

## Original Research Articles

# Metagenomic analysis of lichen-like samples correlated to the mortality of cultured shrimp in aquaculture ponds in Vietnam.

Luong T. Vu<sup>1</sup>†, Huyen T.M. Nguyen<sup>2</sup>✉, Hiep P. Hoang<sup>1</sup>, Linh T. Phung<sup>3</sup>, Sang V. Nguyen<sup>3</sup>, Thanh V.N. Nguyen<sup>4</sup>, Lien T.H. Nghiem<sup>4</sup>, Hao M. Do<sup>5</sup>, Mai Q. Nguyen<sup>6</sup>

<sup>1</sup> Thai Nguyen University of Education, Quang Trung, Thai Nguyen, Vietnam, <sup>2</sup> University of Science and Technology of Hanoi, Vietnam Academy of Science and Technology, Hanoi, Vietnam, <sup>3</sup> VNU University of Science, Vietnam National University, Hanoi, Vietnam, <sup>4</sup> Institute of Physics, Vietnam Academy of Science and Technology, Hanoi, Vietnam, <sup>5</sup> Institute of Science and Technology for Energy and Environment, Vietnam Academy of Science and Technology, Hanoi, Vietnam, <sup>6</sup> Faculty of Chemical and Food Technology, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam

Keywords: Shrimp farming, lichen-like organism, microorganism community, metagenomic analysis

<https://doi.org/10.46989/001c.155208>

## Israeli Journal of Aquaculture - Bamidgeh

Vol. 78, Issue 1, 2026

Farmed shrimp, especially in aquaculture ponds in mainland, faced many risks due to disease and the stunted growth in Vietnam. For easy solving those problems shrimp farmers could use a lot of antibiotics, leading to antibiotic residues accumulating in the shrimp and creating antibiotic resistance in the surrounding organisms and environment. This report focused on analyzing of a fast-growing lichen-like species collected from shrimp farms in the Northern (M01) and Southern (NM1E2) region of Vietnam with a same symptom of stunted growth, respectively. Using shotgun metagenomics and bioinformatics analysis, microorganism community of lichen-like samples were identified. The results showed that *Pseudomonas*, *Halomonas*, *Aeromonas*, *Vibrio*, *Shewanella*, and *Moraxella* were the main bacterial genera contribute to the samples and half of them (*Vibrio*, *Shewanella*, and *Moraxella*) could be potential pathogenic bacteria. The fungal analysis showed that both samples were colonized with *Fusarium oxysporum* and *Fusarium graminearum*, the fungus associated with plant diseases but increasingly implicated in infectious diseases in aquatic animals. The prevalence of other opportunistic fungi such as *Aspergillus fumigatus* also suggested that they could play a larger role in deteriorating shrimp health than originally thought. Our results are preliminary evaluation for the potential influence of the lichen-like organism in shrimp farm. The results suggest targeted environmental management and improved biosecurity measures to mitigate disease risks and enhance shrimp farm sustainability.

## INTRODUCTION

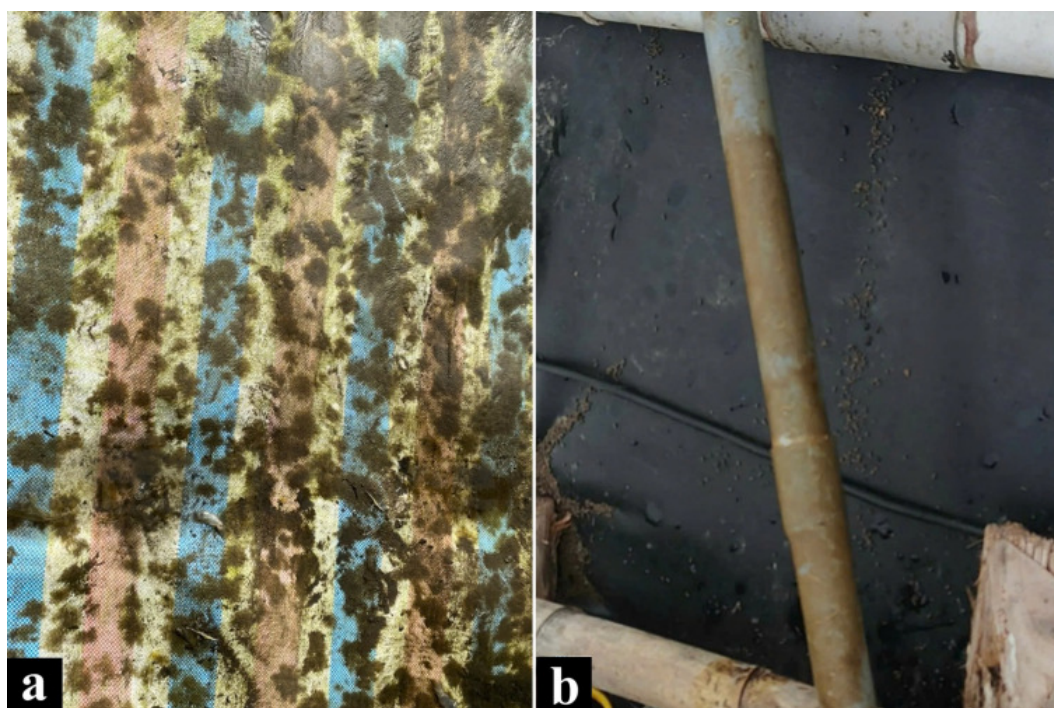
Frozen seawater shrimp exporting from Vietnam in 2024 ranked third in the world's exports.<sup>1</sup> Most of exported products came from shrimp farming, in which two main species *Litopenaeus vannamei* and *Penaeus monodon* of shrimps were cultured in aquaculture ponds in most of coastal provinces in Vietnam.<sup>2</sup> The shrimp culture contributed not only to provide the local requirement but also to export of shrimp production in Vietnam. There are many problems that can affect the health and productivity of farmed shrimp including diseases.<sup>3</sup> According to the report by journal of Vietnamese seafood from Department of Animal Health - Ministry of Agriculture and Rural Development, Vietnam, the total area of brackish water shrimp farming

affected in 2023 was more than 21 thousand hectares (ha) in which about 17 thousand ha infected due to the environment and 155 ha the sources of disease is unknown. In 2024, the infected shrimp farming area was more than 4 thousand hectares, a decrease compared to 2023 but still accounted for 19% of the total shrimp farming area.<sup>4</sup> Antibiotics were always used instead of disease prevention and treatment measures in the way of thinking of shrimp farmers.<sup>5</sup>

In recent years, a new suspected pathogen infecting shrimp has been documented and reported by shrimp farmers, they named coin fungus or dog-foot fungus by Vietnamese shrimp farmers. The coin fungus was considered to be a lichen-like organism, or a combination of filamentous fungi and photosynthetic organisms due to their morphological characteristics. They had a very strong fishy smell,

† Joint first authors

✉ Corresponding author. Huyen Thi Minh Nguyen, e-mail: [nguyen-thi-minh.huyen@usth.edu.vn](mailto:nguyen-thi-minh.huyen@usth.edu.vn)



**Figure 1.** The samples used for analyzing whole genomic sequencing. **a**, M01 sample was collected from a shrimp farm at Duong Kinh district, Hai Phong City, Vietnam after finishing the unsuccessful shrimp farming periods. **b**, NM1E2 sample was collected from a shrimp farm at Ca Mau province, Vietnam during the shrimp farming process which clearly observed the plaque appear on the water tube in the pond.

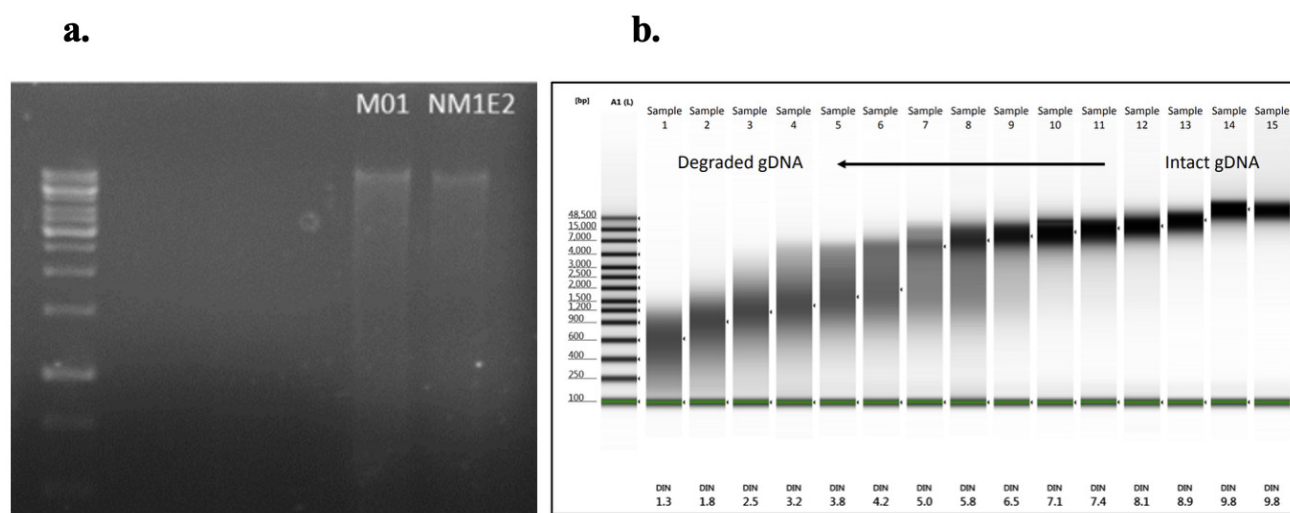
scaly, branch-shaped or in a bunch of threads, often clinging to the tarpaulin at lower 20–30 cm from the water surface or to the equipment in the pond and developed rapidly (figure 1; Microbelift). Depending on the nutritional level, they had different sizes, making lichen-like organism treatment complicated.

The signs for detecting the lichen-like organism are: 1, through the pond environment such as: fishy smell of water; various color the lichen-like green, white, yellow, or brown; scaly shape or branched tree shape or a bunch of fibers clinging to the bottom, the banks of the pond or any objects and tools in the shrimp pond. 2, through the signs of shrimp health: loose, broken intestines; a flat body, stunted and more severe is white feces disease or fall to the bottom due to eating this organism.

Although they have not been exactly identified, the colony form of lichen-like could create an ideal shelter for many parasitic organisms that would be harmful to shrimp.<sup>6</sup> Disease outbreaks in shrimp aquaculture, especially in intensive culture systems such as distributed ponds, became increasingly severe and economically expensive in recent times.<sup>7</sup> Identified pathogens included bacteria, fungi, and parasites that thrive well in poor water quality and adverse environmental conditions. Amongst bacterial diseases, Acute Hepatic Venous Valve Disease (AHPND), also known as Early Mortality Syndrome (EMS), had been widely reported in Vietnam, Thailand, and China, mostly caused by *Vibrio spp.* Similarly, Necrotic Liver Disease (NHP), often associated with the bacterium *Hepatobacter penaei*, induced mass mortality at elevated salinities and temperatures, was typically marked by reduced ap-

petite, stunted growth, and pale livers.<sup>8,9</sup> Vibriosis reduced growth rate, lethargy, obscured of muscles as well as emptiness in gut of shrimps.<sup>10</sup> Furthermore, fusariosis - a disease related to *Fusarium spp.* was informed as a fungal-induced disease. They were not only affecting the tissue but also made the lesion on the cephalothorax and abdomen of shrimps.<sup>11</sup> Larval mycosis, the other fungal disease which affect the early periods like eggs, larvae and post-larvae of shrimps,<sup>12</sup> brought serious lost for aquaculture shrimp because of the regeneration of new batch for shrimp cultivation will not be possible.

Up to date, no scientific publication related to the identification of above lichen-like samples which recognized by Vietnamese farmer from their shrimp ponds. This lichen-like sample was considered a potentially cause of disease in shrimp by the attraction to shrimp as a food with its inherent fishy smell. While the presence of this biological complex was always accompanied with visible damage to shrimp cultured, they need to be studied. Since their structure could be a diverse community, separating them for identification may not cover their entire ecosystem. Therefore, this study used next-generation sequencing - a modern molecular biology technique as a tool to resolve this community structure. Shotgun metagenomics sequencing was applied to analyze the diversity of organisms in the lichen-like samples in shrimp ponds in Vietnam. Our report is preliminary results for the analyzing the potential suspect in shrimp's aquaculture problem.



**Figure 2. Total DNA extracted in the samples by agarose gel electrophoresis (a) and DNA integrity number (b) (Marcus et al., Agilent Technologies).**

## MATERIALS AND METHODS

### LICHEN-LIKE SAMPLES COLLECTION AND MATERIALS

The lichen-like samples were collected at a shrimp farm (20.758013990204127, 106.74954873659055) at Duong Kinh district, Hai Phong City, Vietnam after finishing the unsuccessful shrimp farming periods (mark at M01; [figure 1a](#)) and another shrimp farm (8.767226500193557, 105.01684276862352) at Ca Mau province, Vietnam during the shrimp farming process which clearly observed the plaque appear on the water tube in the pond (mark at NM1E2, [figure 1b](#)). The sample collected was identified as a nearly round shape, accumulated to either the tarpaulin near the bottom of the pond after finishing the culture of shrimps (M01) or the tools and facilities floating inside the shrimp ponds (NM1E2). Sample was kept frozen at  $-80^{\circ}\text{C}$  before extracting total DNA.

### DNA EXTRACTION

Total genomic DNA was extracted using the QIAamp DNA kits for DNA extraction (Qiagen, Germany), following the manufacturers' protocols. DNA purity and concentration were measured by spectrophotometer (NanoDrop 2000, Thermo Scientific) and the integrity was assessed 0.8% agarose gel electrophoresis. Quality of DNA (DIN) had been checked prior sending to sequencing according to Marcus et al., Agilent Technologies ([figure 2](#)).

### DNA SEQUENCING AND DATA ANALYSIS

The total DNA of samples were sent for metagenomic DNA sequencing in Novogene Co., Ltd. under shotgun metagenomic sequencing contract ID H401SC24125304. DNA extraction and library construction were performed with Illumina-compatible methods, followed by paired-end sequencing on the Nova X Plus platform ( $2 \times 150$  bp). The steps of data analysis are following procedure:

### BIOINFORMATICS PIPELINE

In order to maintain the integrity and fidelity of the sequencing output, raw reads derived from the Illumina NovaSeq PE150 platform were submitted to quality control using FastQC v0.11.9 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The parameters ([table 1](#)) were carefully examined to determine whether the reads had sufficient quality for downstream metagenomic analysis:

Given that all FastQC reports indicated consistently high-quality reads across these key metrics, no additional trimming or filtering was required, and the raw reads were retained for downstream analyses including taxonomic classification, assembly, and functional annotation.

For taxonomic profiling, high-quality reads were classified against the NCBI RefSeq database (March 2025 release) using Kraken2 v2.1.2,<sup>13,14</sup> with default settings. Species-level abundance estimations were refined using Bracken v2.6.2.<sup>13</sup>

Metagenomic reads were assembled with MEGAHIT v1.2.9, which employs a de Bruijn graph-based approach that was tailored for complicated communities. Low-frequency k-mers that occurred only once were discarded ( $-\text{min-count } 2$ ), and contigs less than 200 bp in length were discarded for ensuring quality.

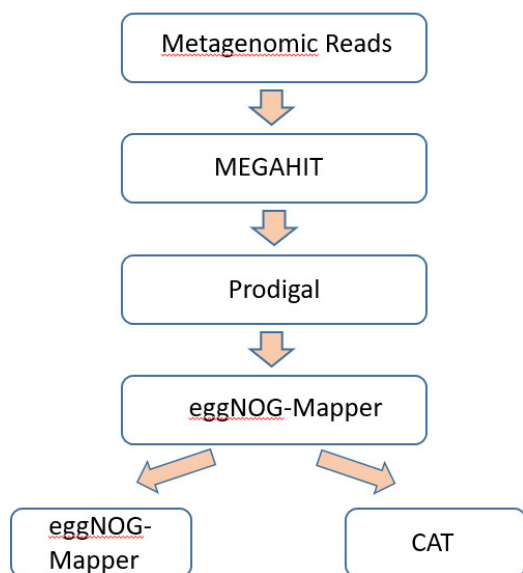
Assembly quality was assessed with QUAST v5.2.0, showing number of contigs, assembly size, N50/N90 (continuity), L50/L90 (efficiency), GC content, and auN (cumulative contiguity). Contigs of length  $> 1,000$  bp were used for downstream analysis.

Functional annotation: Open reading frames (ORFs) were also predicted from contigs using Prodigal v2.6.3 in metagenomic mode and translated protein sequences were also annotated with eggNOG-mapper v2.1.9 against the eggNOG v5.0 database to yield KEGG pathways, EC numbers, and functional categories ([figure 3](#)).

Taxonomic classification was performed using CAT (Contig Annotation Tool) with DIAMOND alignment against the NCBI NR database. Consensus was only attrib-

**Table 1. The parameters of FastQC**

| FastQC Metric                | Description                                                                                                                                                       | Interpretation (for high-quality data)                                                                        |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
| Per Base Sequence Quality    | Evaluates the Phred quality score (quality of the identification of the nucleobases generated by automated DNA sequencing) at each base position across all reads | Phred score consistently > Q30 across read length; stable boxplots; no noticeable decline at 3' ends          |
| Per Sequence Quality Scores  | Calculates the average quality score of each read                                                                                                                 | > 90% of reads have average scores above Q30 → high base-calling accuracy and read reliability                |
| Per Tile Sequence Quality    | Assesses quality variation across spatial tiles of the sequencing flowcell                                                                                        | No significant deviations between tiles → even sequencing run without lane-specific artifacts                 |
| Per Base Sequence Content    | Monitors the proportion of A/T/G/C at each base position                                                                                                          | Balanced base composition; minor variation at the start is acceptable in metagenomic datasets                 |
| Per Sequence GC Content      | Analyzes GC percentage distribution across reads                                                                                                                  | Unimodal distribution as expected for complex metagenomic samples; no abnormal peaks suggesting contamination |
| Per Base N Content           | Detects the presence of undetermined bases ("N")                                                                                                                  | N ~ 0 → precise base-calling and clear signal                                                                 |
| Sequence Length Distribution | Displays the distribution of read lengths                                                                                                                         | Uniform read length (~150 bp) typical of NovaSeq data; indicates successful adapter trimming                  |
| Sequence Duplication Levels  | Estimates the proportion of duplicated reads                                                                                                                      | < 10% duplication → high library complexity and minimal PCR amplification bias                                |
| Overrepresented Sequences    | Identifies any highly frequent sequences that may indicate contamination                                                                                          | No dominant overrepresented sequences → well-prepared library without adapter dimers or contaminants          |
| Adapter Content              | Measures residual adapter sequences                                                                                                                               | Very low adapter content → trimming during preprocessing was effective                                        |

**Figure 3. Map of process of Functional Annotation**

uted if >50% of ORFs on a contig shared the same lineage, thus making it more authentic and preventing horizontal gene transfer mistakes.

**Targeted Read Mapping:** To confirm key taxa, high-quality reads were mapped against reference genomes of *Fusarium oxysporum*, *Vibrio campbellii*, and *Vibrio alginolyticus* using Bowtie2 v2.5.2 in very-sensitive mode. Mapping statistics and genome coverage offered evidence of their presence.

#### STATISTICAL ANALYSIS

Alpha diversity indices, including Shannon, Simpson, and Chao1, were calculated to assess microbial diversity within samples using the scikit-bio Python package. Chao1 Index was a non-parametric index that used to estimate the richness of species in a community with consideration of the abundance of rare taxa. The singletons (species occurring once) and doubletons (species occurring twice) were employed by the estimator to extrapolate the number of unobserved species.<sup>15</sup> Shannon Diversity Index expressed the quantifying richness (number of species present) and evenness (distribution of abundances among species) in a community.<sup>16</sup> It responded to small taxa and indicated how evenly individuals were distributed among species. Larger values indicated more diverse and evenly-balanced communities; Simpson Diversity Index was the probability of two randomly selected individuals in a sample belonging to the same species,<sup>17</sup> The Simpson index were more sensitive to

**Table 2. DNA Quality Requirement for NGS Sequencing**

| DNA quality requirement     |            |
|-----------------------------|------------|
| Concentration               | > 15 ng/μl |
| Volume                      | > 30 μl    |
| Purity                      | > 1.7      |
| Total DNA Content           | > 200      |
| DIN (DNA Integrity Numbers) | > 7.1      |

dominant taxa and less sensitive to rare species compared to the Shannon index. (Table 5)

For taxonomic abundance analysis, raw abundance data were log-transformed ( $\log_{10}[x+1e-9]$ ) and subsequently standardized using z-score normalization to allow cross-sample comparisons. Heatmaps were constructed based on z-scores to visualize the distribution patterns of key bacterial and fungal genera across environments.

All statistical analyses and visualizations were performed using Python (v3.10) libraries, including seaborn, scikit-bio, and matplotlib. Statistical significance was set at  $p < 0.05$  unless otherwise indicated.

## RESULTS AND DISCUSSION

### DNA EXTRACTION

Genomic DNA extracted from both lichen-like samples (M01 and MM1E2) demonstrated high quality and suitability for downstream next-generation sequencing applications. The DNA concentrations were 94.1 ng/μL for M01 and 40.1 ng/μL for MM1E2, exceeding the minimum input requirements for library preparation protocols (table 2 and 3). Purity ratios were within or near ideal ranges, with A260/280 values of 1.95 for M01 and 2.14 for MM1E2, suggesting minimal protein contamination. The A260/230 ratios were 1.52 for M01 and 1.9 for MM1E2, indicating low levels of residual salts and organic contaminants. Notably, the DNA Integrity Numbers (DIN) of 6.5 for M01 and 7.4 for MM1E2 confirmed that both DNA samples were sufficient quality for shotgun metagenomic sequencing (figure 2).

### SEQUENCING ANALYSIS

Shotgun metagenomic sequencing was performed using the Illumina NovaSeq 6000 platform with a paired-end 150 bp (PE150) strategy. The sequencing yielded 64,954,792 raw reads (9.7 Gb) in M01 and 55,135,332 raw reads (8.4 Gb) in NM1E2. Quality filtering and preprocessing resulted in high effective read retention, with 97.68% and 98.24% of reads retained for M01 and NM1E2, respectively. The sequencing

error rate was consistently low at 0.01% for both samples (Table 4).

Base quality metrics were high, with Q20 values exceeding 97% (97.89% for M01 and 98.77% for NM1E2, respectively), and Q30 scores of 94.25% for M01 and 96.37% for NM1E2, indicating that the majority of bases were of high accuracy. Notably, GC content differed between the two samples, with M01 exhibiting a higher GC percentage (62.41%) than NM1E2 (42.60%), which reflected underlying differences in microbial community structure or metagenomic DNA composition. Collectively, these quality metrics demonstrated that the sequencing data were high fidelity and suitable for comprehensive downstream metagenomic analyses.

### DIVERSITY INDICES

A total of 34.4 million and 28.0 million read pairs were selected from the M01 and NM1E2, respectively, for analyze diversity in samples. Taxonomic analysis identified 7,603 species in M01 and 5,802 species in NM1E2, of which 5,617 species were shared between the two samples (Fig. 4, table 5). These demonstrated the presence of a substantial core microbiome while also highlighting niche-specific species. The Alpha diversity indices showed differences between the two microbial communities. The M01 sample exhibited greater richness and diversity, with higher Shannon (6.640) and Chao (8068) than the NM1E2 with Shannon (4.011) and Chao1 (6479), respectively. Thus the M01 sample represented a more complex and evenly distributed community. Similarly, the Simpson index (0.985 vs. 0.932) suggested higher species evenness in M01 (Table 5).

### TAXONOMY PROFILES

#### BACTERIAL COMMUNITY – COMPOSITION OF THE MAIN PHYLUM

Taxonomic profiling was conducted using Kraken2 for initial classification, followed by Bracken to refine abundance estimates based on assigned reads. In total, 16.19% and 16.12% of the reads were successfully classified up to the genus level in samples M01 and NM1E2, respectively. Relative abundance values were calculated as the proportion of classified reads assigned to each taxon.

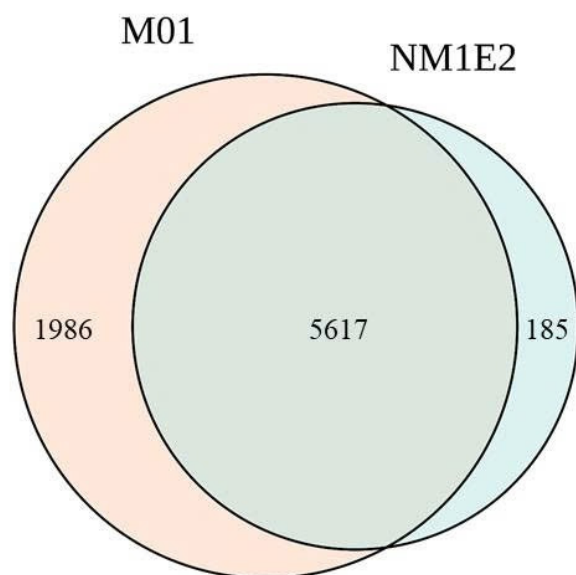
At the phylum level, six dominant bacterial phyla were identified across both samples: Proteobacteria, Actinobacteria, Bacteroidetes, Chordata, Cyanobacteria, Firmicutes (Fig. 5). Among these, Proteobacteria consistently represented the most abundant phylum in both samples, ranging from 9.34% in M01 to 12.98% in NM1E2.

**Table 3. DNA quality and quantity by absorbance measurement**

| Sample | Concentration (ng/ul) | A260/A280 | A260/230 | DIN |
|--------|-----------------------|-----------|----------|-----|
| M01    | 94.1                  | 1.95      | 1.52     | 6.5 |
| NM1E2  | 46.5                  | 1.67      | 1.9      | 7.4 |

**Table 4. Data quality summary**

| Sample | Raw reads  | Effective (%) | Error (%) | Q20 (%) | Q30 (%) | GC (%) |
|--------|------------|---------------|-----------|---------|---------|--------|
| M01    | 64,954,792 | 97.68         | 0.01      | 97.89   | 94.25   | 62.41  |
| NM1E2  | 56,135,332 | 98.24         | 0.01      | 98.77   | 96.37   | 42.60  |



**Figure 4. The Venn diagram representing for shared observe richness between M01 and NM1E2 samples. The number in each space (different by color) represents for the different or similar species in each sample.**

The significant presence of Bacteroidetes in sample NM1E2, mainly due to the genus *Flavobacteria* (0.25%, supplementary document 2), could be associated with the accumulation of organic matter and potential disease risk in aquaculture environments.<sup>18</sup> In contrast, Actinobacteria were dominant in M01, reflecting a microbiota specific to environments rich in organic matter and with an ability to participate in the decomposition and recycling of nutrients. This result consistent with the description of Goodfellow et al.,<sup>19</sup> where Actinobacteria accounted for more than 10%

of the total sediment microbiota. The result suggested that the change in microenvironmental conditions – from a potentially eutrophic state in NM1E2 to a more organically active state in M01 – might have a direct impact on microbial community structure and thus on livestock health.

#### GAMMAPROTEOBACTERIA COMPOSITION

At the class level, taxonomic comparison between the two samples revealed distinct shifts in microbial composition (Fig. 6). The class Gammaproteobacteria, a major subgroup within the phylum Proteobacteria, exhibited a substantial increase in relative abundance from 4.4% in sample M01 to 10.7% in NM1E2. This result suggested a possible expansion of taxa within this class in NM1E2.

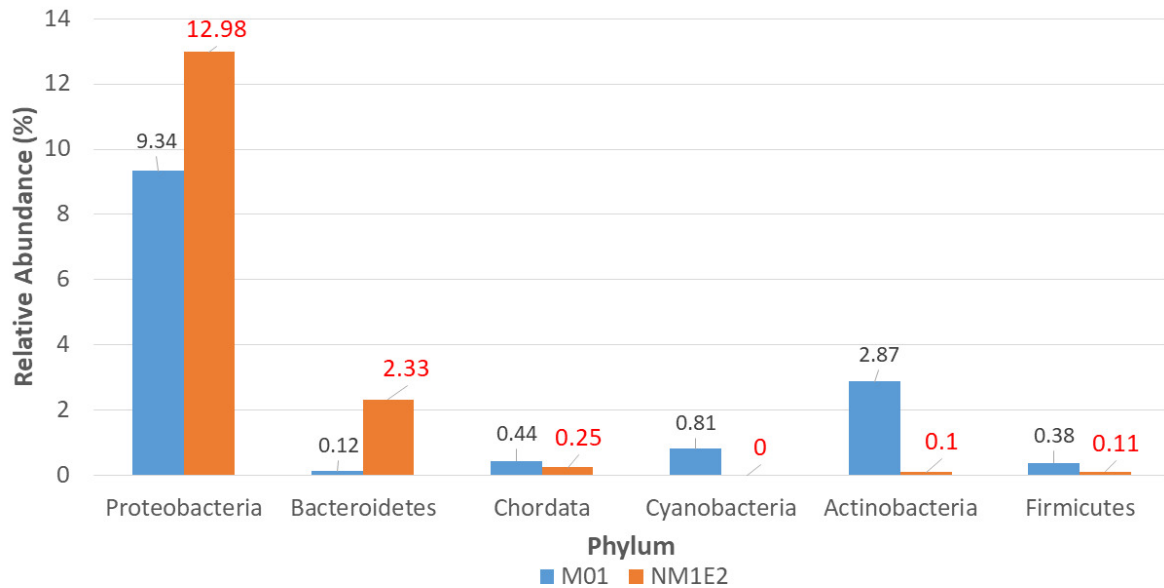
In contrast, other classes within Proteobacteria which were often associated with environmentally stable or mutualistic bacterial groups, such as Alphaproteobacteria (~2.5% in M01 vs. ~0.4% in NM1E2) and Betaproteobacteria (~3% in M01 vs. 2.1% in NM1E2), were relatively more abundant in M01. The decrease of these classes in NM1E2 would reflect ecological disruption of the sample (Pini et al., 2011, Kaushik et al.<sup>20</sup>).

Further analysis in the Gammaproteobacteria – major subgroup in phylum Proteobacteria (Fig. 6), six genera had been consistently detected in both samples (Fig. 7), although the proportions varied. Among them, *Pseudomonas* was the most abundant (0.18% in M01 vs. 0.06% in NM1E2). *Pseudomonas* species had been known for their ability to degrade a wide range of organic matter, from PAHs, hydrocarbons to plastics, thanks to their rich oxidative enzyme network and their ability to form protective biofilms, which help them survive in polluted and organic-rich environments. They were also active in the phosphorus cycle (P-accumulators), supporting nutrient processing in aquatic systems.<sup>21,22</sup>

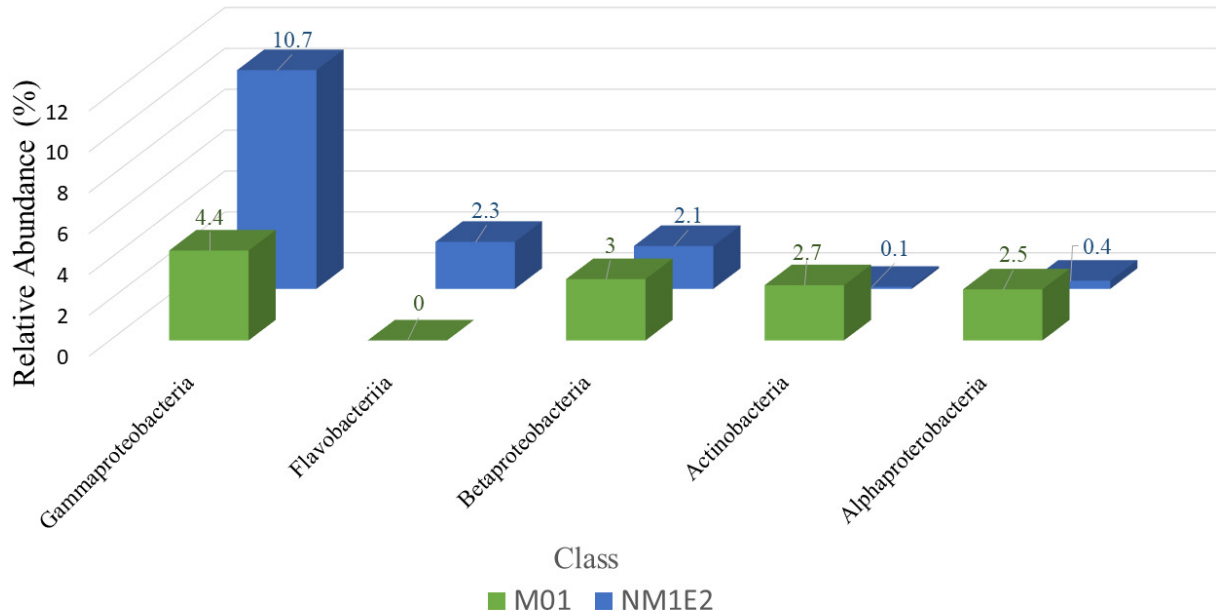
The remaining genera such as *Halomonas*, *Aeromonas*, *Vibrio*, *Shewanella*, and *Moraxella*, although presented at

**Table 5. Summary of diversity indices and read analysis in two samples of study.**

|                        | M01      | NM1E1    |
|------------------------|----------|----------|
| <i>Result analysis</i> |          |          |
| Number of read pairs   | 34477396 | 28067666 |
| Classified species     | 7603     | 5802     |
| <i>Alpha diversity</i> |          |          |
| Shannon                | 6.640    | 4.011    |
| Chao1                  | 8068     | 6479     |
| Simpson                | 0.985    | 0.932    |



**Figure 5. Main bacterial Phylum abundance between the two samples. NM1E2 sample had mainly Proteobacteria and Bacteroidetes while M01 sample had mainly Proteobacteria and Actinobacteria in comparison with other families of microorganisms.**



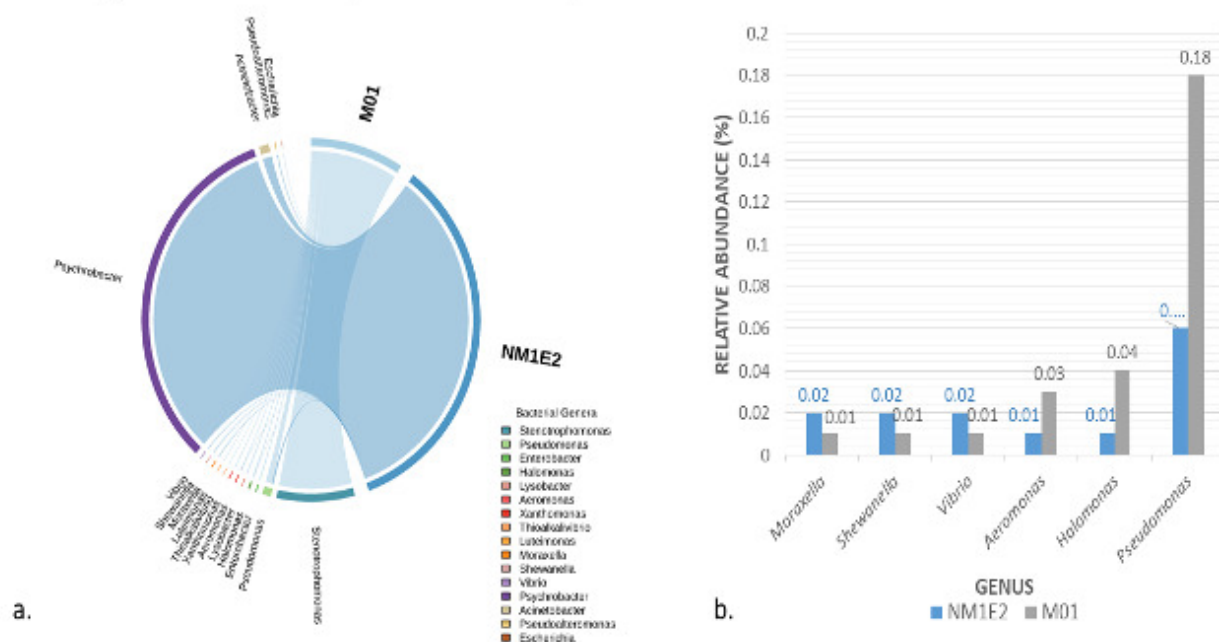
**Figure 6. Relative abundance at class level of common bacterial in the samples. The gammaproteobacteria class (belonging to the Proteobacteria family) is predominant in the two samples M01 and NM1E2.**

low levels (0.01–0.04%) (Fig.7b), all had important ecological roles. *Shewanella* metabolized metals such as manganese, iron, uranium, etc., contributed to geochemical cycles, and helped reduce toxicity in mud ecosystems<sup>23</sup>; *Vibrio* and *Aeromonas* often related with diseases in aquatic animals and were presented in large numbers in

waters rich in organic matter or showing signs of pollution; *Moraxella* also noted to be present in the process of hydrocarbon decomposition; *Halomonas* specialized in high salinity environments.<sup>24,25</sup>

The functional diversity of these genera such as organic matter decomposition, nutrient cycling, metal toxicity re-

Common genera of the class Gammaproteobacteria in two samples



**Figure 7. The genus diversity of class Gammaproteobacteria in phylum Proteobacteria (a) and relative abundance (b) in both samples. a. Psychrobacter, Acinetobacter, Pseudoalteromonas, Escherichia, Pseudomonas, Halomonas, Aeromonas, Moraxella, Shewanella, Vibrio were the most abundant gammaproteobacteria genera in sample NM1E2; Stenotrophomonas, Pseudomonas, Enterobacter, Halomonas, Lysobacter, Aeromonas, Xanthomonas, Thioalkalivibrio, Luteimonas, Moraxella, Shewanella, Vibrio were the most abundant gammaproteobacteria genera in sample M01, respectively. b. The genera Pseudomonas, Halomonas, Aeromonas, Vibrio, Shewanella, and Moraxella were the most abundant and common genera between the two samples, M01 and NM1E2.**

duction and pathogenicity, allowed Gammaproteobacteria to become the dominant microbial group in aquaculture mud environments. The predominance of *Pseudomonas* in M01 would reflect a more organic humus-rich environment, which encouraging the growth of species with efficient biodegradation and nutrient cycling capabilities. Meanwhile, the lower abundance in NM1E2, where organic matter was less concentrated, still maintained the presence of genera necessary for microbial ecosystem balance.

Despite their low relative abundances, the presence of *Aeromonas* and *Vibrio* in both samples raised attention, as these genera were widely recognized as potential pathogens in shrimp aquaculture systems.<sup>26,27</sup> Their detection within the lichen-like biofilm sample (NM1E2) surged questions regarding the possible role of such substrates as reservoirs or vectors for opportunistic bacteria.

#### DIFFERENTIAL TAXONOMIC DISTRIBUTION REVEALED BY HEATMAP ANALYSIS

To assess the variation in bacterial community composition between samples, a heatmap of the top genera based on log-transformed and Z-score standardized relative abundance data were generated. This transformation was performed across rows (i.e., within genera), allowing visualization of deviations relative to each genus's average across samples. Hierarchical clustering was applied to both taxa and samples to reveal compositional relationships. Genera

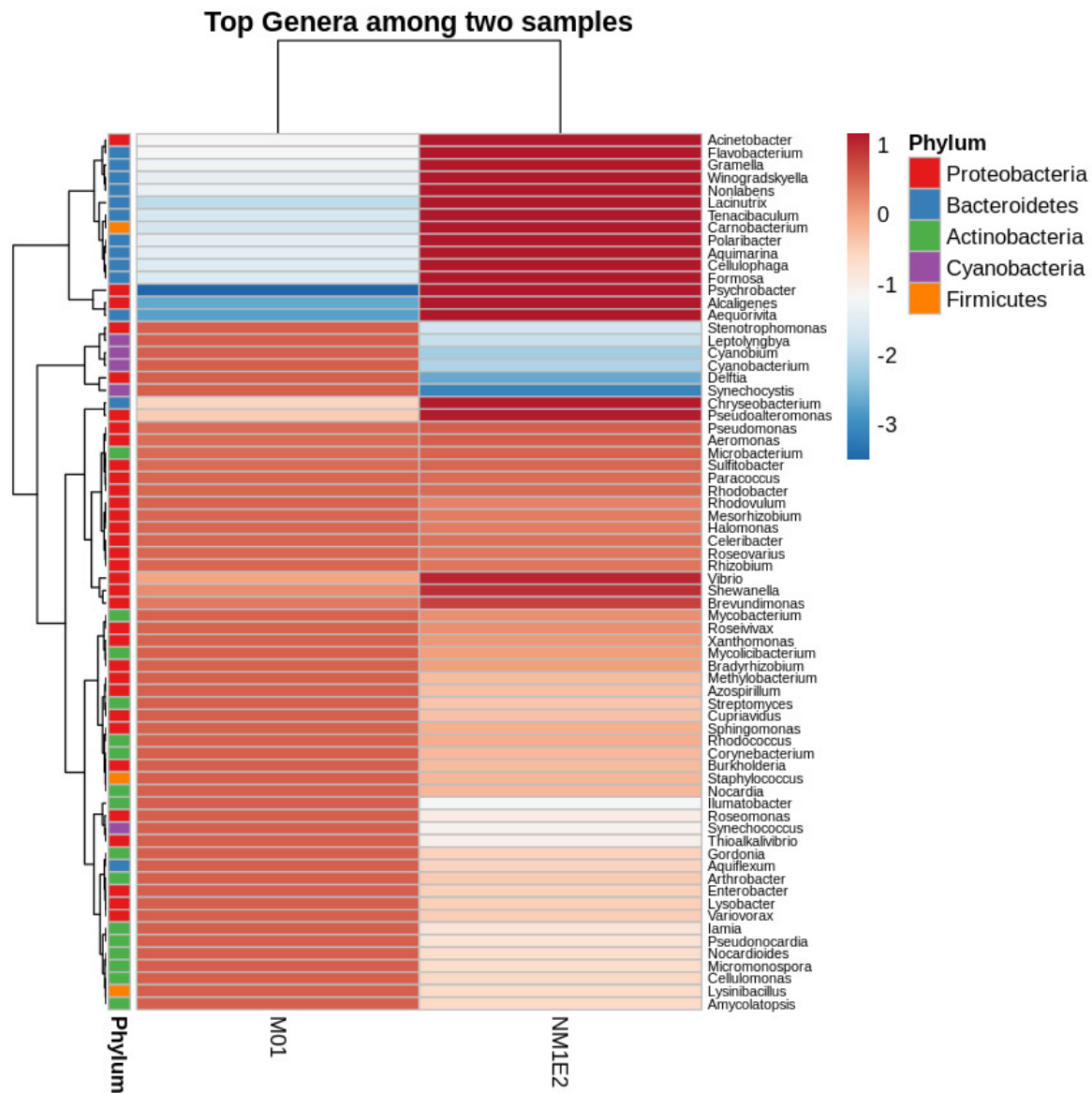
were annotated by phylum to provide taxonomic context (Fig. 8).

To identify genera of particular interest, we applied a threshold on the standardized abundance scores, selecting taxa with values exceeding -0.5 in at least one sample. Based on this criterion, two genera *Aeromonas* and *Vibrio* were retained for further species-level investigation. These genera were known to contain multiple species of relevance to shrimp aquaculture, prompting focused downstream analysis.

#### VIBRIO AND AEROMONAS - POTENTIAL CANDIDATES FOR SHRIMP PATHOGENS

From the genus-level taxonomic heatmap (Fig. 9), species-level refinement was performed to identify potential pathogenic taxa (Z-score > -0.5) present within both samples collected from shrimp aquaculture ponds exhibiting disease signs. This approach yielded several candidate species of interest: *Vibrio alginolyticus*, *V. campbellii*, *V. metschnikovii*, *V. metoecus*, *V. ponticus*, *V. palustris*, *Aeromonas media*, *V. vulnificus*, *V. ruber*, *V. tapetis*, *V. anguillarum*, and *V. fluvialis*.

Among these, *V. alginolyticus* and *V. campbellii* were well-understand shrimp pathogens, frequently implicated in vibriosis and larval mortality.<sup>26,27</sup> *Aeromonas media*, although primarily accompanying with freshwater environments, had been reported as an opportunistic pathogen under stress-inducing conditions. The remaining *Vibrio*



**Figure 8. Heatmap of Top 70 Bacterial Genera Relative Abundance and Hierarchical Clustering in both samples.** Data were log-transformed and Z-score standardized across rows (within each genus) to highlight deviations from the genus's average abundance. Genera are annotated by their respective phyla. Genera with standardized abundance scores exceeding -0.5 in both samples.

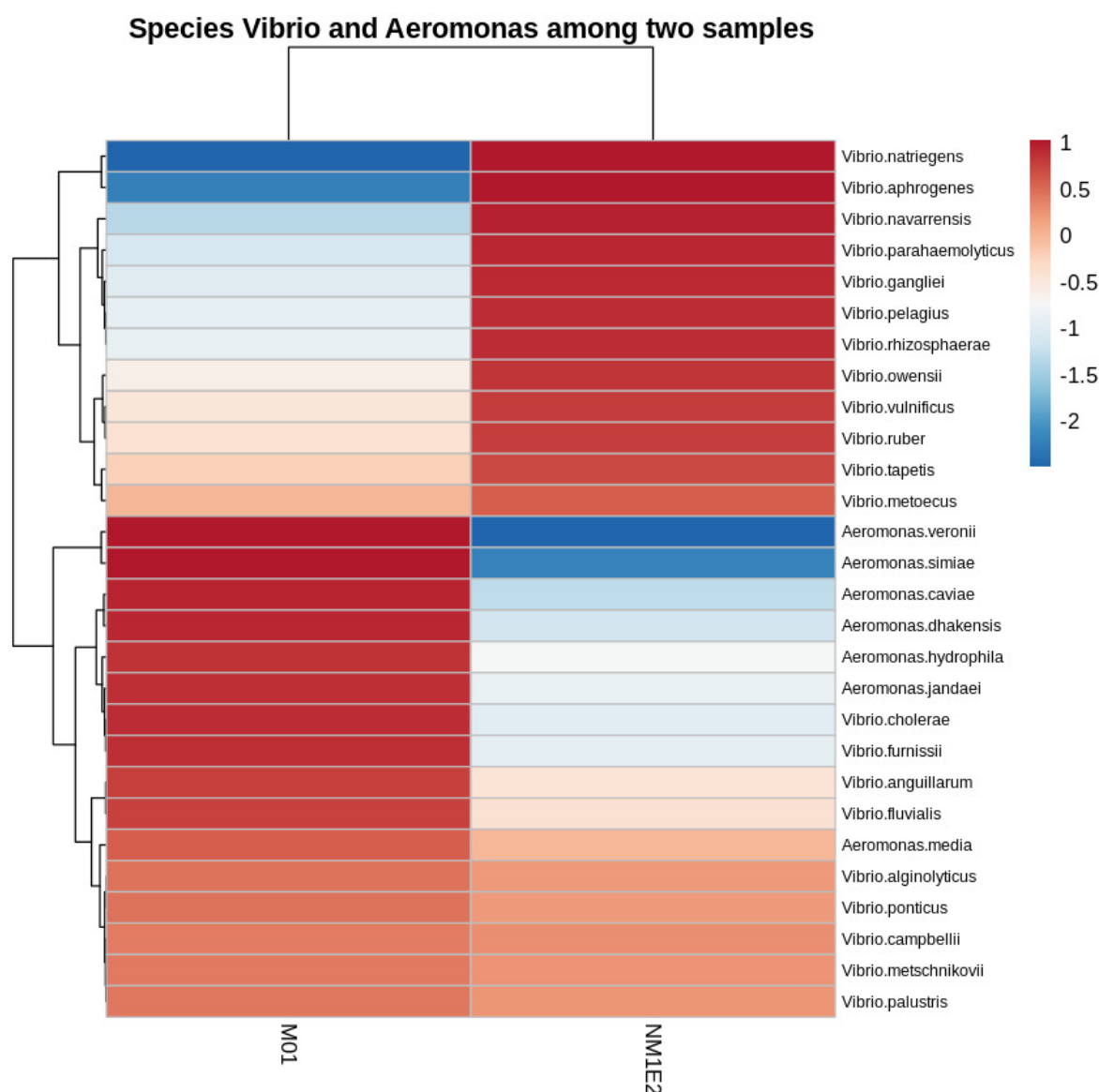
species - *V. metschnikovii*, *V. metoecus*, *V. ponticus*, and *V. palustris* - were not commonly regarded as primary pathogens in shrimp, but their detection in our samples suggested potential ecological or opportunistic roles that merit further investigation.

While abundance-based screening alone could not establish pathogenicity, the use of a Z-score normalization strategy enabled the identification of relatively enriched taxa across both suspected sources. These findings supported the hypothesis that lichen-like biofilm aggregated in shrimp ponds might serve as reservoirs for bacteria of pathogenic potential, providing a focused list of candidates for downstream validation of virulence factors and functional gene content.

#### FUNGUS COMMUNITY - COMPOSITION OF THE MAIN PHYLUM

In addition to bacterial taxonomic profiling, fungal communities were analyzed separately to investigate their potential involvement in the composition and ecological function of the lichen-like samples from shrimp ponds. Taxonomic classification at the phylum level revealed that both samples were dominated by members of the phylum Ascomycota, comprising 0.1542% in M01 and 0.0808% in NM1E2 of classified reads, respectively (Fig. 10). Ascomycota was the largest fungal phylum and included most known lichen-forming fungi, as well as numerous saprotrophic and opportunistically pathogenic species.<sup>28,29</sup>

The second most abundant phylum was Basidiomycota, with lower relative abundances of 0.0255% in M01 and 0.0090% in NM1E2. Additionally, low levels of Mi-



**Figure 9. Species-Level Heatmap of Potential Pathogens in Disease-Affected Shrimp Ponds. Species with standardized abundance scores exceeding -0.5 in both samples, including known shrimp pathogens such as *Vibrio alginolyticus* and *Vibrio campbellii*, along with *Aeromonas media*, an opportunistic pathogen.**

crosporidia (0.0033% in M01; 0.0012% in NM1E2) were detected. Microsporidia were obligate intracellular fungi, had been known to infect aquatic invertebrates, including crustaceans such as shrimp.<sup>30</sup>

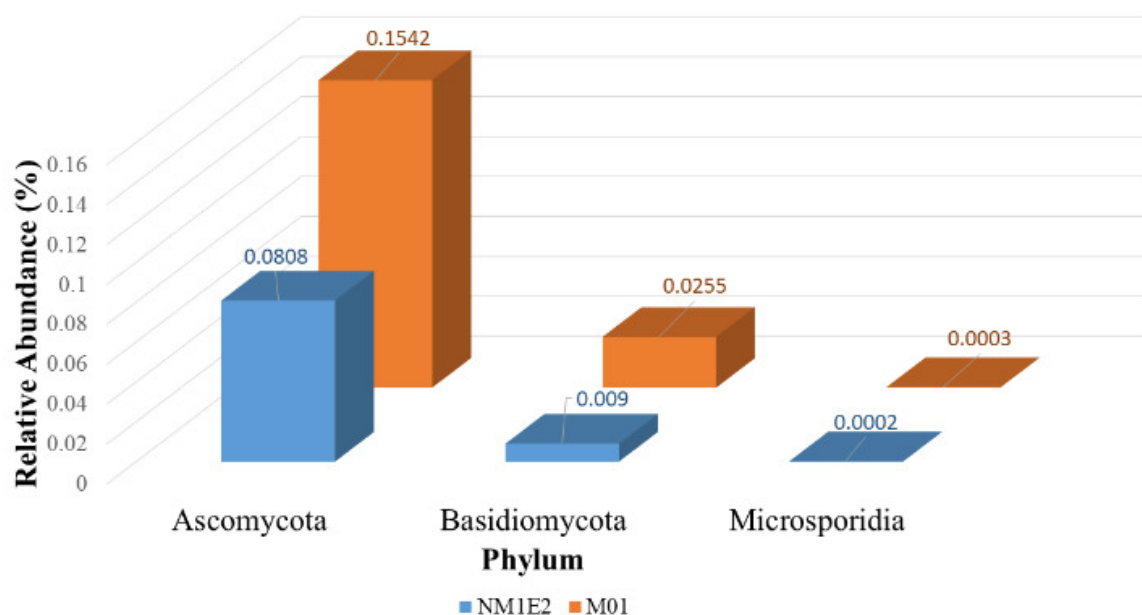
The predominance of Ascomycota across both suspected lichen-like samples supported the hypothesis that they would involve lichenized fungal components. Given their ecological relevance and potential opportunistic behavior, Ascomycota was selected for further taxonomic resolution at the genus and species levels to explore possible links to shrimp health deterioration.

To identify fungal genera with potential biological relevance in the suspected lichen-like samples, standardized abundance data (Z-score) were examined across both samples. A threshold of -0.5 was applied to retain genera with relatively elevated standardized abundance in at least one sample. Based on this criterion, three genera - *Fusarium*, *Encephalitozoon*, and *Aspergillus* - were identified as can-

didates of interest due to their relative enrichment in the fungal community (Fig. 11).

*Fusarium* had been known that they linked to infections in shrimp and other aquatic organisms. Particularly, *F. solani* had been implicated in soft tissue necrosis and shell disease in crustaceans. *Aspergillus*, although more commonly associated with terrestrial hosts, had also been detected in diseased shrimp and hatchery environments, suggesting a possible opportunistic role under stress or poor water conditions.<sup>29</sup>

*Encephalitozoon* belonged to the Microsporidia, a group of obligate intracellular fungal-related pathogens, and had been linked to chronic infections in shrimp, including growth retardation and hepatopancreatic lesions.<sup>30</sup> The consistent detection and elevated abundance of these genera in both sediment and biofilm samples suggested a potential ecological role, either as opportunistic pathogens or



**Figure 10. Composition of the main Fungus Phylum between samples (M01) and (NM1E2). The Ascomycota phylum was the most abundant fungal phylum in both samples.**

as part of dysbiotic microbial communities connected with shrimp health deterioration.

#### FUSARIUM OXYSPORUM WITH POTENTIAL IMPLICATIONS FOR SHRIMP HEALTH

Fungal species-level profiling revealed multiple taxa with elevated standardized abundance in both environmental samples, particularly within the genera *Fusarium*, *Aspergillus*, and *Encephalitozoon* (Fig. 12). Among these, *Fusarium oxysporum* was consistently detected with high standardized abundance across both samples, meeting the pre-defined selection threshold (Z-score > -0.5).

*F. oxysporum* was best known as a plant pathogen but had also been reported in aquatic environments and would linked with diseases in shrimp. Several studies detected *Fusarium spp.*, including *F. oxysporum*, in shrimp exhibiting soft tissue necrosis, exoskeletal lesions, and reduced immune responses.<sup>29</sup> The detection of *F. oxysporum* in our samples raised concern over its potential role as an opportunistic pathogen or indicator of dysbiosis.

Although the presence of *F. oxysporum* was not in itself sufficient to conclude its pathogenic role, the presence of other prominent fungi such as *Aspergillus fumigatus* and *Encephalitozoon spp.* under regular frequency suggested that the fungal component of the pond sludge microbiota deserved greater investigation regarding shrimp health. These fungal genera had also been shown in the past to be ubiquitous in aquaculture environments. For example, *Fusarium*, *Aspergillus*, and *Penicillium* were common fungal genera in shrimp pond water and sludge samples, with some of them being capable of toxin or extracellular enzyme production that could affect the animals. In addition, *A. fumigatus* had been identified to be the fungi able to trigger immune re-

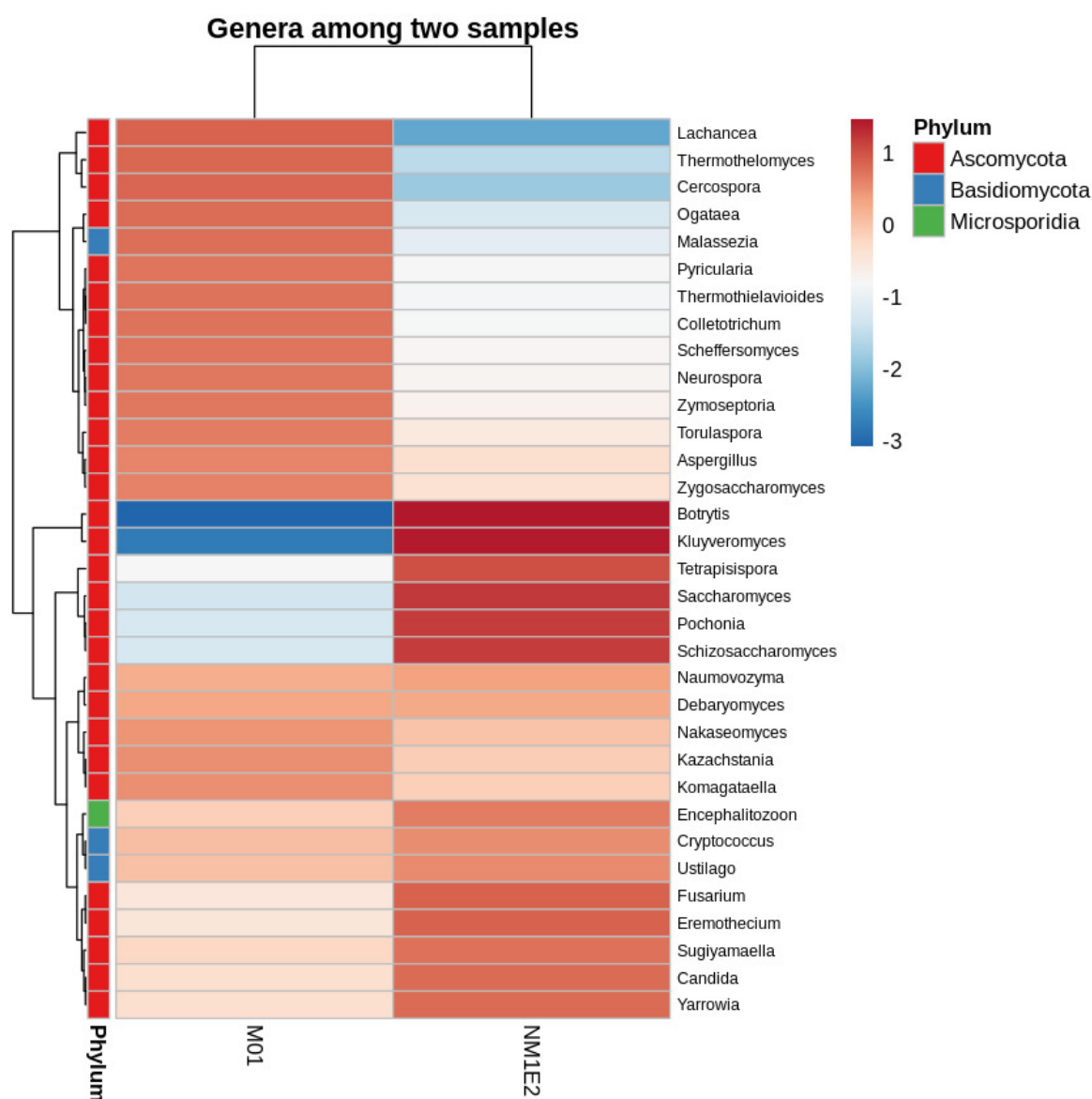
sponses or tissue damage in aquatic animals, especially in conditions where environment were unstable.<sup>31</sup> Alternatively, *Encephalitozoon spp.* belonging to the microsporidia group had been identified in pathology cases of shrimp and other crustaceans.<sup>32</sup>

Combined with the results of the present study, it suggested that fungal members in microbial community would play a more important role in shrimp pond pathology than ever thought before.

#### COMPARATIVE GENOME ASSEMBLY, PATHOGEN DISTRIBUTION, AND FUNCTIONAL POTENTIAL

Shotgun metagenomic assembly statistics highlighted the apparent structural differences between M01 and NM1E2 microbial populations. M01 generated 492,199 contigs of size ≥500 bp with an assembly length of ~593.7 Mbp, while 44,080 contigs and ~58.6 Mbp for NM1E2, suggested higher microbial diversity and genomic complexity in the sediment. In contrast, NM1E2 had greater assembly continuity, a greater N50 (1,584 bp vs. 1,364 bp), greatly greater auN (37,093 vs. 7,015.7), and a largest contig size of 688,721 bp (vs. 422,112 bp for M01) (Table 6). GC content between samples differed sharply (62.65% for M01 vs. 48.33% for NM1E2), consistent with the dominance of GC-rich sediment-dwelling bacteria in M01 and lower-GC species such as *Vibrio spp.* and fungi in NM1E2. Both assemblies contained no ambiguous bases, which guaranteed high downstream data quality analysis.

Bowtie2 read mapping to reference genomes revealed niche-specialized pathogen relative increase in proportion (Table 7). *Vibrio alginolyticus* experienced nearly comprehensive genome coverage (>99.9%) in both samples but was

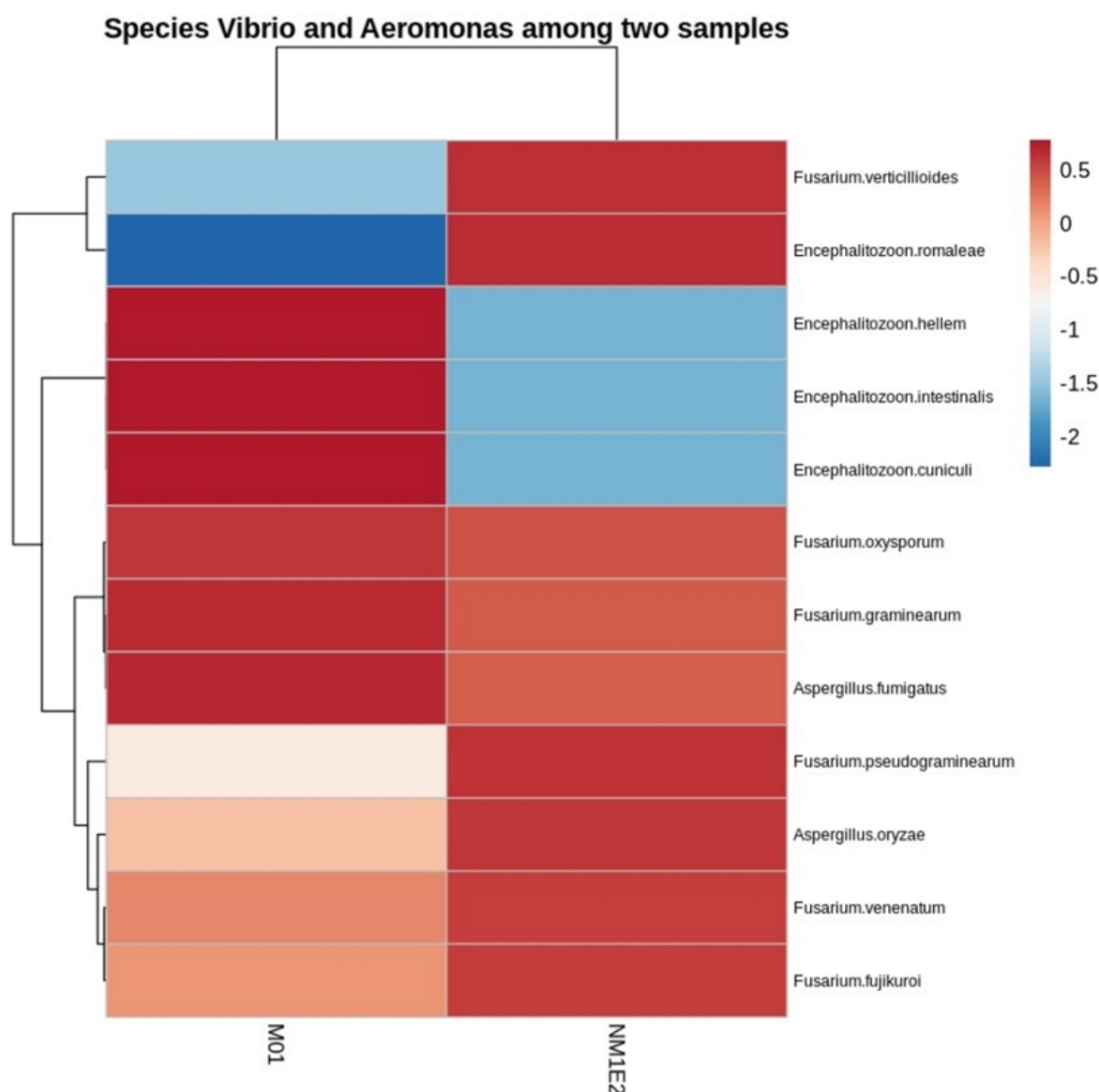


**Figure 11. Potentially Dangerous Fungus Species: *Fusarium*, *Encephalitozoon*, and *Aspergillus*.** This figure identifies *Fusarium*, *Encephalitozoon*, and *Aspergillus* as the most dangerous species in samples suspected of being associated with shrimp disease, based on their Z-score standardized abundance.

drastically more abundant in NM1E2, with chromosome 2 reaching ~80× coverage compared to 18.35× in M01. Similarly, *Vibrio campbellii* was concentrated in NM1E2 (75.28× for chromosome 1 and 40.11× for chromosome 2) over M01 (17.34× and 10.30×, respectively), and low-copy plasmids were present only in NM1E2, suggested carriage of mobile virulence or resistance determinants. Alternatively, *Fusarium oxysporum* was highly enriched in M01 with broad coverage over 14 of 18 chromosomes—particularly accessory chromosomes CP128286.1 (183.69×), CP128294.1 (174.44×), and CP128297.1 (318.29×)—and NM1E2 exhibited lower coverage over five chromosomes. These results supported ecological segregation of pathogens wherein *Vibrio* spp. possess biofilm-like aggregates and *F. oxysporum* dominate sediment.

Functional annotation of *Vibrio*-associated contigs revealed prominent features related to virulence, adaptation, and genome plasticity. Some Mu-like prophage genes

(gp28, gp29, gpT, Mu DNA-binding protein, phage virion morphogenesis proteins) were identified that indicate active lysogenic phage elements with a potential for horizontal gene transfer and enhanced pathogenicity. Motility-associated genes, including flagellar hook protein FlgE and hook-associated scaffolding proteins, were detected—in line with surface colonization behavior in *V. alginolyticus* and *V. campbellii*. Stress adaptation genes such as Na<sup>+</sup>/H<sup>+</sup> antiporters (osmoregulation), universal stress proteins, and DNA repair enzymes were present, suggesting tolerance under changing pond conditions. As a whole, the findings suggested that dominant *Vibrio* populations in NM1E2 had genetic traits that could augment persistence, pathogenicity, and adaptation within aquaculture systems, and *F. oxysporum* was a common sediment-borne fungal pathogen.



**Figure 12. Fungal species-level profiling of both samples. Species with enhanced abundance (Z-score) across both environmental samples, particularly within the genera *Fusarium*, *Aspergillus*, and *Encephalitozoon*. *F. oxysporum* was consistently detected with high standardized abundance in both samples, meeting the selection threshold.**

## CONCLUSION

Our study revealed a rich microbial population in lichen-like samples which collected from disease shrimp farms. The difference in ratio of taxa in each sample due to the position of sampling but they had quite similar types of microorganism. There was dominance of *Vibrio* species—majorly *V. alginolyticus* and *V. campbelli*—in sample collected at floating tool and facilities inside shrimp ponds (NM1E1). The sample collected near the bottom of the pond after finishing the shrimp culture (M01) revealed dense enrichments of *F. oxysporum*, a fungus that would linked with aquatic animal diseases. Its frequent presence along with other opportunistic fungi such as *Aspergillus* and *Encephalitozoon* suggested a potential though under-researched role of the fungal community in the deterioration

of shrimp health. These findings highlighted the need for additional functional and histopathological research to establish the viability of both bacterial and fungal elements to disease in aquaculture systems.

Future studies should expand the number of analyzed samples and include shrimp aquaculture ponds from diverse geographic regions across Vietnam to capture broader microbial variability and environmental influences. Additionally, the insights gained from this study could be applied to develop microbiome-informed management strategies aimed at improving the microbial balance in shrimp ponds. Such approaches might enhance disease resistance and overall productivity in aquaculture systems, contributing to more sustainable shrimp farming practices nationwide.

**Table 6. Contig Qualification**

|                                | M01         | NM1E2      |
|--------------------------------|-------------|------------|
| # contigs                      | 492 199     | 44080      |
| # contigs ( $\geq 0$ bp)       | 1069283     | 113467     |
| # contigs ( $\geq 1000$ bp)    | 150560      | 13296      |
| Largest contig                 | 422 112     | 688 721    |
| Total length                   | 593 715 348 | 58 676 830 |
| Total length ( $\geq 0$ bp)    | 811 711 485 | 82 996 913 |
| Total length ( $\geq 1000$ bp) | 365 125 770 | 38 083 386 |
| N50                            | 1364        | 1584       |
| N90                            | 576         | 585        |
| auN                            | 7015.7      | 37093      |
| L50                            | 91523       | 6226       |
| L90                            | 381.233     | 33197      |
| GC (%)                         | 62.65       | 48.33      |
| Per base quality               |             |            |
| # N's per 100 kbp              | 0           | 0          |
| # N's                          | 0           | 0          |

All statistics are based on contigs of size  $\geq 500$  bp, unless otherwise noted (e.g., “# contigs ( $\geq 0$  bp)” and “Total length ( $\geq 0$  bp)” include all contigs).

#### ACKNOWLEDGMENTS

This study was partial supported by the project TNUE-2024-15 from University of Education, Thai Nguyen University. The authors thank to the colleagues from University of Science and Technology of Hanoi and International Center of Physics, Institute of Physics, Vietnam Academy of Science and Technology for providing the place of storage sample as well as helpful advice in sampling and analysis of data.

#### AUTHORS' CONTRIBUTION

Conceptualization: Luong T. Vu (Equal), Huyen T.M. Nguyen (Equal). Formal Analysis: Luong T. Vu (Equal), Huyen T.M. Nguyen (Equal). Investigation: Luong T. Vu (Equal), Huyen T.M. Nguyen (Equal). Writing – original draft: Luong T. Vu (Equal), Huyen T.M. Nguyen (Equal), Linh T. Phung (Equal). Funding acquisition: Luong T. Vu (Equal), Hiep P. Hoang (Equal). Writing – review & editing: Huyen T.M. Nguyen (Equal), Linh T. Phung (Equal), Sang V. Nguyen (Equal). Supervision: Huyen T.M. Nguyen (Equal), Sang V. Nguyen (Equal). Methodology: Hiep P. Hoang (Equal), Thanh V.N. Nguyen (Equal), Lien T.H. Nghiem (Equal). Resources: Lien T.H. Nghiem (Equal), Hao M. Do (Equal), Mai Q. Nguyen (Equal).

#### COMPETING OF INTEREST – COPE

The authors declare no competing interests in this research.

#### ETHICAL CONDUCT APPROVAL – IACUC

Sample collection was conducted with the consent of farm owners. No endangered or protected species were involved, and no specific ethics approval was required under Vietnamese regulations for environmental sampling in aquaculture ponds.

#### INFORMED CONSENT STATEMENT

All authors and institutions have confirmed this manuscript for publication.

#### DATA AVAILABILITY STATEMENT

All are available upon reasonable request.

Submitted: October 03, 2025 CDT. Accepted: November 12, 2025 CDT. Published: January 20, 2026 CDT.

**Table 7. Comparative Genome Coverage of Key Pathogens Across Samples**

| Species                     |                |                                 | M01         |           |                    | NM1E2       |           |                    |
|-----------------------------|----------------|---------------------------------|-------------|-----------|--------------------|-------------|-----------|--------------------|
|                             |                |                                 | Mapped read | Avg depth | Genome breadth (%) | Mapped read | Avg depth | Genome breadth (%) |
| <i>Vibrio alginolyticus</i> | >NZ_CP098033.1 | Strain E110 chromosome 1        | 382         | 11.39     | 99.93              | 3057        | 26.18     | 99.92              |
|                             | >NZ_CP098034.1 | Strain E110 chromosome 2        | 6226        | 18.35     | 99.91              | 55781       | 80.00     | 99.89              |
| <i>Vibrio campbellii</i>    | >NZ_CP026315.1 | Strain BoB-90 chromosome 1      | 6015        | 17.34     | 99.91              | 49737       | 75.28     | 99.94              |
|                             | >NZ_CP026316.1 | Strain BoB-90 chromosome 2      | 387         | 10.30     | 99.91              | 3315        | 40.11     | 99.92              |
|                             | >NZ_CP026317.1 | Strain BoB-90 plasmid unnamed 1 | 0           | 0         | 0                  | 4           | 1.00      | 100%               |
|                             | >NZ_CP026319.1 | Strain BoB-90 plasmid unnamed 3 | 0           | 0         | 0                  | 1           | 1.00      | 100%               |
|                             | >NZ_CP026320.1 | Strain BoB-90 plasmid unnamed 4 | 0           | 0         | 0                  | 1           | 1.00      | 100%               |
|                             | >CP128299.1    | Strain Fo5176 chromosome 1      | 53          | 38.40     | 100%               | 0           | 0         | 0                  |
| <i>Fusarium oxysporum</i>   | >CP128298.1    | Strain Fo5176 chromosome 2      | 3545        | 138.45    | 100%               | 1672        | 82.25     | 100%               |
|                             | >CP128296.1    | Strain Fo5176 chromosome 3      | 22          | 11.80     | 100%               | 0           | 0         | 0                  |
|                             | >CP128295.1    | Strain Fo5176 chromosome 4      | 42          | 16.84     | 100%               | 0           | 0         | 0                  |
|                             | >CP128294.1    | Strain Fo5176 chromosome 5      | 9042        | 174.44    | 99.89              | 4           | 1.22      | 99.18              |
|                             | >CP128297.1    | Strain Fo5176 chromosome 6      | 834         | 318.29    | 100%               | 0           | 0         | 0                  |
|                             | >CP128292.1    | Strain Fo5176 chromosome 7      | 15          | 8.31      | 100%               | 0           | 0         | 0                  |
|                             | >CP128289.1    | Strain Fo5176 chromosome 8      | 0           | 0         | 0                  | 0           | 0         | 0                  |
|                             | >CP128290.1    | Strain Fo5176 chromosome 9      | 330         | 59.27     | 100%               | 0           | 0         | 0                  |
|                             | >CP128284.1    | Strain Fo5176 chromosome 10     | 11          | 8.40      | 100%               | 0           | 0         | 0                  |
|                             | >CP128288.1    | Strain Fo5176 chromosome 11     | 50          | 6.17      | 100%               | 0           | 0         | 0                  |
|                             | >CP128293.2    | Strain Fo5176 chromosome 12     | 283         | 55.99     | 100%               | 4           | 3.82      | 100%               |
|                             | >CP128285.1    | Strain Fo5176 chromosome 13     | 8           | 4.67      | 100%               | 0           | 0         | 0                  |
|                             | >CP128291.1    | Strain Fo5176 chromosome 14     | 167         | 18.74     | 100%               | 2           | 100%      | 100%               |
|                             | >CP128286.1    | Strain Fo5176 chromosome 15     | 536         | 183.69    | 100%               | 0           | 0         | 0                  |
|                             | >CP128283.1    | Strain Fo5176 chromosome 16     | 0           | 0         | 0                  | 0           | 0         | 0                  |
|                             | >CP128287.1    | Strain Fo5176 chromosome 17     | 7           | 3.79      | 100%               | 0           | 0         | 0                  |
|                             | >CP128282.1    | Strain Fo5176 chromosome 18     | 147         | 48.69     | 100%               | 0           | 0         | 0                  |



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

## REFERENCES

1. Workman D. Big Export Sales for Frozen or Fresh Shrimps. World's Top Exports. 2024. <https://www.worldstopexports.com/big-export-sales-for-frozen-shrimps/>
2. VASEP. Boosting shrimp exports for continued growth in 2025. Vietnam Association of Seafood Exporters and Producers. Seafood from Viet Nam. 2025. <https://seafood.vasep.com.vn/key-seafood-sectors/shrimp/news/boosting-shrimp-exports-for-continued-growth-in-2025-32600.html>
3. Vietnam Scoping Report- What Are Our Findings after Having Visited Eleven Shrimp Farms in Vietnam? Shrimp Welfare project; 2023. <https://www.shrimpwelfareproject.org/vietnam-scoping-report>
4. Vietnamese seafood journal. Aquatic diseases in 2024 decreased compared to the same period last year. 2025. <https://thuysanvietnam.com.vn/dich-benh-tren-thuy-san-nam-2024-giam-so-voi-cung-ky-nam-ngoai/>
5. Farming Sustainable & Traceable Shrimps in Vietnam. *Vietfishmagazine*. Published online 2024;29-31. <https://vietfishmagazine.com/wp-content/uploads/2024/02/VF-15.pdf>
6. Iber BT, Kasan NA. Recent advances in Shrimp aquaculture wastewater management. *Heliyon*. 2021;7:e08283. doi:[10.1016/j.heliyon.2021.e08283](https://doi.org/10.1016/j.heliyon.2021.e08283)
7. Aya S, Susan O, Vu HN, Guenwoo L. Contaminated water spillovers or peer effects? Determinants of disease outbreaks in shrimp farming in Vietnam. *Agricultural Economics*. 2025;56:210-227. doi:[10.1111/agec.12872](https://doi.org/10.1111/agec.12872)
8. del Río-Rodríguez RE, Soto-Rodríguez S, Lara-Flores M, Cu-Escamilla AD, Gomez-Solano MI. A necrotizing hepatopancreatitis (NHP) outbreak in a shrimp farm in Campeche, Mexico: a first case report. *Aquaculture*. 2006;255:606-609. doi:[10.1016/j.aquaculture.2005.12.014](https://doi.org/10.1016/j.aquaculture.2005.12.014)
9. Lightner DV. *A Handbook of Shrimp Pathology and Diagnostic Procedures for Diseases of Cultured Penaeid Shrimp*. World Aquaculture Society; 1996.
10. Karunasagar I, Karunasagar I, Umesha RK. *Microbial Diseases in Shrimp Aquaculture*. Department of Fishery Microbiology, University of Agricultural Sciences, College of Fisheries, Mangalore-575 002, India; 2004:121-134. [https://drs.nio.res.in/drs/bitstream/handle/2264/78/Karunasagar\\_chap13.pdf?sequence=1](https://drs.nio.res.in/drs/bitstream/handle/2264/78/Karunasagar_chap13.pdf?sequence=1)
11. Colorni A. Fusariosis in the shrimp *Penaeus semisulcatus* cultured in Israel. *Mycopathologia*. 1989;108(2):145-147. doi:[10.1007/bf00436066](https://doi.org/10.1007/bf00436066)
12. Lavilla-Pitogo CR, Lio-Po GD, Cruz-Lacierda ER, Alapide-Tendencia EV, De La Pena LD. *Diseases of Penaeid Shrimps in the Philippines*. Aquaculture Department, Southeast Asian Fisheries Development Center; 2000. <https://www.seafdec.org.ph/wp-content/uploads/2012/11/aem16.pdf>
13. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science*. 2017;3:e104. doi:[10.7717/peerj-cs.104](https://doi.org/10.7717/peerj-cs.104)
14. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019;20:257. doi:[10.1186/s13059-019-1891-0](https://doi.org/10.1186/s13059-019-1891-0)
15. Chao A. Nonparametric Estimation of the Number of Classes in a Population. *Scand J Statist*. 1984;11(4):265-270. <http://www.jstor.org/stable/4615964>
16. Shannon CE. A mathematical theory of communication. *Bell Syst tech j*. 1948;27(3):379-423. doi:[10.1002/j.1538-7305.1948.tb01338.x](https://doi.org/10.1002/j.1538-7305.1948.tb01338.x)
17. Simpson EH. Measurement of diversity. *Nature*. 1949;163(4148):688-688. doi:[10.1038/163688a0](https://doi.org/10.1038/163688a0)
18. Bartelme RP, Smith MC, Sepulveda-Villet OJ, Newton RJ. Component microenvironments and system biogeography structure microorganism distributions in recirculating aquaculture and aquaponic systems. *Mosphere*. 2019;4(4):10-1128. doi:[10.1128/msphere.00143-19](https://doi.org/10.1128/msphere.00143-19)
19. Goodfellow M, Williams S. Ecology of actinomycetes. *Ann Rev Microbiol*. 1983;37:189-216. doi:[10.1146/annurev.mi.37.100183.001201](https://doi.org/10.1146/annurev.mi.37.100183.001201)
20. Kaushik R, Balasubramanian R, Dunstan H. Microbial quality and phylogenetic diversity of fresh rainwater and tropical freshwater reservoir. *PLoS One*. 2014;9(6):e100737. doi:[10.1371/journal.pone.0100737](https://doi.org/10.1371/journal.pone.0100737)
21. Li M, Hu X, Ni T, et al. Efficient low-temperature wastewater treatment by *Pseudomonas zhanjiangensis* sp. nov.: a novel cold-tolerant bacterium isolated from mangrove sediment. *Front Microbiol*. 2024;15. doi:[10.3389/fmicb.2024.1491174](https://doi.org/10.3389/fmicb.2024.1491174)

22. Wackett LP, Kwart LD, Gibson DT. Benzylic monooxygenation catalyzed by toluene dioxygenase from *Pseudomonas putida*. *Biochemistry*. 1988;27(4):1360-1367. doi:[10.1021/bi00404a041](https://doi.org/10.1021/bi00404a041)
23. Liu C, Gorby YA, Zachara JM, Fredrickson J, Brown CF. Reduction kinetics of Fe(III), Co(III), U(VI), Cr(VI), and Tc(VII) in cultures of dissimilatory metal-reducing bacteria. *Biotechnol Bioeng*. 2002;80(6):637-649. doi:[10.1002/bit.10430](https://doi.org/10.1002/bit.10430)
24. Zan Z, Mao Q, Han Z, Sun J. Changes in the intestinal microbiota of farmed northern sheatfish (*Silurus soldatovi*) associated with natural bacterial infection. *J World Aquac Soc*. 2023;54(6):1575-1591. doi:[10.1111/jwas.13000](https://doi.org/10.1111/jwas.13000)
25. Gregoracci GB, Nascimento JR, Cabral A d. S, et al. Structuring of bacterioplankton diversity in a large tropical bay. *PLoS ONE*. 2012;7(2):e31408. doi:[10.1371/journal.pone.0031408](https://doi.org/10.1371/journal.pone.0031408)
26. Austin B, Zhang XH. *Vibrio harveyi*: a significant pathogen of marine vertebrates and invertebrates. *Lett Appl Microbiol*. 2006;43(2):119-124. doi:[10.1111/j.1472-765x.2006.01989.x](https://doi.org/10.1111/j.1472-765x.2006.01989.x)
27. Gomez-Gil B, Soto-Rodriguez S, García-Gasca A, et al. Molecular identification of *Vibrio harveyi*-related isolates associated with diseased aquatic organisms. *Microbiology*. 2004;150(6):1769-1777. doi:[10.1099/mic.0.26797-0](https://doi.org/10.1099/mic.0.26797-0)
28. Honegger R. 22 Fungal Evolution: Symbiosis and Morphogenesis. In: *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis*. ; 1991:319-340. doi:[10.1017/S0016672300030391](https://doi.org/10.1017/S0016672300030391)
29. Raghukumar S. *Fungi in Coastal and Oceanic Marine Ecosystems*. Springer; 2017. doi:[10.1039/C9SM02038A](https://doi.org/10.1039/C9SM02038A)
30. Stentiford GD, Feist SW, Stone DM, Bateman KS, Dunn AM. Microsporidia: diverse, dynamic, and emergent pathogens in aquatic systems. *Trends Parasitol*. 2013;29(11):567-578. doi:[10.1016/j.pt.2013.08.005](https://doi.org/10.1016/j.pt.2013.08.005)
31. Powers-Fletcher MV, Jambunathan K, Brewer JL, et al. Impact of the lectin chaperone calnexin on the stress response, virulence and proteolytic secretome of the fungal pathogen *Aspergillus fumigatus*. *PLoS ONE*. 2011;6(12):e28865. doi:[10.1371/journal.pone.0028865](https://doi.org/10.1371/journal.pone.0028865)
32. Ribeiro MFB, Passos L. Natural co-infection of *Babesia caballi* and *Encephalitozoon*-like microsporidia in the tick *Anocentor nitens* (acari: ixodidae). *J Invertebr Pathol*. 2006;93(3):183-185. doi:[10.1016/j.jip.2006.07.003](https://doi.org/10.1016/j.jip.2006.07.003)