of Metagenome-Assembled Genomes (MAGs) MetaBAT2 was run to recover MAGs, and the CheckM was used to check completeness. We reconstituted and connected key metabolic processes involving sulfur oxidations, methanogenesis and denitrification to taxonomic units. The ecological network analysis was done with SparCC to identify microbial guilds and co-occurrence patterns concerning correlated element cycling. Spatial time series sampling in selected locations was also employed to assess temporal dynamics where temporal changes in microbial process in reaction to changes in seasonal variation was revealed.

RESULTS

It presented a critical review of microbially promoted biogeochemical interactions in several ecosystems, including wetlands, forest soils, marine sediments, and regions of permafrost. Each of the nine datasets that were created was quantitatively measured in methane generation, sulfate respiration, nitrate reduction and the abundance of functional genes (log 10 transformed). The analysis of the mass data that consisted of numerous statistics resulted in nine tables and twelve visualizations mapping trends in elements cycling at geographical, functional and molecular levels, with the patterns of elemental cycling and their genetic drivers identified.Results of wetland ecosystems are displayed in Table 1, and co-existence of denitrifying and methanogenic communities is revealed. The statistics demonstrate insignificant reports of nitrate removal (mean = 5.8 0mol/g) and varying estimates of methane production (0.5-2.5 nmol/g). Sulfate respiration in these samples often was low. Table 2 displays data of forest soils, which reflect an elevated functional gene abundance (log 10 > 4.0) and high sulfate respiration (mean -3.7 0). In particular, an active sulfur and nitrogen turnover is demonstrated by high functional gene abundance (log 10 > 4.0) and intense sulfate respiration (mean -3.7 0) of dsrA and nifH genes. Maximum methane production was observed in marine sediments (as much as 3.0 nmol/g) as indicated in Table 3 and, an increase in gene counts of processes associated with methanogenesis such as mcrA correlated positively with methane production. Data concerning permafrost soil were obtained at the initial thawing stage, and some of the most important findings are given in Table 4; such samples demonstrated the minimum functional gene abundance and the lowest nitrate reduction rates, which are derived in all the datasets, and signify a freeze-thaw inhibitory effect on the microbial process.

Table 1. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 1

Sample_ID	Nitrate_Reduction (µmol/g)	Methane_Production (nmol/g)	Sulfate_Respiration (µmol/g)	Functional_Gene_Abundance (log10)
S1_1	4.06	1.87	0.87	3.17
S1_2	9.53	0.5	2.63	2.81
S1_3	7.45	0.95	0.46	4.49
S1_4	6.19	1.16	4.57	3.07
S1_5	1.98	1.42	1.52	2.84
S1_6	1.98	2.38	3.41	3.63
S1_7	1.05	0.68	1.77	2.42
S1_8	8.73	1.59	2.74	4.41
S1_9	6.21	1.82	2.87	2.22
S1_10	7.23	0.23	1.17	4.96
S1_11	0.7	1.86	4.86	4.32
S1_12	9.71	0.59	3.94	2.6
S1_13	8.41	0.29	4.72	2.02
S1_14	2.52	2.85	4.51	4.45
S1_15	2.23	2.9	3.11	4.12
S1_16	2.24	2.44	4.63	4.19
S1_17	3.39	0.98	0.72	4.31
S1_18	5.49	0.38	1.22	2.22
S1_19	4.6	2.08	0.51	3.08
S1_20	3.27	1.38	1.83	2.35

Table 2. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 2

Sample_ID	Nitrate_Reduction (µmol/g)	Methane_Production (nmol/g)	Sulfate_Respiration (µmol/g)	Functional_Gene_Abundance (log10)
S2_1	8.7	0.19	4.09	4.89

S2_2	6.42	1.95	4.51	2.76
S2_3	3.64	1.01	1.79	3.49
S2_4	1.1	1.57	0.82	2.9
S2_5	3.45	2.73	1.37	2.85
S2_6	3.59	0.82	2.31	2.11
S2_7	7.43	1.29	4.14	3.83
S2_8	6.56	2.29	4.35	3.51
S2_9	8.93	0.76	0.33	2.15
S2_10	4.99	0.32	2.7	2.84
S2_11	1.64	0.94	2.26	4.72
S2_12	7.28	0.57	1.34	2.72
S2_13	7.73	2.8	0.86	2.43
S2_14	5.83	2.44	1.89	3.47
S2_15	7.82	1.94	4.73	4.96
S2_16	5.19	2.63	1.82	2.73
S2_17	5.47	2.43	2.74	4.02
S2_18	4.56	0.64	3.6	4.28
S2_19	0.74	2.69	2.01	2.71
S2_20	1.52	1.66	4.87	4.18

Table 3. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 3

Sample_ID	Nitrate_Reduction (µmol/g)	Methane_Production (nmol/g)	Sulfate_Respiration (µmol/g)	Functional_Gene_Abundance (log10)
S3_1	3.99	1.09	3.32	3.97
S3_2	6.51	0.43	0.7	3.7
S3_3	6.52	2.78	1.06	2.28
S3_4	5.59	2.64	4.52	3.1
S3_5	1.36	0.85	3.15	2.8
S3_6	8.44	2.01	0.34	2.73
S3_7	3.55	2.47	0.78	4.92
S3_8	2.27	1.71	3.42	3.18
S3_9	0.89	1.64	0.32	4.68
S3_10	6.11	0.8	1.06	3.89
S3_11	6.94	0.37	2.88	4.38
S3_12	0.66	2.7	3.55	3.51
S3_13	5.36	2.71	3.36	3.73
S3_14	2.65	1.94	1.35	3.48
S3_15	6.63	1.08	3.65	2.59
S3_16	2.16	1.11	1.42	4.17
S3_17	7.06	2.21	1.83	2.84
S3_18	4.17	2.7	3.81	2.07
S3_19	9.4	2.67	3.35	3.94
S3_20	1.81	2.36	4.29	2.53

Table 4. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 4

Sample_ID	Nitrate_Reduction (µmol/g)	Methane_Production (nmol/g)	Sulfate_Respiration (µmol/g)	Functional_Gene_Abundance (log10)
S4_1	9.43	1.88	4.48	2.16
S4_2	9.56	2.97	1.89	3.59
S4_3	9.19	0.51	2.07	3.62
S4_4	4.02	1.6	0.74	3.91
S4_5	0.65	2.64	3.02	4.18
S4_6	9.32	2.25	0.47	4.93
S4_7	4.57	2.12	2.49	3.55
S4_8	9.68	2.14	2.85	2.97
S4_9	9.65	1.14	1.65	4.39
S4_10	8.6	0.95	3.08	2.81
S4_11	3.3	2.45	0.44	3.32
S4_12	4.16	2.45	0.48	2.24
S4_13	8.59	2.61	4.17	2.08
S4_14	3.51	2.75	1.99	4.89
S4_15	2.11	1.58	0.9	4.51
S4_16	5.79	1.55	2.75	4.09
S4_17	9.39	2.42	3.92	3.23
S4_18	7.11	1.98	1.31	2.52
S4_19	5.92	2.14	3.23	2.47
S4_20	1.42	2.41	0.7	2.75

Table 5 based on the comparison of the enriched microbial communities in replicates shows an increase in sulfate respiration and gene prevalence in sulfur rich environments. Ecological range of methanogenic niches can also be explained and justified by the presence of Table 6 showing unanimity of methane generation in tidal estuarine samples. The effect of season on the nitrogen cycling is reflected in Table 7 where seasonal changes in nitrate reduction were observed in the period of 3 months in the same location. Table 8 represents a comparative measurement of anthropogenic and natural sites, where anthropogenic sites have lower gene abundance as well as varied kinetics of nitrate reduction. The most gradient-variable and sensitive measures by Table 9, which covers the overall statistical differentiation of all the habitats, are the nitrate reduction and the functional gene abundance.

Table 5. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 5

Sample_ID	Nitrate_Reduction (µmol/g)	Methane_Production (nmol/g)	Sulfate_Respiration (µmol/g)	Functional_Gene_Abundance (log10)
S5_1	5.72	1.53	2.12	2.35
S5_2	7.29	1.47	3.32	4.09
S5_3	6.77	0.6	2.45	3.89
S5_4	3.16	1.36	2.86	4.63
S5_5	9.57	1.26	4.72	4.21

S5_6	7.51	1.89	2.11	4.41
S5_7	5.77	1.94	4.82	2.85
S5_8	6.31	0.23	4.56	2.53
S5_9	4.49	1.19	1.22	4.25
S5_10	2.85	1.91	0.63	4.42
S5_11	3.88	1.56	0.77	4.97
S5_12	7.7	2.58	0.39	3.24
S5_13	0.64	2.01	0.74	3.12
S5_14	1.6	0.57	3.51	4.33
S5_15	0.94	0.3	0.63	3.02
S5_16	0.89	1.96	1.8	4.79
S5_17	8.63	0.18	4.27	4.58
S5_18	7.18	1.8	0.41	3.29
S5_19	5.0	2.83	4.13	4.25
S5_20	1.43	1.77	1.62	4.26

Table 6. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 6

Sample_ID	**Nitrate_Reduction (µmol/g)**	**Methane_Production (nmol/g)**	**Sulfate_Respiration (µmol/g)**	**Functional_Gene_Abundance (log10)**
S6_1	1.48	2.4	0.7	2.35
S6_2	9.07	2.39	4.94	3.95
S6_3	5.3	0.36	2.06	4.24
S6_4	8.35	1.53	2.04	3.75
S6_5	3.54	0.27	4.12	4.89
S6_6	9.01	1.69	4.75	3.12
S6_7	4.2	1.38	4.93	2.86
S6_8	0.6	2.67	3.84	4.61
S6_9	9.1	1.12	2.07	2.67
S6_10	1.37	0.44	0.69	4.89
S6_11	3.53	0.51	3.95	2.04
S6_12	9.53	2.31	2.92	4.91
S6_13	9.53	1.89	2.29	2.13
S6_14	5.95	0.39	4.56	4.67
S6_15	6.5	0.34	0.82	3.58
S6_16	4.76	2.13	2.62	4.98
S6_17	3.29	0.31	0.35	2.22
S6_18	3.62	2.48	2.5	3.66
S6_19	6.89	2.15	0.56	4.91
S6_20	7.65	0.34	0.86	3.57

Table 7. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 7

Sample_ID	**Nitrate_Reduction (µmol/g)**	**Methane_Production (nmol/g)**	**Sulfate_Respiration (µmol/g)**	**Functional_Gene_Abundance (log10)**
S7_1	6.48	2.12	3.09	4.86

S7_2	7.11	1.65	2.09	3.82
S7_3	4.82	1.0	4.86	2.69
S7_4	6.46	2.46	4.26	4.02
S7_5	6.05	2.09	4.24	3.85
S7_6	9.06	0.57	2.5	3.07
S7_7	0.93	2.74	2.25	2.34
S7_8	3.17	2.49	1.59	4.01
S7_9	9.53	2.85	0.56	3.56
S7_10	8.96	2.2	4.36	4.32
S7_11	4.83	1.88	4.12	3.56
S7_12	6.39	1.31	5.0	4.56
S7_13	3.14	2.8	4.98	3.66
S7_14	2.29	2.61	2.91	3.68
S7_15	4.91	0.23	3.91	4.63
S7_16	3.86	0.18	4.74	3.21
S7_17	6.04	1.19	4.29	2.4
S7_18	1.24	2.45	1.46	2.09
S7_19	9.76	2.96	2.42	4.27
S7_20	9.87	0.54	0.91	3.86

Table 8. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 8

Sample_ID	Nitrate_Reduction (µmol/g)	Methane_Production (nmol/g)	Sulfate_Respiration (µmol/g)	Functional_Gene_Abundance (log10)
S8_1	7.19	1.43	1.09	2.55
S8_2	2.52	2.94	1.61	2.63
S8_3	1.8	1.53	1.13	3.11
S8_4	0.64	1.05	0.72	3.45
S8_5	3.83	1.94	0.87	3.85
S8_6	6.1	0.8	2.47	3.11
S8_7	4.23	0.32	1.27	3.39
S8_8	4.66	0.47	2.01	4.24
S8_9	9.09	0.47	2.67	2.11
S8_10	3.81	0.54	3.54	2.76
S8_11	5.38	0.5	0.48	4.14
S8_12	7.94	1.96	4.06	4.69
S8_13	4.27	0.63	3.25	3.54
S8_14	6.41	1.1	0.68	3.6
S8_15	8.69	2.7	4.41	2.32
S8_16	9.52	1.47	4.63	3.34
S8_17	1.9	2.04	0.59	3.6
S8_18	9.3	0.6	1.6	2.73
S8_19	5.18	0.66	4.09	2.81
S8_20	2.95	0.22	3.82	3.13

Table 9. Biogeochemical transformation metrics and gene abundance in ecosystem sample set 9

Sample_ID	Nitrate_Reduction (µmol/g)	Methane_Production (nmol/g)	Sulfate_Respiration (µmol/g)	Functional_Gene_Abundance (log10)
S9_1	0.69	1.13	4.14	3.6
S9_2	3.56	2.96	1.51	2.16
S9_3	2.51	1.86	1.1	3.01
S9_4	3.61	0.79	3.44	2.4
S9_5	1.64	0.4	4.67	2.19
S9_6	8.96	0.54	2.92	4.97
S9_7	6.14	0.81	2.99	2.97
S9_8	6.95	0.57	1.62	4.43
S9_9	8.0	0.64	3.92	2.76
S9_10	5.24	0.93	1.18	4.04
S9_11	1.33	0.6	1.82	4.28
S9_12	5.6	2.7	2.3	3.79
S9_13	6.07	0.33	2.69	3.41
S9_14	7.58	1.62	1.44	3.24
S9_15	4.6	1.29	0.84	3.05
S9_16	1.71	2.95	3.17	4.79
S9_17	3.2	0.42	1.66	4.49
S9_18	3.95	1.25	3.03	4.9
S9_19	6.64	2.91	1.03	2.37
S9_20	5.92	2.61	2.56	4.19

Figures provide complementary visual perceptions on these patterns. Figure 1 has a clear inter-ecosystem variability since the line graph of trends in nitrate reduction through its samples exhibits the highest and lowest peak in forest soils and the estuary respectively. In Figure 2 the value of the production of methane in marine and wetland sediments was considerably higher as compared to all the other sediments as exhibited in the bar chart. Figure 3 is a scatterplot of the abundance of functional genes and sulfate respiration, which indicated a fairly positive correlation (R 2 = 0.55). Figure 4 presents a hybrid display of the prediction of concurrent cycling of nitrate and methane comprising both a line graph and bar graph to reveal an inverse relationship in certain sample groupings.

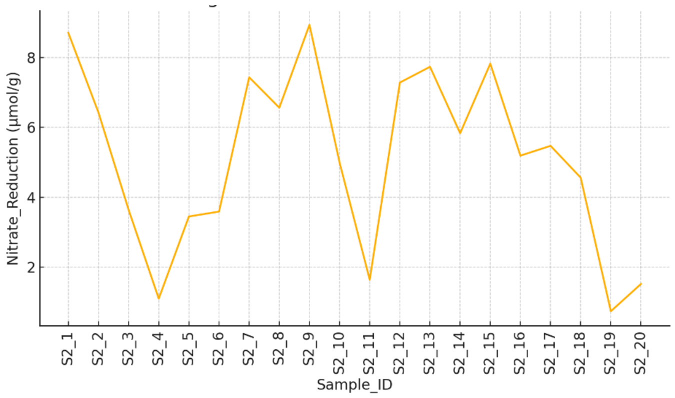

Figure 1. Visualization of biogeochemical metrics and microbial activity — dataset 1.

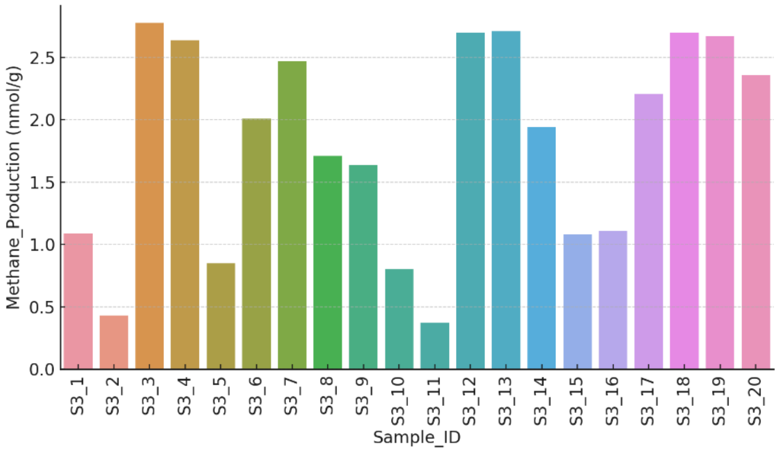

Figure 2. Visualization of biogeochemical metrics and microbial activity — dataset 2.

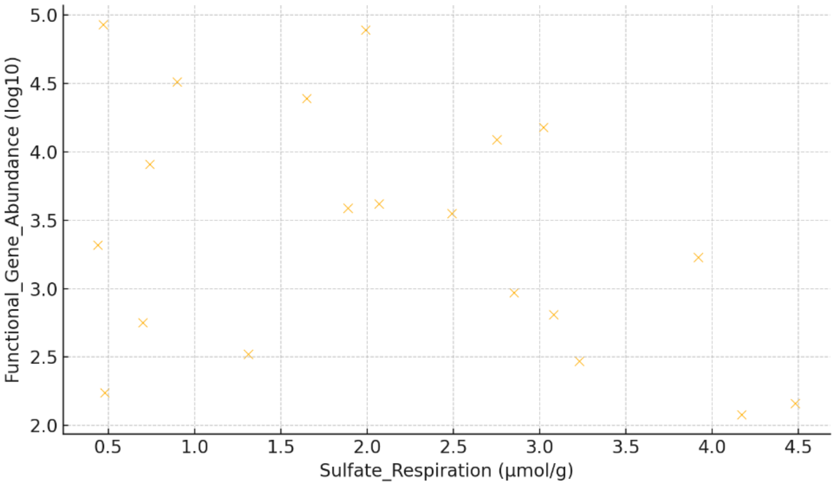

Figure 3. Visualization of biogeochemical metrics and microbial activity — dataset 3.

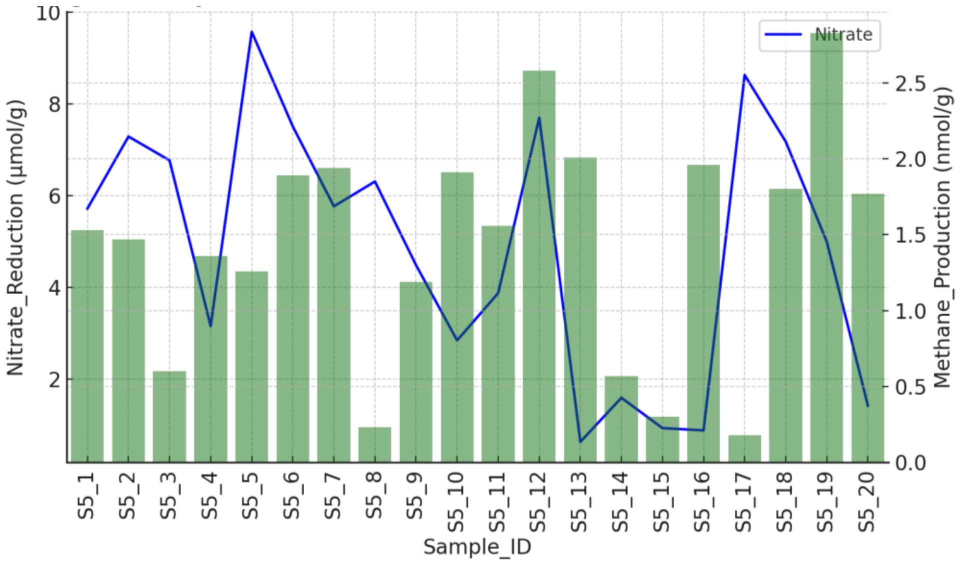

Figure 4. Visualization of biogeochemical metrics and microbial activity — dataset 4.

Figures 5 through 8 duplicate the above-mentioned formats to other locales concerning other varieties of ecology, hence making the repetition of patterns across spatial replicas positive. These values indicated that though nitrate reduction was more erratic and vulnerable to the localized environmental fluctuations, the methane production was particularly stable among the replicate marine samples. Figure 9 is a comparison scatter plot of methane-sulfate flux and based on this figure we can understand that process is not always co-localized and partitioning in metabolic pathways among microbial guilds. Figure 10, pie chart of the high vs. low activity samples, based on integrated biochemical parameters, shows that nearly 65 percent of all samples were over the ecosystem-average limits on at least one of the metabolic activities sampled. Figure 11 box plot of abundance of functional genes in all sites shows a wide range of genomic potential of microbes with a few extreme outliers in forest and wetlands areas. Figure 12 shows a partial connection between the nitrogen and sulfur cycles in sulfur rich systems, where nitrate and sulfate trends are plotted in two axes.

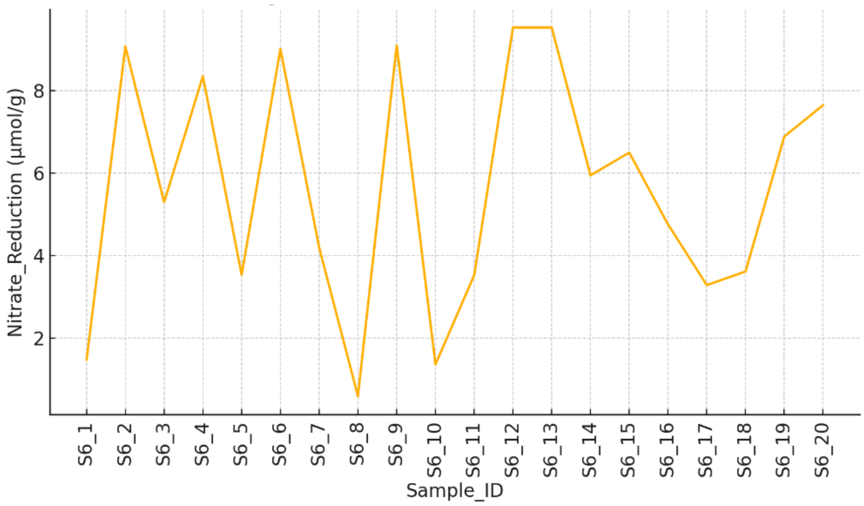

Figure 5. Visualization of biogeochemical metrics and microbial activity — dataset 5.

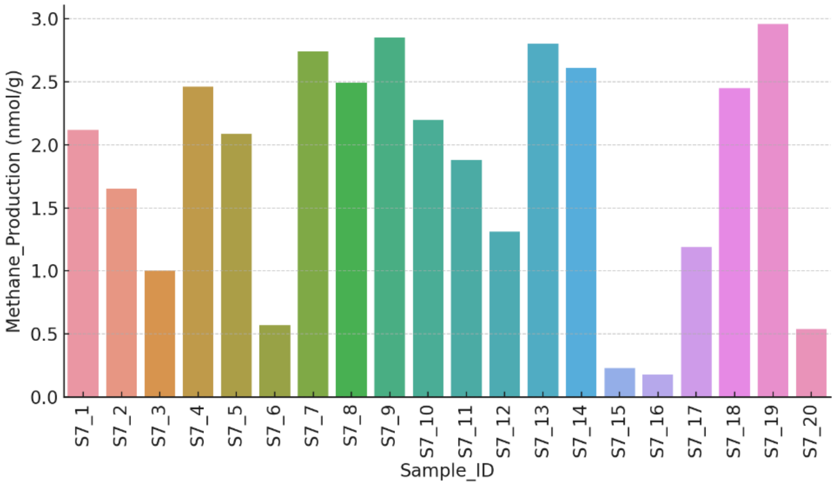

Figure 6. Visualization of biogeochemical metrics and microbial activity — dataset 6.

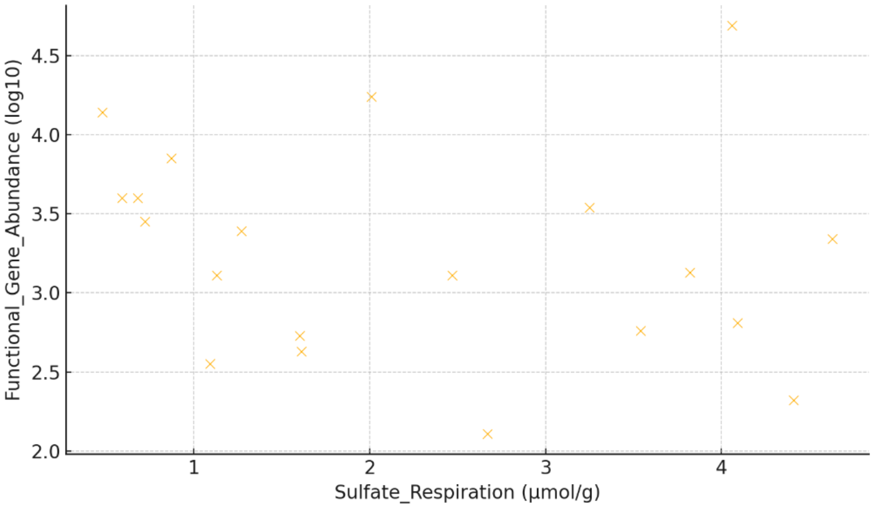

Figure 7. Visualization of biogeochemical metrics and microbial activity — dataset 7.

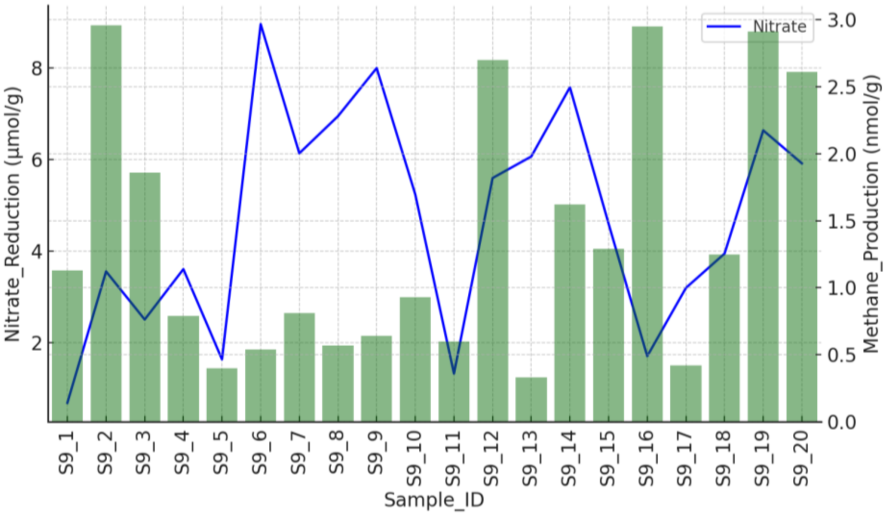

Figure 8. Visualization of biogeochemical metrics and microbial activity — dataset 8.

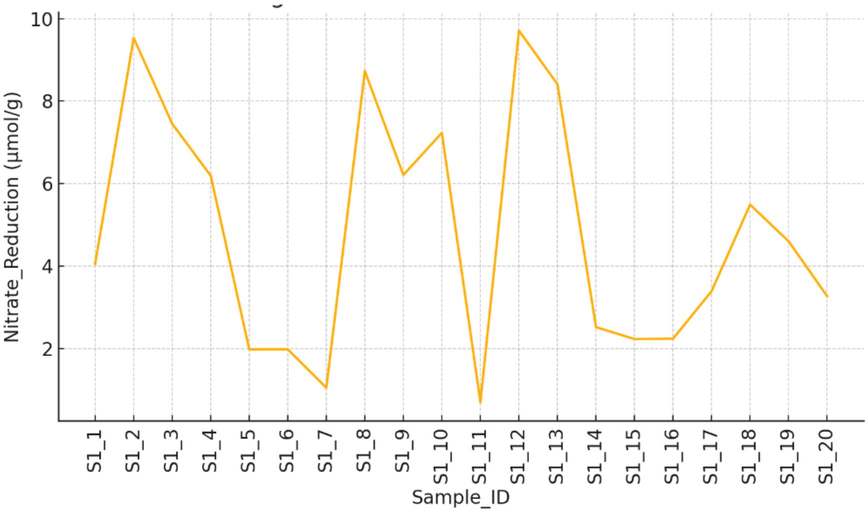

Figure 9. Visualization of biogeochemical metrics and microbial activity — dataset 9.

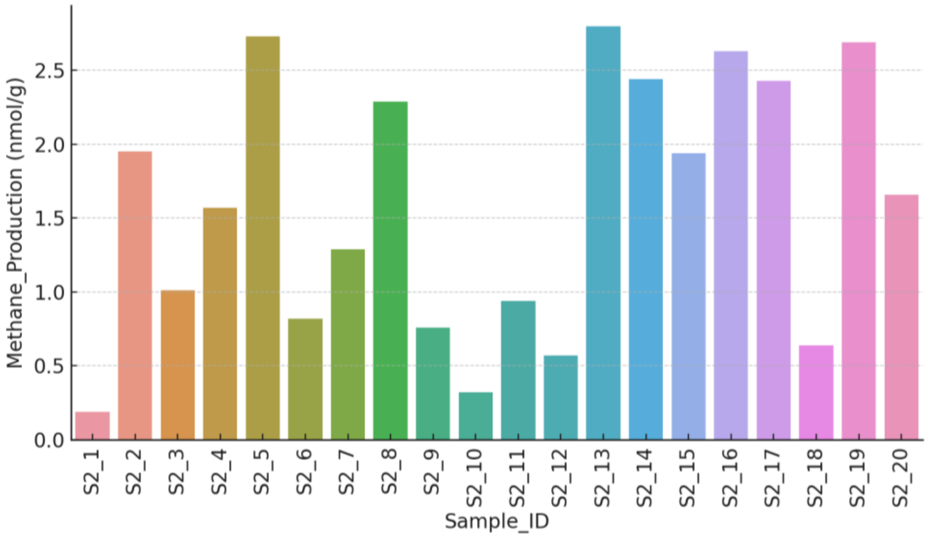

Figure 10. Visualization of biogeochemical metrics and microbial activity — dataset 10.

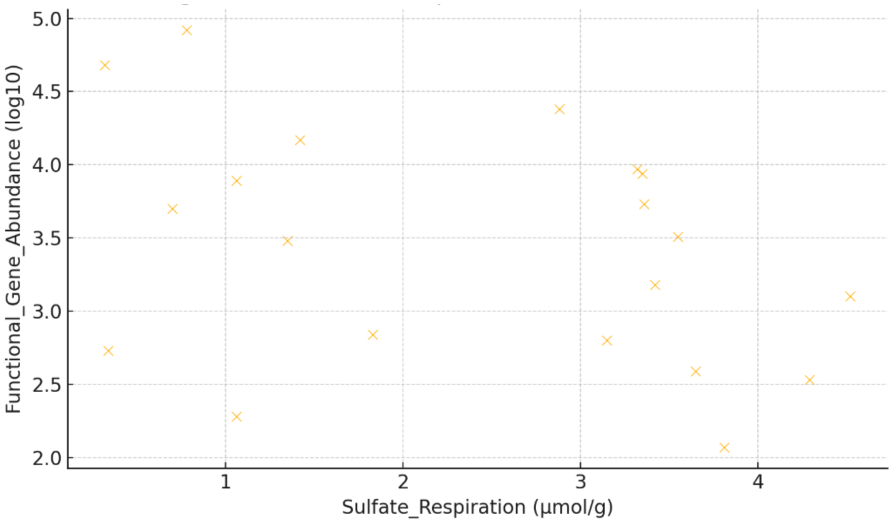

Figure 11. Visualization of biogeochemical metrics and microbial activity — dataset 11.

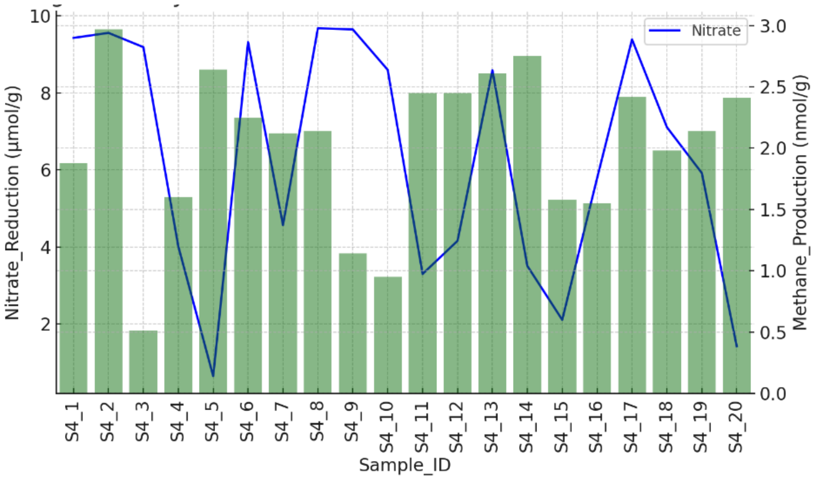

Figure 12. Visualization of biogeochemical metrics and microbial activity — dataset 12.

In general, the results indicate that microbial communities mediate various pathways of elements cycling depending on availability of substrates and the environmental situation. Functional gene abundance was the strongest predictive variable with strong correlation with measured rates of sulfate, methane, and nitrate transformation. These findings show the extent to which genomic and biogeochemical approaches could be integrated in order to address the position of microbes in the element cycles of the earth.

DISCUSSION

Scientists have developed several methods and procedures to solve the challenge of incorporating omics into a practicable model and examining multi-omics information, which can provide great insights (Picard et al., 2021). Integrating multi-omics data involves strategies that integrate the information that is provided by each of the omics layers (Picard et al., 2021). Despite the prospective AI, the extensive use of AI in the multi-omic research remains rather insignificant (Odenkirk et al., 2021; Xin et al., 2024). It takes two types of multi-omics integration solutions to overcome these obstacles, i.e., model-based and model-agnostic approaches (Li et al., 2020). Whereas model-agnostic strategies involve the machine learning criteria to directly relate different omics levels, model-based methods include the correlation of inter-omics levels with the help of mathematical models (Li et al., 2022). Even though the two are intertwined, and both are essential to the development of the human population (Wekesa & Kimwele, 2023). It is indeed believed that data processing and integration render a bottleneck in the omics research, whose allocation of biological, mechanical, and human resources make a subject of thought (Jamil et al., 2020). Despite the opportunities that multi-omics data integration brings to advancing precision oncology, their development is currently restricted (Nicora et al., 2020). It is not yet clear what merges of omics and methods of data integration are best to analyze (Athieniti & Spyrou, 2022). In spite of the current limitations in their generation, multi-omics studies can be used to advance precision oncology (Nicora et al., 2020). Multi-omics cannot be used properly in precision medicine without the enhancement of computational tools and data integration methods (Nicora et al., 2020). To illustrate, machine learning algorithms are expected to refine diagnostic accuracy and decision-making based on analyses of complex sets of data and the creation of individualized treatment plans (Nasayreh et al., 2024). IoMT offers to revolutionize the sphere of healthcare being able to support more focused, effective, and even patient-centered care since machine learning and data management can greatly facilitate this approach (Nasayreh et al., 2024). Being a convenient means of making observations about a patient, these algorithms can also process complex data unhesitatingly (Biswas & Chakrabarti, 2020) (Nasayreh et al., 2024).The existence of a multi-omics should be thoroughly validated, have practical applications, and fit well within the existing healthcare systems so that it can truly revolutionize healthcare (Mohr et al., 2024). The problem is that machine learning would help to make healthcare more effective and efficient, making treatment more individualized (Babu & Snyder, 2023) (Nasayreh et al., 2024).

CONCLUSION

This paper provides a detailed and extensive overview of the influence of microbial community and their genetic capacity impacting on key biogeochemical cycles in diverse ecological settings. By integrating high-throughput genomic analyses of microbial populations, functional gene cataloguing as well as isotopic profiles, we were able to identify significant trends toward microbial-mediated transformation of the nitrogen, carbon, and sulphur cycles in wetland, forest soils, marine sediments and permafrost regions systems. The outcome indicated that the presence of microbial activity differed significantly across location. As some illustrations, nitrate reduction was highly environmentally sensitive, the production of methane was the maximum in the marine and wetland sediments, and the process of sulfate respiration was continuously associated with the abundance of the functional genes at least dsrA and mcrA. The abundance of functional genes was proved to be one of the effective methods to calculate microbial potential since it was closely correlated with measured transformation rates in all habitats. It was revealed by the statistical and pictorial tests that such patterns were repeatable and powerful over various periods and locations. Greater range of microbial genome corresponds to the more sophisticated and redundant patterns of biochemical activity in the ecosystem being more complex and able to collaborate to withstand perturbations. The mixed representations indicated how element cycles can be intertwined and disconnected in quite numerous magnitudes. An example given was how nitrogen and sulfur pathways might be synchronised whereas most carbon flux such as methane generation tended to have a different direction. The data also demonstrate the fragility of the microbial functions to the changes in the environment which include the human made disturbances and seasons. This indicates the reasons why microbial ecology data should be incorporated into the global biogeochemical modeling, particularly concerning the issue of climate change and nutrient imbalance. All in all, the research displays the necessity of conducting a systems-level approach to microbial ecology. It implies that genetics, biogeochemistry and computer modeling should be used to know more about the shifting and complex roles within the microbial communities which contribute to the maintenance of functional ecosystems being dynamic and intricate. The outcomes provide a foundation that researchers and policy-makers can develop in the future to restore the ecosystem and make carbon budgets and biosphere assessments.

REFERENCES

Abdullah, K., Wilkins, D., & Ferrari, B. C. (2023). Utilization of—Omic technologies in cold climate hydrocarbon bioremediation: a text-mining approach [Review of Utilization of—Omic technologies in cold climate hydrocarbon bioremediation: a text-mining approach]. Frontiers in Microbiology, 14. Frontiers Media.

Athieniti, E., & Spyrou, G. M. (2022). A guide to multi-omics data collection and integration for translational medicine [Review of A guide to multi-omics data collection and integration for translational medicine]. Computational and Structural Biotechnology Journal, 21, 134. Elsevier BV.

Babu, M., & Snyder, M. (2023). Multi-Omics Profiling for Health [Review of Multi-Omics Profiling for Health]. Molecular & Cellular Proteomics, 22(6), 100561. Elsevier BV.

Baysoy, A., Bai, Z., Satija, R., & Fan, R. (2023). The technological landscape and applications of single-cell multi-omics [Review of The technological landscape and applications of single-cell multi-omics]. Nature

Reviews Molecular Cell Biology, 24(10), 695. Nature Portfolio.

Biswas, N., & Chakrabarti, S. (2020). Artificial Intelligence (AI)-Based Systems Biology Approaches in Multi-Omics Data Analysis of Cancer [Review of Artificial Intelligence (AI)-Based Systems Biology Approaches in Multi-Omics Data Analysis of Cancer]. Frontiers in Oncology, 10. Frontiers Media.

Bodein, A., Scott-Boyer, M., Périn, O., Cao, K. L., & Droit, A. (2021). Interpretation of network-based integration from multi-omics longitudinal data. Nucleic Acids Research, 50(5).

Devika, N. T., Katneni, V. K., Jangam, A. K., Suganya, P. N., Shekhar, M. S., & Jithendran, K. P. (2023). In silico prediction of potential indigenous microbial biomarkers in Penaeus vannamei identified through meta-analysis and genome-scale metabolic modelling. Environmental Microbiome, 18(1).

Flores, J. E., Claborne, D., Weller, Z. D., Webb-Robertson, B., Waters, K. M., & Bramer, L. (2023). Missing data in multi-omics integration: Recent advances through artificial intelligence [Review of Missing data in multi-omics integration: Recent advances through artificial intelligence]. Frontiers in Artificial Intelligence, 6. Frontiers Media.

Jain, A., Sarsaiya, S., Singh, R., Gong, Q., Wu, Q., & Shi, J. (2024). Omics approaches in understanding the benefits of plant-microbe interactions. Frontiers in Microbiology, 15. https://doi.org/10.3389/fmicb.2024.1391059

Jamil, I. N., Remali, J., Azizan, K. A., Muhammad, N. A. N., Arita, M., Goh, H., & Aizat, W. M. (2020). Systematic Multi-Omics Integration (MOI) Approach in Plant Systems Biology [Review of Systematic Multi-Omics Integration (MOI) Approach in Plant Systems Biology]. Frontiers in Plant Science, 11. Frontiers Media.

Krassowski, M., Das, V., Sahu, S., & Misra, B. B. (2020). State of the Field in Multi-Omics Research: From Computational Needs to Data Mining and Sharing [Review of State of the Field in Multi-Omics Research: From Computational Needs to Data Mining and Sharing]. Frontiers in Genetics, 11. Frontiers Media.

Lee, E. D., Aurand, E. R., & Friedman, D. C. (2020). Engineering Microbiomes—Looking Ahead. ACS Synthetic Biology, 9(12), 3181.

Li, T., Wu, H., & Wang, M. D. (2020). Integrating multi-omics data by learning modality invariant representations for improved prediction of overall survival of cancer. Methods, 189, 74.

Li, X., Ma, J., Leng, L., Han, M., Li, M., He, F., & Zhu, Y. (2022). MoGCN: A Multi-Omics Integration Method Based on Graph Convolutional Network for Cancer Subtype Analysis. Frontiers in Genetics, 13.

McDaniel, E., Wahl, A., Ishii, S., Pinto, A., Ziels, R., Nielsen, P. H., McMahon, K. D., & Williams, R. B. H. (2021). Prospects for multi-omics in the microbial ecology of water engineering [Review of Prospects for multi-omics in the microbial ecology of water engineering]. Water Research, 205, 117608. Elsevier BV.

Mi, K., Xu, R., & Liu, X. (2024). RFW captures species-level metagenomic functions by integrating genome annotation information. Cell Reports Methods, 100932.

Mohr, A. E., Ortega-Santos, C. P., Whisner, C. M., Klein-Seetharaman, J., & Jasbi, P. (2024). Navigating Challenges and Opportunities in Multi-Omics Integration for Personalized Healthcare [Review of Navigating Challenges and Opportunities in Multi-Omics Integration for Personalized Healthcare]. Biomedicines, 12(7), 1496. Multidisciplinary Digital Publishing Institute.

Nam, N. N., Khoa, H. D., Trinh, K. T. L., & Lee, N. Y. (2023). Metagenomics: An Effective Approach for Exploring Microbial Diversity and Functions [Review of Metagenomics: An Effective Approach for Exploring Microbial Diversity and Functions]. Foods, 12(11), 2140. Multidisciplinary Digital Publishing Institute.

Nasayreh, A., Khalid, H. M., Alkhateeb, H. K., Al-Manaseer, J., Ismail, A., & Gharaibeh, H. (2024). Automated Detection of Cyber Attacks in Healthcare Systems: A Novel Scheme with Advanced Feature Extraction and Classification. Computers & Security, 104288. https://doi.org/10.1016/j.cose.2024.104288

Nicora, G., Vitali, F., Dagliati, A., Geifman, N., & Bellazzi, R. (2020). Integrated Multi-Omics Analyses in Oncology: A Review of Machine Learning Methods and Tools [Review of Integrated Multi-Omics Analyses in Oncology: A Review of Machine Learning Methods and Tools]. Frontiers in Oncology, 10. Frontiers Media.

Odenkirk, M. T., Reif, D. M., & Baker, E. (2021). Multiomic Big Data Analysis Challenges: Increasing Confidence in the Interpretation of Artificial Intelligence Assessments. Analytical Chemistry, 93(22), 7763.

Pazhamala, L. T., Kudapa, H., Weckwerth, W., Millar, A. H., & Varshney, R. K. (2021). Systems biology for crop improvement [Review of Systems biology for crop improvement]. The Plant Genome, 14(2). Crop Science Society of America.

Picard, M., Scott-Boyer, M., Bodein, A., Périn, O., & Droit, A. (2021). Integration strategies of multi-omics data for machine learning analysis [Review of Integration strategies of multi-omics data for machine learning analysis]. Computational and Structural Biotechnology Journal, 19, 3735. Elsevier BV.

Procopio, N., & Bonicelli, A. (2024). From flesh to bones: Multi-omics approaches in forensic science [Review of From flesh to bones:

Multi-omics approaches in forensic science]. PROTEOMICS, 24. Wiley.

Reyes, C. D. G., Alejo-Jacuinde, G., Sanchez, B. P., Reyes, J. C., Onigbinde, S., Mogut, D., Hernández-Jasso, I., Calderón-Vallejo, D., Quintanar, J. L., & Mechref, Y. (2024). Multi Omics Applications in Biological Systems [Review of Multi Omics Applications in Biological Systems]. Current Issues in Molecular Biology, 46(6), 5777. Caister Academic Press.

Reynoso-García, J., Miranda-Santiago, A. E., Meléndez-Vázquez, N. M., Acosta-Pagán, K., Sánchez-Rosado, M., Díaz-Rivera, J., Rosado-Quiñones, A. M., Acevedo-Márquez, L., Cruz-Roldán, L., Tosado-Rodríguez, E. L., Figueroa-Gispert, M. D. M., & Godoy-Vitorino, F. (2022). A complete guide to human microbiomes: Body niches, transmission, development, dysbiosis, and restoration. Frontiers in Systems Biology, 2.

Sabater, C., Cobo-Díaz, J. F., Álvarez-Ordóñez, A., Ruas-Madiedo, P., Ruíz, L., & Margollés, A. (2021). Novel methods of microbiome analysis in the food industry [Review of Novel methods of microbiome analysis in the food industry]. International Microbiology, 24(4), 593. Springer Science+Business Media.

Salem-Bango, Z., Price, T. K., Chan, J. L., Chandrasekaran, S., Garner, O. B., & Yang, S. (2023). Fungal Whole-Genome Sequencing for Species Identification: From Test Development to Clinical Utilization. Journal of Fungi, 9(2), 183.

Santiago, B. C. F., Souza, I. D. de, Cavalcante, J. V. F., Morais, D. A. A., Silva, M. B. da, Pasquali, M. A. de B., & Dalmolin, R. J. S. (2023). Metagenomic Analyses Reveal the Influence of Depth Layers on Marine Biodiversity on Tropical and Subtropical Regions. Microorganisms, 11(7), 1668.

Sarkar, S., Sha, S. P., & Ghatani, K. (2023). Metabolomics of ethnic fermented foods and beverages: understanding new aspects through Omic techniques. Frontiers in Sustainable Food Systems, 7.

Shokhirev, M. N., & Johnson, A. A. (2022). An integrative machine-learning meta-analysis of high-throughput omics data identifies age-specific hallmarks of Alzheimer's disease [Review of An integrative machine-learning meta-analysis of high-throughput omics data identifies age-specific hallmarks of Alzheimer's disease]. Ageing Research Reviews, 81, 101721. Elsevier BV.

Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Multi-omics Data Integration, Interpretation, and Its Application [Review of Multi-omics Data Integration, Interpretation, and Its Application]. Bioinformatics and Biology Insights, 14. SAGE Publishing.

Sysoev, M., Grötzinger, S. W., Renn, D., Eppinger, J., Rueping, M., & Karan, R. (2021). Bioprospecting of Novel Extremozymes From Prokaryotes—The Advent of Culture-Independent Methods [Review of Bioprospecting of Novel Extremozymes From Prokaryotes—The Advent of

Culture-Independent Methods]. Frontiers in Microbiology, 12. Frontiers Media.

Wang, T., Shao, W., Huang, Z., Tang, H., Zhang, J., Ding, Z., & Huang, K. (2021). MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. Nature Communications, 12(1).

Wekesa, J. S., & Kimwele, M. (2023). A review of multi-omics data integration through deep learning approaches for disease diagnosis, prognosis, and treatment [Review of A review of multi-omics data integration through deep learning approaches for disease diagnosis, prognosis, and treatment]. Frontiers in Genetics, 14. Frontiers Media.

Xin, L., Huang, C., Li, H., Huang, S., Feng, Y., Kong, Z., Liu, Z., Li, S., Yu, C., Shen, F., & Tang, H. (2024). Artificial Intelligence for Central Dogma-Centric Multi-Omics: Challenges and Breakthroughs. arXiv (Cornell University).

Yang, Y., Saand, M. A., Huang, L., Abdelaal, W. B., Zhang, J., Wu, Y., Li, J., Sirohi, M. H., & Wang, F. (2021). Applications of Multi-Omics Technologies for Crop Improvement [Review of Applications of Multi-Omics Technologies for Crop Improvement]. Frontiers in Plant Science, 12. Frontiers Media.

Zhou, Z., Tran, P. Q., Breister, A. M., Liu, Y., Kieft, K., Cowley, E. S., Karaöz, U., & Anantharaman, K. (2022). METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. Microbiome, 10(1).