UNIVERSITY OF VETERINARY MEDICINE BUDAPEST

CENTRE FOR BIOINFORMATICS

A metagenomic study of carp skin mucosa and fishpond water

Niamh Cox

Supervisor: Dr Norbert Solymosi & Dr Márton Papp
Centre for Bioinformatics

2024

Abstract

The skin mucus bacteriome of fish plays an important role in the health of their hosts. Despite the economic importance of the common carp (*Cyprinus carpio*), research on its skin bacteriome composition is still missing. To date, most studies on the composition of fish skin bacteriome have used amplicon sequencing, despite the limitations associated with this method. In our study, a shotgun metagenomic approach was applied to characterise the external mucus bacteriome of 8 carp specimens from two different ponds on a fish farm in Hungary. Besides the carp samples, water was also sequenced from the two corresponding ponds. Each carp skin sample was dominated by the phylum *Proteobacteria*, followed by *Actinobacteria, Bacteroidota, Firmicutes, Cyanobacteria* and *Planctomycetota.* Additionally, we have found strong concordance between the water and carp skin mucus samples, despite most studies describing an opposite relationship. Furthermore, shotgun metagenomics allowed us to apply functional annotation to the metagenomes, which revealed several metabolic functions. We present, to our knowledge, one of the first description of the common carp (*Cyprinus carpio*) skin mucus bacteriome. Even though our results showed a high level of host genome contamination, we could still provide valuable insight into the external bacterial community of this species. The presented data can provide a basis for future metagenomic studies of carp or other fish species.

Absztrakt

A halak bőrnyálkahártya-bakteriómája fontos szerepet játszik gazdájuk egészségében. A közönséges ponty (*Cyprinus carpio*) gazdasági jelentősége ellenére a bőrbakterióma összetételének kutatása még mindig hiányzik. A halak bőrbaktériomjának összetételével kapcsolatos legtöbb tanulmány eddig amplikon-szekvenálást használt, a módszerrel kapcsolatos korlátok ellenére. Tanulmányunkban shotgun metagenomikai megközelítést alkalmaztunk egy magyarországi halgazdaság két különböző tavából származó 8 ponty példány külső nyálkabaktériomjának jellemzésére. A pontyminták mellett a két megfelelő tóból származó vizet is szekvenáltuk. Minden pontybőrmintában a Proteobacteria törzs dominált, amelyet az *Actinobacteria, Bacteroidota, Firmicutes, Cyanobacteria* és *Planctomycetota* törzsek követtek. Emellett erős egyezést találtunk a víz és a pontybőr nyálkamintái között, annak ellenére, hogy a legtöbb tanulmány ellentétes kapcsolatot ír le.

Továbbá a shotgun metagenomika lehetővé tette számunkra, hogy funkcionális annotációt alkalmazzunk a metagenomokra, ami számos metabolikus funkciót tárt fel. Tudomásunk szerint a közönséges ponty (*Cyprinus carpio*) bőrnyálkahártya bakteriómájának egyik első leírását mutatjuk be. Bár eredményeink nagyfokú gazdagenom-szennyezettséget mutattak, mégis értékes betekintést tudtunk nyújtani e faj külső baktériumközösségébe. A bemutatott adatok alapot nyújthatnak a ponty vagy más halfajok jövőbeli metagenomikai vizsgálataihoz.

Table of contents

1. List of abbreviations

DNA- Deoxyribonucleic acid

RNA- Ribonucleic acid

rRNA- Ribosomal RNA

NGS- Next Generation Sequencing

PCR- Polymerase chain reaction

AMP- Antimicrobial peptides

PEAR- Paired-end read merger

NCBI- National centre for biotechnology information

KEGG- Kyoto Encyclopedia of Genes and Genomes

2. Introduction

Aquaculture, fish farming, is a crucial sector in the economies of many countries worldwide. With the global challenge of feeding 8 billion people on this planet, food security is an ever-growing concern. The escalating demand for fish can be helped through fish farming. By shifting reliance from natural fish stocks to farmed fish, will decrease the strain on the depleting fish stocks suffering from over-fishing. Aquaculture helps by providing a predictable and consistent supply of seafood for consumers worldwide.

Ensuring the health and well-being of farmed fish and the environments that they inhabit is paramount in optimising production output. Carefully monitoring water quality, implementation of disease prevention, and optimization of feed management can help maximise productivity. Studying the microbiome of the water itself will give greater insight into the health of the aquatic environment and thus the well-being of the fish themselves as it is believed that the surrounding aqua habitat influences the microbiome of the fish. Other factors such as stress, diet, and water pH also influence the microbiome of fish and these factors must be monitored to maintain the health of the fish.

The microbiome of the skin and mucosa itself of the fish is an indicator of the well-being of the fish population. Through studying the microbiome of the water and the fish mucosa, dysbiosis, a microbial imbalance, can be observed. The fish farmers can use this information to make informed decisions and adjustments to their husbandry practices to promote optimal conditions for fish growth and development.

Despite the advancement in microbiome research in aquaculture, there remains a notable gap in the study of the common carp- a species belonging to one of the largest families of freshwater fish. A study into the common carp is needed as it is a massive contributor to many European countries' economies. This research will give insight into the factors influencing the health and productivity of the common carp.

The creation of Next Generation Sequencing in 2006 has offered unprecedented insight into genomic studies. NGS has gained traction in different research fields as it has become more

accessible. By using NGS the microbiome of the lake water and the mucosa of the fish themselves can be studied in more detail.

3. Literature review

3.1 Metagenomics

Metagenomics encompasses the analysis of genetic DNA or RNA from microbial sources, enabling the sequencing and examination of nucleotides. In recent years, nucleic acid sequencing has become increasingly common in laboratories worldwide, owing to its accessibility and time efficiency (Grada and Weinbrecht, 2013). This technique is applied in various areas of research, including antimicrobial resistance, cancer research, and microbiome studies (Chiu and Miller, 2019). With metagenomics, a collection of sampled genetic material can be studied simultaneously. This is useful instead of, for example, traditional microbial culturing where the results from a sample are limited.

3.2 Next Generation Sequencing

Between 2004-2006, Next Generation Sequencing (NGS) was introduced, revolutionising the rapid sequencing of DNA or RNA. Before this, there was the Sanger sequencing method, which was a first-generation sequencing method. This method could only analyse one sequence at a time (Hu et al., 2021). The Sanger method was used for the Human Genome Project which took over a decade to map, sequence, and publish the entire human genome (Frost, n.d.). NGS can sequence thousands of genes in a significantly shorter time frame (Qin, 2019), thus producing huge amounts of data, at a fraction of the previous costs, making it more cost and time effective for researchers. NGS involves extracting nucleic acid from a DNA/RNA sample. The extracted material then undergoes 16S rRNA gene PCR amplification. These amplicons are then sequenced, and the data is processed (Boers, Jansen and Hays, 2019). This is useful for discovering rare variants, and disease-causing mutations, and for diagnosing pathological conditions (Grada and Weinbrecht, 2013).

3.3 Amplicon Sequencing

A part of NGS is Amplicon sequencing. Amplicon sequencing is “a highly targeted approach that enables researchers to analyse genetic variation in specific genomic regions.” (emea.illumina.com, n.d.). This is used when researchers want to study specific regions of a genome, like the 16S rRNA gene for bacteria, with high sensitivity and specificity. In this method, DNA is extracted from a sample and undergoes PCR amplification of the targeted regions of interest. The amplicons are then sequenced and can be analysed (Boers, Jansen and Hays, 2019).

3.4 Shotgun Sequencing

Shotgun sequencing entails fragmenting and cloning an entire genome to establish a library, where each fragment undergoes partial sequencing (Clark and Nanette Jean Pazdernik, 2013). Unlike Amplicon sequencing, specific regions are not targeted in shotgun sequencing; rather, the genome is randomly fragmented into smaller segments and then sequenced. The advantage of this approach lies in its ability to provide a comprehensive overview of a genome and to uncover unknown genes. In this method, nucleic acids are extracted and fragmented into small pieces, which are subsequently sequenced independently.

One of the benefits of using shotgun sequencing over amplicon sequencing is that a higher taxonomic resolution can be found, identifying a wider range of microorganisms. Amplicon sequencing provides a deeper understanding of targeted specific regions, but shotgun metagenomics gives an overall understanding of the entire microbial genome present.

3.5 Illumina Sequencing

Illumina sequencing is one of the many next generation sequencing platforms. Nowadays, it is more accessible to labs and thus the research including nucleic acid sequencing is rapidly growing using this type of NGS. This is why Illumina sequencing is one of the most commonly used platforms in research today (Grada and Weinbrecht, 2013).

3.5.1 Method of Illumina sequencing

- The DNA sample is prepared by cutting it into small pieces and adding special sequences, called adapters, to the ends of the DNA molecules.
- The DNA fragments with adapters are attached to a surface and copied many times over, creating clusters of identical DNA molecules.
- One of the DNA strands is removed, leaving single-stranded copies. A sequencing primer is then attached to the adapters.
- Differently labelled nucleotides are added one at a time, and the DNA strands are extended by a DNA polymerase. The process stops after each nucleotide is added.
- The flow cell is imaged to detect the incorporated nucleotides.
  The images are analysed to determine the sequence of bases in the DNA fragments. (Kircher, Heyn and Kelso, 2011)

3.6 Next Generation Sequencing in Human Medicine

NGS has become integral to research in human medicine. It enables the analysis of an individual's genetic composition, facilitating the identification of genetic disorders and the development of personalised treatment approaches. NGS also allows for the identification of cancer genomes and mutations, aiding in the formulation of targeted treatment strategies and the monitoring of disease progression. Moreover, NGS is invaluable for the precise detection of bacteria and viruses, diagnosing infectious diseases, and tracking antimicrobial resistance. Additionally, NGS has revolutionised non-invasive prenatal screening for foetal Down syndrome, achieving a detection rate of 100% (Gregg et al., 2014). The use of NGS in biomedical research has provided unprecedented insights into human genetics and, consequently, human health.

3.7 Next Generation Sequencing in Veterinary Medicine

Veterinary medicine has begun exploring the potential applications of Next-Generation Sequencing (NGS), although with the factor of cost, the possibilities are not yet as explored as with human medicine. Before NGS, culturing was the main way of finding microbial origin in a sample. Different culture conditions can be used to identify more specific pathogens, such as specialised media, temperature control and oxygen exposure, however, there are limitations to this. PCR, a nucleic acid base test, can be used in place of culturing nowadays. The advantages of using PCR are that it is cheap, time-efficient, and sensitive. However, this comes with limitations, such that PCR can only identify targets that have been predefined (Lecuit and Eloit, 2014). NGS is becoming more available, time efficient and the price is decreasing.

NGS offers numerous advantages over standard testing methods, particularly in virus identification. It enables simultaneous detection and identification of multiple virus types without requiring prior knowledge of the genome or specific targeting (Kubacki, Fraefel, and Bachofen, 2020). This approach is significantly more time-efficient and effective than routine tests. This could have enormous benefits for farming as well as veterinary medicine. For example, pork production is of substantial economic importance, and outbreaks of known or new viruses within pig farms could have a severe economic impact, as well as endangering human health with zoonotic diseases such as the swine flu pandemic (H1N1) (Kubacki, Fraefel and Bachofen, 2020). Having a method to identify novel diseases will help with future outbreaks and understanding the disease.

In 2011, a previously unidentified disease emerged among cattle in Germany and the Netherlands. After ruling out all known diseases causing symptoms such as fever, diarrhoea, and decreased milk production, researchers turned to metagenomics to analyse blood samples from the affected cattle. This led to the discovery of a novel orthobunyavirus, now recognized as the Schmallenberg virus (Hoffmann et al., 2012). The detection of this virus is incredibly significant to the veterinary community as it has continued to spread over Europe since its discovery. With this, it has brought premature, still-births and foetal malformations which have economically affected farmers. Using NGS the genomes of the virus were analysed between years to note any

variabilities, but only a few amino acid substitutions were found (Wernike and Beer, 2017). This is useful to note any mutations of the virus of outbreaks and modify treatment plans.

3.8 Next Generation Sequencing limitations

NGS offers numerous benefits, such as reduced costs, high output, and rapid results. Nonetheless, NGS also presents limitations.

- Firstly, sampling bias may occur when the sample-to-unit ratio is small, thus constraining the full potential of NGS.
- The quality of the DNA and RNA samples can also impact the accuracy and reliability of the results.
- Additionally, there exists a minimum input requirement for NGS to function effectively.
- Although the cost of NGS has significantly decreased over time, it is still expensive, especially if multiple samples need to be sequenced for a research project.

3.9 Microbiome

A microbiome is defined as a community of microorganisms that are found in an environment (Segre, 2023), including yeast, viruses, and bacteria. The bacteria are collectively known as a bacteriome. A diverse bacteriome plays a crucial role in maintaining animal health by contributing to nutrient metabolism and disease defence mechanisms. Any imbalance within this bacteriome, termed dysbiosis, heightens the animal's susceptibility to diseases. In fish, the microbiome inhabits various sites such as the skin, mucosa, gastrointestinal tract, and gills. Numerous factors, including season and diet, exert influence over the microbiome (Xavier, Severino, and Silva, 2023).

3.9.1 Microbiome of a lake

Before metagenomics, the ability to research the microbiome fully in an environment such as a lake was poor as this was carried out mainly through culturing. Targeted metagenomic

methods such as 16S rRNA Amplicon sequencing or shotgun sequencing are mainly used nowadays to research the biodiversity of a microbiome to its full extent (Mangrola et al., 2015). Having a clear and broad understanding of the microbiome of an environment such as a lake is beneficial for ensuring the environment remains healthy. With the rise in the human population, there is an ever-growing demand for food. Wild fishing isn't as sustainable as farmed fish, as well as the scalability of fish farming can more effectively keep up with the demand. With aquaculture, there is a greater possibility of infectious diseases that can impact the farm chronically (Pulkkinen et al., 2009). Having a greater understanding of the microbiological environment of the water in fish farms can give us an insight into the general health and well-being of the fish.

3.9.2 Microbiome of a fish

Microorganisms present within a lake can have both positive and negative effects on the health of aquatic organisms. Generally, the intestinal microbiome of fish is much more researched and understood than the skin mucosa of fish. The gut microbiome of a fish is critical to understand for farms as this can maximise productivity and growth. The skin and mucous membranes of fish serve as their initial defence against microorganisms. Epidermal cells help maintain osmotic equilibrium and physically prevent microorganisms from infiltrating healthy fish (Shephard, 1994). Fish mucosa, which is exposed to the environment, contains antimicrobial peptides (AMPs) like piscidin, exhibiting broad-spectrum antibacterial activity against microbes (Silphaduang and Noga, 2001). Additionally, mast cells on gill surfaces produce AMPs such as chrysophsin and pleurocidin, which also possess antibacterial properties (Murray, Gallant, and Douglas, 2003). Lysozymes can be found both on the external mucosa, as well as internal mucosa of fish. The lysosomes can lyse the peptidoglycan wall of some gram-positive and gram-negative bacteria (Paulsen, Engstad and Robertsen, 2001). The mucosa of fish can easily enter dysbiosis due to stress such as high stocking density and hypoxia (Boutin et al., 2013). Dysbiosis of the mucosa makes the fish more susceptible to becoming infected with other secondary infections.

3.9.3 Common Carp

The common carp (*Cyprinus carpio*) is a species of freshwater fish that belongs to the family Cyprinidae which is one of the largest families of freshwater fish (Rahman, 2015). In some European countries, approximately 80% of the total fish production is attributed to the common carp (Woynarovich, Moth-Poulsen and Péteri, 2010) Furthermore, it holds the distinction of being ranked as the third most cultivated freshwater fish species globally, with carp production reaching a staggering 28.9 million tonnes in 2018(FAO, 2020). The significant contribution of carp farming to the economies of various countries and the substantial demand for meeting the needs of their populations underscores the critical importance of maintaining a healthy lake microbiome. This includes ensuring the well-being of the microbial community within the aquatic environment where carp are farmed, as well as the microbiome associated specifically with the common carp species. A healthy lake microbiome plays a vital role in supporting the overall ecosystem health, which directly impacts the growth and productivity of common carp populations. The microbiome associated with carp influences various aspects of their health, including digestion, immunity, and disease resistance. Therefore, maintaining a balanced and diverse microbiome is essential for promoting the robust growth and well-being of carp populations, which in turn contributes to the sustainability and profitability of carp farming operations.

4. Materials and Methods

4.1 Samples

All carp (*Cyprinus carpio*) skin mucus samples were taken from two separate ponds at a fish farm in Hungary. Sampling was performed on 15.09.2021. The skin mucus was collected in restricted to the lateral line region of the fishes so faecal matter wouldn't impair our results. Furthermore, great care was taken during handling to avoid contamination of the samples. Four samples were taken at each pond.

At the farm where samples were collected, both scaly and mirror carp phenotypes are kept. During the sample collection, we could sample two of each at one pond, however, only one scaly and three mirror carp at the other. Furthermore, it is worth mentioning that two specimens from Pond 1 had ulcers on their skin, otherwise, all sampled fish were sterile tubes with plastic tools that were changed between each subject. The mucus collection appeared to be healthy. Details on the metadata on each sample, along with the number of reads used for classification, can be found in Supplementary File 1.

In addition to the skin mucus samples, water was collected from each pond. Water and mucus samples were frozen immediately after collection on dry ice and were subjected to shotgun metagenomic sequencing.

4.2 Sequencing

DNA purifications from the carp skin mucus samples were performed in triplicates and the resulting total DNA extracts were pooled together. All extractions were carried out using ZymoBIOMICS DNA/RNA miniprep kits (R2002, Zymo Research, Irvine, USA). For efficient lysis of the mucus samples, bead homogenization was performed using a Vortex-Genie 2 with a bead size of 0.1 mm, a homogenization time of 15 min at maximum speed, after which the Zymo Research kit's DNA purification protocol was followed. Total DNA qualities were assessed with

an Agilent 2200 TapeStation instrument (Agilent Technologies, Santa Clara, USA), and DNA quantities were measured using a Qubit Flex Fluorometer.

We closely followed all manufacturer recommendations when preparing sequencing libraries for Illumina sequencing platform (Illumina Inc., San Diego, USA). Pooled total DNA samples were used to construct libraries using the NEBNext Ultra II Library Prep Kit (NEB, Ipswich, MA, USA). Paired-end shotgun metagenome sequencing was performed on a NextSeq 550 (Illumina) sequencer using the NextSeq High Output Kit v2 sequencing reagent kit. Primary data analysis (i.e., base-calling) was performed using “bcl2fastq” software (version 2.17.1.14, Illumina).

Picture 1. Of Czikkhalas halbolt, the location where the samples were taken from.

4.3 Bioinformatic Analysis

Quality control of the raw reads was performed with FastQC v0.11.934 and MultiQC v1.1135. Read pairs were merged with PEAR v0.9.1136 before quality filtering. However, to retain as much information as possible for downstream analysis forward reads of those pairs that

couldn't be merged were also used. TrimGalore v0.6.737 was used for quality trimming of the merged and forward unmerged (see above) reads. After these steps, reads were dereplicated with VSEARCH v2.18.038. Taxonomic classification of the reads was performed with Kraken v2.1.239 to the NCBI nt database (built on: 26.12.2022).

Following classification, the bacteriome of samples was analysed in R v4.1.2 environment. The taxonomic composition of samples was examined with the aid of the phyloseq v1.38.041 R Bioconductor package. Rarefaction curves were calculated with the vegan v2.6-242 package at the species level. Good's coverage estimator was adapted based on Lin et al. at the species level by the following formula: $(1-(n/N)) * 100$, where n is the number of species with only one classified read, and N is the number of all reads classified at the species level.

Shannon index was calculated with the phyloseq package to infer alpha diversity of samples at species level classification. To avoid bias caused by the unequal sequencing depth between samples, Shannon indices were calculated as an average of 1000 iterations of random rarefications of each sample to an even depth (number of reads after rarefication: 6274). Alongside the mean Shannon indices from the random rarefications, the average number of observed species was also calculated.

Furthermore, bacteriome was analysed at the phylum level composition. However, as it can be important for a broad understanding of the bacterial community at hand and might be beneficial when comparing different hosts, deeper knowledge and inferences on its function can only be made by analysing the structure at finer taxonomic levels as well. We aimed to obtain a comprehensive description of the bacterial composition of the common carp skin mucus and consequently decided to analyse all of the dominant phyla at the genus level in more detail. We believe that a more even description can be provided by this method on the bacterial genera present in the external mucus of the carp than would be possible by a global threshold on genus-level relative abundances. Dominant phyla were selected if their relative abundance had reached at least 1% in any of the fish samples analysed.

5. Results

5.1 Skin mucus bacterial community

Due to the high contamination by the host genome in our fish skin mucus samples (the percentage of the kingdom Bacteria (mean ± SD) was 0.12 ± 0.12 in fish skin mucus, whereas it was 70.65 ± 0.47 in water samples) the rarefaction curves did not reach their plateau. Even though this might limit our conclusions on the bacteriome composition of the common carp skin mucus, our samples still provide valuable insight into the main constitution of fish skin mucus bacteriome. Rarefaction curves are presented in Supplementary Figure 1 (in Supplementary File 2), while Good's coverage values can be found in Supplementary File 1.

Shannon index was calculated for each sample (either skin mucus or water) to obtain a broader understanding of the communities by their alpha-diversity. Shannon index values and the number of observed species are summarised for each sample in Figure 1.

The bacterial composition of samples was analysed at different taxonomic levels to gain a deeper understanding of the skin bacteriome of carp. Dominant bacterial phyla and their relative abundance are presented in Figure 1 for each carp skin sample and the two water samples as well. A phylum was considered dominant if its abundance had reached at least 1% in any of the samples analysed. According to this threshold *Proteobacteria, Actinobacteria, Bacteroidota, Firmicutes, Cyanobacteria* and *Planctomycetota* were regarded as dominant and were selected for further genus level description.

Relative abundances of bacterial genera found within these phyla are presented in Supplementary Figures 2-7 (in Supplementary File 2). For better readability of the plots and to highlight the more substantial genera based on relative abundance, those that did not reach 1% relative abundance in at least one sample were aggregated in the artificial category termed "Others". The distribution of the relative abundance values associated with this category for each bacterial phylum included in the genus-level analysis is presented in Supplementary Figure 8 (in Supplementary File 2). Planctomycetota and Cyanobacteria have the lowest values for the

relative distribution for the category "Others", while in the case of *Proteobacteria*, even the lowest value is above 40%, indicating the high number of low abundance genera considering that phylum. Relative abundance values used for 1 and Supplementary Figures 2-7 (in Supplementary File 2) are presented in Supplementary File 3.

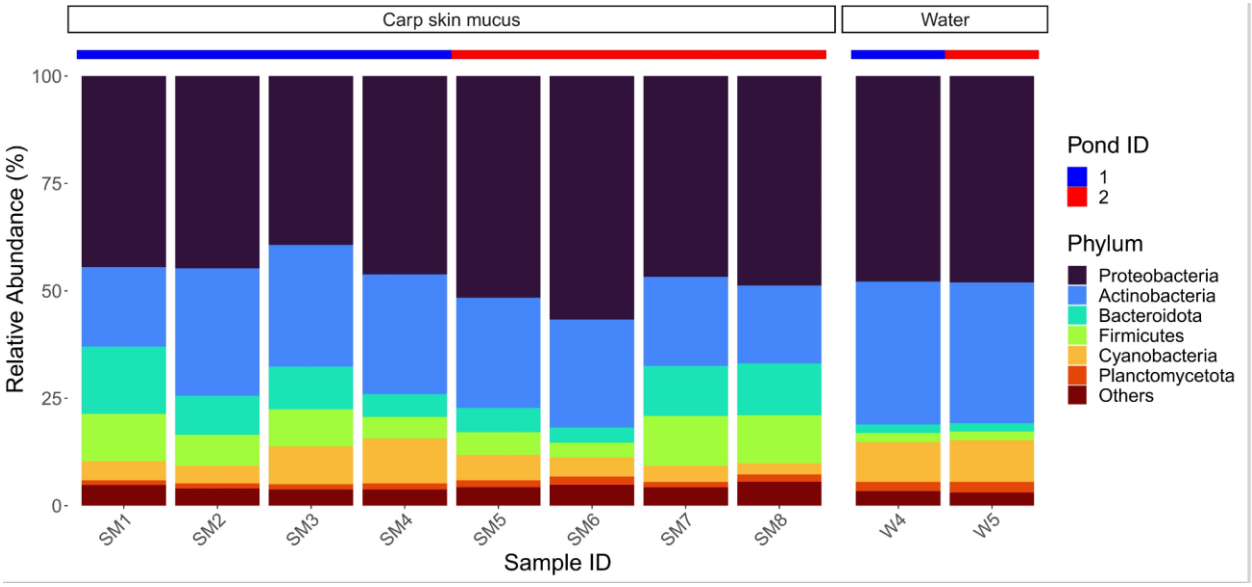

Figure 1. Relative abundance of the dominant bacterial phyla in each sample. Samples are presented on the x-axis, and corresponding bars show the relative abundance of phyla found in our analysis. Water and carp skin mucus samples are separated, and the ID of the pond the samples are originating from is indicated above the bars. Phyla that did not reach at least 1% in at least 1 sample were aggregated in the artificial category termed "Others".

5. Discussion

The external mucous membranes of fish, and more specifically, the common carp, function as a barrier to the aquatic environment. The microbiome acts as a layer for defence, homeostasis, and development of the fish (Meng et al., 2021). Dysbiosis of this microbiome can lead to an unhealthy fish and route for secondary infections. This is critical to understand given the economic importance of common carp to many countries worldwide, especially in Europe where it holds considerable commercial value.

In this study, we focused on sequencing the bacteriome found on the external mucosa of common carp, a freshwater fish species, as well as the lake water in which these fish are farmed. This research represents a significant contribution as it is one of the few studies to specifically examine the bacterial communities present on the external mucosa of common carp. The microbiome of the internal mucous membranes of fish is a much more widely researched topic as opposed to the external mucous membrane microbiome. This research is also important to understand the relationship between the aquatic environment and stressors, to the microbiome of the skin of the fish. This would help to improve the production in aquaculture.

The data generated from this study holds the potential for informing and guiding future research. By serving as a comparative baseline dataset, the findings obtained here can offer valuable insights into the microbial ecology of common carp and its aquatic environment. Understanding the healthy external microbiome of the common carp can aid in disease prevention and treatment.

The samples collected underwent shotgun metagenomic sequencing. Shotgun sequencing is becoming more widely available for research and the cost is decreasing over time. This research into the skin microbiome of the carp can be a gateway for aquaculture disease management. Disease diagnosis is being done using metagenomics in not only human medicine but now veterinary medicine too and as it continues to advance, the areas of research will also broaden.

The data was then analysed using the Shannon index, a widely used metric for quantifying the alpha diversity of species within a given community. A higher Shannon index value indicates greater diversity, with many different species present and a more even distribution of individuals among them. A lower Shannon index value suggests lower diversity, with fewer species and a less even distribution of individuals among them.

Analysis of the lake water from the two different ponds was very similar in their bacterial diversity. Specifically, Pond 1 had a Shannon index of 7.32, while Pond 2 had a slightly higher value of 7.35. Similarly, the observed species count in Pond 1 was 2707.35, and in Pond 2, it was 2668.29. Further examination of the bacterial composition revealed that *Proteobacteria* and *Actinobacteria* were the most abundant phyla present in both ponds, with *Proteobacteria* being the dominant group. The other bacteria found *were bacteroidota, firmicutes, cyanobacteria, planctomycetota*, and others. Potential environmental factors such as temperature, pH, salinity, seasonality, dissolved oxygen levels, and water depth may play significant roles in shaping the microbial communities within these lake ecosystems (Newton et al.).

The carp skin mucosa samples showed high host genome contamination. However, the results of the skin mucosa samples showed similar levels and types of bacteria as seen in the lake water samples. *Proteobacteria* dominated, followed *by actinobacteria, bacteroidota, firmicutes, cyanobacteria, planctomycetota*, and others. The ratios of relative abundance of each bacteria are similar on the mucosa of the carp, as the pond water. With only one exception of SM3 showing a higher abundance of *actinobacteria* than the pond water.

The relative abundance on each carp mucosa can be compared based on the pond number that they live in, Pond 1 and Pond 2. There does not seem to be a pattern or difference between the fish found in the two ponds. The Shannon index of the carp found in Pond 1 range from 6.98 to 7.40, and the Shannon index of the carp in Pond 2 range from 6.80 to 7.49.

Studies on the external mucous membranes of fish, including species like zebrafish (*Danio Rerio*), have used Next Generation Sequencing (NGS) to research the microbial communities residing on their mucous membranes. Zebrafish are freshwater fish belonging to the

family Cyprinidae. In the case of zebrafish, this study revealed that *Actinobacteria* emerged as the predominant phylum colonising their external mucosal surfaces, with *Proteobacteria* following in abundance. However, what stood out in this study was the stark contrast in bacterial diversity between the fish's external mucous membranes and the surrounding aquatic environment. The water in which the fish resided exhibited a significantly higher degree of bacterial diversity (Wakeman et al., 2021). The discrepancy in bacterial diversity between fish mucosal surfaces and ambient water could be attributed to several factors. Factors such as host immune responses, mucin production, and interspecies interactions might influence the composition and stability of the microbial population residing on fish external membranes.

In another notable study focusing on the Common Carp, both the internal and external mucous membrane microbiomes were sampled, sequenced, and analysed using Illumina MiSeq. The purpose of this study was to gain a better understanding of the changes that occur in the microbiome of the mucous membranes when the fish are infected with a less lethal strain of the Spring Viraemia of Carp Virus (SVCV), a pathogen that poses a severe risk to aquatic life (Meng et al. 2021). They found that on the control samples, the external mucous membranes were dominated by *Proteobacteria* and *Bacteriodetes*. However, the comparison between gut and skin samples showed notable differences in bacterial abundance, reflecting the specialised functions of these distinct anatomical sites. While the gut microbiome primarily contributes to digestion and the immune system, the skin microbiome plays a pivotal role in host defence mechanisms and environmental interactions. These researchers believe that the environment has a profound influence on the microbiome of the external mucous membranes (Meng et al. 2021). They found that the infected carp had a decreased relative abundance of *Proteobacteria* on their external mucous membranes, as found in another study of rainbow trout (Zhang et al., 2018). They also found that the levels of *Bacteroidetes* increased on the skin but decreased on the internal mucous membranes. This study gives valuable insight into the health management of aquaculture and disease mitigation (Wakeman et al., 2021).

Due to the minimal amount of research done on the microbiome of the external mucous membranes of the common carp, this research has given future research a baseline, onto which results can be compared. This can also be used by fish farmers of not only the common carp but

other species, as a baseline of the overall health of the aquatic environment and the fish themselves.

6. Conclusion

To the best of our knowledge, our results are the first to describe the bacterial community of the mucus that covers the skin of *C. carpio*. Considering the economic importance of this species and the important role that the microbiome can play in understanding and detecting various pathological conditions, studies of the bacteriome could provide valuable data. Our results have demonstrated that shotgun metagenomics can provide useful information on the skin bacteriome of this species, with particular emphasis on the potential for functional studies. A particularly important observation of our analysis is that the host genome can significantly contaminate the samples, and even though it still can provide useful information, it is preferable to be prepared in advance when designing experiments. Furthermore, our results may provide a useful basis for comparison for future studies in common carp.

7. Summary

This study is one of the very few studies that describes the microbiome found on the external mucous membranes of the Common Carp (*C. Carpio).* This species has great economic importance for many countries worldwide, and fish farming is growing with importance due to the overfishing of the natural fish populations, to cope with the demand of the growing worldwide population.

The microbiome of fish, in general, can tell researchers a lot about the overall health of not only the fish populations but also the aquatic environment in which they live. This study found a direct correlation between the microbiome of the aquatic water habitat and the external mucosa of the common carp.

In both ponds, *Proteobacteria* and *Actinobacteria* were the most abundant phyla. The other bacteria present were *bacteroidota, firmicutes, cyanobacteria, planctomycetota,* and others. Both ponds contained similar Shannon Indexes, suggesting that the amount and types of bacteria present in both ponds were very similar.

All samples from the common carp contained the same bacteria as the pond water, as well as similar Shannon indexes as the pond water. The bacteria most prevalent in the pond were also most prevalent on the skin of the carp.

This data, collected using shotgun genomics, can be used in future as a baseline for future research into the microbiome of the common carp. Fish farm producers could also use this research and future research to understand more about the well-being of their fish populations, and how to treat and mitigate disease within their fish farms.

8. References

Boers SA, Jansen R, Hays JP (2019) Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. European Journal of Clinical Microbiology & Infectious Diseases 38:1059–1070. https://doi.org/10.1007/s10096-019-03520-3

Boutin S, Bernatchez L, Audet C, Derôme N (2013) Network Analysis Highlights Complex Interactions between Pathogen, Host and Commensal Microbiota. PLoS ONE 8:e84772. https://doi.org/10.1371/journal.pone.0084772

Chiu CY, Miller SA (2019) Clinical metagenomics. Nature reviews Genetics 20:341–355. https://doi.org/10.1038/s41576-019-0113-7

Clark DP, Pazdernik NJ (2012) Molecular Biology : Understanding the Genetic Revolution. Elsevier Science

Frost A The Human Genome Project — Knowledge Hub. In: GeNotes. https://www.genomicseducation.hee.nhs.uk/genotes/knowledge-hub/the-human-genome-project/

Grada A, Weinbrecht K (2013) Next-Generation Sequencing: Methodology and Application. Journal of Investigative Dermatology 133:1–4. https://doi.org/10.1038/jid.2013.248

Gregg AR, Van den Veyver IB, Gross SJ, Madankumar R, Rink BD, Norton ME (2014) Noninvasive Prenatal Screening by Next-Generation Sequencing. Annual Review of Genomics and Human Genetics 15:327–347. https://doi.org/10.1146/annurev-genom-090413-025341

Hoffmann B, Scheuch M, Höper D, Jungblut R, Holsteg M, Schirrmeier H, Eschbaumer M, Goller KV, Wernike K, Fischer M, Breithaupt A, Mettenleiter TC, Beer M (2012) Novel Orthobunyavirus in Cattle, Europe, 2011. Emerging Infectious Diseases 18:469–472. https://doi.org/10.3201/eid1803.111905

Hu T, Chitnis N, Monos D, Dinh A (2021) Next-generation sequencing technologies: An overview. Human Immunology 82. https://doi.org/10.1016/j.humimm.2021.02.012

Kircher M, Heyn P, Kelso J (2011) Addressing challenges in the production and analysis of illumina sequencing data. BMC Genomics 12. https://doi.org/10.1186/1471-2164-12-382

Kubacki J, Fraefel C, Bachofen C (2020) Implementation of next-generation sequencing for virus identification in veterinary diagnostic laboratories. Journal of Veterinary Diagnostic Investigation 33:235–247. https://doi.org/10.1177/1040638720982630

Lecuit M, Eloit M (2014) The diagnosis of infectious diseases by whole genome next generation sequencing: a new era is opening. Frontiers in Cellular and Infection Microbiology 4. https://doi.org/10.3389/fcimb.2014.00025

Mangrola A, Dudhagara P, Koringa P, Joshi CG, Parmar M, Patel R (2015) Deciphering the microbiota of Tuwa hot spring, India using shotgun metagenomic sequencing approach. Genomics Data 4:153–155. https://doi.org/10.1016/j.gdata.2015.04.014

Meng K-F, Ding L-G, Wu S, Wu Z-B, Cheng G-F, Zhai X, Sun R-H, Xu Z (2021) Interactions Between Commensal Microbiota and Mucosal Immunity in Teleost Fish During Viral Infection With SVCV. Frontiers in Immunology 12. https://doi.org/10.3389/fimmu.2021.654758

Merrifield DL, Rodiles A (2015) 10 - The fish microbiome and its interactions with mucosal tissues. In: ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/B9780124171862000108?via%3Dihub

Murray HM, Gallant JW, Douglas SE (2003) Cellular localization of pleurocidin gene expression and synthesis in winter flounder gill using immunohistochemistry and in situ hybridization. Cell and Tissue Research 312:197–202. https://doi.org/10.1007/s00441-003-0723-3

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S (2011) A Guide to the Natural History of Freshwater Lake Bacteria. Microbiology and Molecular Biology Reviews 75:14–49. https://doi.org/10.1128/mmbr.00028-10

Paulsen SM, Engstad RE, Robertsen B (2001) Enhanced lysozyme production in Atlantic salmon (Salmo salar L.) macrophages treated with yeast β-glucan and bacterial lipopolysaccharide. Fish & Shellfish Immunology 11:23–37. https://doi.org/10.1006/fsim.2000.0291

Peteri A (2005) FAO Fisheries & Aquaculture. In: www.fao.org. https://www.fao.org/fishery/en/culturedspecies/cyprinus_carpio?lang=en

Pulkkinen K, Suomalainen L-R ., Read AF, Ebert D, Rintamäki P, Valtonen ET (2009) Intensive fish farming and the evolution of pathogen virulence: the case of columnaris disease in Finland. Proceedings of the Royal Society B: Biological Sciences 277:593–600. https://doi.org/10.1098/rspb.2009.1659

Qin D (2019) Next-generation Sequencing and Its Clinical Application. Cancer Biology & Medicine 16:4–10. https://doi.org/10.20892/j.issn.2095-3941.2018.0055

Rahman MM (2015) Role of common carp (Cyprinus carpio) in aquaculture production systems. Frontiers in Life Science 8:399–410. https://doi.org/10.1080/21553769.2015.1045629

Segre J Talking Glossary of Genetic Terms | NHGRI. In: www.genome.gov. https://www.genome.gov/genetics-glossary/Microbiome.. Accessed 7 Mar 2024

Shephard KL (1994) Functions for fish mucus. Reviews in Fish Biology and Fisheries 4:401–429. https://doi.org/10.1007/bf00042888

Silphaduang U, Noga EJ (2001) Peptide antibiotics in mast cells of fish. Nature 414:268–269. https://doi.org/10.1038/35104690

Wakeman W, Long A, Estes AM, Jozwick AKS (2021) Zebrafish, *Danio rerio*, Skin Mucus Harbors a Distinct Bacterial Community Dominated by Actinobacteria. Zebrafish 18:354–362. https://doi.org/10.1089/zeb.2021.0040

Wernike K, Beer M (2017) Chapter Three - Schmallenberg Virus: A Novel Virus of Veterinary Importance. In: ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0065352717300131

Woynarovich A, Moth-Poulsen T, Péteri A (2010) Carp polyculture in Central and Eastern Europe, the Caucasus and Central Asia: a manual. Food and Agriculture Organization of the United Nations

Xavier R, Severino R, Silva S (2023) Signatures of dysbiosis in fish microbiomes in the context of aquaculture. Reviews in Aquaculture. https://doi.org/10.1111/raq.12862

Zhang X, Ding L, Yu Y, Kong W, Yin Y, Huang Z, Zhang X, Xu Z (2018) The Change of Teleost Skin Commensal Microbiota Is Associated With Skin Mucosal Transcriptomic Responses During Parasitic Infection by Ichthyophthirius multifillis. Frontiers in Immunology 9. https://doi.org/10.3389/fimmu.2018.02972