Application of Molecular Tools to Update Aquatic Ecotoxicology Monitoring

By: Valeria Manfredi, BSc, MScR

Faculty of Medicine and Health Sciences

SVMS, University of Nottingham

Sutton Bonington Campus

Thesis submitted to The University of Nottingham for the degree of Doctor of Philosophy

September 2024

Supervised by:

Professor Kevin Gough (University of Nottingham)

Dr Helen Rees (ADAS)

Professor Tania Dottorini (Un University of Nottingham)

Dr Nadine Taylor (Cambridge Environmental Assessment)

Declaration

I hereby declare that the work present in this thesis has been written by myself and has not been previously submitted for any other degree. All procedures and work described were performed by me.

Valeria Manfredi

Date: 14th of September 2024

Table of Contents

Acknowledgements

"Happiness can be found, even in the darkest of times, if one only remembers to turn on the light."

A.D.

First and foremost, I would like to thank my primary supervisors Professor Kevin Gough and Dr. Helen Rees for their unwavering support, invaluable guidance, and insightful feedback throughout this research. Without them, this work would have not been possible. I would also like to thank my second supervisor Professor Tania Dottorini for her guidance and support during my research.

I want to express my sincere gratitude to CEA for sponsoring my PhD research. Special thanks go to Dr. Nadine Taylor for her invaluable guidance and to the Head of CEA, Adrian Terry, for their support and mentorship throughout my academic journey as well as the other members of CEA who have assisted me with sampling collections and analysis, including Marie Brown, Zoe Parker-Crosse, and Hanna Schuster.

An immense thanks to the ADAS team: Ben, Jon, Claire, and Keith for allowing me to use their facility and for their advice, support and belief in me.

A special thanks to my friends and colleagues: Shannon for spoiling me with flowers and chocolates whenever I was sad, Caleb for always finding the kindest words and Kieran for his great taste in music and constant support in the lab.

I am deeply touched by the support of my SVMS friends Nisha and Andu, who have not only been trusted colleagues but cherished companions on this journey of life. Their constant encouragement and presence have infused each day with warmth and meaning, for which I am profoundly grateful.

I am incredibly grateful to my dear friends Andrea, Livia, and Simone for being the pillars of laughter and understanding in my life. Their ability to make me smile and their unwavering support in listening to my complaints have been invaluable, and I cherish their friendship deeply.

To Ahmad, words cannot express my gratitude for being my rock throughout my Ph.D. journey. Having you by my side has been the greatest gift I could have ever received. Your presence has been a constant source of strength, and I deeply appreciate your willingness to listen and help whenever I needed it.

And finally, want to express my heartfelt gratitude to my family: my mum, my dad, Lelfo, Poppy, Gabriele and Mawi. Thank you for believing in me, even during times when I doubted myself. Your unconditional love and unwavering support have been my anchor, and I am forever grateful for your presence by my side. Thank you also for being brave enough to let me leave my nest and pursue my dreams it's because of your courage that I've been able to spread my wings and soar.

Abstract

The field of aquatic ecotoxicology studies the structural and functional disturbances induced in the short, medium, and long-term by contamination factors on ecological aquatic systems. Traditionally, organisms selected for the ecotoxicity studies are identified morphologically. Zooplankton communities are frequently employed as indicators of ecological shifts in water quality, serving to monitor such changes (Hemraj et al., 2017; Parmar et al., 2016). Because of their quick and prolific reproduction, zooplankton respond promptly to environmental alterations (Parmar et al., 2016), although they are still poorly studied.

This study aimed to apply conventional microscopy and metabarcoding detection to identify and quantify the zooplankton genera that are known indicators within mesocosms. Protocols for the identification of zooplankton genera by metabarcoding were developed. Within a mesocosm system, both methods were used to monitor seasonal changes in communities and also applied to an ecotoxicity trial using the herbicide glyphosate. The final part of the study compared metabarcoding analysis between NGS platforms: Illumina vs Nanopore sequencing.

The comparison between morphological and metabarcoding identification showed a similar number of genera when monitoring zooplankton over a season, even though both methods have detected genera absent in the other. In the pilot-study ecotoxicity trial, the selected doses of glyphosate were expected to reduce zooplankton populations (Hebert et al., 2020). However, the effects seen by both morphological and metabarcoding analysis methods were relatively subtle. The effects of glyphosate were found to be different when applying the two methodologies used to monitor community composition and significant differences were found only in metabarcoding analyses.

However, both methodologies established a NOEC lower than 5.5mg a.i./L, based on effects at the community and/or taxa level. When Illumina and Nanopore sequencing data were compared, it was demonstrated that, as expected, Nanopore metabarcoding did not generate metabarcodes with Illumina-like quality.

Overall, it was deduced that by combining both morphology and metabarcoding techniques, more extensive descriptions of zooplankton populations can be obtained. However, metabarcoding detected more genera than morphological analyses and an effect was found as expected for the Roundup™ doses on community composition, indicating that metabarcoding could potentially replace traditional morphological analysis.

List of Figures

List of Tables

Abbreviations

AMPA	-	Aminomethylphosphonic acid
˚C	-	Degree Celsius
A.I	-	Active ingredient
ANOSIM	-	Analysis of variance
ANOVA	-	Analysis of variance
BLAST	-	Basic Local Alignment Search Tool
BOLD	-	Barcode of life data system
BWA	-	Burrows–Wheeler Aligner
C1	-	Concentration of starting solution
C2	-	Volume of starting solution
CA	-	Community analysis
CEA	-	Cambridge Environmental Assessment
COI	-	Cytochrome c oxidase I
C_T	-	Cycle threshold
DCA	-	Detrended Correspondence Analysis
DCS	-	DNA control sample
DIWS	-	Depth Integrated Water Samplers
EC50	-	half maximal effective concentration
EDCs	-	Endocrine-disrupting compounds
eDNA	-	Environmental DNA
EFSA	-	European Food Safety Authority
EPSPS	-	enolpyruvylshikimate-3-phosphate synthase
FLASH	-	Fast Length Adjustment of Short reads
GATK	-	Genome Analysis Toolkit
GBH	-	Glyphosate-based herbicide
GLP	-	Good Laboratory Practice
H	-	Shannon index
HCD	-	Historical control data
HD	-	High dose
JMAFF	-	Forestry and Fisheries
LD	-	Low dose
LOEC	-	Lowest Observed Effect Concentration
M.A.D	-	Mutual Acceptance of Data
MDD	-	Minimum Detectable Difference
MOTU	-	Molecular operational taxonomic units
mtDNA	-	Mitochondrial DNA
NCBI	-	National Center for Biotechnology Information
NGS	-	Next generation sequencing
NMDS	-	Non metric multidimensional scaling
NOEAEC	-	No observed ecologically adverse effect concentration

NOEC	-	No observable effect Concentration
NTU	-	Nephelometric Turbidity unit
OCSPP	-	Office of Chemical Safety and Pollution Prevention
OECD	-	Organisation for Economic Co-operation and Development
ONT	-	Oxfor Nanopore Technology
OTU	-	Operational Taxonimc Units
PCA	-	Principal component analysis
PCR	-	Polymerase chain reaction
PGM	-	Personal Genome Machine
PRC	-	Principal response curve
qPCR	-	Quantitative PCR
R1	-	Repeat 1
R2	-	Repeat 2
rmANOVA	-	Repeated measures analysis
RT	-	Room Temperature
SDS	-	Safety data sheet
SMRT	-	Single Molecule Real Time
SOP	-	Standard Operating Procedure
uS	-	MicroSiemens
V1	-	final concentration of new solution
V2	-	final volume of the new solution

Chapter 1: General Introduction and aims

1.1 Ecotoxicology

Contamination of the environment with chemicals, which often have unintended effects on ecosystems, has reached regional and global scales (Schmitt-Jansen et al., 2008). The field of ecotoxicology which is a combination of ecology and toxicology, has developed methods and techniques to meet the resulting challenges. Ecotoxicology has been defined as 'the study of the harmful effects of chemicals upon ecosystems' (Traas and Leewen, 2007). This field is considered relatively young and originated in the middle of the 20th century when public health bodies raised concerns about the possible effects of toxic substances on the environment. This modern discipline is directly associated with the need to identify, predict, control, and reduce the negative consequences of human activities on the environment (Tarazona and Ramos-Peralonso, 2014).

The main objective of ecotoxicology is to study the structural and functional disturbances induced in the short, medium, and long-term by contamination factors on ecological systems. These factors, including all physical, chemical, and sometimes biological agents, result essentially from the direct and indirect effects of anthropogenic activities (Boudou and Ribeyre, 1997). Ecotoxicology applies toxicology evaluation to predict the effects of contaminants in biological systems (Moiseenko, 2008). It also investigates the effects of toxic compounds on wildlife (Tarazona and Ramos-Peralonso, 2014) and therefore the indirect effects of anthropogenic activities (Boudou and Ribeyre, 1997). This approach is then related to the basic principle of ecology, which aims to comprehend the relationships between organisms and their environment through a comprehensive analysis using functional units of different sizes and levels of complexity. These units could be either terrestrial or aquatic microsystems (Boudou and Ribeyre, 1997).

1.1.1 Aquatic Ecotoxicology

The field of aquatic ecotoxicology is the study of the toxic effects of chemicals at the cellular, individual, population and community levels (Hoffman et al., 2002). Water is the prime natural resource, and its quality is affected by both biotic (e.g. geological and

climatic) and abiotic factors (e.g. industrial and agricultural) (Bartram, 1996; Schmitt-Janses et al., 2008). The freshwater supply has greatly decreased due to the increase in population, urbanisation, and agricultural factors (Poonam et al., 2013). It is of fundamental importance to regularly monitor water resources and constantly assess water quality levels (Poonam et al., 2013). According to Kazi et al. (2009), the major factor determining the water quality of both marine and freshwater systems is human activities. Water quality monitoring can be achieved through chemical analyses of water; physical water measurements (pH, conductivity, oxygen levels, temperature, and light penetration); the use of bioindicators (such as plankton and other selected organisms); biological tests such as the description of abundance, density, and diversity of aquatic organisms; and with toxicity tests (Boudou and Ribeyre, 1997).

1.2 Chemicals in the water

As stated, water bodies are often contaminated with many anthropogenic toxic chemicals that can affect natural community composition. An example of how anthropogenic factors affect community composition can be seen in Fig. 1.1. The diagram shows the complex interactions within a food web and how it is influenced by both chemical and environmental factors. Various chemicals, including endocrine disruptors, insecticides, metals, pharmaceuticals, herbicides, and fungicides, affect different levels of the food web. The food web itself includes producers and decomposers at the base, herbivores and detritivores in the middle, and predators at the top. Additionally, environmental factors such as temperature, light, pH, flow velocity, organic matter, nutrients, and food quality impact the food web. The arrows indicate how these factors influence each other, highlighting the interconnected and dependent nature of these interactions within an ecosystem. This diagram emphasises the delicate balance between living organisms and their physical environment in maintaining ecological stability.

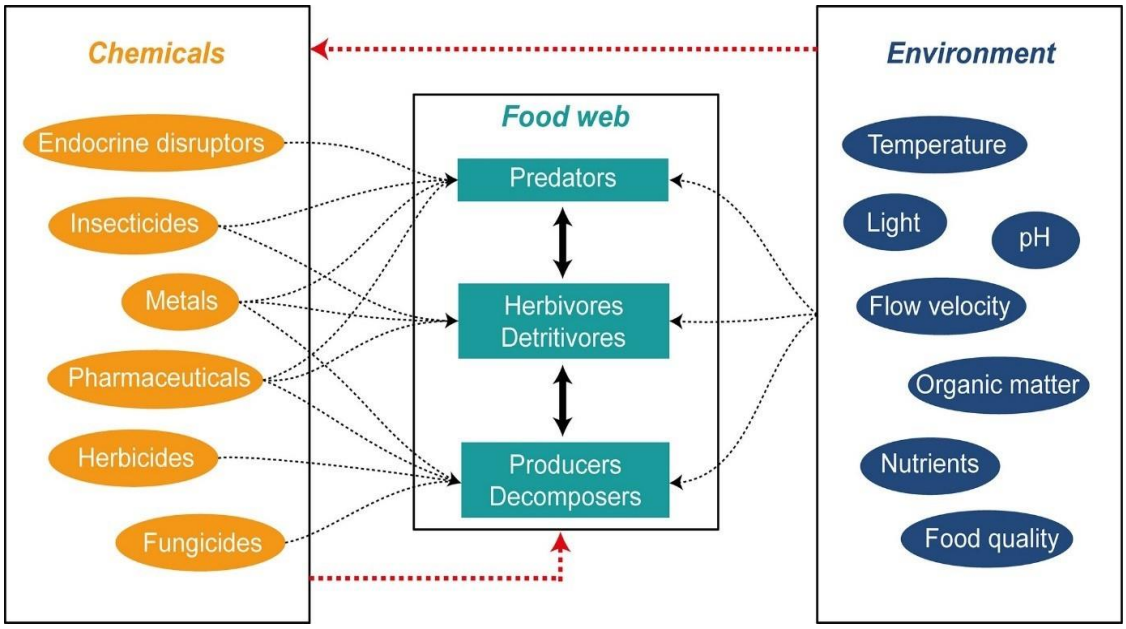


Figure 1. 1: How chemicals combined with environmental factors influence the food web.

Schematic representation of the potential effects of chemicals and environmental factors acting in concert on different trophic levels of food webs. Reproduced from (Gessner and Tlili, 2016).

Heavy metals can enter water bodies via mining activities, industrial discharges, and atmospheric depositions (Miretzku and Cirelli, 2010). This can lead to significant risks to aquatic organisms and thus, human health as once in water, the heavy metals can accumulate over time in both sediments and biota (Min et al., 2022). Heavy metals cause effects at the cellular level. Pharmaceuticals, and personal care products (PPCPs) are still an emerging group of contaminants whose effects are not well studied. They enter water bodies through wastewater discharges and incorrect disposal (Ślósarczyk et al., 2021). Depending on the compound, pharmaceuticals in water can affect aquatic organisms in different ways. One of the most common, antibacterial resistance, is in part caused by the dispersal of antibiotics in the environment. Endocrine-disrupting compounds (EDCs) are synthetic chemicals that enter water bodies through many different routes such as agricultural runoff and wastewater discharges (Jobling et al., 1998). They affect the hormonal system of aquatic organisms interfering with their development and reproduction (Diamandi-Kandarakis et al., 2009). Apart from the heavy metals, the most common chemical affecting waters are pesticides. Pesticides are man-made chemicals that are directly or indirectly inserted in the environment used to

control insects, weeds, fungi, and other pests (Khan et al., 2010; Xiao et al., 2010). Pesticide is the general term used that includes a large number of compounds like fungicides, nematicides, insecticides, herbicides, and plant growth hormones.

1.2.1 Pesticides

Pesticides are toxic chemicals used to control pathogens. They are mainly used in agricultural production to mitigate crop damage caused by other organisms (Sánchez-Bayo, 2011). Pesticides can cause serious problems as they are specifically designed to kill targeted organisms, sometimes affecting also non-targeted organisms, and they are intentionally released into the environment (Hanazato, 2001). Pesticides can enter aquatic systems either from their direct application or from terrestrial runoff. Several studies have demonstrated that pesticide concentrations in the natural environment are often high enough to kill certain kinds of organisms and therefore can affect the structure and function of natural communities (Hanazato, 2001, Helgen et al., 1988, Hatakeyama et al., 1991, Kreutzweiser et al., 2004). Because there is a great variety of pesticides in the world, specific data on aquatic contamination for any pesticide is often limited. In natural habitats, environmental surveys frequently reveal lower pesticide concentrations compared to those employed in experimental trials, while still causing toxicity to organisms that haven't been subjected to testing (Relyea and Hoverman, 2006).

It is important to consider that experimental evaluation of pesticides differs from pesticide applications in the natural environment (Gessner and Tlili, 2016). For example, in the natural environment, many pesticides are applied at a specific time of the year while others are sprayed throughout a growing season. Therefore, pesticides have specific application and frequency-time factors that will often introduce the so-called ‘pesticide-pulse’ into a system. The crucial dogma of pesticide assessment in ecotoxicology testing is to examine the effects of pesticides on ecosystems by assessing the resilience and ability of the ecosystem to return to its original state after the application, a process which can take weeks to months depending on the toxicant (Kreutzweiser *et al.* 2002).

1.2.1.1 Insecticides

Insecticides are toxic substances that are used to control insect populations including eggs and larvae (Araujo et al., 2023). They are primarily used to control pests in agricultural fields. Recently, a class of neuro-active insecticides (the neonicotinoids) have become extensively used. They are very similar to nicotine, and they act by spreading throughout all the plant tissues. The most commonly used neonicotinoids are acetamipirid imidacloprid, nithiazine and thiacloprid. Imidacloprid has been the most widely used insecticide in the past 20 years (Araujo et al., 2023).

1.2.1.2 Herbicides

Herbicides are chemical substances used to control and reduce the growth of targeted unwanted plants. They reach aquatic bodies through various pathways. Agricultural field runoff is one of the most common, followed by spray drift during chemical application and leaching through the soil. Herbicides interfere with plant processes in different ways. They are classified into different categories based on target plants, mechanism of action, and chemical composition (Matozzo et al., 2018). Some of the most common herbicides target enzymes which are essential for amino acid synthesis thus affecting plant growth. Others disrupt cell division, leading to eventual plant death (Matozzo et al., 2018). Often, non-target plants are affected by herbicide exposure, creating potentially catastrophic consequences in the aquatic ecosystem (Zhang et al. 2024).

1.2.1.2.1 Glyphosate

Glyphosate is one of the most commonly used herbicides in the world. It is a wide-ranging herbicide that was first synthesized in 1950 and subsequently, in 1974, was further developed and patented under the name Roundup™ (Martins-Gomes et al., 2022). Glyphosate works by inhibiting the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), which is fundamental in plants for the biosynthesis of aromatic amino acids. This process affects the plant's growth and eventually leads to its death (Martens et al, 2018). Glyphosate targets a broad range of weeds, and it is considered relatively safe for agricultural purposes as it has minimal risk to non-target plants and animals when appropriately used. Glyphosate can be spread into the water

directly as an herbicide, but it can also enter aquatic ecosystems indirectly through runoff from agricultural fields or urban areas (Zhang et al, 2018). Glyphosate can degrade into aminomethylphosphonic acid (AMPA) or other metabolites and persist for a long period of time (Tresnakova et al., 2021). Degraded glyphosate can also absorb to suspended particles in the water which can result in sediment accumulation (Battaglin et al, 2014). Prolonged exposure of aquatic organisms to herbicides can result in acute and chronic effects that could change aquatic systems' internal composition. One of the most sensitive aquatic organisms to herbicides is plankton. Plankton is at the base of the food chain and negative effects on them can ultimately lead to catastrophic effects throughout the ecosystem (Sarkar and Chattopadhyay, 2005). There is currently a lack of studies on glyphosate's toxicity mechanisms in planktonic organisms. According to Duke and Powles et al, (2009), one possible mechanism of action is that glyphosate can inhibit the shikimate pathway. This pathway is essential for aromatic amino acid synthesis in plants and bacteria. Glyphosate has been proven to have a toxic effect on phytoplankton and thus an indirect effect on the rest of the entire ecosystem as they are the primary producers (Wang et al., 2016).

The mechanism and time period of Roundup breakdown in the environment are crucial for understanding its long-term ecological impacts. When applied directly to water, the primary mechanism for the degradation of glyphosate is microbial activity within the aquatic environment. Aquatic microorganisms, including bacteria and fungi, metabolise glyphosate through biodegradation, converting it into aminomethylphosphonic acid (AMPA) and eventually carbon dioxide and inorganic phosphate (Franz et al., 1997). The rate of microbial degradation can be influenced by various factors, such as water temperature, pH, nutrient availability, and microbial community composition (Zabaloy et al., 2008).

The time period for the breakdown of Roundup in water can vary widely. Under favourable conditions, such as warm, nutrient-rich, and microbially active waters, glyphosate can degrade relatively quickly, with half-life values ranging from a few days to several weeks (Giesy et al., 2000). However, in less optimal conditions, such as in colder or nutrient-poor waters, the breakdown process can be significantly slower, with half-lives extending up to several months (Torstensson et al., 1989). Additionally, the persistence of glyphosate and its primary degradation product, AMPA, can be influenced by factors such as adsorption to suspended particles and sediment

accumulation, which can reduce their bioavailability for microbial degradation (Borggaard & Gimsing, 2008). Overall, while the breakdown of Roundup is primarily facilitated by microbial activity and can occur relatively rapidly under optimal conditions, various environmental factors can extend the persistence of glyphosate and its degradation products in aquatic environments.

1.3 Ecotoxicology Testing

Conventional methods for evaluating the quality of water are based on the comparison of experimentally determined parameter values with those found within the existing guidelines (Poonam et al., 2013). These guidelines are needed to establish a legal framework for hazard assessment of contaminants in water resources (Schmitt-Jansen et al., 2008).

1.3.1 Gold Standard Ecotoxicology Testing

There are several organisations that test and release policies and guidelines for ecotoxicology testing. These include Forestry and Fisheries (JMAFF); Office of Chemical Safety and Pollution Prevention (OCSPP); Organisation for Economic Co-operation and Development (OECD) and the European Food Safety Authority (EFSA) (Hanson et al., 2017).

The OECD is the leading organisation in methods experimentation of safety testing chemicals (Gourmelon and Ahtiainen, 2007). All the methodologies present in the OECD guidelines are covered by the principle of the Mutual Acceptance of Data (M.A.D) built on Good Laboratory Practice (GLP) (Gourmelon and Ahtiainen, 2007). These regulatory guides provide requirements for anyone testing and assessing chemical substances.

OECD testing has been widely applied in ecotoxicity testing of freshwater organisms. Usually, for freshwater analysis, parameters are generally required to be within the following range: temperature 20 (±1°C); pH (6.0-9.0) and oxygen level higher than 60% (Hoffman et al., 2002). For freshwater organisms, standardised acute and chronic tests have been performed on Cladocerans, particularly the species *Daphnia* (Hanazato, 2001). The OECD not only provide sanctioned methods for assessing a chemical's

effect on an organism but also provides parameters for culturing them. An example of this can be found in Table 1.1.

Table 1. 1: Daphnia culture parameters based on OECD No. 202 and 211.

Temperature (°C)	Dissolved Oxygen (%ASV)	pH	Photoperiod	Lighting (LUX)
20 ± 2	>60%	6 – 9	16 hours light - 8 hours dark	1000 - 15000

According to Hoffman et al. (2002), all toxicity tests are based on the principle that the response of living organisms to their exposure to a toxic agent is dependent on the dose and exposure level. By using this principle, aquatic ecotoxicology tests are designed to obtain a concentration-response relationship. Testing in ecotoxicology is mainly divided into four categories: acute, chronic, static and flow-through testing (Hoffman et al., 2002).

Acute toxicity tests are short-term tests designed to measure the effect of chemicals on the survival of the aquatic species during a certain period of their life-span. They are usually performed using five different concentrations of the chemical, a control and a vehicle control (e.g. a solvent) when needed (Hanazato, 2001, Hoffman et al., 2002). Chronic tests evaluate the toxic effect on aquatic organisms on growth, reproduction and behaviour. They are designed to measure the long-term effects of toxic compounds (Eggen et al., 2004, Hoffman et al., 2002). Two other toxicity tests that are less widely used are the static and the flow-through tests. Static tests are designed to test the organisms in still water. In these tests, the toxic compound is added until it reaches the required concentration, and the water is not changed during the course of the experiment. The flow-through toxicity test is usually used to check if the toxicant concentration remains at a constant level by replacing the toxicant and water either continuously (continuous-flow-through test) or at regular intermittent intervals (intermittent-flow tests) (Hanson et al., 2017). Generally, flow-through tests are thought to be more representative of natural water systems (Hanson et al., 2017).

Ecotoxicity tests are performed on either “mono-species”, “multi-species” or a combination of both. Mono-species toxicity tests are appropriate when the experiment is being carried out to observe and monitor the toxicological effects on individual characteristics such as mortality, behaviour, growth, and reproduction (Boudou and Ribeyre, 1997). There are a few disadvantages to mono-species tests: individual responses are not often sufficient to represent feedback of toxic effects in nature; identifying a single-species is expensive and time-consuming and finally, it excludes biotic interactions that would happen in a natural environment (Hanazato 1998). Multispecies tests, on the other hand, are potentially more complete, realistic, and feasible. They have ecosystem processes of nutrient cycling and energy transfer that are only present in multispecies systems. Despite multispecies testing being considered more robust than single-species testing, many industrial and governmental hazard assessment schemes use mono-species testing (Suter II, 1983).

1.3.2 Ecotoxicity Testing in Mesocosms

In the European pesticide registration process, microcosm and mesocosm experiments represent the most advanced aquatic testing tier to evaluate their environmental impact. Assessments of these studies heavily depend on determining the concentrations at which no observable effect (NOEC) occur for various population-level parameters. It is recommended that a power analysis be included for the concentration-response relationships that determine these NOECs, as well as for measurement endpoints where significant effects are not evident (Brock et al., 2015). The application of pesticides in the natural environment is affected by environmental factors such as temperature, pH, conductivity, and nutrient and organic supply. However, these conditions are not often easily replicated in the laboratory or in mesocosm experiments (Gessner and Tlili, 2016). However, mesocosm experiments do provide one of the most accurate representations of real conditions in ecotoxicity testing and can be done outdoors or indoors. Mesocosm-based experiments offer a valuable approach for studying ecological processes under controlled yet realistic conditions, bridging the gap between laboratory and field studies. Mesocosms vary in shape and size (from 1 L up to 10,000 L) and are considered an important link between field studies and laboratory experiments (Boudou & Ribeyre, 1997). Their design can vary according to the final aim of the experiment, facilitating the study of interactions in aquatic bodies long-term

by minimising environmental variability (Köhler & Triebskorn, 2013). Figure 1.2 illustrates two types of mesocosms: a flat-bottomed mesocosm (A) and a sloped mesocosm (B). The flat-bottomed mesocosm is designed to provide a uniform depth, making it particularly suitable for insecticide trials as it supports a greater variation in invertebrate species. In contrast, the sloped mesocosm creates a gradient of depths, accommodating a larger variety of plant species and is therefore predominantly used in herbicide trials. This variety allows for a more accurate simulation of natural environments where herbicides might impact multiple plant species across different water depths. This controlled setting allows researchers to manipulate specific variables, enabling the investigation of causal relationships and ecological interactions with greater precision (Carpenter, 1996). Additionally, mesocosms can be used to simulate future environmental scenarios, such as climate change or pollution, providing insights into potential ecosystem responses (Gonzalez & Bradley, 1998).

However, mesocosm-based experiments also have limitations. One significant drawback is the potential for artefacts due to the artificial nature of the setup, which may not fully capture the complexity and variability of natural ecosystems (Petersen et al., 2009). The scale of mesocosms, though larger than laboratory microcosms, remains smaller than natural systems, potentially limiting the extrapolation of findings to larger spatial and temporal scales (Petersen et al., 2009). Furthermore, the enclosed environment of mesocosms can alter species behaviour and interactions, possibly leading to results that differ from those observed in the wild (Carpenter, 1996). Maintenance of mesocosms can also be resource-intensive and technically challenging, particularly for long-term studies (Stewart et al., 2013).

In summary, while mesocosm-based experiments provide a powerful tool for ecological research, offering controlled conditions to study complex interactions and predict future changes, researchers must carefully consider their limitations. Ensuring that experimental designs minimise artefacts and acknowledging the potential constraints on scalability and natural variability are essential for deriving meaningful and applicable insights from mesocosm studies. Several studies have demonstrated the efficacy of mesocosm use in aquatic ecotoxicology to study the effects on aquatic plants, algae, invertebrates, plankton, and fish (Carpenter & Lodge, 1986; Loerracher et al., 2023).

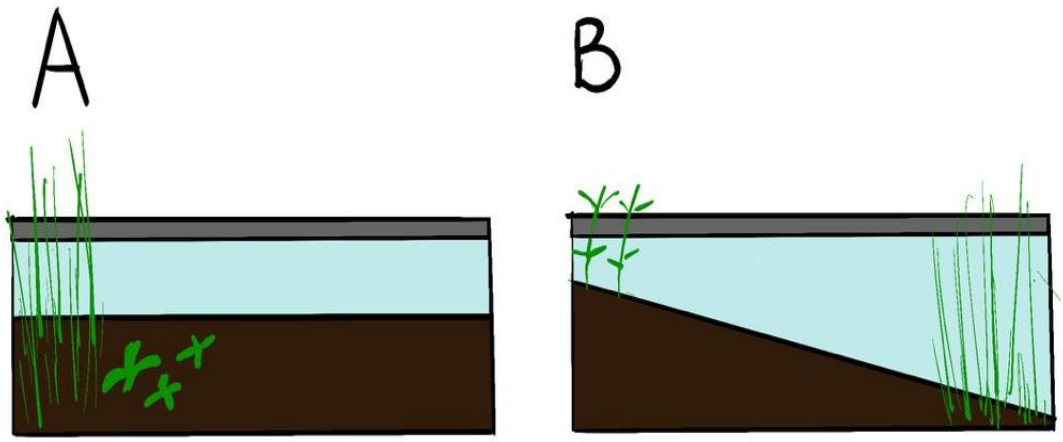


Figure 1. 2: Representation of mesocosm structure.

A represents a flat-bottomed mesocosm and B is a sloped mesocosm. The sloped mesocosms are designed to incorporate a larger variety of plant species and thus mostly used in herbicide trials. The flat bottom mesocosms are mostly used in insecticide trials as they allow a greater variation in invertebrates species.

1.4 Bioindicators of Water Quality

Usually, when a mono-species is selected for ecotoxicity testing, the organism is a so-called biomarker or bioindicator (Fossi, 1994). The term 'bioindicator' is used to refer to both biotic and abiotic factors that react to environmental changes (Parmar et al., 2016). For a species to be considered a bioindicator it should have a set of specific characteristics (Singh et al., 2013): the species must provide a representative measurable change likely to affect the whole community, thus the ecosystem; the species has to be taxonomically well documented; the species has to be feasible to survey; and finally, the species should be relatively easy to identify.

1.4.1 Plankton as Bioindicator of Water Quality

Plankton communities are often used as bioindicators to monitor ecological changes in water quality (Hemraj et al., 2017, Parmar et al., 2016). Plankton reacts rapidly to ecological changes due to their short and rapid reproduction rate (Parmar et al., 2016) but are still poorly studied. The word 'plankton' means to 'drift' and it is used to

describe the passively drifting of small plants (phytoplankton) and animals (zooplankton) in water bodies (Hays et al., 2005). Plankton is fundamental to the systems community as it decomposes organic matter. Planktonic organisms are also extremely important in carbon transportation and nutrient cycling in water bodies (Figure 1.3). This diagram illustrates the dynamics of carbon and nutrient cycling in a pelagic habitat, highlighting the role of zooplankton. At the surface, gas exchange occurs between the atmosphere and the water, with oxygen (O_2) entering and carbon dioxide (CO_2) being released. Phytoplankton near the surface acquire nutrients and contribute to biomass through photosynthesis. This biomass is transferred through the food web via processes such as predation and sedimentation. Zooplankton, positioned centrally in the diagram, play a crucial role in this system by consuming phytoplankton and being preyed upon by higher trophic levels, contributing to the biomass of these levels. Zooplankton contributes to the carbon sediment through their death and excretion, which sinks to the bottom, highlighting their role in biogeochemical cycles. Additionally, nutrient recycling occurs, supporting the ongoing productivity of the ecosystem. This diagram underscores the interconnected processes that sustain the pelagic food web and the vital role of zooplankton in maintaining ecological balance.

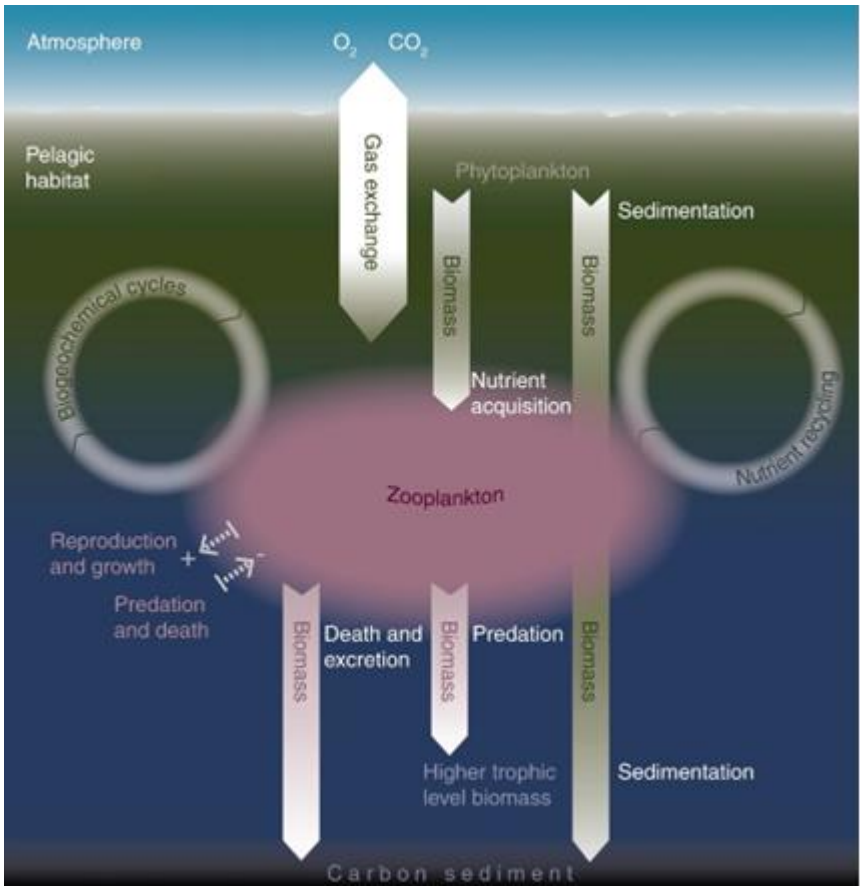


Figure 1. 3: Planktonic role in carbon transportation.

Schematic representation of the roles of phytoplankton and zooplankton in carbon transportation, nutrient cycling, and food chains. Reproduced from (Brierley, 2017).

Plankton assemblages, are fundamental components of aquatic ecosystems, playing crucial roles in biogeochemical cycles, food web dynamics, and primary production. The composition and abundance of these communities are highly influenced by various environmental, abiotic, and seasonal factors. Environmental conditions such as temperature, salinity, and nutrient availability affect plankton dynamics; for instance, temperature influences metabolic rates and reproductive cycles of plankton species, affecting their population dynamics and community structure (Edwards & Richardson, 2004). Salinity variations, often resulting from freshwater inflows or oceanic currents, can alter species composition due to the varying tolerance levels of planktonic organisms (Cloern & Jassby, 2010). Nutrient availability, driven by processes such as upwelling, river discharge, and anthropogenic inputs, is a key determinant of phytoplankton productivity, which in turn supports higher trophic levels (Anderson et

al., 2002). Additionally, abiotic factors including light availability, pH, and dissolved oxygen significantly influence plankton communities. Light availability affects photosynthetic activity in phytoplankton, influencing their growth rates and distribution patterns (Behrenfeld et al., 2006). Changes in pH, often associated with ocean acidification, can impact calcifying organisms like certain species of zooplankton, potentially altering community composition and ecosystem function (Fabry et al., 2008). Dissolved oxygen levels, influenced by thermal stratification and eutrophication, can create hypoxic conditions detrimental to many plankton species (Diaz & Rosenberg, 2008). Seasonal changes also bring about significant shifts in plankton assemblages due to variations in temperature, light, and nutrient cycling. In temperate regions, spring and autumn are typically associated with phytoplankton blooms driven by increased nutrient availability and favourable light conditions (Sommer et al., 1986). During summer, stratification can limit nutrient mixing, often leading to reduced productivity or shifts towards species adapted to low-nutrient conditions (Reynolds, 2006). Winter periods are generally characterised by lower temperatures and light levels, resulting in decreased plankton activity and biomass (Smayda, 1997).

1.4.1.1 Zooplankton

Zooplankton are animals that vary in size: microzooplankton (<200 μm), mesozooplankton (0.2-2 mm), microzooplankton (2-20 mm) and finally megazooplankton (>20 mm) (Parmar et al., 2016). They usually are not efficient swimmers, and they often rely on water movement as a moving mechanism. They are essential organisms in water bodies as they occupy a central position in the food chain by transferring energy from primary producers (phytoplankton and bacterioplankton) to trophic organisms such as fish (Parmar et al., 2016, Hanazato, 1998). Zooplankton not only support trophic organisms, but also microbial communities. Microbes colonise both zooplankton's faeces and corpses making them rich sources for carbon detrital feeders as shown in Fig. 1.4 (Richardson, 2008). The figure showed how zooplankton ingested phytoplankton and microzooplankton, incorporating particulate organic carbon (POC) into its system for growth and reproduction. The zooplankton's metabolic activities included excretion, which released dissolved organic carbon (DOC), and respiration, which emitted carbon dioxide (CO_2). Additionally, egestion resulted in the

expulsion of undigested material as POC. Some organic matter also leaked from faecal pellets as DOC.

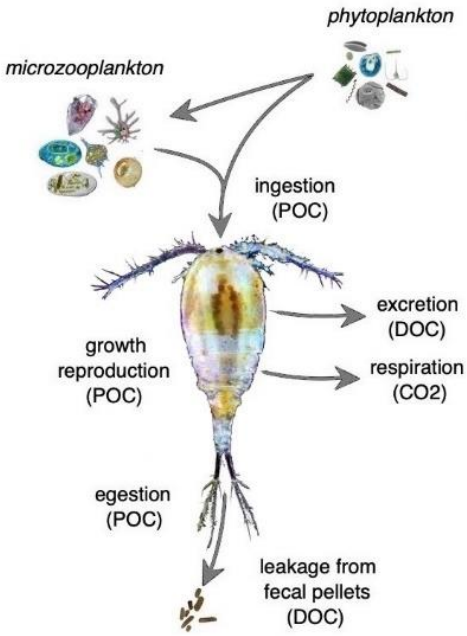


Figure 1. 4: The feeding cycle of a zooplankton.

Zooplankton feeding cycle where DOC = dissolved organic carbon, POC = particulate organic carbon. Reproduced from (Møller et al., 2003).

Zooplankton communities are highly diverse (Pennak, 1946) and are divided into 3 orders: Cladocera, Copepoda and Rotifera (Fig. 1.5).

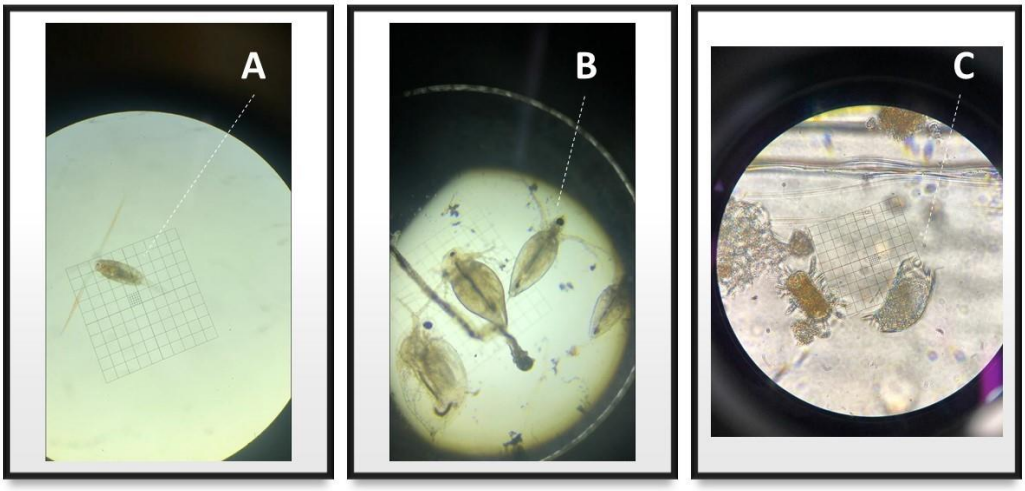


Figure 1. 5: Example of morphological diversity across the three zooplankton groups.

Where A- showed a *Copepoda (Cyclopoidae); B-* a *Cladocera (Daphnia* and *Simoncephalus*) and finally, C- *Rotifera (Mytillina*).

Copepods are present in almost all aquatic bodies. They have developed adaptive mechanisms to survive even in the most unstable habitats (Heuschele and Selander, 2014). Migration and dormancy are the two most common behaviours found in copepods to avoid both biotic and abiotic conditions that threaten their growth and reproduction. Utilising dormancy, however, the whole Copepods population can hibernate for a specific period allowing them to avoid harmful conditions (Hanazato, 2001).

Rotifera are amongst the smallest freshwater metazoans, and they are extremely varied in morphological features even though they are bilaterally symmetrical (Sládeček, 1983). Rotifers have the highest reproduction rates amongst the zooplankton groups, reaching over 1000 individuals per litre. Because of this, rotifers often dominate the zooplankton community making them one of the best groups for aquaculture and ecotoxicological analyses (Wallace and Snell, 2006).

Cladocerans are the second most abundant group of zooplankton and consist of around 620 known species (Forró et al., 2007), they inhabit most types of water bodies. (Forró et al., 2007). Cladocerans act as an important link in the food chain as most of them are considered herbivorous (feeding on phytoplankton) and are preyed upon by fish and

invertebrates. Thus, they have an important role in the transfer of energy in the aquatic food-chain (Sarma et al., 2005). Most of the detailed studies on Cladocerans have been made on the genus *Daphnia* and because of that, daphnids are used as model organisms, including in ecotoxicity tests.

The *Daphnia* genus is probably one of the most studied in ecology. It has been studied due to its role as a primary consumer in aquatic bodies, behaviour, evolution, and sexual and asexual reproduction cycle (Stollewerk, 2010). The genus *Daphnia* has an enormous variety of species (more than 100 known) and whilst they occupy most types of aquatic bodies, they are absent from extreme habitats as they are incredibly sensitive to environmental factors (Ebert, 2005).

Sexual reproduction in zooplankton is a critical aspect of their life history, influencing genetic diversity and population resilience. Many zooplankton species exhibit complex reproductive cycles, alternating between asexual and sexual reproduction depending on environmental conditions (Allan, 1976). For instance, during favourable conditions, asexual reproduction allows for rapid population growth, while sexual reproduction is often triggered by environmental stressors such as changes in temperature, food availability, or population density (Stelzer, 2011). This switch to sexual reproduction results in the production of resting eggs or cysts that can withstand adverse conditions, ensuring species survival during periods of environmental stress (Hairston & Kearns, 2002). An example of this coping mechanism can be seen in the *Daphnia* genera.

When not under stress, *Daphnia* reproduces asexually (parthenogenesis) by generating diploid eggs which are released into the water after approximately 3 days. However, if triggered by external stimuli, *Daphnia* can reproduce by meiosis generating haploid resting eggs that require fertilisation and a dormancy period to develop and resume in response to external stimuli. The cycle between parthenogenesis and sexual reproduction is called cyclic parthenogenetic and is very common in most *Daphnia* species (Figure 1.6A and 1.6B).

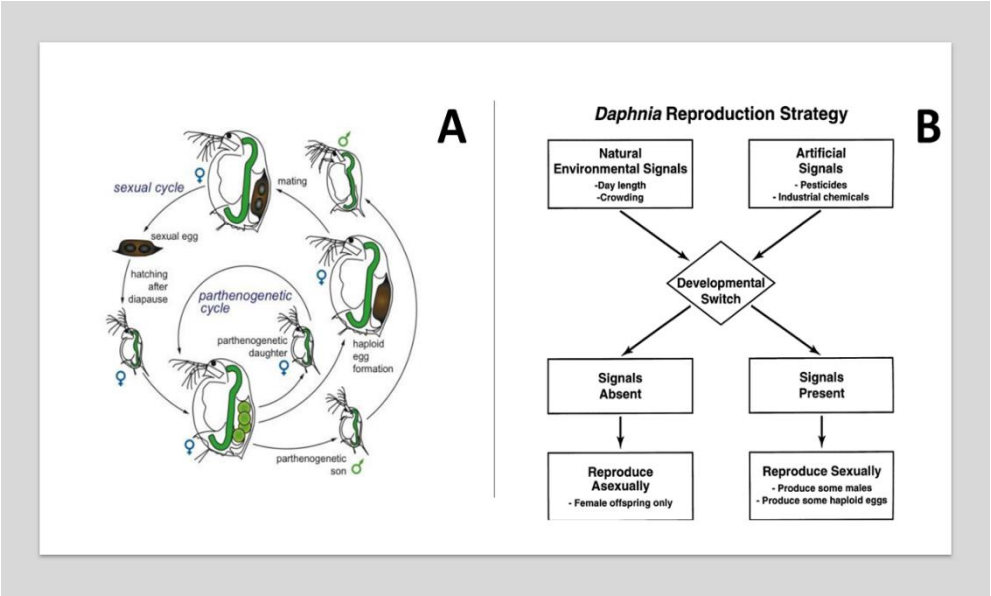


Figure 1. 6: *Daphnia* reproduction cycle.

A) Sexual and the asexual (parthenogenetic) life cycle of a *Daphnia*; B) *Daphnia* reproductive strategy. Adult females can produce three different kinds of offspring, depending on environmental conditions. Diploid eggs are produced asexually and develop into females or males. Adapted from (Dodson et al., 2010).

In addition to reproductive strategies, zooplankton have developed a range of behavioural adaptations to cope with environmental stress. Vertical migration is one such strategy, where zooplankton move to deeper waters during the day to avoid visual predators and harmful UV radiation, ascending to the surface at night to feed (Lampert, 1989). This diel vertical migration helps reduce predation risk and optimises feeding efficiency, contributing to the survival of zooplankton populations. Moreover, some zooplankton species exhibit changes in their activity levels and feeding behaviours in response to environmental cues, thereby reducing metabolic rates and conserving energy during periods of low food availability (Ringelberg, 2010).

1.4.1.1.1 Indirect Effects of Food Source Loss on Zooplankton Populations

The indirect effects of losing a food source on zooplankton populations can vary significantly depending on several factors, including the type of food source, the adaptability of the zooplankton species, and the overall stability of the ecosystem. Typically, the immediate response to a loss of food source is a decline in zooplankton reproductive rates and survival, observable within a few days to weeks (Kiørboe, 2008).

This initial impact is followed by more profound effects on population dynamics over the ensuing months. Species with flexible feeding strategies may switch to alternative food sources, mitigating the impact to some extent, but specialised feeders are likely to experience more severe and prolonged declines (Hansen & Bech, 1996). Furthermore, the cascading effects through the food web can lead to longer-term alterations in community structure and function, potentially taking several seasons to years to stabilise (Persson et al., 2007). In ecosystems where primary producers are affected by environmental changes, the indirect effects on zooplankton can be exacerbated, leading to prolonged periods of instability and reduced biodiversity (Sommer et al., 1986). Therefore, while the immediate effects of food source loss are often rapidly apparent, the full extent of indirect impacts on zooplankton populations and their recovery can span extended periods, highlighting the complexity of trophic interactions and ecosystem resilience.

1.4.1.2 Phytoplankton

There are seven phytoplankton orders: *Charophyta, Chlorophyta, Choanozoea, Cryprophyta, Cyanophyta, Euglenophyta* and *Ochrophyta*. Phytoplankton is mostly found in the upper portion of water bodies as they contain chlorophyll and require light to live. Although they compose less than 1% of the earth's biomass, they are responsible for 50% of the world's oxygen production (Borics et al., 2020). It is thought that there are over 20,000 different species of phytoplankton. Phytoplankton is strongly affected by nutrient distribution and environmental changes. As stated before, temperature is one of the main environmental factors that affects the plankton community, the most significant being the change in thermal stratification leading to a weaker vertical mixing process (Sommer, 2012).

1.4.1.3 Plankton Identification Methods

Both zooplankton and phytoplankton have traditionally been identified through morphological identification based on the phenotypic characteristics of the individuals (Friedheim, 2016, Savin et al., 2004). The study of freshwater planktonic organisms has been challenging as the taxonomical identification requirements are often expensive and time-consuming, usually requiring a specialist in the field (Elías-Gutiérrez et al., 2018). Friedheim (2016) argues that morphological identification has been the basis of

identification and can be accurate, although appropriate taxonomical expertise may be too rare for the number of studies needed (Galan et al., 2012). Galan also stated that morphological identification is extremely complicated as animals go through different life stages where often, during the larval stage the morphology is completely different from the adult one. Thanks to the use of molecular techniques, plankton identification is becoming easier and more accurate, although misidentification can occur or be difficult due to a lack of data on reference databases (Djurhuus et al., 2018). Many researchers have compared the results of morphological and molecular identification methods on zooplankton to check differences and establish which one is more accurate. The majority of these studies showed similar results indicating that molecular methods take less time and theoretically do not require taxonomic expertise (Djurhuus et al., 2018, Savin et al., 2004, Elías-Gutiérrez et al., 2018). Despite the successful amplification, some limitations were still found in these studies such as the rapid DNA degradation due to the standard storage method (as seen in Elías-Gutiérrez et al., 2018) or the scarcity of the reference databases (as seen in Djurhuus et al., 2018).

1.5 Molecular Biology Classification Methods

Molecular tools have been extremely useful in assessing ecosystem diversity in various communities. This can be especially helpful in the case of the meiofauna as, due to their small size, organisms belonging to this group are hard to identify morphologically (Feio et al., 2020). Molecular barcoding uses molecular markers in an organism's DNA to identify species. Molecular markers are sections of an organismal genome that can be obtained by the polymerase chain reaction (PCR) (Mullis et al., 1986). There are several types of molecular markers such as microsatellites, and minisatellites. The identification of organism-specific molecular markers allows scientists to quantify genetic diversity, track individual's movement, measure inbreeding, characterise new species, identify species from mixed samples, and retrace historical patterns of dispersal (Kirk and Freeland, 2011).

1.5.1 PCR Analysis

PCR can vary depending on the experiment. A type of PCR that has been used in plankton sequences analysis is quantitative PCR (qPCR) (Endo et al., 2010, Loh et al.,

2014). The innovative aspect of qPCR is that the standard gel electrophoresis to visualize PCR products after the completion of the reaction is not needed anymore as qPCR shows in real-time the quantification of the PCR product (VanGuilder et al., 2008). qPCR reactions are performed inside a thermocycler that detects double-stranded DNA products labelled by a fluorescent probe. qPCR is divided into three main phases as the number of cycles and the amount of product increase (Figure 1.7). The initial phase is the exponential phase in which there is exponential growth, the linear phase and the plateau phase where the accumulation of the product ends. The fluorescence is monitored during the whole process. The higher the number of DNA molecules in the sample, the faster the fluorescence will increase during the cycles. The cycle threshold (C_T) is defined as the number of cycles taken to obtain an arbitrary amount of fluorescence (VanGuilder et al., 2008). Lower Cycle threshold (CT) values mean a higher initial copy number of the target gene. The qPCR method has been extensively used in ecotoxicology studies (Juan-Garcia et al., 2023 and Liu et al., 2017). However, this method has limitations when analysing mixed samples as it includes target sequence specificity depending on the design of the primers and probes, which may limit their applicability to multiple sequences and taxa (Bonacorsi et al., 2021)(Aerie et al., 2009).

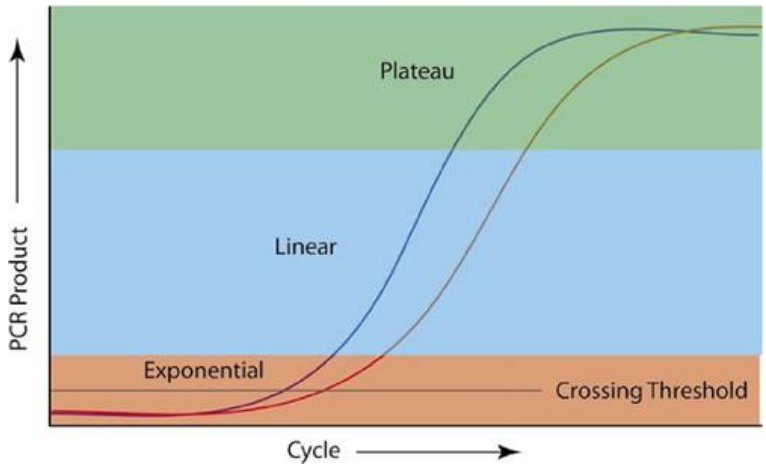


Figure 1. 7: The three phases of qPCR: Reproduced by (VanGuilder et al., 2008).

1.5.2 DNA-Barcoding

Many studies are still limited to a narrow region of the genome, making it difficult to generalise about organisms and their evolutionary history (Narum et al., 2013). DNA-barcoding allows the identification of unknown specimens using standardised species-specific genomic regions called DNA barcodes (Shokralla et al., 2012). The mitochondrial gene cytochrome c oxidase I (COI) has been proven to be the core of a global bio identification system for animals (Hebert et al., 2003). It is important to stress that the quality of the DNA in samples has a crucial role in the analysis outcome. Different protocols for preserving DNA and preventing it from degradation have been developed and vary based on the studied organism (Liu et al., 2020).

Genetically, each species and individual are unique, meaning that no one genome is identical to the others (Galan et al. 2012). By targeting areas of high sequence conservation, mitochondrial DNA for example, that show little change in terms of individuals within a species but identifiable changes between species, DNA-barcoding can be used to identify species.

The DNA-barcoding method consists of Sanger sequencing, a standardised short DNA-fragment (barcode) that is unique for each species and therefore used for species identification (Bucklin et al., 2016). 'DNA-metabarcoding' refers to multi-species identification and extends DNA-based species identification to communities of individuals belonging to different species. This method uses parallel-sequencing of bulk samples or environmental samples (eDNA) to generate faster and more detailed data using high-throughput sequencing technologies such as Illumina or Ion-proton (Cristescu, 2014). Both methods rely on a comparison between the DNA sequences obtained and a reference database for the species of interest (Corell and Rodríguez-Ezpeleta, 2014).

The term "barcode gap" refers to the difference in genetic variation between species and within species in DNA barcode sequences. The barcode gap is observed when there is a clear separation between the genetic distances within species (intraspecific variation) and between species (interspecific variation) (Hebert et al., 2003; Meier et al., 2006). The concept of the barcode gap is crucial for species identification using DNA barcoding techniques. It suggests that genetic distances between sequences of individuals within the same species are typically smaller than the distances between

sequences of individuals from different species. This clear distinction allows for the reliable identification of species based on their DNA barcode sequences. The barcode gap has implications for various fields such as taxonomy, ecology, and biodiversity assessment, as it provides a molecular tool for rapid and accurate species identification, especially in cases where traditional morphological identification is challenging (Hebert et al., 2003; Meier et al., 2006

DNA barcoding has probably been the most applied molecular method in both marine and freshwater plankton identification studies (McManus and Katz, 2009). However, a considerable barrier to effectively using DNA barcodes for freshwater zooplankton is the limited amount of sequences in the reference databases: Barcode of Life Data System (BOLD) (Ratnasingham et al., 2007) and The National Center for Biotechnology Information (NCBI) (Elías-Gutiérrez et al., 2018). The DNA-barcoding method (Figure 1.8) consists of sequencing a standardised short DNA fragment (barcode) that is unique for each species and therefore used for species identification (Bucklin et al., 2016). Barcoding primers are used to amplify the gene from genomic DNA by PCR. For example, the Folmer COI primers are designed to target a specific region of the COI gene (approximately 650 bp), allowing for its amplification through PCR (Folmer et al., 1994). The product obtained by PCR is usually then analysed using Sanger sequencing. The method relies on a comparison between the product obtained and a reference database (Corell and Rodríguez-Ezpeleta, 2014).

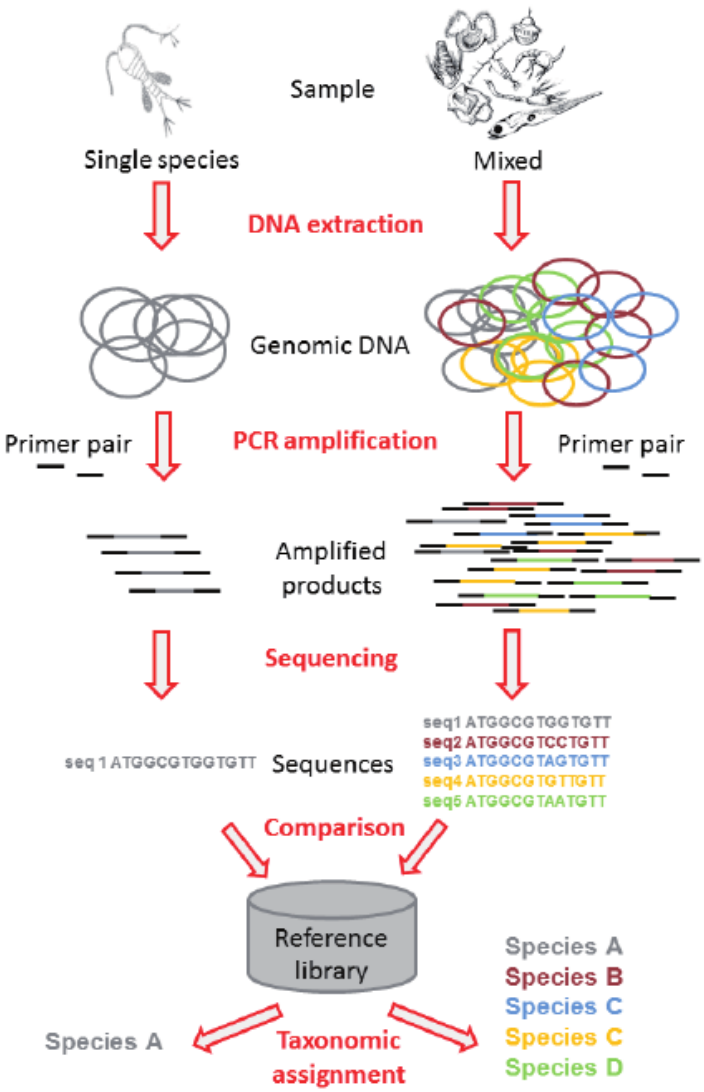


Figure 1. 8: Comparison between DNA barcoding and metabarcoding.

Schematic representation of the processes of DNA barcoding (left) and metabarcoding (right). Reproduced from (Corell and Rodríguez-Ezpeleta, 2014).

A barcode gene must meet certain conditions: firstly, it must be distinct between as many taxa as possible, but with defined barcode gap, and it must have conserved regions to facilitate primer design. For zooplankton species, two commonly used DNA-barcodes are the cytochrome c oxidase subunit I (COI) and the nuclear small-subunit ribosomal RNA (18S rRNA) (Corell and Rodríguez-Ezpeleta, 2014). COI is one of the most used genes for barcoding metazoans. According to Goodall-Copestake et al. (2012), CO1 popularity as a DNA-barcode derives from the availability of several sets of conserved PCR primers (such as Folmer et al. 1994), and from its dual purpose for

estimating both intra-specific variation and inter-species identification. Similarly, the 18S rRNA gene has been previously used as a gene for zooplankton barcoding thanks to its highly conserved regions. However, 18S is still not considered as efficient as COI possibly due to its lower evolutionary rate which makes it harder to differentiate closely related taxa (Corell and Rodríguez-Ezpeleta, 2014).

An alternative marker to barcode phytoplankton is the 16S rDNA region (Kitamura et al., 2016, Valenzuela-González et al., 2016). However, many researchers have encountered problems in designing primers for this region as it is hypervariable (Valenzuela-González et al., 2016). The 23S rDNA region can also be used, researchers have been able to design better primers for phytoplankton taxa as this region is more conserved and allows high levels of coverage (Yoon et al., 2016).

1.5.3 Metabarcoding

As previously stated, metabarcoding has been used in zooplankton composition analysis as it better reflects any alteration of zooplankton communities influenced by environmental factors (Yang et al., 2017). DNA metabarcoding is a developing approach that can identify communities of species from a mixed sample through direct sampling (which can either be bulk DNA e.g. organisms isolated from sediment collected with sediment cores or environmental DNA (eDNA), based on high-throughput sequencing (HTS) of a specific DNA-barcode (Liu et al., 2020).

eDNA describes the genetic material in faeces, urine, blood, sperm, skin, hair etc or naked DNA that is present in environmental samples (e.g. sediment, water and air) (Barnes et al., 2014). eDNA can be used in metabarcoding analyses. Even though eDNA assays are relatively new to the field, they have already proved to have great potential in biological monitoring. One of the major positive impacts of using an eDNA sample is the fact that sample collection does not always cause habitat disturbance or destruction contrary to normal sampling methods such as the use of traps or nets. It offers a non-invasive and sensitive approach to biodiversity monitoring and can detect species that are difficult to survey using traditional methods. However, eDNA methods cannot be used to determine sex ratios or body conditions of organisms (Goldberg et al., 2016). Also, eDNA degradation has an impact on the quality of analysis as eDNA is subject to biotic and abiotic factors especially in warm and tropical regions (Goldberg

et al., 2016). With traditional direct sampling metabarcoding, DNA quality is higher. Metabarcoding is considered the future of monitoring as it alleviates the expense and the amount of time and effort taken for morphological identification of zooplankton communities (Cristescu, 2014). However, there are still some limitations and weaknesses (Yang and Zhang et al. 2018). For example, multi-PCR reaction replicates are needed to reduce the biases of PCR that can result in the preferential amplification of particular sequences.

1.5.3.1 Next generation sequencing

Next-generation sequencing (NGS) is a fast and cost-effective method of high-throughput DNA sequencing that allows for the analysis of large quantities of data. Unlike other, more traditional methods, which require large amounts of time and labour; NGS can sequence millions of DNA fragments simultaneously, revolutionising the amount of genomic data produced (Goodwin et al, 2016). NGS has a wide scope of applications including the study of genetic variation, identification of harmful mutations, and analysis of gene regulation and expression.

NGS was first developed in the mid-1990s when parallel methods of DNA sequencing were first explored. The 454 Sequencer, introduced by 454 Life Sciences (now part of Roche) in 2005, was the first NGS platform commercially available (Metzket et al, 2010). Since then, various other NGS platforms have been developed, including Illumina, Ion Torrent PGM (Personal Genome Machine), Oxford Nanopore Technology (ONT) and Pacific Biosciences (PacBio) Sequencing, each with its advantages. Illumina sequencers produce over 80% of sequencing data worldwide, making it the most widely used NGS platform. Illumina uses reversible terminator sequencing where fluorescently labelled reversible terminators are added, allowing for individual nucleotide additions to be detected. Illumina is associated with high accuracy, low cost, and high throughput; but short read lengths and high error rates in homopolymer regions can limit its use (Gupta and Verma, 2019). Ion Torrent sequencing (2010) NGS technique uses a semiconductor-based detection system. Nucleotide additions are detected through hydrogen ion release during nucleotide incorporation. PacBio sequencing relies on single-molecule, real time (SMRT) sequencing technology; involving real-time detection of DNA polymerase activity as individual nucleotides are incorporated into the growing DNA strand. PacBio has excellent accuracy, long read lengths, and can

detect real-time modifications such as methylation; but it has a relatively high cost and low throughput (Andrew, 2010).

1.5.3.1.1 Illumina Sequencing

Illumina sequencing works by performing a series of steps and chemistry analysis as shown in figure 1.9. Briefly, the steps are fragmenting the DNA or RNA sample into small pieces, ligating adapter sequences to the fragment's end and then using bridge amplification to amplify the fragments on a flow cell (Mardis, 2008). Once amplified, the fragments then go through sequencing-by-synthesis chemistry where the DNA strand has fluorescently labelled nucleotides added. Finally, the fluorophores are cleaved off before the next cycle of nucleotide incorporation (Goodwin et al, 2016). This whole process usually generates millions of short reads that are then either assembled *de novo* or aligned to a reference genome to create a consensus sequence.

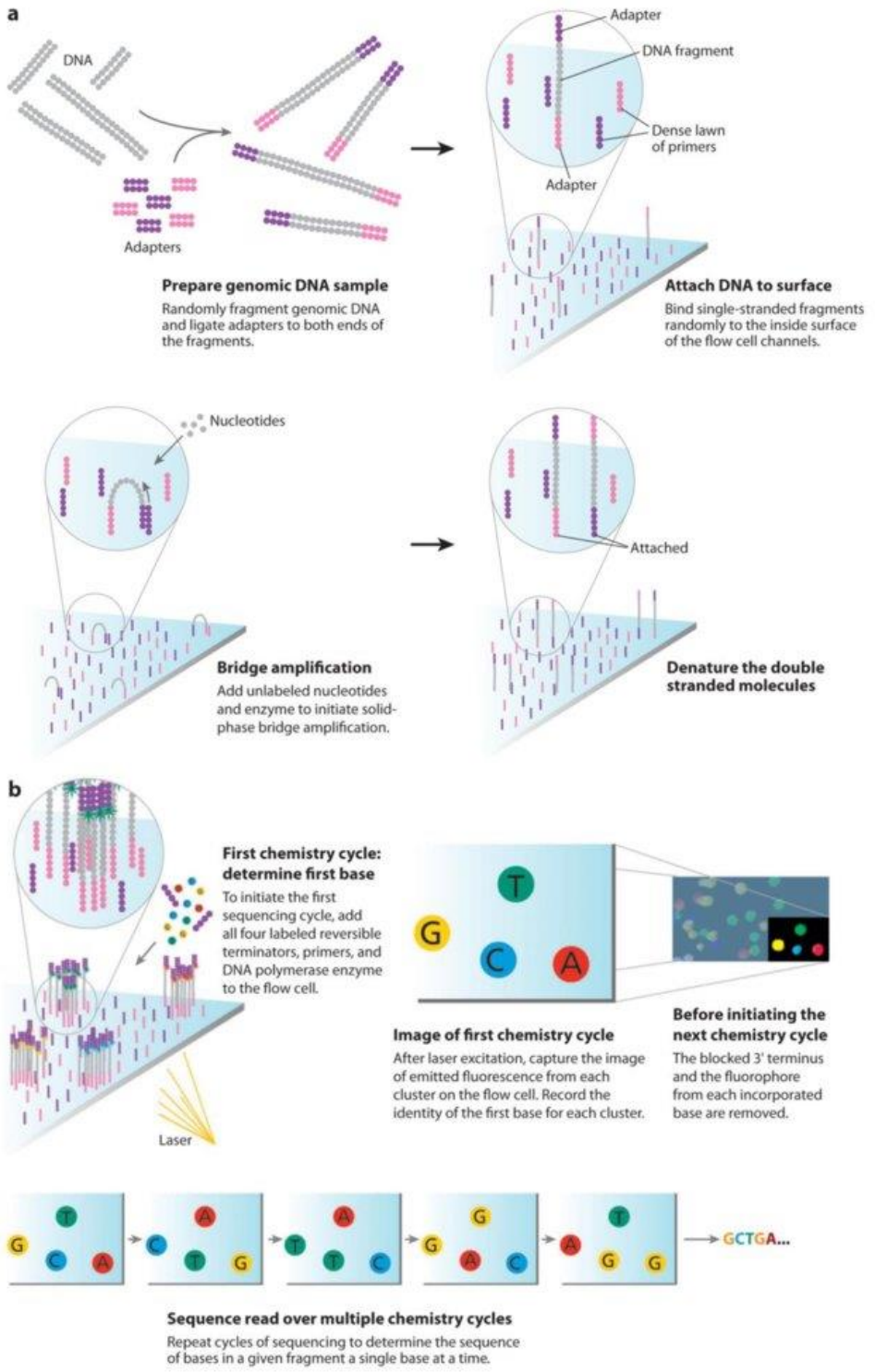


Figure 1. 9: The principle of Illumina sequencing process.

(A) DNA is converted into an Illumina adapter library and amplified by "bridge amplification" on the surface of the flow cell. (B) Amplified molecules are sequenced by the cycle of reversible termination chemistry (Knapik, K. (2013) modified from Mardis, et al., 2008).

Illumina sequencing can be applied to sequence amplicons for metabarcoding. Metagenomic studies commonly employ the analysis of the prokaryotic 16S ribosomal RNA gene (16S rRNA) (Figure 1.10), recognized for its approximately 1,500 bp length and nine variable regions interspersed between conserved regions (Woese, 1987). These variable regions serve as pivotal markers for phylogenetic classifications, such as genus or species, within diverse microbial populations (Pace, 1997). The selection of which 16S rRNA region to sequence is subject to debate and contingent upon factors such as experimental objectives, design, and sample type (Klindworth et al., 2013). This protocol delineates a method for sample preparation aimed at sequencing the variable V3 and V4 regions of the 16S rRNA gene while remaining adaptable for other regions using region-specific primers (Kozich et al., 2013). It integrates seamlessly with benchtop sequencing systems, offering a comprehensive workflow for 16S rRNA amplicon sequencing. The workflow encompasses primer ordering, library preparation through limited cycle PCR, and sequencing on the MiSeq platform, generating high-quality reads in a single run (Caporaso et al., 2012). Subsequent analysis options, such as the Metagenomics Workflow on MiSeq Reporter or BaseSpace, facilitate taxonomic classification, showcasing genus or species-level data graphically (Schloss et al., 2009). This versatile protocol extends beyond 16S rRNA sequencing, accommodating alternative regions and other targeted amplicon sequences of interest, with additional options available for secondary analysis (Kuczynski et al., 2012).

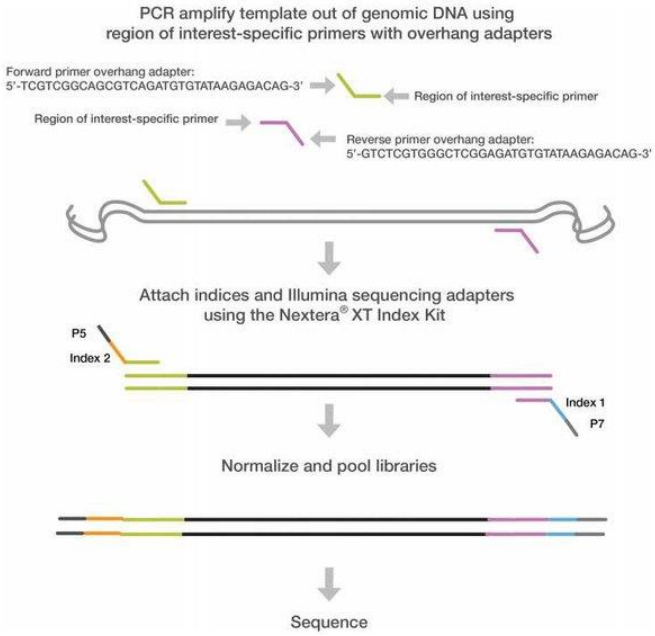


Figure 1. 10: 16S V3 and V6 amplicon workflow.

Modified according to the guideline of Illumina for V3/V4 amplicon sequencing (16S Metagenomic Sequencing Library Preparation; Illumina 2013).

1.5.3.1.2 Oxford Nanopore Sequencing

Nanopore sequencing is another NGS sequencing technique that uses nanopore-based sensing of individual nucleotides to generate long reads from either DNA or RNA samples (Wang et al., 2021). This technology is still being improved but it is already widely used in clinical diagnostics, microbiology, genomics research and ecology surveys as it potentially performs real-time sequencing with enough computer power (Jain et al, 2016). Nanopore sequencing works by performing a series of steps and chemistry analysis as shown in Figure 1.11. Nanopore works by passing either DNA or RNA molecules through a pore called “nanopore” that is on a membrane able to measure the electrical current’s changes as each nucleotide moves through the pore (Branton et al, 2008). Then, changes in the ionic current correspond to each nucleotide sequence present in the sensing region which will be then decoded with computational algorithms allowing the sequencing to be in real-time (Wang et al., 2021). As with Illumina sequencing, nanopore sequencing requires library preparation. The library preparation is divided into several steps: DNA/RNA extraction, PCR amplicons, end repair, adapter ligation and sequencing (Jain et al, 2016). Beginning with DNA or RNA extraction, genetic material is liberated from the sample matrix. Following this, end

repair ensures the genetic material's ends are uniform and suitable for subsequent processing. Adapter ligation then attaches short DNA or RNA sequences to the repaired ends, facilitating recognition by the nanopore sequencing platform. Finally, sequencing occurs as the prepared molecules traverse the nanopore, with changes in electrical current recorded in real-time. (Wang et al, 2021).

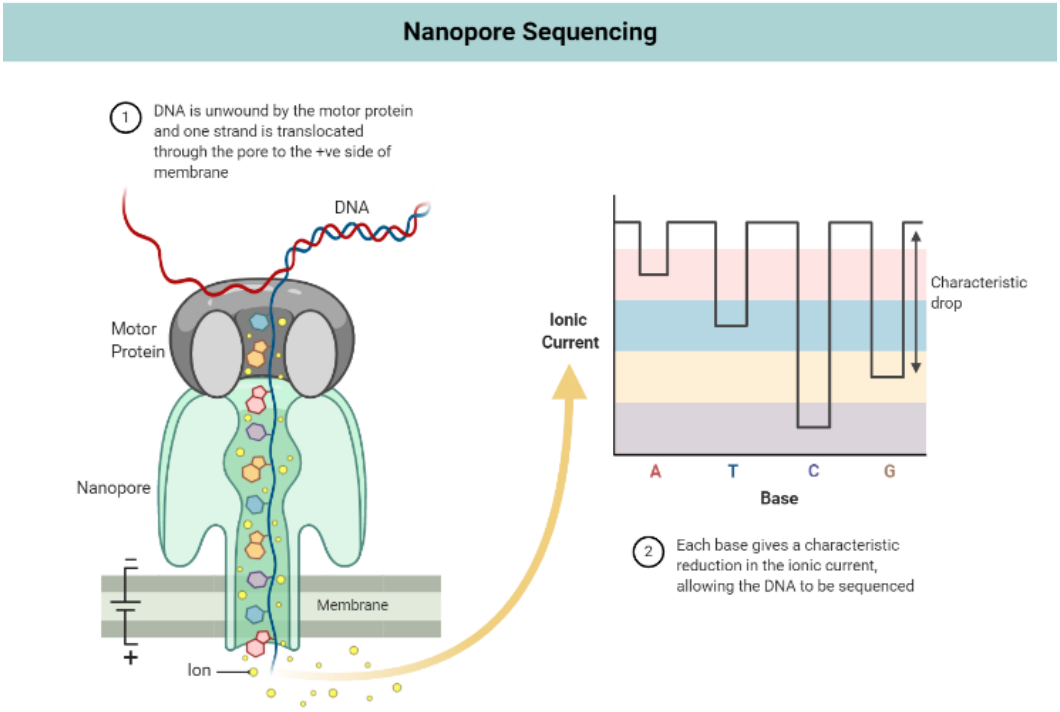


Figure 1. 11: Outline of how the Nanopore sequencing methods works.

Biological pores, integrated into artificial membranes, are subjected to controlled voltages, enabling the passage of DNA strands facilitated by motor proteins. Due to the DNA's negative charge, it is drawn towards the positively charged side of the membrane, while ions concurrently flow through the pore. As the DNA traverses the narrowest part of the pore, known as the constriction, each nucleotide base induces a distinct reduction in ionic current, facilitating the sequencing of the DNA molecule (Reproduced from Kasianowicz et al., 1996)

1.6 Bioinformatics as a Taxonomic Classification Method

Molecular biology tools would not be as effective without the complementary use of bioinformatics analyses. Bioinformatics is defined as the application of tools of computation and analysis to the interpretation of biological data, and it is an interdisciplinary field fundamental to molecular biology and molecular ecology (Bayat,

2002; Excoffier et al, 2005). Bioinformatics is being applied to the analysis of gene variation and expression, gene and protein structure prediction and function, sequences analysis and presentation, and analysis of molecular pathways to understand and study gene-disease interactions (Bayat, 2002).

1.6.1 Reference Libraries

Taxonomically and geographically complete reference databases of DNA sequences for each species are essential in both barcoding and metabarcoding techniques (Yang et al., 2017). The incompleteness and sometimes inaccuracy of databases are often believed to be the cause of some of the difficulties in metabarcoding.

The two main public databases of DNA barcode data for animals, plants and fungi are The Barcode of Life Data Systems (BOLD) and GenBank (Ratnasingham et al., 2007). BOLD currently contains sequences for ~296,000 formally described species, mostly for the COI gene, and each uploaded sequence must pass a quality check run by the administrator. GenBank is even bigger than BOLD, containing over 212 million sequences performing basic quality checks (Meiklejohn et al., 2019). Although GenBank contains many COI sequences, there is still a shortage of DNA barcode sequences for small body organisms such as zooplankton but this is improving year after year (Figure 1.12). The relatively low coverage for zooplankton could be caused by the low yield and low quality of genomic DNA from zooplankton specimens which often leads to inefficient PCR and a low rate of successful sequencing (Yang et al., 2017). This figure underscores the progress in sequencing efforts while also highlighting the need for continued expansion and improvement in the genomic databases for small-bodied organisms like zooplankton. The identification of zooplankton by molecular methods combined with bioinformatics is therefore challenging. This is not only based on the lack of sequences on the databases but also due to zooplankton divergence. The genetic data for phytoplankton is at an even earlier stage compared to zooplankton, there are relatively few sequences that are available in databases, and they are often misidentified (Yang et al., 2017). Once the sequences are retrieved and the databases are created, this will underpin the future analysis of plankton communities by metabarcoding. However, more and more studies are slowly emerging with successful zooplankton identification through metabarcoding (Bucklin et al., 2022, Creer et al. (2010), Pawlowski et al. (2018), and Zhan et al. (2013) demonstrating the

utility of metabarcoding in biodiversity assessment, particularly in the analysis of meiofaunal biospheres and the detection of rare species in aquatic communities.

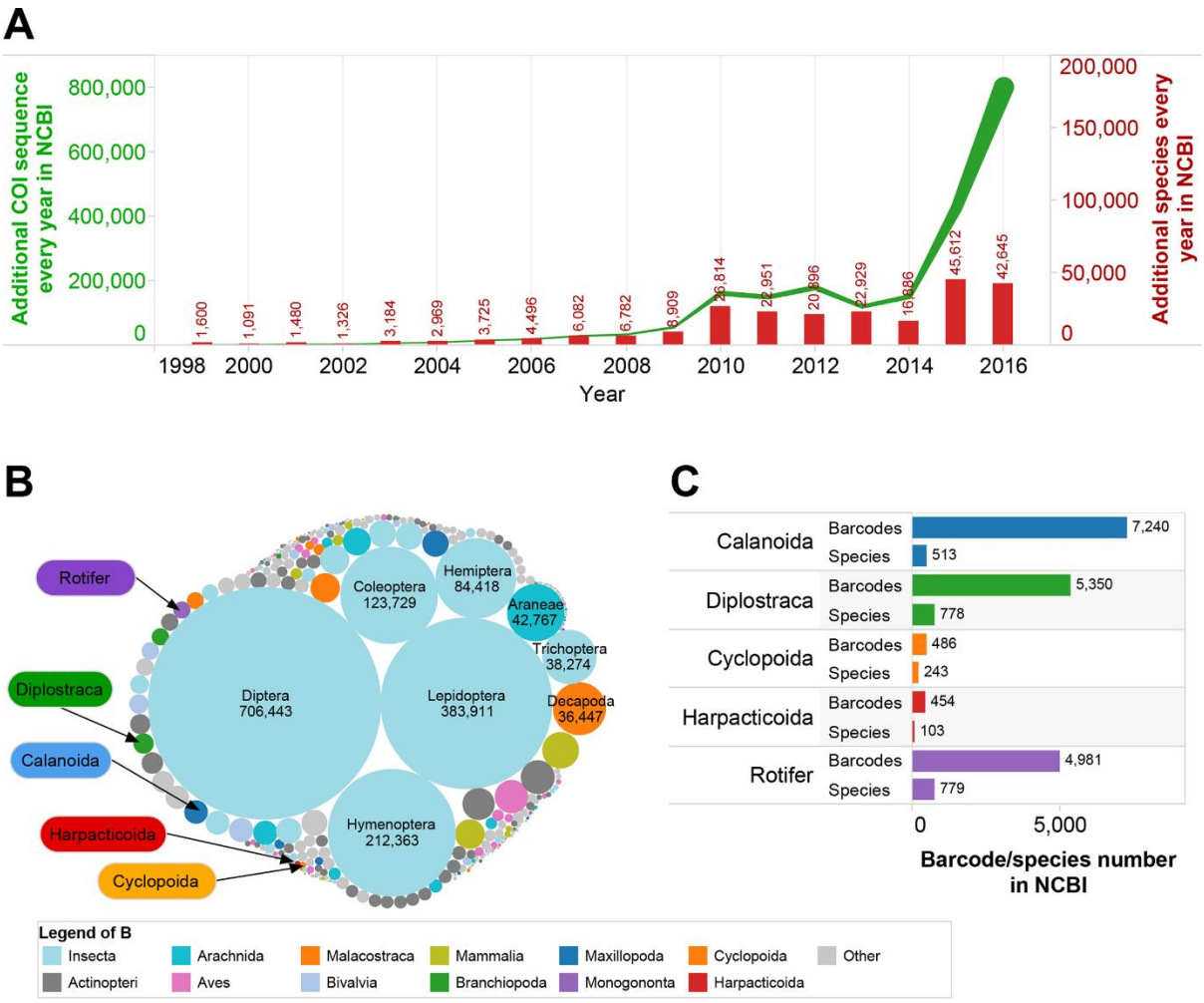


Figure 1. 12: Barcode sequences number on GenBank (NCBI) for the COI gene.

(A): Growth trend of COI sequences. (B): Taxa composition of COI sequences. (C): Composition of zooplankton COI sequences. Reproduced from (Yang et al., 2017b).

1.6.2 Data-processing analysis in NGS

Data-processing analysis in NGS generally begins with data pre-processing to improve the quality of the reads and their reliability, this is usually done by using a bioinformatic script. The bioinformatics pipelines vary based on the different sequencing platforms. Analysis of Illumina sequencing data often begins with quality control and trimming using Trimmomatic (Bolger et al., 2014), followed by read alignment to a reference genome using Burrows–Wheeler Aligner (BWA) (Li & Durbin, 2009). Variant calling is commonly performed using the Genome Analysis Toolkit (GATK) (McKenna et al., 2010), while differential expression analysis can be conducted using tools such as

DESeq2 (Love et al., 2014). Alternatively, some researchers prefer using STAR for alignment (Dobin et al., 2013), featureCounts for read quantification (Liao et al., 2014), and DESeq2 for differential expression analysis. For Nanopore sequencing data, basecalling is typically performed using Guppy (Oxford Nanopore Technologies) (ONT), followed by alignment with Minimap2 (Li, 2018). Consensus polishing of the aligned reads is often done using Medaka (ONT) (ONT). Another approach involves using the MinION Analysis Pipeline (MAP) provided by ONT, which integrates basecalling, alignment, and variant calling processes (ONT). These pipelines have been widely employed in genomic studies, including investigations into the diversity and ecology of zooplankton communities. There are several quality control tools which assess per-base sequence quality and detect adapter contamination (Andrews, 2010). After quality control, the adapters get trimmed, with the use of tools such as Cutadapt (Bolger et al., 2014) and Trimmomatic (Bolger et al., 2014). Further read filtering and trimming are then used to discard low-quality reads that failed the defined threshold so that there is a higher variant calling accuracy and a better alignment efficiency (Martin, 2011; Bolger et al., 2014). The filtered reads are then aligned and/or mapped to a reference database using e.g. Basic Local Alignment Search Tool (BLAST) (Lindeque et al., 2013).

Aims of the project

The field of aquatic ecotoxicology has developed validated gold standard methods to identify, assess and reduce relevant contamination in aquatic ecosystems (Schmitt-Jansen et al. 2008). While the application of metabarcoding techniques to identify indicator-species has been broadly studied, to the best of our knowledge the application of metabarcoding to identify plankton in ecotoxicity studies remains poorly investigated. Both zooplankton and phytoplankton are used in ecotoxicological studies as they are bioindicators of water quality (Parmar et al. 2016) and anthropogenic stressors can adversely impact community assemblages and ecosystem function. To date, the study of freshwater planktonic organisms using established microscopy methods is challenging, as taxonomic identification requirements are frequently expensive, time-consuming, and often require a specialist in the field (Elias-Gutierrez et al. 2018). Using molecular techniques, plankton identification is becoming easier and more accurate (Djurhuus et al. 2018) and has the clear potential to be applied to metabarcoding. Within this study, it is hypothesised that metabarcoding techniques can be developed and applied to both identify and quantify the relative abundance of indicator-genera in in-field ecotoxicity studies to replace or complement traditional microscopy techniques.

The aims of this project are to:

- Identify and quantify the zooplankton genera that serve as indicators within mesocosms to monitor community seasonal changes.
- Establish protocols and develop bioinformatic tools for the identification of zooplankton genera by metabarcoding
- Assess the effectiveness of the metabarcoding protocol in an ecotoxicity trial using the herbicide glyphosate.
- Evaluate and compare the performance of different NGS platforms (Illumina and Nanopore sequencing) for metabarcoding analysis.

Chapter 2: Materials and Methods

All sampling analyses and plankton culturing were carried out at Cambridge Environmental Assessment (CEA), Boxworth, UK. All the laboratory analyses were carried out in the laboratory of the University of Nottingham.

2.1 Methods used in Results Chapter 3: Protocol design for the molecular analysis of phytoplankton and zooplankton genera

2.1.1 Target genera

A list of expected phytoplankton organisms was made based on previous studies carried out at CEA mesocosms trials (Table 2.1)

Table 2. 1: List of most detected phytoplankton in mesocosms at CEA.

Division	Class	Order	Lowest taxon
Bacillariophyta	Bacillariophyceae	Pennales	Pennate diatoms
Charophyta	Coleochaetophyceae	Coleochaetales	Coleochaete spp.
	Conjugatophyceae	Desmidales	Closterium spp.
			Cosmarium spp.
			Gonatozygon spp.
			Staurastrum spp.
		Zygnematales	Mougeotia spp.
			Spirogyra spp.
Chlorophyta	Chlorophyceae	Chaetophorales	Aphanochaete spp.
			Protoderma spp.
			Stigeoclonium spp.
		Chlamydomonadales	Apiocystis spp.
			Carteria spp.
			Chlamydomonas spp.
			Gonium spp.
			Pandorina spp.
			Paulschulzia spp.
			Sphaerocystis spp.
		Oedogoniales	Oedogonium spp.

		Sphaeropleales	Ankistrodesmus spp.
			Ankyra spp.
			Characium spp.
			Coelastrum spp.
			Coenochloris spp.
			Gloeocystis spp.
			Golenkinia spp.
			Monoraphidium spp.
			Pediastrum spp.
			Planktosphaeria spp.
			Scenedesmus spp.
			Tetraedron spp.
	Trebouxiophyceae	Chlorellales	Crucigeniella spp.
			Dictyosphaerium spp.
			Geminella spp.
			Nephrocytium spp.
			Oocystis spp.
Choanozoa	Choanoflagellatea	Craspedida	Codosiga spp.
Cryptophyta	Cryptophyceae	Cryptomondales	Cryptomonas spp.
		Pyrenomonadales	Rhodomonas spp.
Cyanophyta	Cyanophyceae	Chroococcales	Chroococcus spp.
		Nostocales	Anabaena spp.
			Gloeotrichia spp.
		Oscillatoriales	Lyngbya spp.
			Oscillatoria spp.
		Synechococcales	Merismopedia spp.
			Snowella spp.
Euglenophyta	Euglenophyceae	Euglenales	Euglena spp.
			Trachelomonas spp.
Miozoa	Dinophyceae	Peridiniales	Peridinium spp.
Ochrophyta	Chrysophyceae	Chromulinales	Dinobryon spp.
	Synurophyceae	Synurales	Mallomonas spp.

2.1.2 Reference database creation for phytoplankton

A database was created by retrieving all the available sequences in NCBI GenBank for the selected phytoplankton genera. For the phytoplankton, a total of 2177 23S and a total of 902 18S sequences were downloaded. Sequences were aligned using Clustal-W multiple sequence alignment in MEGA version 6 (Tamura et al., 2013).

2.1.3 Phytoplankton monoculture growth and analysis

Phytoplankton monocultures were carried out to grow the targeted genera and proceed with both morphological and molecular analysis. This was to provide controls for testing primers for the amplification of target genes.

2.1.3.1 Phytoplankton staining, and morphological identification

The phytoplankton genera *Chlorella*, *Spirogyra*, *Desmus desmus* and *Ankistrodesmus* were selected as target genera. They were selected based on the algae culture availability found at CEA. These were cultured using the Bolds Basal Medium (BBM; Sigma) on-site under the following conditions: room temperature at 20 ±2 °C, Lux 3000-8000 and maintenance in a photoperiod of 16 hours light and 8 hours dark and stored at -20°C in a light-tight sampling bottle, ensuring the creation of a dark environment, aimed at minimizing the risk of photodegradation and preserving sample integrity. Samples were stained using Lugol's iodine solution and left to set for 24 hours before starting the morphological identification. The sample was then transferred into a combined plate chamber (Figure 2.1). This method allows for sample settling, staining, and subsequent microscopic examination without the need for transferring the sample to a separate microscope slide. The combined plate chamber simplifies the process by enabling all steps to be conducted within the same chamber. The chamber is then used for sample settling and subsequent microscopic observation. This unified approach ensures that the sample remains undisturbed during the entire process, reducing the chances of cell disruption or contamination. The stained sample can be observed directly under a microscope, eliminating the need for transfer to a separate slide, and simplifying the procedure while maintaining sample integrity for accurate analysis of phytoplankton morphology. The morphology was carried out using an SP 98 inverted microscope. The number of individuals and number of rows counted were entered into a database that calculates the number of individuals per litre in the following way:

- **Calculating the dilution factor:** Dilution factor = total volume of diluted sample/volume of sample aliquot diluted.
- **Find organisms per sample:** (No. organisms counted x (Total no. of rows / no. rows counted) x preserved sample volume x dilution factor)
- **Find organisms per litre (L):** Organisms per sample/volume of water sieved.

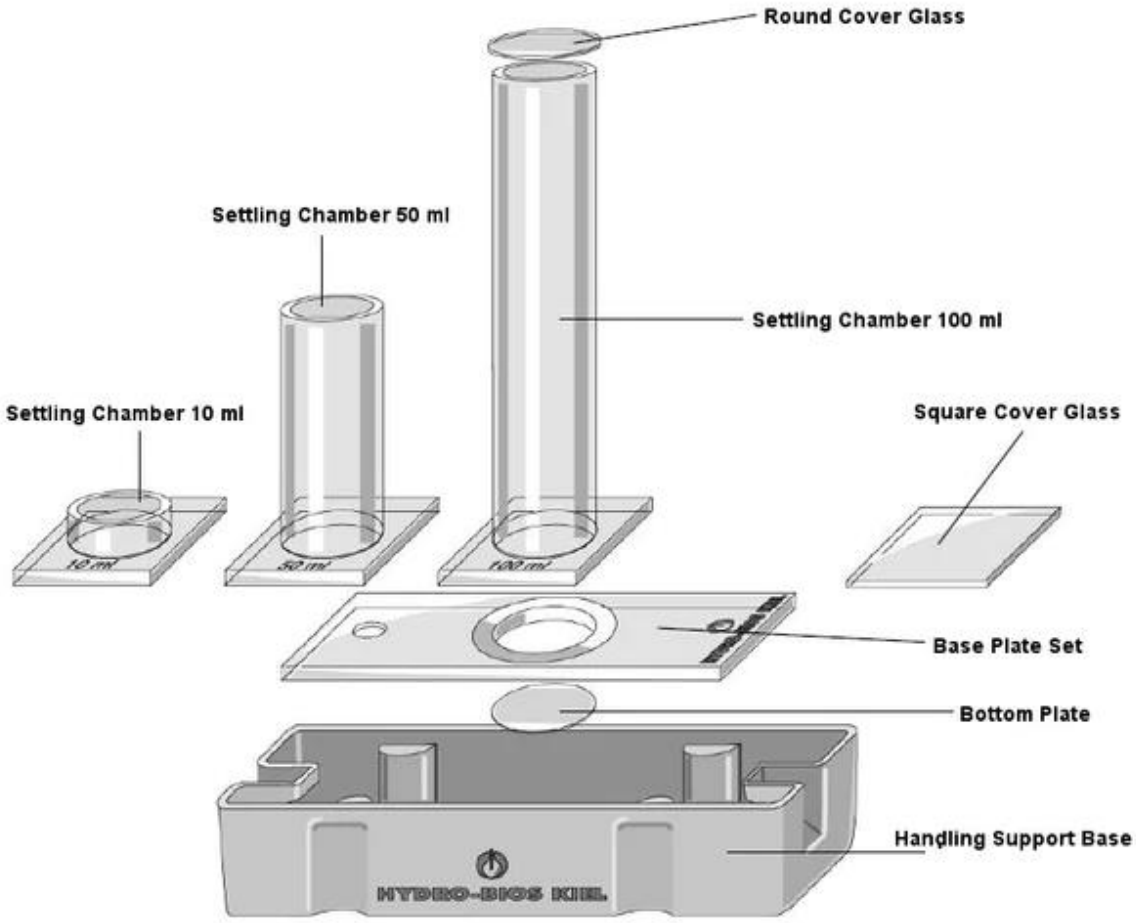


Figure 2. 1: Example of a Combined Plate Chamber © 2023 BIOWEB Global.

This method facilitates sample settling, staining, and microscopic examination without transferring the sample to a separate slide. It uses a combined plate chamber, streamlining the process by allowing all steps to be performed within the same chamber, which is then used for both settling and microscopic observation.

2.1.4 Phytoplankton DNA-extraction

The DNeasy Plant Mini Kit (Qiagen, Maryland) was used to extract the DNA from all phytoplankton genera according to the manufacturer's instructions. Briefly, 1 mL of samples from *Desmus desmus*, *Chlorella*, *Spirogyra*, and *Ankistrodesmus* were centrifuged at 20,000 x g for 3 minutes and the supernatant was removed from the tube leaving the pellet only. 400 µl buffer AP1 and 4 µl RNase A were then added, vortexed and incubated for 10 min at 65°C. Afterwards, 130 µl Buffer P3 were added and incubated for 5 min on ice. The lysate was centrifuged for 5 minutes at 20,000 x g, transferred into a QIAshredder spin column, and centrifuged again for 2 minutes at 20,000 x g. The flow-through was then transferred into a new tube without disturbing the pellet and 1.5 volumes of buffer AW1 was added. The mixture was then transferred into a DNeasy Mini spin column and centrifuged for 1 minute at 6000 x g with the flow-through discarded. Subsequently, 500 µl of buffer AW2 was added and centrifuged at 20,000 x g and transferred into a new microcentrifuge tube. Finally, the solution was eluted with 100 µl of buffer AE, and incubated at room temperature for 5 minutes with a final spin for 1 minute 6,000 x g.

2.1.5 Metabarcoding Primer Design and qPCR

All phytoplankton were tested with two different metabarcoding primer sets (Table 2.2). A nested qPCR was carried out for the gene 23s using the A23SrVF primers in each of the two rounds of PCR (Yoon et al., 2016). For the 18s gene, Uni18S primers were used (Zhang et al., 2018).

Table 2. 2: Phytoplankton primers used for qPCR analysis.

Where A23SrVF1/2 and R1/2 are used together in the nested qPCR.

Phytoplankton qPCR primer-set		
ID		PRIMER
A23SrVF1	Forward	5'-GGACARAAAGACCCTATG -3'
A23SrVF2	Reverse	5'-CARAAAGACCCTATGMAGCT 3'
A23SrVR1	Forward	5'-AGATCAGCCTGTTATCC -3'
A23SrVR2	Reverse	5'-TCAGCCTGTTATCCCTAG -3'
Uni18S	Forward	5'- AGGGCAAKYCTGGTGCCAGC-3'

Uni18Sr	Reverse	5'- GRCGGTATCTRATCGYCTT -3'

qPCR analyses were performed using a Light-Cycler 480 (Roche) using the following mix: 12.5µl SYBR green (Bio-Rad), 1 µl of each reverse and forward primers (Sigma), 7.5 µl ddH20 and 3 µl DNA; total 25 µl. Primers (Sigma) were prepared in stock solution of 100 pmol/µl and working solution of 5 pmol/µl.

A touch-down qPCR amplification program was used (Don et al., 1991) for all reactions as follows: initial denaturation 95℃ for 3 minutes followed by 40 cycles of denaturation at 95℃ for 30 seconds, annealing for 30 seconds starting from a temperature of 65℃ reaching a temperature of 55 and finally changing per round, a temperature of 72 ℃ for 30 seconds. Finally, 1 cycle of 95 ℃ for 5 seconds and 65℃ for 1 min was run for the melting curve.

2.1.6 Gel electrophoresis

Agarose gel electrophoresis was used to run the qPCR products to check size and for PCR clean-up and gel extraction. Agarose gel (3% (w/v)) were made up of 3 g of agarose powder, 100 ml 1XTAE buffer and 3 µL SyberSafe (Invitrogen). The gel was run at 100v for 45 minutes and checked against a Quick-Load Purple 100 bp DNA ladder (Bio Labs).

2.1.7 Sanger sequencing

DNA was extracted from agarose gel using the NucleoSpin Gel and PCR Clean-up kit (Thermo Scientific) according to the manufacturer's instructions. Briefly, the selected DNA fragments were cut from the agarose gel. For each 100 mg of gel, 200 µl of buffer NTI were added and incubated at 50℃ until completely dissolved. Once dissolved, the solution was transferred into a NucleoSpin Gel and PCR Clean-up Column and centrifuged for 30 seconds at 11,000 x g and the flow-through was discarded. Subsequently, 700 µl of buffer NT3 was added and the solution was centrifuged for 30 seconds at 11,000 x g with the flow-through discarded (this step was repeated). Finally, the solution was eluted by adding 15 µl of buffer NE and incubated at room temperature

for 5 minutes prior to centrifugation at 11,000 x g. Sanger sequencing was conducted via Source BioScience.

The DNA sequences were then manually checked and identified and assigned a "putative" identification name (either to species, genus, or any other taxonomic level) based on the first hit by BLAST (NCBI). In order to have a high level of confidence, putative genus name ≥ 97% match were considered a good match with the GenBank samples.

2.1.8 Zooplankton analysis

A list of expected Zooplankton organisms was made based on previous studies carried out at CEA mesocosms trials (Table 2.3)

Table 2. 3: List of most detected zooplankton in mesocosms at CEA together with the GenBank availability for the 18s gene.

Group	Family	Lowest taxon	GenBank availability
Cladocera	*Chydoridae*	*Acroperus sp.*	2
		Alona sp.	1
		Alonella sp.	1
		Chydorus sp.	3
		Graptoleberis testudinaria	1
		Peracantha truncata	NA
	Daphniidae	*Ceriodaphnia sp.*	7
		Daphnia sp.	191
		Simocephalus sp.	33
	Eurycercidae	*Eurycercus sp.*	5
Copepoda	Cyclopoida	*Cyclopoida*	60
	Diaptomidae	*Diaptomidae*	200
	Nauplia (Juvenile of Cyclopoida/Diaptomidae)	*Nauplia (Juvenile of Cyclopoida/Diaptomidae)*	NA
Gastrotricha	Gastrotricha	*Gastrotricha*	NA
Rotifera	Asplanchnidae	*Asplanchna sp.*	5
	Brachionidae	*Brachionus sp.*	701
		Keratella sp.	5
		Notholca sp.	4
	Euchlanidae	*Euchlanis sp.*	4
	Gastropodidae	*Gastropodidae*	1
	Lecanidae	*Lecane sp.*	10
	Lepadellidae	*Colurella sp.*	4
		Lepadella sp.	3
		Squatinella sp.	3
	Mytilinidae	*Mytilina sp.*	5
	Notommatidae	*Cephalodella sp.*	2

		Monommata sp.	2
	Philodinidae	*Rotaria sp.*	107
	Synchaetidae	*Polyarthra sp.*	1
		Synchaeta sp.	54
	Testudinellidae	*Testudinella sp.*	6
		Trichocerca sp.	8
	Trichotriidae	*Trichotria sp.*	3

2.1.8.1 Reference database creation for zooplankton

A reference database was generated for the 18S gene by retrieving all available sequences (1,432) for the selected zooplankton genera from the list provided by CEA (Table 2.3) using GenBank. As is standard practice, identification was at the lowest taxon available (genera level and species when available). However, for specific analyses, focus was given to the taxonomic groups Cladocera (Phylum: Arthropoda, Subphylum: Crustacea, Class: Branchiopoda, Order: *Cladocera*), *Copepoda* (Phylum: Arthropoda, Subphylum: Crustacea, Class: *Maxillopoda*, Subclass: *Copepoda*), and *Rotifera* (Phylum: *Rotifera*). Throughout this study, these will be referred to as "groups" according to the CEA table (2.3). These taxonomic classifications are particularly relevant for ecotoxicity studies, aligning with their common usage in such analyses and ensuring a comprehensive approach to capturing ecological dynamics.

2.1.8.2 Zooplankton collection

Sloped mesocosms were used in this study. The used mesocosms were 2.6 m long, 1 m wide and 0.7 m deep. The sloped mesocosms are designed to enable the incorporation of a wide variety of plant species, providing a habitat which mimics the plant community found in edge-of-field environments (© CEA. All rights reserved). The water samples were collected using a depth-integrated water sampler (DIWS) tube (Figure 2.2B). The tube was inserted in the water, then the first bung was inserted into the top of the tube, and the tube was lifted through the water column. While the bottom end of the tube was still submerged, the second bung was inserted into the sampler. The tube was then positioned over a bucket, and the water was emptied by removing the lower bung. As the collection mesocosms were sloped, this procedure was repeated until a 1 L sample had been taken from each of the four corners of the deep (50 cm) zone (twice) and one from the centre of the deep-water zone (Figure 2.2 A). In addition,

a calibrated jug was used to take six 1-litre samples from the sloped area (three in the 20 - 30 cm depth zone and three in the 0 - 20 cm depth zone) for a total of 15 water samples (Figure 2.2A). All samples were pooled together into a bucket. Before proceeding with the zooplankton filtering, 750 mL of the pooled water was collected and stored in a light-sensitive bottle. The pooled water (15 L) was then filtered through a coarse sieve (2 mm) into a second bucket to remove any large organisms and detritus. Finally, the coarse-sieved water was filtered through a fine sieve (30 µm) and the retained organisms were re-suspended in 20 mL 70% ethanol in a Falcon tube (Figure 2.2C). Samples that were considered too cloudy because of sediments, were diluted to 40 mL using 70% ethanol. The final volume of each sample was noted on the label and then recorded as needed for the final morphological identification calculations. Samples were stored under ambient conditions until the preserved organisms were counted microscopically.

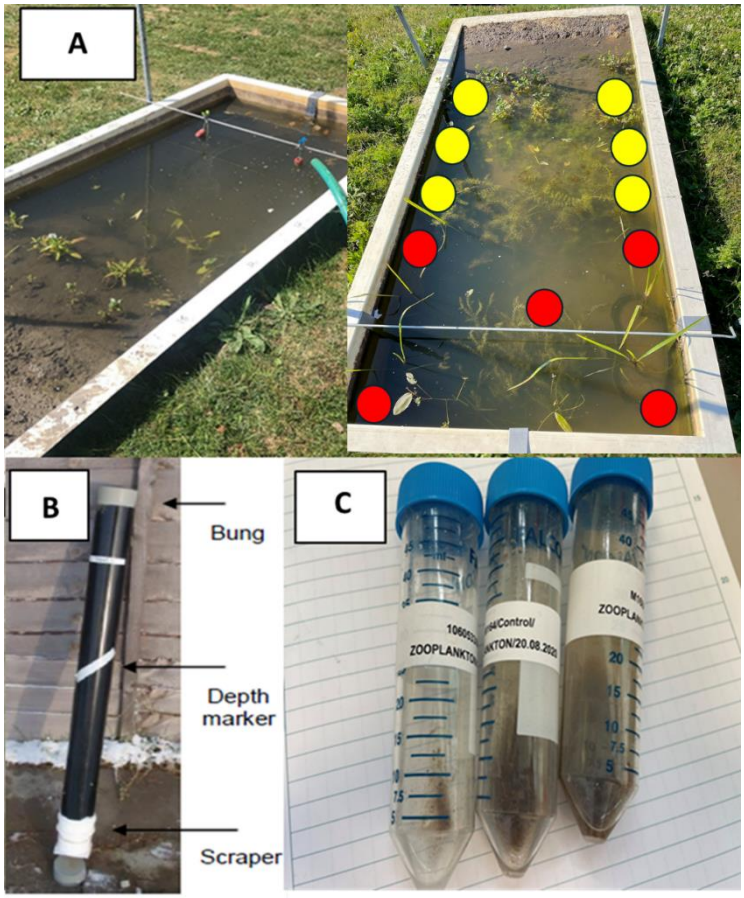


Figure 2. 2: Process of phytoplankton and zooplankton sampling.

A: Sloped mesocosm examples. Water collection was collected in the circled areas where red represents water collected with the DIWS tube and yellow water collected

with the glass jar, B: Depth integrated water samplers (DIWS) Edge sample and C: Sieved Samples stored in falcon tubes.

2.1.8.3 Morphological analyses

Samples were morphologically analysed using an SP 98 inverted microscope. An aliquot (1 mL) of the sample was transferred to a Sedgewick-Rafter counting cells and the organisms (excluding insects and other large organisms not associated with zooplankton) were counted. The number of individuals and the number of rows counted were entered into a database that calculated the number of individuals per litre in the following way:

Calculating the dilution factor: Dilution factor = total volume of diluted sample/volume of sample aliquot diluted.

Find organisms per sample: (No. organisms counted x (Total no. of rows / no. rows counted) x preserved sample volume x dilution factor)

Find organisms per litre (L): Organisms per sample/volume of water sieved.

2.1.8.4 Methods for zooplankton isolation

Zooplankton isolation was the first method used to group the zooplankton based on its genera. The isolation was performed to differentiate the zooplankton and to facilitate the molecular analysis when designing genera-specific primers. During the process of morphological identification, genera were isolated in different labelled 1.5 mL tubes containing 70% ethanol. This process was performed using a disposable glass Pasteur pipette for counting live plankton out of a dish. The Pasteur pipette was cut and bent as desired, and the end fire-polished until thin enough for the selected zooplankton genera dimensions.

2.1.8.5 Methods used in zooplankton monoculture growth and analysis

Zooplankton monocultures was the second method applied to isolate genera. Following the CEA Standard Operating Procedure (SOP) protocols, the most abundant and easiest genera to culture were selected to provide control zooplankton for analysis. Individual zooplankton genera were required as controls for the development of primer sets for genera-specific PCR and metabarcoding.

2.1.8.5.1 Maintaining monocultures

Cultures of the largest and most common zooplankton (*Daphnia, Diaptomidae, Cyclopoidae* and *Simocephalus*) were made and kept in the laboratory in a specific environment (Figure 2.3). The algae *Desmodesmus subspicatus* was cultured in the laboratory as a food source for the zooplankton cultures. Cultures were made based on a provided SOP from CEA. In summary: cultures were maintained under aerated conditions or in static flasks with gentle shaking to re-suspend the algae. The environmental conditions included room temperature at 20 ± 2°C, overhead fluorescent lighting with a Lux range of 3000-8000 lux, and a photoperiod of 16 hours of light and 8 hours of darkness. Cultures were allowed to grow until the algae reached a dark shade of green, at which point they were considered ready to be used for feeding or sub-cultured for continued growth. When not in use, algae cultures were preserved in the fridge at 3-8°C. Periodic spectrophotometer calibration was performed.

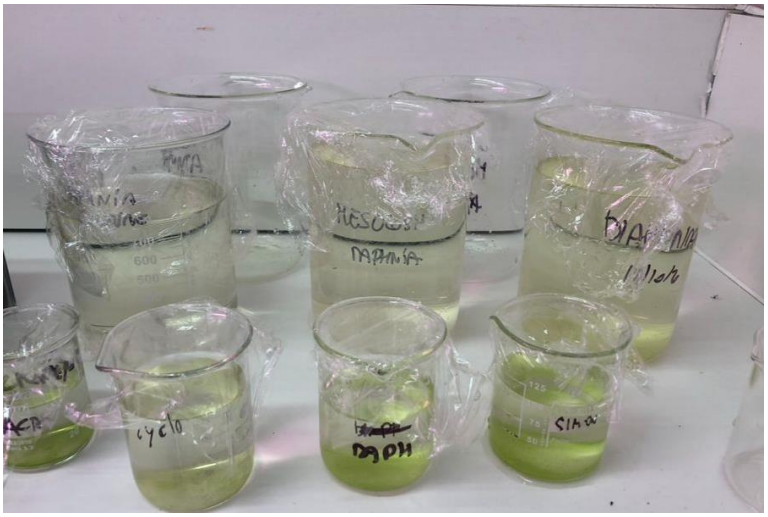


Figure 2. 3:Cultures of the most abundant Zooplankton genera found at CEA mesocosms.

The zooplankton cultures were maintained slightly differently depending on the genera. All the maintenance SOPs were provided by CEA. For *Daphnia Magna and Pulex: Daphnia* were cultured in 1L glass beakers and covered with Clingfilm to minimise contamination. Visual assessments of *Daphnia*'s health were performed approximately

five times a week, with the culture medium changed three times weekly. *Daphnia* were transferred between culture vessels using a plastic pipette with an internal diameter of >5mm. Neonates (<24 hours old) for culture initiation were sourced from the third offspring or later of the adult culture. Neonates, unless required for a new culture initiation or experimental purposes, were removed from culture vessels via a filtration tube and discarded. Adult *Daphnia* were isolated before this process by collecting them using a Pasteur pipette and placing them in a new beaker filled with mesocosm-filtered water using a 30 μm sieve to remove zooplankton and large phytoplankton. Equipment was designated solely for *Daphnia* culturing, and if multiple *Daphnia* strains were maintained, the apparatus was labelled and kept separately to prevent strain mixing.

The culture environment involved maintaining cultures under semi-static conditions in a suitable room, was set at a temperature of 20 ± 2°C, a controlled photoperiod of 16 hours light and 8 hours dark using overhead fluorescent strip lights, and Lux measurements recorded weekly in the range of 1000-1500 lux. The culture medium consisted of mesocosm water, initially filtered through a 30 μm sieve to eliminate zooplankton and large phytoplankton. Subsequent filtration through a 0.7 μm filter, using a vacuum pump system, removed the remaining phytoplankton. Fully filtered media was aerated for at least 24 hours before use in *Daphnia* culturing vessels, with water replaced thrice weekly (Monday, Wednesday, and Friday). Water quality measurements for filtered aerated mesocosm water were taken every Friday and every Monday for expired media, with specific ranges for temperature, pH, and dissolved oxygen. Filtered water was not stored for more than seven days, and aeration did not exceed three days.

In terms of culture maintenance, each culture consisted of 800 mL of filtered mesocosm water with approximately 15 *Daphnia*. Daily feeding occurred using the age of *Daphnia* to determine the amount of required algae. Daily counts were conducted, noting live adults, neonates, and the amount of food. Media replacement, carried out three times per week, involved transferring *Daphnia* into new vessels with fresh, aerated, filtered mesocosm water. For culture renewal, new vessels were initiated with 800 mL of fresh, aerated filtered mesocosm water and approximately 15 neonates, synchronized to within 24 hours and less than 24 hours old. Removal of neonates without media replacement was achieved by isolating adult *Daphnia* using a Pasteur pipette into a smaller vessel with filtered pond water and transferring neonates via a filtration tube.

Observations related to *Daphnia* health and other factors were recorded daily. Culture vessels were labelled with culture identity, and a new record sheet was started for each new culture initiated, detailing the culture reference. The *Daphnia* feeding regime involved daily feeding with a suspension of unicellular green algae. The amount of algae provided was based on *Daphnia* age and was calculated in mg per day, with knowledge of mg per algal cell used as a food source. Prepared suspensions of algal feed were stored in a refrigerator at 5 ± 3°C for two weeks or in the freezer at -20°C for up to six months. The feeding regime varied with *Daphnia* age and was specified in the SOP provided by CEA, with special considerations for Fridays where cultures received twice the volume of algae compared to Monday to Thursday.

2.1.8.6 DNA-extraction

The *Daphnia* DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Maryland), according to the manufacturer's instructions. Briefly, most of the ethanol was removed from the samples by spinning and pipetting out the ethanol leaving only the pellet. Then 20 µl of proteinase K was added to the pellet and incubated at 56°C for 20 minutes. Once the product was completely lysed, 200 µl of buffer AL was added and the product was incubated again at 56°C for 10 minutes. Subsequently, 200 µl of 96% ethanol was added and the sample vortexed. The mixture was transferred into a DNeasy Mini spin column and centrifuged at 6,000 x g for 1 minute with the flow-through discarded. Afterwards, 500 µl of Buffer AW1 was added and centrifuged for 1 minute at 6,000 x g with the flow-through discarded. Then, 500 µl of Buffer AW2 was added, and centrifuged for 3 minutes at 20,000 x g with the flow-through discarded. Finally, the DNA was eluted by adding 200 µl Buffer AE, incubating for 5 minutes at room temperature with a final 1-minute spin at 6,000 x g. The concentration of *Daphnia* DNA was checked using Qubit 3.0 Kit (Invitrogen) following the manufacturer's instructions.

2.1.8.7 Test of metabarcoding primer on Daphnia

Daphnia samples were first tested with a general metabarcoding primer set and then tested with designed genus-specific primers.

The metabarcoding primer Citochrome Oxidase I (COI) (Folmer et al., 1995) and 18s (Zhan et al., 2013) were first tested on *Daphnia* samples taken from the CEA culture as

previously reported (Bucklin et al., 2022 and Djurhuus et al., 2018). The qPCR analyses on *Daphnia* DNA were performed in triplicates (D1, D2 and D3) using a Light-Cycler 480 (Roche) using the following mix: 12.5µl SYBR green (Bio-Rad), 1 µl of each reverse and forward primers (Sigma), 7.5 µl ddH20 and 3 µl DNA; total 25 µl. Primers (Sigma) were prepared in a stock solution of 100 pmol/µl and a working solution of 5 pmol/µl. Positive control pike used initially was Pyke (*Esox lucius*) provided by ADAS.

A touch-down qPCR amplification program was used (Don et al., 1991) for all reactions as follows: initial denaturation 95°C for 3 minutes followed by 40 cycles of denaturation at 95°C for 30 seconds, annealing for 30 seconds starting with an initial temperature of 65°C decreasing by one degree over 11 rounds reaching a temperature of 55 and finally, an extension temperature of 72 °C for 30 seconds. Finally, 1 cycle of 95 °C for 5 seconds and 65°C for 1 minute was run for the melting curve.

2.1.8.8 Genera-specific qPCR

As an alternative or complementary approach to metabarcoding, genus-specific PCRs that target the main indicator zooplankton genus were investigated using *Daphnia* as the exemplar.

Four genus-specific primer sets (three for the COI and one for the 18s gene) were designed using Primer 3 plus software (Untergasser et al., 2012) for the *Daphnia* genera summarised in Table 2.4. To design the 18s genus-specific primer for the *Daphnia* genera, 215 previously aligned sequences were uploaded on Primer 3 plus software using the following parameters:

- Product size between 200-250 with optimal product size as 200 bp.
- Melting temperature between 59-65°C with an optimal melting temperature of 60°C with a maximum difference of 3°C between the forward and reverse primers,
- GC content between 40-60%.

Three genus-specific primers for *Daphnia* COI were designed. The 1869 *Daphnia* sequences were filtered by similarity (99%) using CD-HIT-EST (http://weizhongli-lab.org/cd-hit/), and 66 sequences were obtained. The filtered sequences were then uploaded on Primer 3 Plus software using the same parameters as above. For all the

primers consensus sequences were obtained in MEGA version 6 (Tamura et al., 2013) and then uploaded on Primer 3 Plus to create a primer set. Results of the genus-specific primer design for *Daphnia* are summarised in Table 2.4. The same approach was carried out for the design of the genus-specific 18s primer.

Table 2. 4: Daphnia genus-specific primers characteristics including PCR product size, and primer annealing temperature.

Daphnia Genus-specific primers				
ID	**PRIMER**		**PRODUCT SIZE**	**ANNEALING TEMPERATURE**
Daphnia_Specific_COI	F	5'-CAGCCCTTACCCTCTTGCTC-3'	200 bp	61.4 ℃
Daphnia_Specific_COI	R	5'-TGAGGCGAGGAAACAAT-3'		57.4 ℃
Daphnia_Specific_COI_2	F	5'-AATATTACCCCCTGCGTGA-3'	170 bp	57.3℃
Daphnia_Specific_COI_2	R	5'-TGAGGCGAGGAAACAATAGC-3'		57.3℃
Daphnia_Specific_COI_3	F	5'-GTNATNCCNATNATNATNGG-3'	201 bp	40℃
Daphnia_Specific_COI_3	R	5'-ACNGANGCNCCNGCNTGNGC-3'		60℃
Daphnia_Specific_18S	F	5'-AGATGCGAGACCGCAAAAT-3'	204 bp	55.3℃
Daphnia_Specific_18S	R	5'-CGAGGCTCGAGTGCATGTAT-3'		58.4℃

The same touchdown PCR program (Don et al., 1991) as section 2.1.8.7 was performed on the same *Daphnia* triplicates (D1, D2 and D3) using both Daphnia_Specific_COI (forward and reverse) and Daphnia_Specific_18S Primers.

Then a temperature gradient was tested to establish the most suitable annealing temperature for the primers (ranging from 65 ℃ to 55℃). A genomic test was also made on *Daphnia* to check the DNA quality by mixing 3 µl of *Daphnia* DNA with 3 µl of Loading Dye and separating them on an agarose gel.

The same temperature gradient was then tested on the second COI genus-specific primer set Daphnia_Specific_COI_2.

A comparison of the two COI genus-specific primer sets (Daphnia_Specific_COI and Daphnia_Specific_COI_2) was made by performing a qPCR using both.

Finally, the last genus-specific designed primer set (Daphnia_Specific_COI_3) was tested on the *Daphnia* triplicates.

2.1.8.9 Gel electrophoresis

The agarose gel electrophoresis was made as per 2.1.6

2.1.8.10 Sanger sequencing

Sanger sequencing preparation was performed as 2.1.7

2.1.8.11 Metabarcoding analysis protocol design and optimisation

Metabarcoding was the method taken forward to complete the study. With this method, there was no need to isolate each selected genera as it was possible to mass-sequence each zooplankton mixed sample.

2.1.8.11.1 DNA-Extraction optimisation

Each mesocosm sample's DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Maryland), according to the manufacturer's instructions. All the samples were diluted to 5 ng/µl to reduce bias (Questel et al., 2021; and Bucklin et al., 2019). All the optimisation and troubleshooting carried out to find the best DNA-extraction protocol are summarised in Table 2.5.

Table 2. 5 : Optimisation of DNA extraction.

Each storage method was tested with and without washing steps, at each final elution and at each incubation time. RT=Room temperature.

DNA-extraction optimisation			
STORAGE METHOD	**FINAL ELUTION**	**WASHING STEPS**	**INCUBATIONS**
NO ETHANOL, RT	200 UL ELUTION BUFFER	ACCORDING TO BUCKLIN ET AL. (2019)	15 MINUTES AT 56°C
70% ETHANOL, RT	100 UL ELUTION BUFFER	NONE	30 MINUTES AT 56°C

90% ETHANOL, RT	50 UL ELUTION BUFFER	NONE	OVERNIGHT AT 56°C
NO ETHANOL, -20 °C	100 UL WATER	NONE	NONE
70% ETHANOL, -20°C	100 UL WATER	NONE	NONE
90% ETHANOL, -20°C	100 UL WATER	NONE	NONE

2.1.8.11.1.1 Final established DNA-Extraction protocol

The pelleted samples (in 70% ethanol stored at room temperature) were washed with distilled water, re-pelleted and then placed into a 50-ml Falcon tube above a 33-µm mesh - which served to suspend and dry the material - and were centrifuged at 4000 g for 3 min. The pellet was transferred to a new 50-ml Falcon tube, and SDS buffer (Tris–HCl, 10 mM; EDTA, pH 8.0, 100 mM; NaCl, 200 mM; SDS 1%) equal to the pellet volume was added. The samples were homogenised using a disposable pestle and mortar. Proteinase K (Fisher Scientific) was added (0.2 mg/ml of sample) and the tubes were incubated for 30 minutes at 6˚C. Once the product was completely lysed, DNA extraction was carried out following the manufacturer's instructions (DNeasy Blood & Tissue Kit, Qiagen, Maryland). DNA concentration was measured using a Nanodrop 8000 spectrophotometer. All samples were then diluted to 5 ng/µl to reduce bias in PCR amplification (Questel et al., 2021; and Bucklin et al., 2019).

2.1.8.11.2 Primer design optimisation

The targeted regions for the zooplankton metabarcoding were both the Cytochrome Oxidase I (mICOIintF/jgHCO2198 and HexCOIF4/HexCOIR4) and the 18S (1380F_EU/1510R_EU and 1391F/ EukBr) as both regions have been shown to perform well in plankton metabarcoding studies (Djurjuus et al. 2018; Bucklin et al. 2019; Schroeder et al. 2021). Primer sets were chosen for both regions (Table 2.6). An overhang adapter sequence was added at the 5' end of some primers for compatibility with Illumina index and sequencing adapters (Illumina 2011).

Table 2. 6: Metabarcoding primer set summary.

Sequences marked in blue are Illumina overhang adapter sequences, Index 1 and 2 sequences are in green and are marked with X's as this sequence is variable for each different sample, those in red are the P5 and P7 sequences and those in black are locus. ndex 1 (i7) and Index 2 (i5) are examples of the type of primers used with the Index sequence itself being altered for different samples

Metabarcoding primers' set					
Primer ID	**Sequence**	**Region**	**T**	**Size**	**Reference**
mICOIintF	TCGTCGGCAGCGTCAGATGTGTATAAGAGAC AGGGWACWGGWTGAACWGTWTAYCCYCC	COI	50.8	480 bp	Leray *et al.* (2013)
jgHCO2198	GTCTCGTGGGCTCGGAGATGTGTATAAGAGA CAGTAIACYTCIGGRTGICCRAARAAYCA	COI	47.5	480 bp	Leray *et al.* (2013)
LCO1490	GGTCAACAAATCATAAAGATATTGG	COI	55	670 bp	Folmer *et al.*, (1994)
HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	COI	55	670 bp	Folmer *et al.*, (1994)
HexCOIF4	TCGTCGGCAGCGTCAGATGTGTATAAGAGAC AGHCCHGAYATRGCHTTYCC	COI	51.9	460 bp	Marquina *et al.* (2018)
HexCOIR4	GTCTCGTGGGCTCGGAGATGTGTATAAGAGA CAGTATDGTRATDGCHCCNGC	COI	51.9	460 bp	Marquina *et al.* (2018)
1380F_EU	TCGTCGGCAGCGTCAGATGTGTATAAGAGAC AGCCCTGCCHTTTGTACACAC	18S	54.6	275 bp	Bucklin et al. (2019)
1510R_EU	GTCTCGTGGGCTCGGAGATGTGTATAAGAGA CAGCCTTCYGCAGGTTCACCTAC	18S	52.6	275 bp	Bucklin et al. (2019)
EukBr	GTCTCGTGGGCTCGGTGATCCTTCTGCAGGTT CACCTAC	18S	58.9	229 bp	Djurjuus et al. (2018)
1391F	TCGTCGGCAGCGTCGTACACACCGCCCGTC	18S	70	229 bp	Djurjuus et al. (2018)
Index 1	CAAGCAGAAGACGGCATACGAGATXXXXXX XXGTCTCGTGGGCTCGG	-	-	-	Illumina (2011)
Index 2	AATGATACGGCGACCACCGAGATCTACACXX XXXXXXTCGTCGGCAGCGTC	-	-	-	Illumina (2011)

2.1.8.11.3 PCR optimisation

PCR reactions were performed using a PCRmax Alpha cycler (Fisher) using the following mastermix: 12.5 µl of Environmental Mastermix (BioRad); 1 µl of both forward and reverse primer; 7.5 µl of nuclease-free water; and 3 µl of template DNA. A touchdown PCR (Don et al. (1991)) was used to amplify the extracted DNA and included: an initial incubation for 5 minutes at 95ºC; then 17 cycles (denaturation at

95ºC for 30 seconds, annealing temperature for 30 seconds, and extension at 72ºC for 60 seconds) where the annealing temperature is reduced by 1ºC each cycle from 62ºC down to 47ºC; followed by 30 cycles at an annealing temperature of 46ºC and a final extension step at 72ºC for 30 seconds before holding at 4ºC until collection of PCR products for analysis. PCR optimisation is summarised in Table 2.7.

Table 2. 7: PCR optimisation summary.

PCR OPTIMISATION				
MASTERMIX	**DNA DILUTIONS**	**POSITIVE CONTROL**	**INHIBITION TEST**	**PCR PROTOCOL**
Environmental master	5 ng/ µl	Cultured *Daphnia*	Mesocosm samples spiked into positive control (*Daphnia Magna)*	Gradient
Q5® High-Fidelity 2X Master Mix (NEB	Original extracted DNA	Environmental samples provided from previous studies	Mesocosm samples spiked into positive control (provided from ADAS*)*	Touchdown
2× KAPA HiFi HotStart ReadyMix	NONE	NONE	NONE	Second Round
NONE	NONE	NONE	NONE	Bucklin et al. (2019) protocol

A positive control was used in each PCR reaction to check the reliability. The positive controls used were pure cultured *Daphnia* (*Daphnia magna*), and saltwater copepods *Tigriopus Californicus* ordered from Reeph*yto (https://reefphyto.co.uk/products/live-copepods).*

Three commercial mastermix were compared to obtain the best amplification results using the samples at both 5 ng/ µl and at the original concentration: Environmental master mix (Thermo Fisher), Q5® High-Fidelity 2X Master Mix (NEB) and 2× KAPA HiFi HotStart ReadyMix (NEB) (Table 2.8).

Table 2. 8: Summary of the three Mastermix protocols.

Environmental master mix (25 µl reaction):	Q5® High-Fidelity 2X Master Mix (NEB) (25 µl reaction):	2× KAPA HiFi HotStart ReadyMix (25 µl reaction):
12.5 µl Mastermix	5 µl Q5 reaction buffer	7.5 µl Mastermix
7.5 µl Water	0.5 µl (10 mM) dNTPs	5 µl F/R primer (1 µM)
1 µl F/R Primer	0.25 µl Q5 Polymerase	2.5 µl DNA template (5 ng/ µl)
3 µl DNA template (5 ng/µl)	5 µl Q5 GC enhancer	-
-	1.25 µl F/R Primer (10 µM	-
-	8.75 µl Water	-
-	3 µl DNA template (5 ng/µl)	-

Two inhibition tests for both Hex and COI primers were made to check if any organism or substance present in the sample was inhibiting the PCR. The first inhibition test consisted of spiking three concentrations (1 µl, 2 µl and 3 µl) of the mesocosm DNA into 1 mL of the *Daphnia magnia* culture extraction to check if the positive control would be affected by the mesocosm samples. The second inhibition test consisted of spiking 3 µl of mesocosm DNA sample into an environmental sample (provided by ADAS) known to PCR amplify COI amplicon from their studies, to check if amplification was going to be inhibited by the mesocosm samples.

Finally, the PCR protocols were optimised by first performing a gradient PCR (65°C – 55 °C) to check for the best annealing temperature. Then, three different PCR protocols were compared (summarised in Table 2.9). A touchdown protocol (Leray 2013) was used to increase the specificity, sensitivity, and yield. The touchdown PCR was first tested with the primers containing the tag compatible with Illumina sequencing. Secondly, a 2-stage PCR was carried out to attach the illumina tags. This was done by doing a touchdown PCR using the primers without the tag, the PCR product was then purified using the Machery-Nagel PCR clean up and gel extraction according to manufacturer instruction, and finally a second round PCR was done on the purified

product using the primers containing the Illumina tag. Both the touchdown and the 2-stage PCR were compared with the PCR protocol used in Bucklin et al. (2019).

Table 2. 9: Summary of PCR protocols.

PCR protocols	
Gradient	56°C for 2 min, 30 cycles of: 95 °C for 30s, gradient: 65°C – 55 °C and a final extension of 72°C for 1 minutes.
Touchdown (Don et al., 1991)	95°C for 5 min, a 17-cycle touchdown phase (denaturation at 95°C for 30 s, annealing temperature for 30 seconds, and extension at 72ºC for 60 seconds where annealing temperature is reduced by 1ºC each cycle from 62ºC down to 47ºC) 30 cycles of 46°C, and a final extension at 72°C for 30 s.
Second Round (Performed after a touchdown PCR)	95°C for 3 min, 12 cycles of: 95°C for 30 s, 63°C for 30 s, 68°C for 30s; and a final extension for 5 minutes.
Bucklin et al. (2019)	98°C for 30 s; 10 cycles of: 98°C for 10 s, 56°C for 30 s, 72°C for 15 s; 15 cycles of: 98°C for 10 s, 62°C for 30 s 72°C for 15 s; and a final extension of 72°C for 7 minutes.

As the Hex and COI primers were giving inconsistent amplification results across a range of mesocosm samples from different sampling dates (day 1- 4), two different 18S primer sets from two published papers were tested to check for reproducibility (Djurhuus et al. 2018; Bucklin et al. 2019). These primers were first tested on *Daphnia Magna* cultures and then on mixed zooplankton samples.

2.1.8.11.3.1 Final PCR protocol

The PCR master mix (per sample) consisted of 2.5 µl genomic DNA (5 ng/µl); 5 µl forward PCR primer 1380F_EU (1 µM); 5 µl reverse PCR primer 1510R_EU (1 µM); 12.5 µl 2× KAPA HiFi HotStart ReadyMix; for a total volume per sample of ultrapure water to a final volume of 25 µl. The PCR protocol was 98°C for 30 s; 10 cycles of: 98°C for 10 s, 56°C for 30 s, 72°C for 15 s; 15 cycles of: 98°C for 10 s, 62°C for 30 s 72°C for 15 s; and 1 cycle of 72°C for 7 min.

2.1.8.11.4 Optimised Gel-Electrophoresis

Agarose gel electrophoresis (1.5 %) was used to run PCR products to check size and for PCR clean-up and gel extraction. DNA was extracted from agarose gels using the NucleoSpin Gel and PCR Clean-up kit (Thermo Scientific) according to the manufacturer's instructions. The purified sample's DNA concentration was measured using a Nanodrop 8000 spectrophotometer.

2.2 Methods used in Results Chapter 4: Assessment of zooplankton metabarcoding as a monitoring approach compared to taxonomic identification by microscopy

For this chapter, the following methods (Database design, 2.1.8.1; DNA-extraction, 2.1.8.11.1 and agarose gel extraction, 2.1.8.14) have already been described.

2.2.1 Mesocosm experiment design

Zooplankton samples were collected as per 2.1.8.2. A general overview of the workflow utilised in the mesocosm experiment design is given in Figure 2.4. Water samples were collected from 5 different mesocosms (162, 163, 168, 173 and 183) located at CEA, Boxworth, as shown in Figure 2.5. The sampling was performed during the main zooplankton abundance season (May-September 2020), on seven different collection dates: Week 22=25th May, Week 24=8th June, Week 26=22nd of June, Week 28=6th July, Week 30=20th July, Week 33=10th August, Week 37=1st September (Figure 2.6), resulting in a total of 35 zooplankton samples. These were preserved in 20- or 40-mL ethanol (depending on sample turbidity) at room temperature.

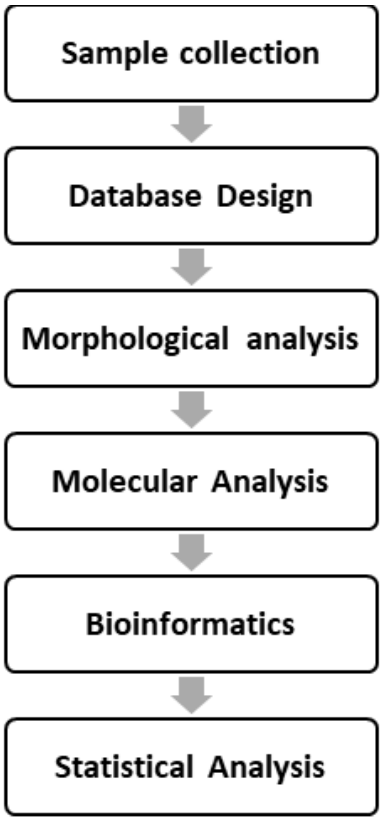


Figure 2. 4: Overview of major steps in the mesocosm experiment design.

Figure 2. 5: Selected mesocosms for the study.

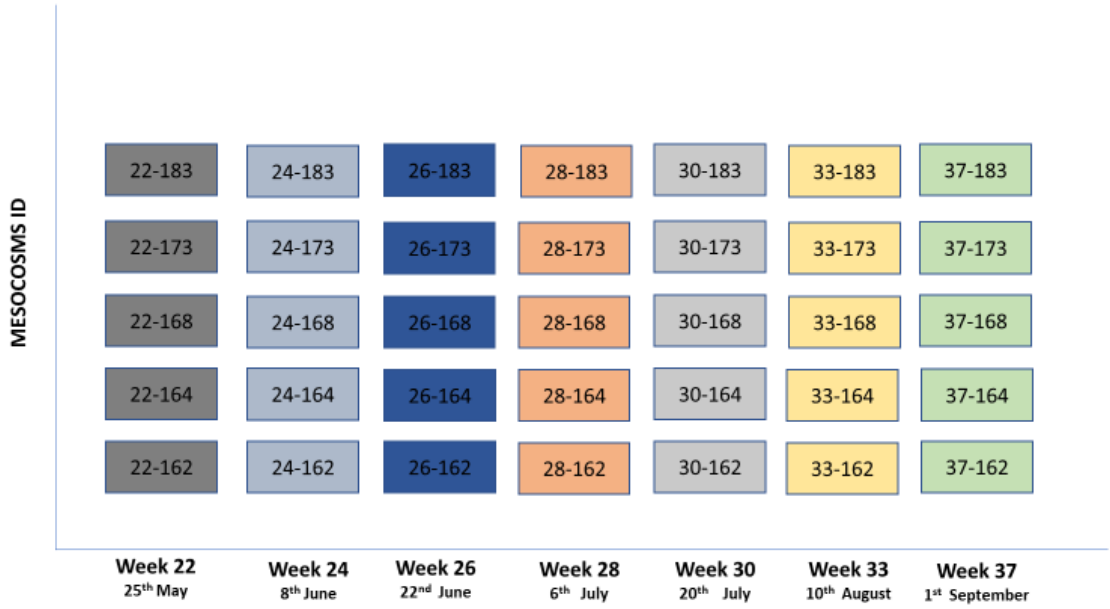


Figure 2. 6: Collection dates and selected mesocosm ID.

Samples were collected between the 25th of May and the 1st of September. ID, where the first number represents the collection week and the second number week, presents the mesocosm number.

2.2.1.1. Water Chemistry Analysis

Water quality parameters (temperature, pH, dissolved oxygen, conductivity, turbidity, and total chlorophyll) measurements were collected from each mesocosm at a water depth of approximately 25 cm using a YSI 6600 V2 Multiparameter Water Quality Sonde on each zooplankton sampling day. The six environmental variable's data were plotted to give a visual representation of the variation across the collection dates.

2.2.2 Morphological analysis

Morphological analyses were carried out according to subsection 2.1.8.3. A scatter plot graph was made for each of the eight most abundant zooplankton genera found during the seven collection days. The graphs were compared with the historical control data (HCD) (Brooks et al., 2019) which includes zooplankton abundance data recorded over seven years from the same location at CEA, Boxworth, to check for reproducibility.

2.2.3 Molecular analysis

2.2.3.1 PCR of zooplankton samples

PCRs were performed in triplicates using the final established protocol according to subsection (2.1.8.13). Two repeats of PCRs were made (R1 and R2) to compare NGS reliability in successive analyses. For two randomly selected samples (1A and 7A), PCRs were also performed in 3, 6, 9 and 12 replicates to check for optimisation of PCR assay analysis. *Tigriopus Californicus* was used as a positive control for all analyses.

2.2.4 Illumina sequencing

2.2.4.1 Sample preparation

96 samples including a positive (*Tigriopus Californicus*) and a negative control (freshly purified ultrapure water) were prepared for Illumina sequencing. The sequencing was conducted by DeepSeq, a next-generation sequencing external facility at The University of Nottingham. Samples were divided into two different repeats to check for reproducibility (as shown in 4.6.2). The final 96 concentrated DNA samples were analysed using Qubit dsDNA Quantification, High Sensitivity Assay kit. The sequencing methodology involved the processing of 96 amplicons targeting the 18S region. Firstly, PCR was conducted with Illumina adapters. The chosen polymerase for the first PCR step was Taqman Environmental Mastermix 2.0 (Fisher Scientific), aligning with the approach employed in previous RSK ADAS projects on environmental samples. Subsequently, at DeepSeq a subset of approximately 4-8 amplicons underwent quality check before all samples were progressed to index incorporation, a cleanup step, quality control of the library preparations, and pooling of samples. The final step involved sequencing the libraries on a MiSeq run, aiming for approximately 80-100,000 250PE reads per sample for a full run. PCRs were set up in a total volume of 50 µL consisting of:

a. 25 µl 2x KAPA HotStart ReadyMix
b. 5 µl Nextera XT Index 1 Primers
c. 5 µL Nextera XT Index 2 Primers
d. 10 µL PCR grade water
e. 5 µL DNA

PCR cycling was as follows: an initial incubation for 3 minutes at 95°C; followed by 8 cycles with denaturation at 95°C for 30 seconds, annealing at 55°C for 30 seconds, and

extension at 72°C for 30 seconds; and a final extension step at 72°C for 5 minutes before holding at 4°C until collection of PCR products for analysis.

The indexed amplicons were quantified via a fluorometric method involving QuantiFluor dsDNA assay (Promega); and qualified using electrophoretic separation on the Agilent TapeStation 4200. This concentration and sizing information was used to calculate the molarity of each sample. Libraries were then pooled in equimolar amounts to create one library for Illumina sequencing. The amplicon library pool was spiked with 20 % PhiX Control v3 library (Illumina) and run on the Illumina MiSeq using a MiSeq Reagent Kit v2 500 cycle kit (Illumina), to generate 250-bp paired-end reads. PhiX DNA is derived from the small, well characterized bacteriophage PhiX genome, it is a concentrated Illumina library (10 nM in 10 µl) that has an average size of 500 bp and consists of balanced base composition at ~45% GC and ~55% AT and serves as an in-run QC for the Illumina sequencing.

2.2.5 Bioinformatic analysis

2.2.5.1 Bioinformatic pipeline for Illumina analysis

A pipeline provided by Dr Jon Owen (RSK ADAS, personal communication, (https://github.com/svxvm1/zooplankton_metabarcoding_IlluminaNanopore_pipelines/blob/main/Illumina.sh)) was used as a template and modified to analyse the Illumina output file (FASTQ) to obtain the blasted sequences (Figure 2.7). Data processing was performed on an Intel i7 PC, running Ubuntu Linux 20.04.3 LTS. The program FLASH 1.2.11 (Fast Length Adjustment of Short reads; Magnoc and Saltzberg 2011) was used to convert paired-end reads Read 1 and Read 2 in the MiSeq platform to a single merged read, using a minimum overlap length of 10 nucleotides (standard) and a maximum of 180 nucleotides to calculate the alignment. Reads were then trimmed reading from the 5’ end using Trimmomatic 0.38 (Bolger et al., 2014) to truncate the sequence if the average phred score of a 5nt sliding window dropped below 30.

Those reads that matched the template-specific primers at the 5’ and 3’ ends (maximum error rate of 0.1% within target-specific primer site, i.e., 2 bp variants allowed) and had a target region of >120bp were then extracted from the data using Cutadapt 1.18 (Martin 2011). Degeneracy within the primer sequences was accounted for when identifying primer sequences within the dataset. Data was then converted from fastq to fasta format using seqtk-1.3 (r106) (GitHub).

Before taxonomic assignment, standard Linux tools were used to identify 100% identical reads and condense them into a single read to minimise time-consuming repetitive BLAST searches. However, a record of the frequency of replicate sequences was maintained. Any reads with less than 3 replicates were excluded from the BLAST search.

A custom Zooplankton BLAST database was created from the Barcode of life data system (BOLD) and GenBank databases using each zooplankton genera known to be found at the CEA mesocosms (Table 2.3 in section 2.1.8) and '18s' before downloading the records in FASTA format for a total of 1,432 sequences.

BLAST searching was performed using the "megablast" program whit an e-value of 1e-15, which is optimised to identify alignments in highly similar sequences and return the top hit for each query sequence in a custom tabulated format.

A custom-written Perl script (https://github.com/svxvm1/zooplankton_metabarcoding_IlluminaNanopore_pipelines/blob/main/Perl.pl) filtered the BLAST output, identifying hits sharing an accession number and passing a set of criteria covering the percentage similarity between the query sequence and the database sequence (99%), and having a query alignment length difference less than 6 bp. Note that ≥ 99% similarity indicates an approximately three-base pair difference between query and reference sequences because the maximum sequence length subjected to taxonomic assignment was around 300 bp. Different percentages of similarity (98, 97 and 96%) were tested. Read counts for each sequence passing the similarity and query alignment length filters were pooled based on the accession number to generate a final frequency count for each accession. As many reads did not match zooplankton's genera in the database, three extra databases were made ("Phytoplankton", "Arthropoda" and "Amphibia") to establish which other components may be present in the samples. Results were then statistically analysed.

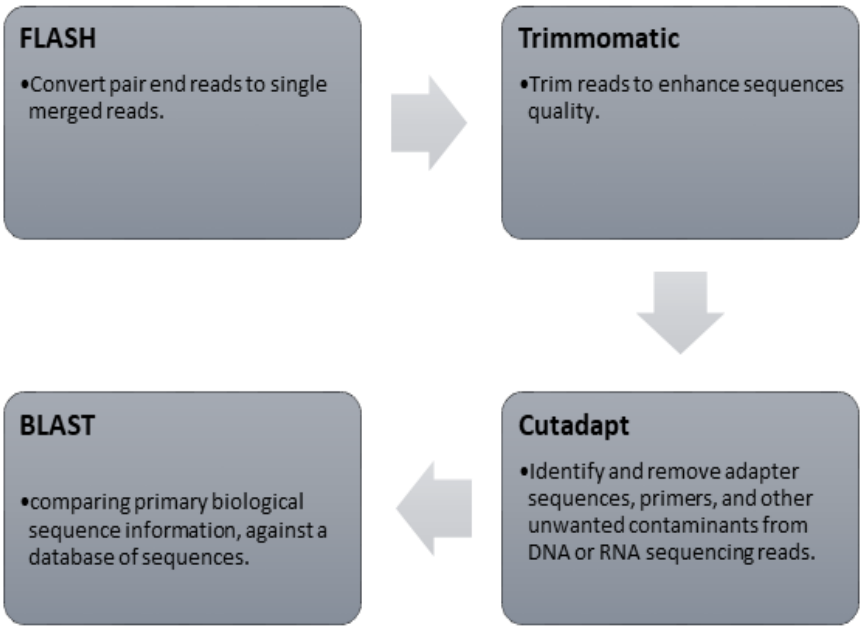


Figure 2. 7: Pipeline of the major steps in the bioinformatics process used to analyse NGS data.

2.2.6 Statistical analysis

2.2.6.1 Water Quality Analysis

All statistical analyses were carried out using R-4.3.2, Prism 10.2.1 (395) and PAST 4.03 (Hammer et al., 2001).

The six environmental variables data were plotted to visually represent the variation across the collection dates.

A one-way Analysis of Variance (ANOVA) with Tukeyy post-hoc test was performed on each environmental variable (Temperature, pH, dissolved oxygen, conductivity, and turbidity) to ascertain if there were any significant differences between the seven sampling dates.

A Mantel test was performed to assess the correlation between the environment variables and the zooplankton genera, using the Bray-Curtis dissimilarity matrix (Dutilleul et al., 2000).

2.2.6.2 Molecular and Morphology Analysis for Zooplankton genera

The primary aims were to 1) identify the most prevalent genera of zooplankton by the 'gold standard' microscopy method and then determine whether these genera could also be identified by molecular analysis and 2) for genera that were detected by both methods across the season (at least 3 samples times) determine whether their relative abundance correlated when measured by the two methods.

2.2.6.3 Data transformation and normalisation

Data from the morphological identification was reported as relative abundance, this was done by dividing the abundance of a genus (organism per litre) by the total number of organisms in the sample and then multiplying by 100. All metabarcoding data for the selected genera was transformed into percentage abundance by dividing the individual species read count by the total number of reads and multiplying by 100 to produce a ranking list from the most to the least abundant genera, per mesocosm at each sampling date.

2.2.6.4 Metabarcoding reproducibility

PCR reproducibility using data from 3, 6, 9 and 12 repeats was tested using ANOVA to check for significant differences among the datasets. Additionally, the two metabarcoding repeats (R1 and R2) were tested against each other to check for reproducibility using a paired T-test.

2.2.6.5 Abundance comparison in both morphology and metabarcoding analyses

To compare the two methodologies, a bar chart for each sampling date was created for Morphology and Metabarcoding % abundance, including all the zooplankton genera found in each mesocosm. A Venn Diagram was produced showing the highest abundance genera identified by each methodology and the genera identified by both. As the majority of genera were not represented well across the season., a selection of the communal genera for both methodologies that were present in more than two mesocosms (replicate) per collection day was made.

A scatter plot graph, including error bars, for all the eight selected genera was generated for both methodologies including abundance (%) over each sampling date. A test was made to check for correlation between the two methodologies by making a Spearman Rank correlation test. This method is used to quantify the degree to which two variables are related or associated.

Finally, a Similarity test using the Bray-Curtis Similarity and Analysis of Similarities (ANOSIM) test, to explore relationships between the sampling dates for both methodologies, was conducted. These tests are often used in ecological research to analyse and interpret community dissimilarity or similarity patterns. These methods are

particularly useful when dealing with multivariate data, such as community composition data, where there are taxa across different samples or sites.

2.3 Methods used in Results Chapter 5: Ecotoxicity study using Roundup™

For this chapter, the following methods: Water quality analysis, 2.2.1.1; Zooplankton sampling, 2.1.8.2 and Illumina sequencing preparation, 2.2.4.1, have already been described.

2.3.1 Herbicide selection

The initial aim was to select a pesticide known to have a direct effect on zooplankton. However, the two originally selected insecticides (Imidacloprid and Acetamiprid) were not feasible due to the amount of test item, and associated costs, required to dose all the replicate mesocosms. The herbicide, Roundup™ Super Concentrate (Monsanto®), containing 480g of glyphosate per litre of liquid, as an active ingredient and other Surfactant blend (proprietary) was selected as the test item. Glyphosate has been widely used in ecotoxicity studies (Hébert et al., 2020), and therefore, the aim here was to investigate its indirect effects on zooplankton. The approximate area of the mesocosms was 86.11 m2 (775 L per six mesocosms), which, according to the product calculator on the website where Roundup™ was ordered, meant that 0.03 L were needed for a price of £84.24.

Glyphosate target doses were 0 mg a.i./L (control), 5.50 mg a.i (active ingredient)/L (low dose), and 22.0 mg a.i./L as high dose (Figure 2.8). These doses were selected to have a minimal effect on the zooplankton at the low dose, with the high dose selected to have a comprehensive reduction and/or decline in the zooplankton community. These concentrations were based on the EC_{50} (effective concentration at which 50% of organisms are affected) of the active ingredient glyphosate on *Daphnia magna* found on the safety data sheet (SDS) of the product:

(https://labelsds.com/images/user_uploads/Roundup%20Pro%20Conc%20SDS%208-12-20.pdf).

The data in the Roundup™ safety sheet provided insights into the toxicity of active ingredient glyphosate, on *Daphnia magna*. In a static test lasting 48 hours, the EC50 for glyphosate was found to be 11 mg a.i./L, indicating the concentration at which adverse effects on *Daphnia* magna become significant within a relatively short exposure period. Additionally, chronic toxicity assessments conducted over 21 days revealed an EC50 of 12.5 mg a.i./L, suggesting that glyphosate continues to exert adverse effects on *Daphnia* magna over a longer duration. The value mentioned relates to the active ingredient glyphosate. These findings underscore the importance of understanding both short-term and long-term impacts of glyphosate on aquatic invertebrates for effective environmental risk assessment and management.

Final concentrations of glyphosate $C1\ x\ V1 = C2\ x\ V2$ (where: V1 = volume of starting solution needed to make the new solution, C1 = concentration of starting solution, V2 = final volume of the new solution, C2 = final concentration of new solution).

- **Application for low dose:** $480 g/L\ x\ V1 = 5.5 mg/L\ x\ 775\ L$

V1= $\frac{775 \times 0.0055}{480} = 8.8\ mL$

- **Application for high dose:** $480 g/L\ x\ V1 = 22.0 mg/L\ x\ 775\ L$

V1= $\frac{775 \times 0.022}{480} = 35.5\ mL$

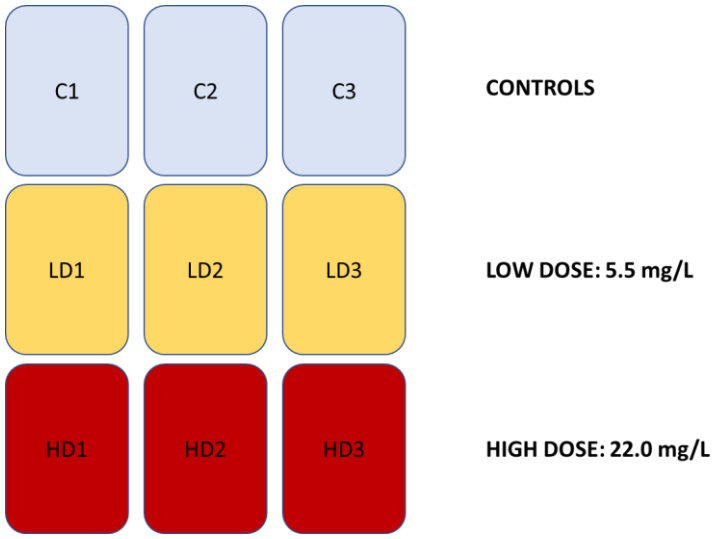


Figure 2. 8: Scheme of herbicide selection at each dose.

Where C1-3 were controls, LD1-3 were Low-dose at 5.5 mg a.i./L and finally HD1-3 were High-dose at 22.0 mg a.i./L.

2.3.2 Mesocosm Experiment Design, sampling, and data analysis

Mesocosms were set up with nine different plants (Table 2.10), which would be targeted by the glyphosate together with an established phytoplankton community (food source of the zooplankton).

Table 2. 10: List of plants present in the mesocosms.

Species	Family	Embryonic class	Morphological characteristics	Planting depth in cm (below water)
Potomogeton natans	Potomogetonaceae	Monocot	Rooted, floating leaved	50
Elodea canadensis	Hydrocharitaceae	Monocot	Rooted, submerged	5
Sparganium erectum	Typhaceae	Dicot	Rooted, emergent	40-50
Myriophyllum spicatum	Haloragaceae	Dicot	Rooted, emergent	30-40
Glyceria maxima	Poaceae	Monocot	Rooted, emergent	30-40 (in pots)
Hippuris vulgaris	Plantaginaceae	Dicot	Rooted, emergent	20-30
Saggitaria sagittifolia	Alismataceae	Monocot	Rooted, emergent	20-30

Myosotis scorpioides	Boraginaceae	Dicot	Rooted, emergent	10-20
Veronica beccabunga	Scrophulariaceae	Dicot	Rooted, emergent	0-10

Sampling of zooplankton, and the water analyses, was divided into two phases, pre- and post-application, and took place from mid-May until the beginning of September 2022 on the following dates: Sampling 1 pre-application on week 18 of the year (03/05/22) pre-application day 43, Sampling 2 pre-application on week 20 (17/05/22) pre-application day 29, Sampling 3 pre-application on week 22 (31/05/222) pre-application day 15, Application Day (before herbicide application) on week 24 (15/06/22) day 0, Sampling 1 post-application on week 25 (23/06/22) post-application 8, Sampling 2 post-application on week 26 (30/06/22) post-application 15, Sampling 3 post-application on week 27 (07/07/22) post-application 22, Sampling 4 post-application on week 28 (14/07/22) post-application 29, Sampling 5 post-application on week 30 (28/07/22) post-application day 43, Sampling 6 post-application on week 32 (11/08/22) post-application day 57, Sampling 7 post-application on week 34 (26/08/22) post-application day 72, Sampling 8 post-application on week 36 (08/09/22) post-application day 85, and finally Sampling 9 post-application on week 38 (22/09/22) post-application day 99. The total duration of the experiment was 20 weeks. Of thirty sloped mesocosms at CEA for ecotoxicity trials, nine were randomly selected for this pilot study before application (Figure 2.9).

Figure 2. 9: Location of nine selected Mesocosms within the CEA facility, respective glyphosate concentration scheme is shown.

2.3.2.1 Molecular analysis

All the molecular analyses, except for adding a filtering step were made according to section 2.2.3. The samples that were found turbid at 20 mL were diluted to 40 mL (according to a CEA SOP). If still turbid, they were re-filtered in the laboratory. The filtering step used a coarse mesh of 10 mm before DNA extraction to remove large plant material and algae as samples were too cloudy.

2.3.2.2 Bioinformatics analysis

All the scripts used are stored in

https://github.com/svxvm1/zooplankton_metabarcoding_IlluminaNanopore_pipelines/tree/main.

The zooplankton database was updated by retrieving a total of 1,596 (164 sequences more than the one retrieved in 2019), sequences from GenBank for the 18s gene based on the genera list provided by CEA (Table 2.3, Section 2.1.8).

Finally, the same pipeline used in section 2.2.5.1 was used to analyse the data.

However, another pipeline was also applied to check for the reproducibility of the results when different methods were used. This comparative analysis aimed to investigate how different methods might impact the results' robustness and reliability, providing insights into the overall reproducibility of the experimental outcomes. The script chosen was Banzai which is a shell script that links together the different programs needed to process the raw results from Illumina sequencing

(https://github.com/jimmyodonnell/banzai). This script was selected after thorough research of the literature analysing ecological environmental DNA data on zooplankton produced using Illumina sequencing, such as Pitz et al. (2020) and Min et al. (2022).

2.3.3 Statistical analysis

All statistical analyses were performed using R-4.3.2, Prism 9.0, Canoco 5.10, Community Analysis V.4.3.16 and PAST 4.03. Prior to statistical analysis, data obtained from NGS were normalised as per section 2.2.6.3. The eight most abundant genera (*Chydorus, Cyclopoidae, Diaptomidae, Daphnia, Keratella, Lepadella, Trichocerca* and *Trichotria*) present with both methodologies were selected as per section 2.2.6.5.

2.3.3.1 Water Chemistry Analysis

A scatter plot was made to visualise the differences in the five environmental variables (temperature, oxygen, pH, conductivity, and turbidity) across the season. An ANOVA test was performed on each environmental variable at each concentration to check how the different doses changed between pre- and post-application. A repeated measures analysis (RM ANOVA) test, with Dunnett's post-hoc test, for both pre-and post-application were also performed to check for significant differences within the five environmental variables between the control, low and high concentrations of glyphosate.

2.3.3.2 Investigating Zooplankton Dynamics: Seasonal Pesticide Variability, Statistical Analyses, and Abundance Comparison

A Detrended Correspondence Analysis (DCA) non-linear model was performed using CANOCO to identify taxonomic groupings of samples distributed across environmental gradients, in this study case (temperature, oxygen, pH, conductivity and turbidity) for the morphological data. In the analysis, the control and the two glyphosate concentrations were included (5.5 mg a.i./L and 22.0 mg a.i./L). In the settings, rare taxa were down-weighted to avoid the use of taxa present only once influencing the results, empty taxa were omitted (only 0s) and all data were log-transformed (log(x+1)).

On the other hand, a linear model was fitted for the metabarcoding data as data was presented as percentage values. A Principal Component Analysis (PCA) was performed using CANOCO to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set. In this analysis, all the environmental variables (temperature, oxygen, pH, conductivity and turbidity) and the two glyphosate concentrations were included (5.5 mg a.i./L and 22.0 mg a.i./L). CANOCO did not allow the downweight of rare taxa, empty taxa were omitted, and no log transformation was made as data were percentage values rather than absolute numbers.

Analysis of glyphosate effects on community composition were performed according to Szöcs et al., (2014) based on a CEA report template. A Principal Response Curve (PRC) analysis was performed on data from both methodologies. The Principal Response Curves (PRCs) method, introduced by van den Brink and ter Braak (1998, 1999), is commonly used for analysing community-level responses in mesocosm experiments. PRC is a multivariate constrained ordination technique that represents a specialised form of Redundancy Analysis (RDA), similar to the multivariate extension of linear regression as expounded by Legendre (2012). By relegating temporal changes to a secondary role, termed as 'partialling out', PRC effectively isolates the impact of time from the response, thereby allowing for a more focused analysis (partial RDA). The subsequent application of a RDA model incorporates treatment and its interaction with time as predictors, yielding an ordination diagram where the primary axis encapsulates the maximum variation attributed to treatment and treatment × time interaction.

A Non-metric Multi-dimensional Scaling (NMDS) (De Puelles et al., 2019, Song et al., 2021) was created for each concentration for both methodologies to analyse and compare community composition data at each sampling point. This was also reinforced by creating a hierarchical phylogenetic tree for each sampling day by using Bray-Curtis dissimilarity. ANOSIM was tested on the 11 time points pre- and post-application of Roundup™ to assess whether there were significant differences between groups of samples based on a dissimilarity matrix (Song et al., 2021). The final test for community composition was to calculate both Shannon Diversity Index (to measure biodiversity) and Chao-1 (to estimate the total number of species in a community by incorporating rare and unobserved species) both pre- and post-application of Roundup™ and to visually represent it using a boxplot for both methodologies (Honggang et al., 2012, Song et al., 2021). Both indices for both methodologies were then analysed using RM ANOVA, with Dunnett's test for post hoc comparison, to assess the difference between both low and high glyphosate doses compared to the control.

2.3.3.2.1 Data analysis of the effects of glyphosate applied to mesocosms at different concentrations

The population impacts were assessed for each group of organisms, treatment level, and sampling instance using a univariate analysis, specifically Dunnett's T-test, along with Minimum Detectable Difference (MDD) using the Community Analysis (CA) software V.4.3.16. Organisms were selected for analysis based on their frequency of occurrence (i.e., having an abundance greater than zero) and the total number of individuals observed throughout the research. Organisms with highly erratic and variable data were excluded from the statistical examination. The determination of Effect Classes for each population was guided by the results obtained from the univariate and MDD analyses, following the recommendations outlined in the EFSA Aquatic Guidance Document (2013) and the work of Brock et al. (2014). Furthermore, the NOECpopulation (no observed effect concentration) and NOEAECpopulation (No observed ecologically adverse effect concentration) were calculated based on the Effect Classification published by EFSA (2013) and adapted by Brock et al. (2015).

The NOEC is a critical parameter in environmental risk assessment, particularly in the context of microcosm and mesocosm studies evaluating the effects of pesticides on

aquatic ecosystems. It is commonly calculated for various measurement endpoints to determine the concentration at which no statistically significant effects are observed compared to control conditions. The NOEC is essential for understanding the treatment-related effects of pesticides in experimental ecosystems and is used to derive effect classes for different taxa and concentrations (Brock et al., 2015) while the NOEAEC takes into consideration the effects that are ecologically permissible and recovery properties from the symptoms of effects (EFSA, 2013).

2.3.3.3 Morphology and Metabarcoding Comparison

The Pearson coefficients was calculated to perform a correlation test on the eight genera and the three groups between the two methodologies (Schroeder et al., 2020). Furthermore, the performance of metabarcoding during an ecotoxicological trial was assessed by comparing step-by-step results from both methodologies for each analysis.

2.4 Methods used in Results Chapter 6: Oxford Nanopore sequencing

2.4.1 Library preparation

The library used for Nanopore sequencing was the same library used for Illumina sequencing (section 2.3). The library was retrieved from DeepSeq after the Illumina analysis and stored at -20 °C before being prepared for Nanopore sequencing. Library preparation and DNA sequencing were performed using the Ligation Sequencing amplicons - Native Barcoding 96 V14 protocol with the SQK-NBD114.96 kit (Oxford Nanopore Technologies, version NBA_9170_v114_revH_15Sep2022, last updated 05/04/2023), with 200 fmol (130 ng for 1 kb amplicons) of DNA per sample used as input for library preparation. Sample concentration was analysed using Qubit™ dsDNA Quantification Assay Kits (Invitrogen). The target concentration (to reach 200 fmol required for the protocol) was calculated using the Promega Biomaths calculator

(https://ita.promega.com/en/resources/tools/biomath/). This was calculated using the 270bp size from the primer used for the PCR (Buckley et al., 2018). The concentration obtained was: 35 ng/ 11.5 µl, or 3.04 ng/µl. The formula used to calculate how much DNA was needed for each sample was the following: $C_1 \times V_1 = C_2 \times V_2$ where C_1= results

from Qubit, V_1= is the DNA concentration required, C_2= Final concentration and V_2= volume required.

In this case: $C_1 \times X = 3.04$ ng/µl × 11.5 µl. This did not result in a large amount of data loss as other sample replicates were still included in the Nanopore sequencing. Samples were already barcoded with the Illumina barcode, so the samples were pooled instead of analysing them separately to reduce the analysis time. Based on the concentrations obtained and the final concentration requested by the Nanopore protocol, it was decided to pool samples into ten pools of eight samples to then demultiplex using bioinformatics tools. The DNA concentration of nine samples was too low to work with, so they were discarded (PA26, PA29, PA31, PA36, PA38, AD3, AD4, AD8 and PO32). As pools had to have the same number of samples to facilitate calculations, the samples with the highest DNA concentration were also excluded to obtain an equal number of samples (PA37, AD5, PA11, AD1 and PO67). Ten pools with samples chosen at random were made (Table 2.11). Samples had DNA concentrations in the range of 0.82 to 76. Each sample was diluted to give a final concentration in the pool of 35 ng in 11.5 µl (Table 2.12). To each pool, a Nanopore barcode was assigned (Table 2.13) according to the 96 well plate barcode organisation shown in Figure 2.10.

Table 2. 11: Mixed Samples pooled together to create the ten pools (1-10) in preparation for Nanopore sequencing.

Each colour represents a sampling time point thus illustrating that each pool was made from a mixture of time points.

1	2	3	4	5	6	7	8	9	10
PO76	PA18	PO38	PA16	AD9	PO77	PO11	PA32	AD7	PA15
PA39	PO66	PA23	AD6	PA35	PO81	PO28	PO88	AD2	PO65
PO75	PO24	PO56	PO25	PO78	PO13	PO71	PA28	PA22	PA13
PA19	PO39	PO34	PO35	PO26	PO61	PO55	PA34	PO18	PO82
PO58	PO68	PA21	PO74	PO64	PO19	PO73	PO21	PO72	PA17
PO59	PO15	PO33	PO17	PO87	PO22	PO63	PO79	PO51	PO16
PO23	PO27	PO37	PO53	PO83	PO52	PA14	PO62	PO57	PO12
PA27	PO14	PO31	PO36	PO89	PO69	PA24	PA12	PA25	PA33

Table 2. 12: Calculation to obtain the requested final pool concentration.

C1 × V1=C2 × V2 where C1= results from Qubit, V1= is the DNA concentration needed to find, C2= Final concentration and V2= volume needed.

DILUTED CONCENTRATIONS							
Pool	**Sample name**	**C1 (ng/ul)**	**C2 (ng/ul)**	**V1 (ul)**	**V2 (ul)**	**Pool DNA (ul)**	**Nuclease-free water (ul)**
1	PO76	0.82	0.38	5.33	11.5	11.64	-0.14
	PA39	3.43	0.38	1.27	11.5		
	PO75	4	0.38	1.09	11.5		
	PA19	4.4	0.38	0.99	11.5		
	PO58	4.66	0.38	0.94	11.5		
	PO59	5.1	0.38	0.86	11.5		
	PO23	6.58	0.38	0.66	11.5		
	PA27	8.92	0.38	0.49	11.5		
2	PA18	0.91	0.38	4.8	11.5	11.49	0.01
	PO66	3.47	0.38	1.26	11.5		
	PO24	3.82	0.38	1.14	11.5		
	PO39	4.37	0.38	1	11.5		
	PO68	4.44	0.38	0.98	11.5		
	PO15	5.1	0.38	0.86	11.5		
	PO27	4.56	0.38	0.8	11.5		
	PO14	6.78	0.38	0.64	11.5		
3	PO38	1.01	0.38	4.33	11.5	11.28	0.22
	PA23	3.18	0.38	1.37	11.5		
	PO56	3.52	0.38	1.24	11.5		
	PO34	4.07	0.38	1.07	11.5		
	PA21	4.47	0.38	0.98	11.5		
	PO33	4.9	0.38	0.89	11.5		
	PO37	5.5	0.38	0.78	11.5		
	PO31	7.16	0.38	0.61	11.5		
4	PA16	1.12	0.38	3.9	11.5	10.78	0.72
	AD6	3.2	0.38	1.37	11.5		
	PO25	3.54	0.38	1.23	11.5		
	PO35	4.12	0.38	1.06	11.5		
	PO74	4.5	0.38	0.97	11.5		
	PO17	4.9	0.38	0.89	11.5		
	PO53	5.58	0.38	0.78	11.5		
	PO36	7.6	0.38	0.58	11.5		
5	AD9	1.67	0.38	2.62	11.5	9.4	2.1
	PA35	3.32	0.38	1.32	11.5		
	PO78	3.6	0.38	1.21	11.5		
	PO26	4.13	0.38	1.06	11.5		
	PO64	4.5	0.38	0.97	11.5		

	PO87	4.94	0.38	0.88	11.5		
	PO83	5.68	0.38	0.77	11.5		
	PO89	7.64	0.38	0.57	11.5		
6	PO77	1.69	0.38	2.59	11.5	9.23	2.27
	PO81	3.36	0.38	1.3	11.5		
	PO13	3.7	0.38	1.18	11.5		
	PO61	4.2	0.38	1.64	11.5		
	PO19	4.56	0.38	0.26	11.5		
	PO22	4.96	0.38	0.88	11.5		
	PO52	5.8	0.38	0.75	11.5		
	PO69	8.28	0.38	0.55	11.5		
7	PO11	1.76	0.38	2.48	11.5	9.06	2.44
	PO28	3.36	0.38	1.3	11.5		
	PO71	3.76	0.38	1. 16	11.5		
	PO55	4.22	0.38	1.04	11.5		
	PO73	4.6	0.38	0.95	11.5		
	PO63	5.02	0.38	0.87	11.5		
	PA14	5.94	0.38	0.74	11.5		
	PA24	8.3	0.38	0.53	11.5		
8	PA32	2.06	0.38	2.12	11.5	871	2.79
	PO88	3.38	0.38	1.29	11.5		
	PA28	3.77	0.38	1.16	11.5		
	PA34	4.35	0.38	1	11.5		
	PO21	4.63	0.38	0.94	11.5		
	PO79	4.7	0.38	0.92	11.5		
	PO62	5.96	0.38	0.73	11.5		
	PA12	8.36	0.38	0.52	11.5		
9	AD7	2.54	0.38	1.72	11.5	8.67	2.83
	AD2	2.84	0.38	1.54	11.5		
	PA22	3.52	0.38	1.24	11.5		
	PO18	4.07	0.38	1.07	11.5		
	PO72	4.64	0.38	0.94	11.5		
	PO51	4.67	0.38	0.94	11.5		
	PO57	6.24	0.38	0.7	11.5		
	PA25	8.5	0.38	0.51	11.5		
10	PA15	2.76	0.38	1.58	11.5	8.68	2.82
	PO65	3.04	0.38	1.44	11.5		
	PA13	3.47	0.38	1.26	11.5		
	PO82	4.03	0.38	1.08	11.5		
	PA17	4.47	0.38	0.98	11.5		
	PO16	5.17	0.38	0.85	11.5		
	PO12	5.2	0.38	0.84	11.5		
	PA33	6.72	0.38	0.65	11.5		

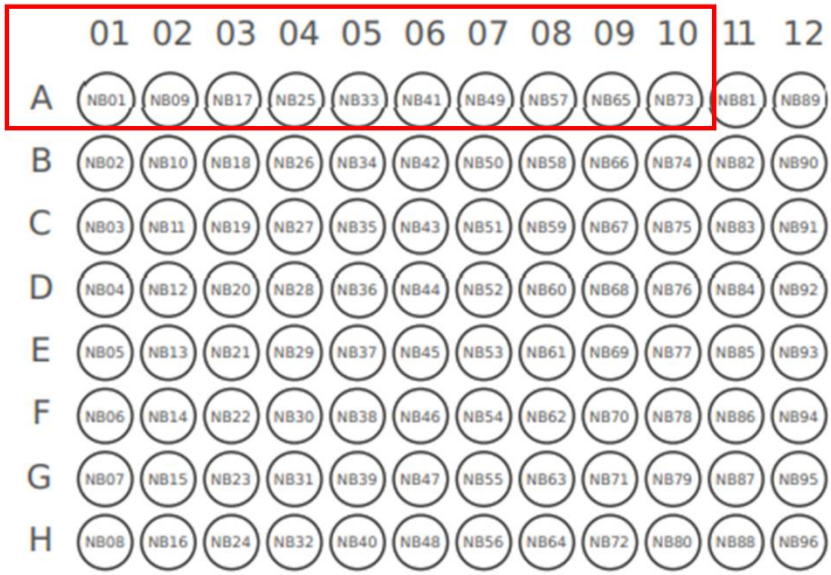


Figure 2. 10: Highlighted in red are the 10 selected barcodes out of the 96 for each pool in this study.

Table 2. 13: Forward and reverse sequence of the selected ten Nanopore barcodes.

Pool	Barcode ID	Barcode Sequence (F)	Barcode Sequence (R)
1	1	CACAAAGACACCGACAACTTTCTT	AAGAAAGTTGTCGGTGTCTTTGTG
2	9	AACCAAGACTCGCTGTGCCTAGTT	AACTAGGCACAGCGAGTCTTGGTT
3	17	ACCCTCCAGGAAAGTACCTCTGAT	ATCAGAGGTACTTTCCTGGAGGGT
4	25	GTAAGTTGGGTATGCAACGCAATG	CATTGCGTTGCATACCCAACTTAC
5	33	CAGACTTGGTACGGTTGGGTAACT	AGTTACCCAACCGTACCAAGTCTG
6	41	GGAGTTCGTCCAGAGAAGTACACG	CGTGTACTTCTCTGGACGAACTCC
7	49	ACTGGTGCAGCTTTGAACATCTAG	CTAGATGTTCAAAGCTGCACCAGT
8	57	GCTAGGTCAATCTCCTTCGGAAGT	ACTTCCGAAGGAGATTGACCTAGC
9	65	TTCTCAGTCTTCCTCCAGACAAGG	CCTTGTCTGGAGGAAGACTGAGAA
10	73	AAGAAACAGGATGACAGAACCCTC	GAGGGTTCTGTCATCCTGTTTCTT

The library preparation was performed according to the kit instructions. In brief: the DNA Control Sample (DCS), which is part of the kit, was subjected to a thawing process at room temperature, followed by vertexing and subsequent placement on ice to maintain sample integrity. To dilute the DCS, 105 µl Elution Buffer (EB) was added directly to one DCS tube, gently mixed by pipetting and thawed on ice. Each pooled DNA was added to each of the ten wells of the 96 well plates. The subsequent assembly

of the reaction mix in each well involved combining 11.5 µl of 200 fmol amplicon DNA, 1 µl diluted DCS, 1.75 µl Ultra II End-prep Reaction Buffer, and 0.75 µl Ultra II End-prep Enzyme Mix. These components were thoroughly mixed by pipetting and a spin-down step in a centrifuge. The incubation of the reaction mix in a thermal cycler was conducted at 20°C for 5 minutes and 65°C for 5 minutes, marking the completion of the end-prepping phase. The subsequent transition to the native barcode ligation step involved preparing the NEB Blunt/TA Ligase Master Mix, with specific thawing steps and mixing procedures for various reagents, including barcodes. Each barcode (oligonucleotide) was individually mixed, and the barcoded samples were pooled in a 1.5 ml Eppendorf DNA LoBind tube. A 1.8X clean-up of the pooled reaction was conducted using resuspended AMPure XP Beads (AXP), followed by an incubation period on a Hula mixer for 10 minutes at room temperature.

The subsequent ethanol wash steps, including pellet drying and resuspension, were performed. The elution of the DNA from the beads involved a 10-minute incubation at 37°C, with periodic agitation to facilitate DNA elution. The resulting eluate was quantified using a Qubit fluorometer. Then, the Native Adapter (NA) was used for the adapter ligation step. The Native Adapter contains sequencing adapters, which, through a multi-nucleotide sticky end, are ligated to the barcoded DNA fragments obtained in the Native Barcode ligation step. The NEBNext Quick Ligation Reaction Module was prepared according to the manufacturer's instructions, and a stepwise ligation reaction involving a Short Fragment Buffer (SFB) was executed. A subsequent AMPure XP Beads (AXP) clean-up, involving incubation, washing, and elution steps, was performed to ensure the purification of the DNA library.

2.4.2 DNA sequencing using the MinION platform

Before initiating the experimental protocol, a crucial preliminary step involved assessing the availability of pores on the MinIon flow cell R10 Version (ONT-08-00592-13-1, FLO-MINI4). This assessment ensured the integrity and functionality of the flow cell, with a specific focus on confirming the presence of accessible and viable nanopores essential for the sequencing process. The flow cell passed the test by having 1,279 pores available. The final steps of the experimental protocol involved the priming and loading the SpotON flow cell, adhering to the compatibility specifications with R10.4.1 flow cells (FLO-MIN114). The priming mix, including Bovine Serum Albumin (BSA) for

enhanced sequencing performance, was prepared, and loaded onto the flow cell. The buffer was drawn back from the priming port opening, ensuring continuous buffer coverage over the sensor array. The loading of the prepared library onto the flow cell was performed, the flow cell was inserted into the MinION Mk1C Sequencing Device (MIN-101C). The run was processed for a total of 24 hours without interruptions. Finally, data processing is carried out by the MinKNOW software, which carries out data acquisition and analysis.

2.4.3 Bioinformatic analysis

The workflow of the Nanopore reads analysis is summarised in Table 2.14. All the scripts used are stored in https://github.com/svxvm1/zooplankton_metabarcoding_IlluminaNanopore_pipelines/tree/main.

Table 2. 14: Nanopore pipeline summary in order of application.

Steps	Tool
Base calling	Guppy
Adapter and Nucleotide trimming	Chopper
Filtration step	Seqtk
Demultiplexing	Customised script
Cutadapt	Adapters and indexes trimming
BLAST	Compare sequences against a database
BLAST output ranking	Sprank

The basecalling process was performed using Guppy which is a data processing toolkit integrated into the sequencing instrument MinKNOW. MinKNOW is the operating software that drives nanopore sequencing devices. It provides device control including selecting the run parameters, sample identification and tracking, and ensuring that the platform chemistry is performing correctly to run the samples. It also carries out several core tasks such as data acquisition, real-time analysis and feedback, local base-calling, and data streaming. The Chopper software (https://github.com/wdecoster/chopper, https://github.com/svxvm1/zooplankton_metabarcoding_IlluminaNanopore_pipelines/blob/main/run_chopper.sh) was used to filter the data from the Fastq (passed) files provided by the MinKNOW software, to filter and trim reads before demultiplexing of

the nanopore barcodes with a phred score (Q) of 8, replacing the Trimmomatic software applied in the previous script utilised in section 2.2.5.1. The following step was Seqtk (https://github.com/svxvm1/zooplankton_metabarcoding_IlluminaNanopore_pipelines/blob/main/run_seqtk.sh), utilised to convert the data from Fastq to Fasta, with additional filtration steps to remove low-quality sequences. Subsequently, a custom python script to match and retrieve the sequences that contain Illumina indices and sort them into different files

(https://github.com/ahmadazd/MatchandWrite_Indexer-MAWI/tree/main)

Briefly, forward and reverse reads were binned into their corresponding samples by fetching the index of each sample using one of two different criteria: one allowing 100% match with the indexes and the other one allowing one mismatch to the indexes. The script searched for indexes, including both P5 and P7 (where Index 2 was associated with P7, and Index 1 was associated with P5) taken from the Metagenomics retrieved by DeepSeq in Appendix S6. If the indexes were found, the corresponding sequence was extracted and named after the respective samples, with the sequences arranged sequentially.Subsequently,Cutadapt (https://github.com/svxvm1/zooplankton_metabarcoding_IlluminaNanopore_pipelines/blob/main/run_cutadapt.sh) was applied in the analysis process to trim the Illumina adapters, indexes, and primers, resulting in 'pure' sample sequences without adapters. The process was then continued using the same steps as the Illumina pipeline in section 2.2.5.1 using the script in

https://github.com/svxvm1/zooplankton_metabarcoding_IlluminaNanopore_pipelines/blob/main/run_aftercutadapt_filtering.sh. Briefly, this script arranges and filter the sequences, blast the sequences against a database (Section 2.1.8.1) using the "megablast" program with an e-value of $1e^{-15}$, which is optimised to identify alignments in highly similar sequences and returns the top hit for each query sequence in a custom tabulated format and rank the blast output by using sprank.pl software with the custom-written Perl script (from section 2.2.5.1.

Chapter 3. Protocol design for the molecular analysis of phytoplankton and zooplankton genera

3.1 Introduction

The identification of planktonic organisms has been a challenge in the aquatic ecotoxicology field. This is because, as stated before, plankton is extremely varied and small and identification through traditional methods always requires an expert. However, plankton identification has gone through a dramatic transformation with the development of metabarcoding, a high-throughput molecular technique. Metabarcoding offers the simultaneous assessment of multiple plankton taxa, enabling the detection of rare, cryptic, or unculturable species that may escape conventional morphological techniques. This technological innovation accelerates the process of plankton identification and, in doing so, has expanded our understanding of planktonic diversity and community dynamics, thus redefining their role in aquatic ecosystems (Lentendu et al., 2018; Pawlowski et al., 2014).

In the first stage of the study, the aims were to identify the most feasible, reliable, and time-effective method to sequence both phytoplankton and zooplankton genera. The approach was to develop and optimise metabarcoding or genera-specific qPCR protocols.

3.2 Results

3.2.1 Molecular Analysis of Phytoplankton Monocultures

For the phytoplankton database, a total of 2,177 sequences for the 23s gene and a total of 902 sequences for the 18s gene were downloaded from both GenBank and BOLD databases. For the 23s gene sequences were found for seven divisions: *Clorophyta* 42.49%; *Charophyta* 34.41%; *Cyanophyta* 15.11%; *Euglenophyta* 6.02%; *Cryptophyte* and *Ochrophyta* 0.96%; and *Choanozoan* 0.05%. For the 18s gene, sequences were found for only 4 divisions: *Chlorophyta* 28.46%; *Cryptophyte* 33.26%; *Charophyta* 23.50%; and *Euglenophyta* 4.88% (Figure 3.1). However, most genera selected based on the CEA list provided were found to be unavailable or unidentified on the databases.

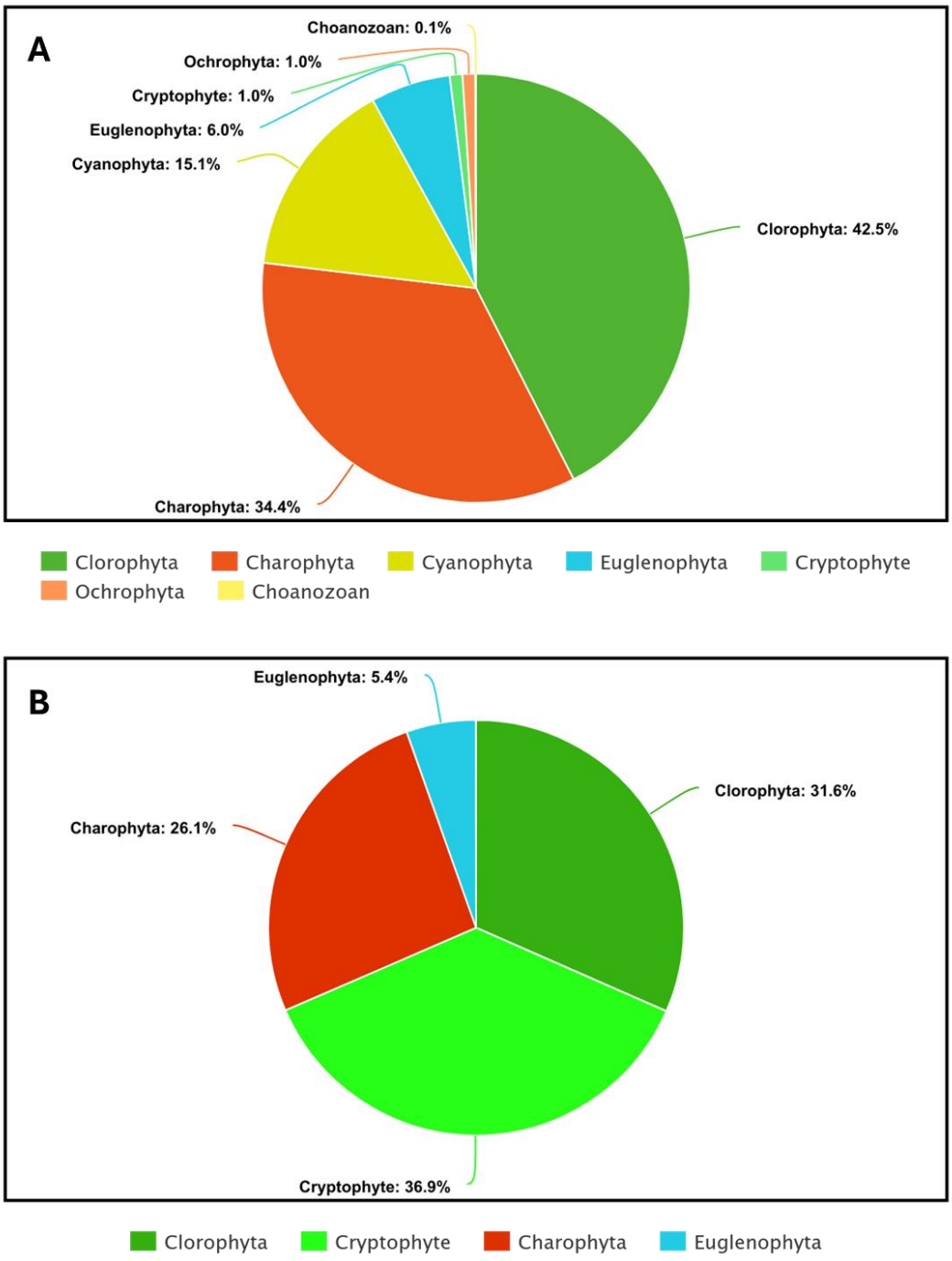


Figure 3. 1: Phytoplankton availability on databases for both targeted genes (23s – A and 18s -B).

The first step in phytoplankton metabarcoding was to specifically target the main genera present in CEA samples by qPCR (Table 2.1, Chapter 2). Monocultures of four genera were used as templates for PCR analysis.

The DNA extracted from 500 µl of a 14-day-old monoculture of four genera of phytoplankton were 2.29 ng/µl, 2.67 ng/µl, 3.23 ng/µl and 1.97 ng/µl for *Desmus desmus, Chlorella, Spirogyra,* and *Ankistrodesmus,* respectively.

Nested qPCR of phytoplankton using the metabarcoding primer A23SrVF1 and A23SrVF2 (Table 2.2, Chapter 2) showed no amplification for *Chlorella* nor for two out of three replicates of *Desmus desmus* (Table 3.1). Although the qPCR showed no amplification for *Chlorella,* the gel electrophoresis showed a single band in one of the replicates. The *Ankistrodesmus* genera showed amplification in all three replicates, however, only two replicates contained a band in the gel electrophoresis. *Desmus desmus* qPCR and gel electrophoresis showed amplification in only one replicate. *Spirogyra* was the only genera that showed amplification in both qPCR and gel electrophoresis (Figure 3.2). When the nested qPCR was repeated using an annealing temperature of 60 °C, all the genera apart from *Ankistrodesmus* amplified in the qPCR. However, bands in the gel electrophoresis were found for both *Chlorella* and Spirogyra but not for *Ankistrodesmus* and *Desmus desmus* (Table 3.2). Although the qPCR showed no amplification for *Chlorella,* the gel electrophoresis showed a single band in one of the replicates. The *Ankistrodesmus* genera showed amplification in all three replicates. However, only two replicates contained a band in the gel electrophoresis. *Desmus desmus* qPCR and gel electrophoresis showed amplification in only one replicate. *Spirogyra* was the only genera that showed amplification in both qPCR and gel electrophoresis. Results of the qPCR using the metabarcoding primer *Uni18S* on the four genera of the phytoplankton (*Desmus desmus*, *Chlorella*, *Spirogyra* and *Ankistrodesmus*) showed no gene amplification for the genera *Spirogyra* and a generally high CT value amplification was observed for *Desmus desmus*, *Chlorella* and *Ankistrodesmus* (Table 3.3).

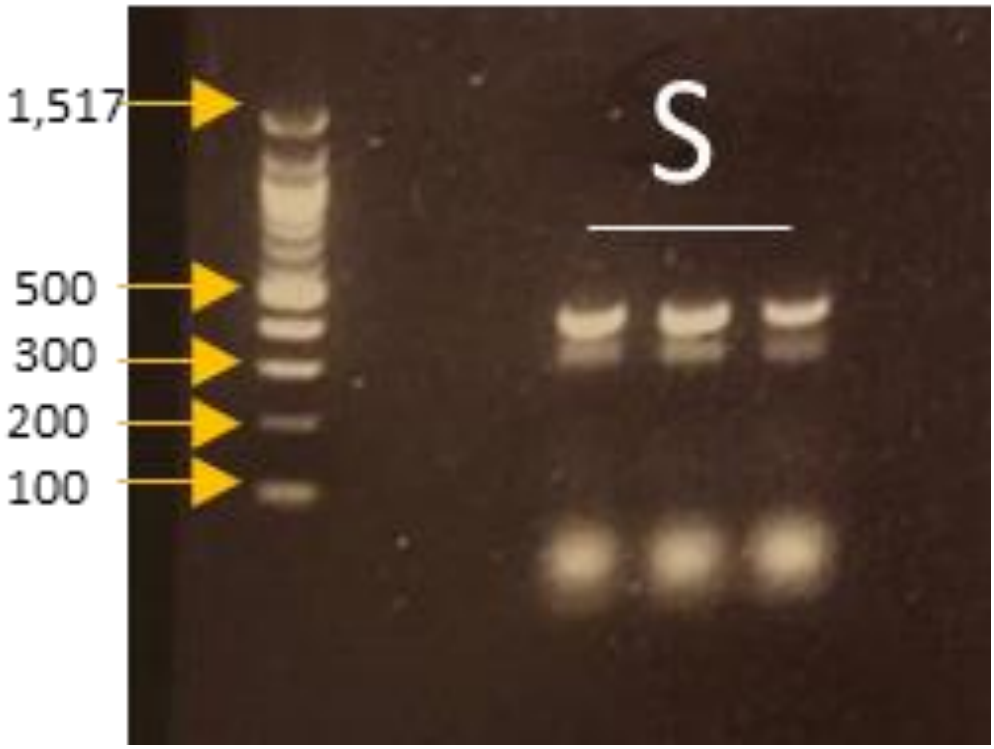


Figure 3. 2: Example of successful amplification at the right size for the three Spirogyra replicates for the 23s gene.

Nested qPCR using the A23SrVF1 and the A23SrVF2 primers (expected fragment size 400bp).

Table 3. 1: Metabarcoding primers for nested qPCR.

A23SrVF1 and 2 were tested on samples of *Chlorella* (C), *Desmus desmus* (DS), *Spirogyra* (S) and *Ankistrodesmus* (AB) (where "X" represents no amplification). All controls (no template, n=12) did not amplify.

phytoplankton nested qPCR			
Primer set	**Sample ID**	**qPCR CT value**	**Gel analyses**
A23SrVF 1 and 2	C	X	X
A23SrVR1 and 2	C	X	X
	C	X	positive
	AB	13.50	X
	AB	12.12	
	AB	12.95	positive
	S	16.91	positive

	S	16.44	positive
	S	16.22	positive
	DS	19.21	positive
	DS	X	X
	DS	X	X

Table 3. 2: Repeats of phytoplankton metabarcoding nested qPCR using the A23SrVF1 and 2 primers.

Tested on samples of *Chlorella* (C), *Desmus desmus* (DS), *Spirogyra* (S) and *Ankistrodesmus* (AB) (where "X" represents no amplification).

Repeated phytoplankton nested qPCR			
Primer set	**Sample ID**	**qPCR CT value**	**Gel analyses**
A23SrVF 1 and 2	C	21.15	positive
A23SrVR1 and 2	C	20.87	positive
	C	21.74	positive
	AB	X	X
	AB	X	X
	AB	X	X
	S	17.87	positive
	S	17.83	positive
	S	16.85	positive
	DS	27.63	X
	DS	23.12	X
	DS	21.97	X

Table 3. 3: Metabarcoding primers (18s) tested on phytoplankton selected genera.

The 18s primers were tested on *Chlorella* (C), *Desmus desmus* (DS), *Spirogyra* (S) and *Ankistrodesmus* (AB) (where "X" represents no amplification).

Phytoplankton qPCR			
Primer set	**Sample ID**	**CT value**	**CT value (-control)**
Uni18s	C	14.41	X
	C	14.46	X

	C	14.58	X
	AB	12.87	X
	AB	16.86	X
	AB	17.91	X
	S	X	X
	S	X	X
	S	X	X
	DS	15.51	X
	DS	16.56	X
	DS	17.35	X

Sequencing results for the amplicons produced with the four genera of phytoplankton (*Desmus desmus*, *Chlorella*, *Spirogyra* and *Ankistrodesmus*) using the nested metabarcoding primer-set A23SrVF1 and F2 showed that the primers successfully amplified the 23S gene (Table 3.4). The *Chlorella* sequence was found to be 97.67% similar to a *Chlorella* sequence on Genbank. Similarly, the suspected genera *Spirogyra* was found to be 99.33% similar to a *Spirogyra* sequence. The *Desmus desmus* genera was matched 98.66% to an uncultured organism which was sequenced within the 23s region. Finally, the *Ankistrodesmus* genera was found to be 98.53% similar to the Brevundimonas genera, which is not related to the phytoplankton category as it is a gram-negative bacterium.

Table 3. 4: Blast results of phytoplankton using the metabarcoding primer-set A23SrVF1 and F2.

The phytoplankton genera blasted were: *Chlorella* (C), *Desmus desmus* (DS), *Spirogyra* (S) and *Ankistrodesmus* (AB).

Phytoplankton BLAST search							
Sample ID	**Blast match**	**Max Score**	**Total Score**	**Query Cover**	**E-value**	**% identity**	**Accession**
C_A23SrV1/2	Chlorella sp. Complete genome	518	518	98%	9.00E-143	97.67%	KJ718922.1
DS_A23SrVF1/2	Uncultured organism 23S	529	529	99%	4.00E-146	98.66%	KF803897.1
S_A23SrVF1/2	Spirogyra sp.	540	540	98%	2.00E-149	99.33%	DQ629184.1
AB_A23SrVF1/2	Brevundimonas	601	601	100%	1.00E-167	98.53%	CP039435,1

Overall, the phytoplankton reference library was incomplete for the majority of genera provided by CEA and preliminary molecular analysis using two published primer sets targeting 23S and 18S genes did not amplify all four target genera. In addition, sequence

analysis of products indicated that only two of four PCR products matched the known organisms in GenBank.

In the initial stage of the study, the aim was to develop a feasible, reliable, and time-effective method to analyse both phytoplankton and zooplankton genera. The molecular analysis of the phytoplankton monocultures highlighted numerous challenges. These results underline the challenges in phytoplankton identification as sequences are still relatively scarce on databases and in publications. This scarcity has directed the study towards zooplankton analysis instead of phytoplankton analysis where sequences on databases are more available.

3.2.2 Molecular Analysis of Zooplankton

For the zooplankton, a total of 9,818 sequences for the COI gene and 1,432 sequences for the 18S gene were retrieved from the GenBank database. Sequences were aligned using Clustal-W multiple sequence alignment in MEGA version 6 (Tamura et al., 2013). While creating the database, the focus was given to the most common zooplankton genera found at Cambridge Environmental Assessments (CEA) in Boxworth, Cambridgeshire. The predominant zooplankton sequences are those of the *Rotifera* group for both genes (58.33% in COI and 78.35% in 18s), followed by the *Cladocera* family (39.32% in COI and 20.51% in 18s), and finally, the smallest portion was found for the *Copepoda* family (2.35% in COI and 1.14% in 18s) (Figure 3.3). Details of the total genera found on databases from the CEA list are shown in Table 2.3.

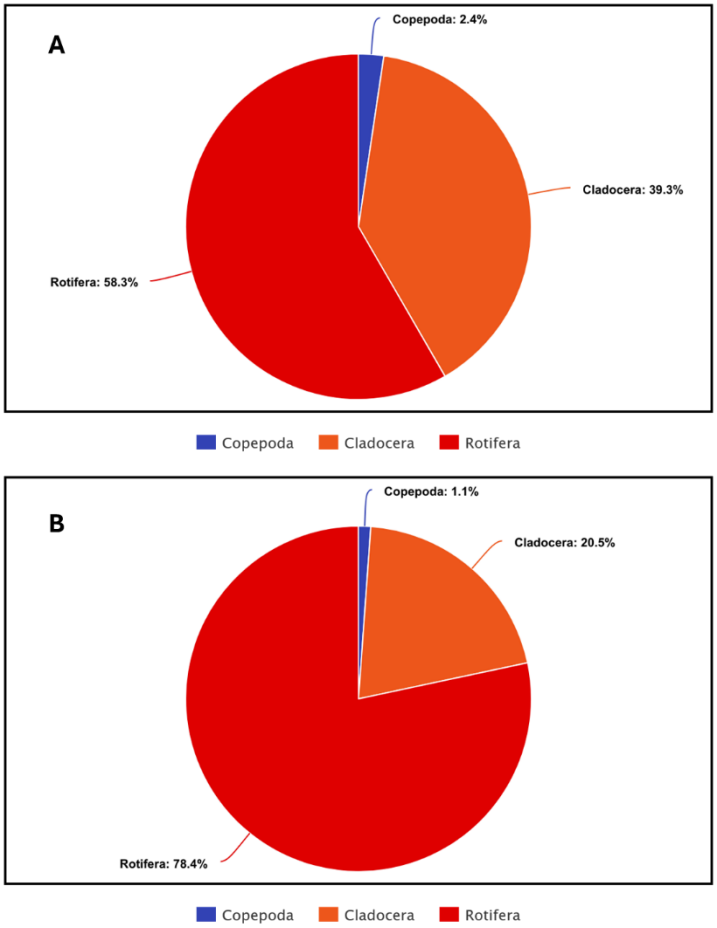


Figure 3. 3: Zooplankton availability on databases for both targeted genes (COI – A and 18s – B).

3.2.2.1 Isolation of zooplankton genera

The isolation method of zooplankton genera collected from mesocosms and isolated under a microscope using a needle were mostly unsuccessful due to contamination during the extractions, and the considerable time required to achieve effective extractions. Therefore, after multiple attempts, this method was discarded.

Monocultures of zooplankton worked well for *Daphnia* as SOPs were provided by CEA; had slight efficacy for Copepods and *Diaptomidae*; and were unsuccessful for *Simocephalus* - reproduction rate was not high, and the majority of the culture were found dead after less than a week. This could be because not many SOPs are present for these genera and therefore the environment was probably not correct for the cultures. Also, the small organism size does not easily allow the genera separation leading to poor reproduction and thus culture continuity. This method required specific laboratory

equipment and enough space for all the cultures in the laboratory. As the number of genera we could isolate was limited, this method was also discarded. However, *Daphnia pulex* was successfully cultured and was used in this capacity. As *Daphnia* is the most studied zooplankton, and cultures were provided by CEA, all the initial analyses were performed on *Daphnia* as an exemplar.

3.2.2.2 Daphnia qPCR using metabarcoding primers.

A summary of the *Daphnia* qPCR analysis using the metabarcoding primers is summarised in Figure 3.4.

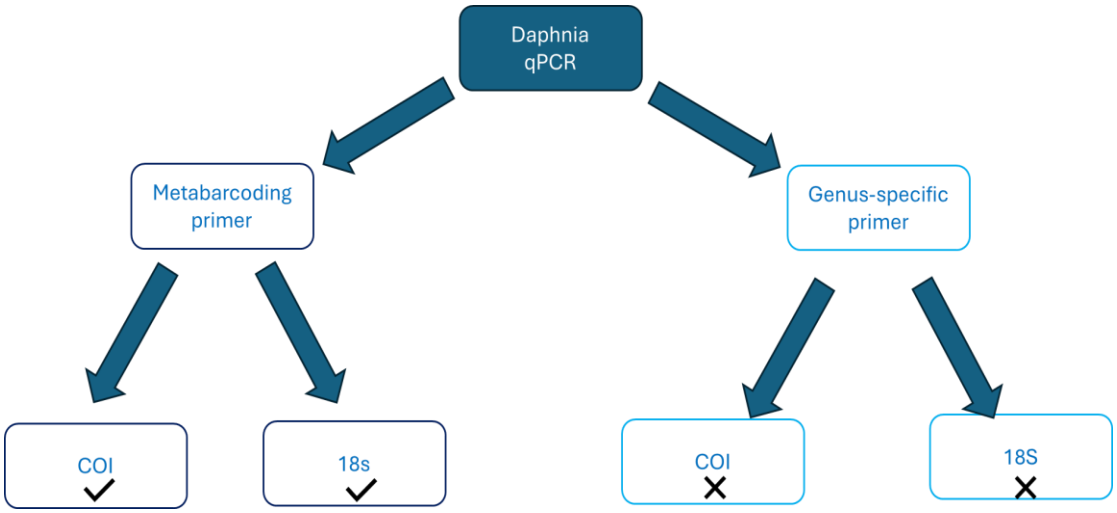


Figure 3. 4: Daphnia qPCR analyses summary.

Where the tick means that amplification worked while the "X" represents no amplification.

Daphnia samples were first tested with a general metabarcoding primer set and then tested with the designed genus-specific primers. Variable DNA concentrations were obtained from each sample of cultured and extracted *Daphnia pulex*. For example, 1.31 ng/µl was retrieved from seven adults of *Daphnia pulex*, while for 50 individuals of mixed ages, 23.9 ng/ µl was obtained. *Daphnia* DNA was tested for degradation prior qPCR. The results showed single high molecular weight bands thus no degradation, although the bands were found to be faint (Appendix S1).

The results of the gel electrophoresis from the PCR of *Daphnia* extracted DNA using the metabarcoding primer for COI (Folmer et al., 1994) showed bands at the expected size (680 bp), which confirms that COI primers can be used to detect *Daphnia* genera (Figure 3.5).

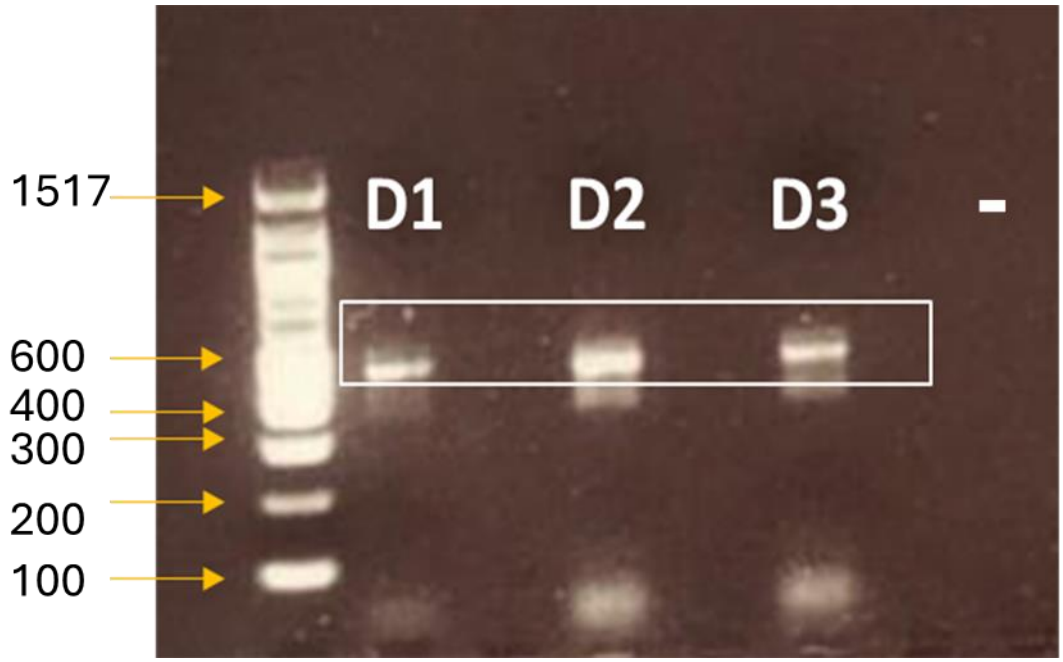


Figure 3. 5: Amplification of *Daphnia* using COI genus-specifc primers (expected fragment size 680 bp, boxed).

The COI (Folmer et al.,1994) and 18S (Zhan et al., 2013) primer sets consistently amplified *Daphnia* DNA. The Ct values for the *Daphnia* samples with the general zooplankton metabarcoding primers for COI and 18s were higher than with the positive control pike (*Esox lucius*) as expected (Table 3.5). The difference in Ct values between the two primer-sets could be due to target specificity as each primer targets a specific genome region; gene copy number; and the number of target gene's copy varies. For example, if the mICOIint primer gene region was present in a higher copy number than the Uni18S one, it may resulted in lower Ct values. Finally, it could be caused by the primer efficiency as some primers have higher binding efficiency to target DNA resulting in a better amplification and thus lower Ct values. *Daphnia* was shown to amplify correctly with the selected metabarcoding primers thus, from this point onwards, *daphnia* was considered to be a positive control for the rest of the analyses.

Table 3. 5: Metabarcoding primers tested on three replicates of *Daphnia* samples.

Daphnia qPCR			
Primer set	**Sample ID**	**CT value (Daphnia)**	**CT value (+ control)**
mICOIintF	D1	6.96	2.55
(Folmer et al., 1994)	D2	5.09	3.00
	D3	6.20	3.40
Uni18S	D1	15.35	4.69
(Zhan et al., 2013)	D2	15.86	5.71
	D3	15.40	5.23

3.2.2.3 Genus-specific PCRs

3.2.2.3.1 *Daphnia* genus-specific primers design and optimisation

Details of the designed genus-specific primers can be found in Table 2.4 in Chapter 2. When the qPCR was performed with the first set of genus-specific primers (COI Primer_F/R and 18S Primer) no amplification was obtained for the 18s primer (no Ct value) and amplification was achieved with the *COI* primer (Table 3.6).

Table 3. 6: First set of designed genus-specific primers tested on *Daphnia* (where "X" represents no amplification).

Daphnia qPCR (1st designed genus-specific primer set)		
Primer set	**Sample ID**	**CT value (Daphnia)**
Daphnia_Specific_COI	D1	25.83
	D2	21.25
	D3	23.40
Daphnia_Specific_18S	D1	X
	D2	X
	D3	X

A gradient annealing temperature was carried out to optimise the amplification of COI from *Daphnia* using the previously tested genus-specific Daphnia_Specific_COI and produced the most efficient amplification at higher temperatures (Table 3.7) This showed a direct relationship between the temperature and Ct values i.e. as the temperature increases the Ct values also increased. However, the gel electrophoresis

(Figure 3.6) showed multiple bands. In addition, no correct size band (200bp) for any of the temperature points were found.

Table 3. 7: *Daphnia*'s annealing temperature optimisation qPCR of the first set of genus-specific primer-set "COI Primer" using a gradient.

Daphnia qPCR COI Primer_F/R optimisation			
Primer set	**Temperature (°C)**	**Sample ID**	**CT value (Daphnia)**
Daphnia_Specific_COI	65.00	D1	32.47
	65.40	D2	31.87
	63.30	D3	32.51
	61.10	D4	29.81
	59.00	D5	26.86
	57.00	D6	23.94
	55.70	D7	22.50
	55.00	D8	22.01

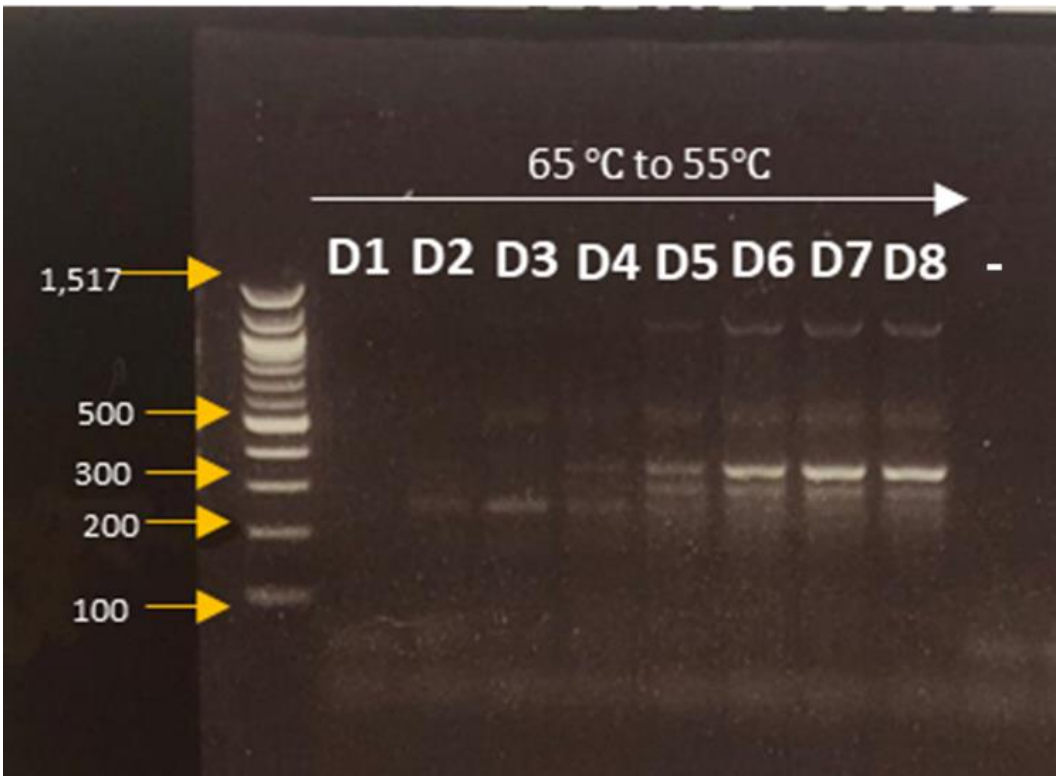


Figure 3. 6: Amplification of Daphia using Daphnia_Specific_COI genus-specifc primer.

Gel electrophoresis of the gradient qPCR (from 65°C to 55°C) on 8 *Daphnia* replicates (expected fragment size 200 bp) using the first Daphnia_Specific_COI genus-specific primer-set and 1Kb ladder.

The second COI primer-set "Daphnia_Specific_COI_2" was found to amplify COI similarly at both higher and lower temperatures of the annealing temperature gradient (Table 3.8). However, the corresponding gel (Appendix S2), showed multiple faint bands none of which were the expected size (170 bp).

Table 3. 8: Second set of genus-specific primers tested on Daphnia.

Daphnia qPCR COI Primer_2 optimisation			
Primer set	**Temperature (°C)**	**Sample ID**	**CT value (Daphnia)**
Daphnia_Specific_COI_2	65.00	D1	23.64
	65.40	D2	23.65
	63.30	D3	23.48
	61.10	D4	23.76
	59.00	D5	22.55
	57.00	D6	21.47
	55.70	D7	21.67
	55.00	D8	20.40

Repeat analysis showed that for both primers (Daphnia_Specific_COI and Daphnia_Specific_COI _2) the *COI* gene was amplified (Table 3.9). However, the gel electrophoresis in Figure 3.7 showed multiple faint bands, and again no bands of the right size (200 bp for Daphnia_Specific_COI and 170 for Daphnia_Specific_COI_2).

Table 3. 9: Comparison between the first two sets of genus-specific primers tested on *Daphnia*.

Daphnia qPCR COI primers comparison		
Primer set	**Sample ID**	**CT value (Daphnia)**
Daphnia_Specific_COI	D1	7.61
	D2	9.52
	D3	8.54
Daphnia_Specific_COI_2	D1	10.38
	D2	6.35
	D3	6.35

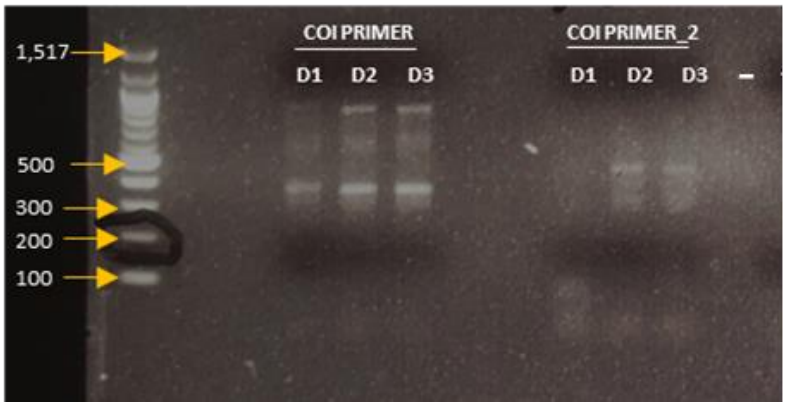


Figure 3. 7: Amplification of three *Daphnia* replicates using COI genus-specific primer gel electrophoresis of the two primer sets (Daphnia_Specific_COI and Daphnia_Specific_COI _2).

A 1Kb ladder was used and the expected fragment size 200 and 170 respectively.

Ct values for *Daphnia* using the third primer set Daphnia_Specific_COI _3 (degenerate) were generally low among the replicates due to the high starting DNA concentrations (Table 3.10). One of the three replicate PCRs for *Daphnia* did not show any amplification. Again, the gel electrophoresis (Appendix S3) showed multiple bands, none of which were the correct size (204 bp). Overall, genus-specific primers could not be designed for the successful amplification of *Daphnia*.

Table 3. 10: Third set of genus-specific primers (degenerate) tested on Daphnia (where "X" represents no amplification).

Daphnia qPCR		
Primer set	**Sample ID**	**CT value (Daphnia)**
Daphnia_Specific_COI_3	D1	X
	D2	10.26
	D3	9.74

3.2.2.4 Validation of qPCR and PCR amplification using DNA-sequencing

As the designed genus-specific primers were unsuccessful, the previously used metabarcoding mCOInf primer set (COI) (section 3.2.2.1) were tested again on cultured *Daphnia* before sequencing. All three replicates of *Daphnia* were successfully amplified using the mCOInf primer set as shown in both the qPCR (Table 3.11) and the gel electrophoresis (Figure 3.8)

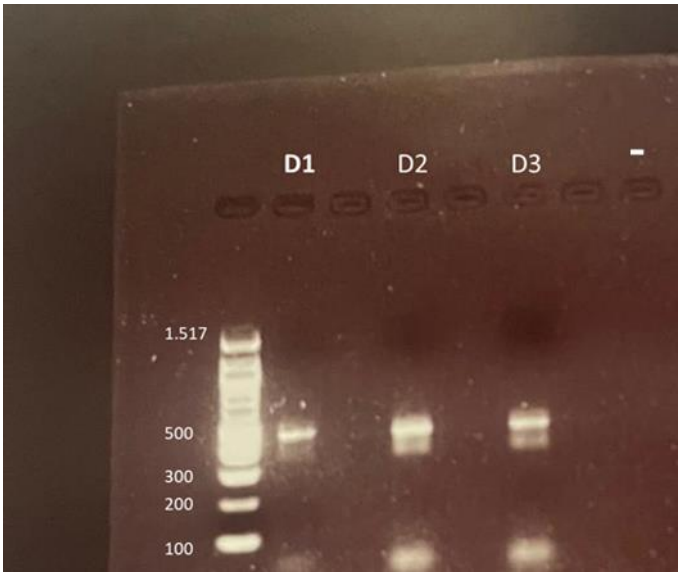


Figure 3. 8: Amplification of Daphnia using mCOInf metabarcoding primer on three *Daphnia* replicates (expected fragment size 480 bp) using 1Kb ladder.

Table 3. 11: Metabarcoding mCOInf primer set tested on Daphnia.

Daphnia qPCR using metabarcoding primer		
Primer set	**Sample ID**	**CT value (Daphnia)**
mCOInf	D1	7.02
	D2	7.59
	D3	7.04

The sequencing of the amplicon from the amplification of *Daphnia* DNA using the metabarcoding mCOInf primer set (COI) are summarised in Table 3.12 where the uploaded sequence was found to be 96.68% similar to *Daphnia magna* isolate on

GenBank. As shown in Figure 3.9, the BLAST results indicated a high similarity between the *Daphnia* sequence and the database reference sequence. As the sequence quality dropped off at the end of the chromatogram, the last 20 bp were trimmed and thus identity dropped.

Table 3. 12: Blast results of *Daphnia* sequence.

Sequenced Sample ID	GenBank ID	Max Score	Total Score	Query Cover	E-Value	% Identification	Accession
D1_mlCOIintF	*Daphnia magna* isolate CZ-N1	547	547	97%	1.00E-151	96.68%	MH683621.1

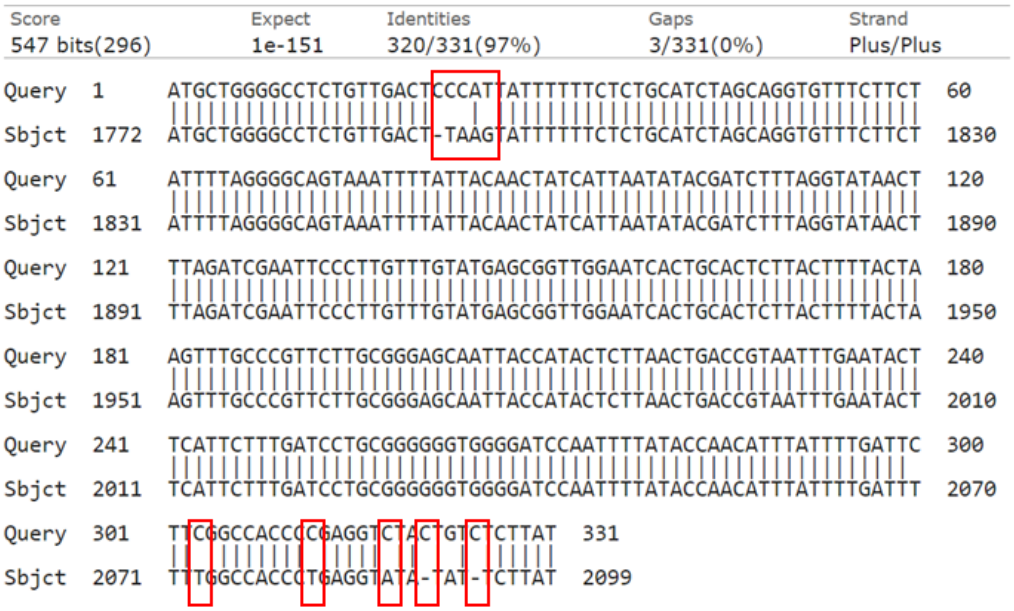


Figure 3. 9: Daphnia Blast results on BLAS NCBI software.

Where Query 1 is the sequences obtained from Sanger and Subject is the BLAST hit for the uploaded *Daphnia* sequence (97% similarity). Red boxes indicate the areas where the two sequences do not match. As the sequence quality dropped off at the end of the chromatogram, the last 20 bp were trimmed and thus identity dropped.

Due to the difficulties in culturing and isolating the zooplankton genera in addition to all the optimisation for *Daphnia*, the qPCR approach was discarded and the approach of metabarcoding the sequence samples in bulk was tested.

3.2.3 Design and Optimisation of a Metabarcoding Analysis Protocol

3.2.3.1 Phase 1, PCR with Q5 mastermix and mCOInf, HEX and 18S primer sets

A summary of the initial Mixed zooplankton metabarcoding analyses is summarised in Figure 3.10.

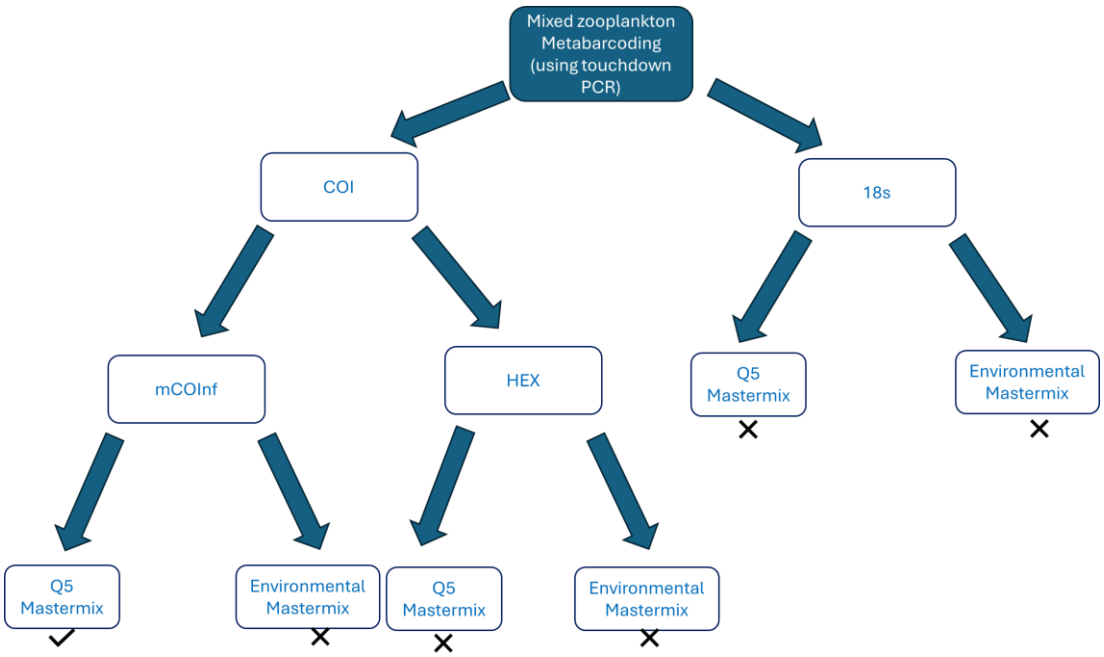


Figure 3. 10: Mixed zooplankton samples analyses using a touchdown PCR summary.

Where the tick means that amplification worked while the "X" represents no amplification.

Initial experiments on mesocosm mixed samples (A-B) from Mesocosm sampling season 2020 day 1 involved COI primer sets (mCOInf and HEX in table 3.13) using a touchdown PCR approach (Don et al 1991) with both environmental master mix and Q5 master mix. The Q5 master mix produced successful amplification (Figure 3.11) for the samples but not for the positive control (*Daphnia*), while the environmental master mix did not show amplification (not shown). The positive control possibly resulted in no amplification due to degradation caused by multiple freeze/thaw cycles. A new

approach of aliquoting the positive control sample was taken from this stage onwards. Amplification products of the correct size were also found for the HEX primer-set when using both mastermixes, however, multiple bands were found at every experiment repeat (not shown). When the same conditions were applied to the 18S primers, no amplification products at the correct size were found (not shown). Further testing of the 18s primers using a gradient PCR (65-55 °C) showed no amplification. Therefore, the two primers chosen for metabarcoding were the mCOInf and Hex primer sets (both target COI regions).

Table 3. 13: Metabarcoding primers used for mixed zooplankton samples.

Primer Name	Oligonucleotides (5'-3')	%GC	Tm	Size	Reference
mICOIintF	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGWACWGGWTGAACWGTWTAYCCYCC	50.8	>75	480 bp	Folmer *et al.*, (1994)
jgHCO2198	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTAIACYTCIGGRTGICCRAARAAYCA	47.5	>75	480 bp	Geller *et al.* (2013)
HexCOIF4	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGHCCHGAYATRGCHTTYCC	51.9	>75	460 bp	Marquina *et al.* (2018)
HexCOIR4	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTATDGTRATDGCHCCNGC	51.9	>75	460 bp	Marquina *et al.* (2018)

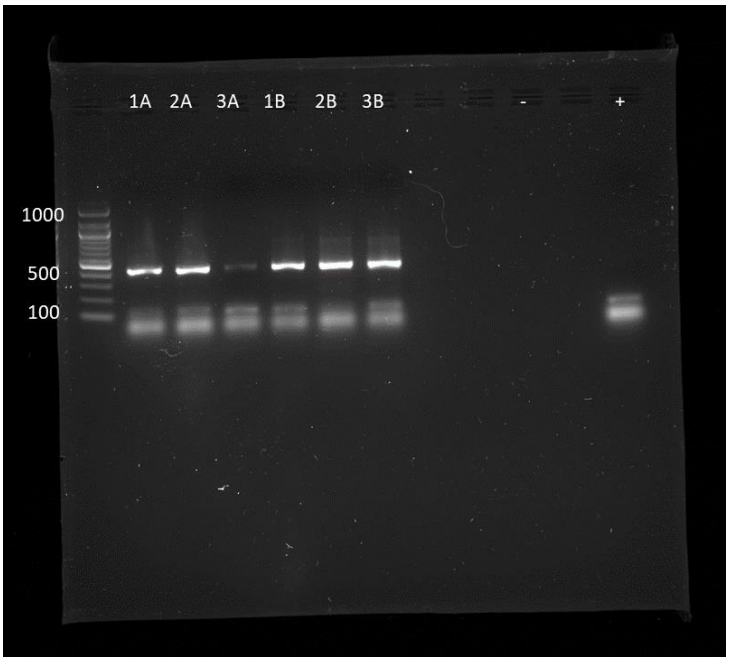


Figure 3. 11: Touchdown PCR amplification of mesocosm sample (day 1) using the mCOInf primers sets using Q5 Mastermix.

2 DNA extracts (A and B, analysed in triplicate) of mixed zooplankton samples collected from the mesocosm at CEA amplified producing a product of the correct size, 480bp, 10 µL PCR product loaded per well; 3 µL 100 bp Ladder loaded. PCR negative controls (-) were negative for amplification. Positive control DNA (+) did not successfully amplify (*Daphnia Magna*) while the negative control resulted negative as expected.

Results of the mesocosm samples from sampling days 1 and 2 using the mCOInf metabarcoding primer set resulted in a correct size amplicon (example shown in Figure 3.11). On the other hand, results from sampling day 3 onwards resulted in no amplification.

Results of the mesocosm samples from sampling day 1 using HEX metabarcoding primer set showed extremely faint amplicons and only in sample A (results not shown). From sampling day 2 onwards, no amplification at the correct size was found. These results showed an inconsistent protocol for preparing zooplankton samples for metabarcoding.

3.2.3.2 Phase 2, testing for matrix inhibitors and optimisation of sample storage

To troubleshoot this unexpected outcome, various optimisation steps were carried out. Fresh primers were ordered to avoid no amplification due to degradation; several template DNA volumes were tested (5, 3, 2, 1 and 0.5 ul), 'clear' and 'less clear' (due to sediments and algae) mixed zooplankton samples were compared, and DNA extractions were assessed i.e. both freshly extracted and the extractions stored at -20°C for a week. Additionally, a melting curve analysis with qPCR was conducted where the primers did not yield the expected results.

Mesocosm and environmental samples for use as positive controls were provided by ADAS from a previous study and were tested with both Hex and COI primers. Amplification products at the correct size were found for all of these positive controls (Figure 3.12). The inhibition test performed using DNA from cultured *Daphnia* spiked with mesocosms samples showed amplification and therefore no inhibition.

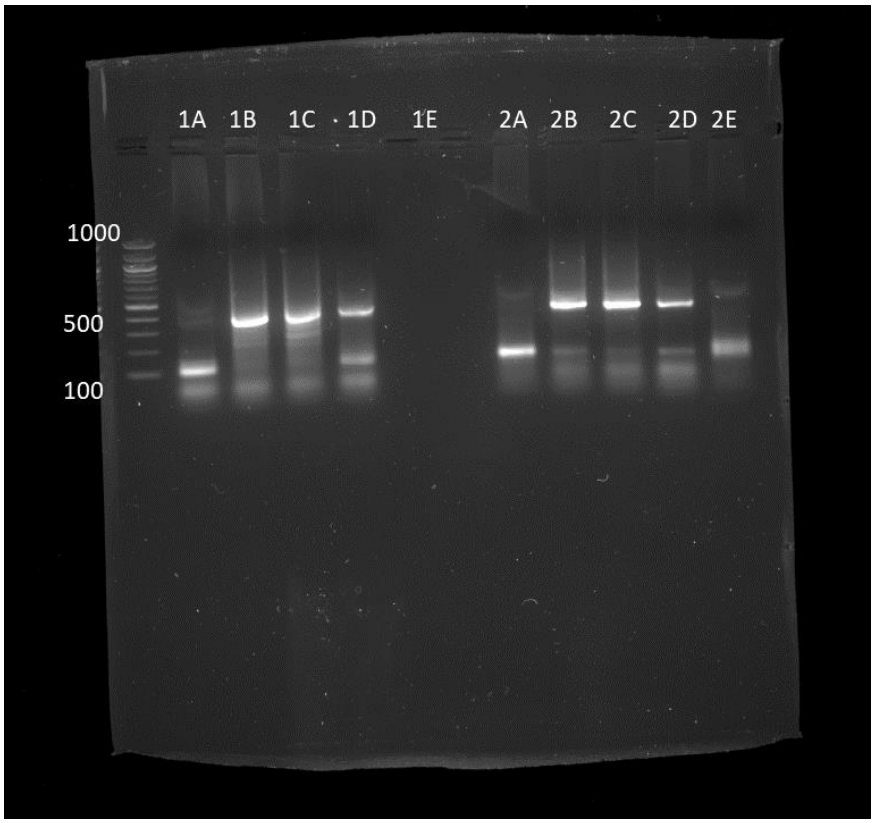


Figure 3. 12: Touchdown PCR amplification of positive controls provided by ADAS.

Where: 1A: single species snail DNA, 1B: environmental sample, 1C: environmental samples, 1D: mesocosm sample from day 1) analysed using the HEX primer and the environmental mastermix. Sample 2 (A-D) were analysed using mCOInf primers sets

and environmental mastermix. Letter E represent the negative control. In both cases, 3 DNA extracts amplified producing a product of the correct size, 480bp (for mCOInf) and (for HEX), 10 µL PCR product loaded per well; 3 µL 100 bp Ladder loaded. PCR negative controls (-) were negative for amplification when mCOInf primers sets were used. Primer size summarised in Table 2. 6.

The optimisation of the results starting from the elution during the DNA-extraction, showed a better amplification when 100 µl of buffer was used (Table 2.4, Chapter 2). When different sample storage methods were tested (Table 2.4, Chapter 2), a better amplification was found from fresh samples in 70% ethanol, stored at room temperature than those previously extracted and stored at -20°C. However, sample amplification with both primer sets was still found to be inconsistent.

3.2.3.3 Phase 3, optimising 18S primer sets and PCR conditions

As the Hex and COI primers were giving inconsistent amplification results across a range of mesocosm samples from different sampling dates (day 1- 4), two different 18S primer sets from two published papers were tested to check for reproducibility (Djurhuus et al. 2018; Bucklin et al. 2019). These primers were first tested on *Daphnia Magna* cultures. Sequencing results for both primer sets (Appendix S4) and both PCR protocols showed a correct match when Blasted for samples 1A,1C and 1E (Table 3.14).

Table 3. 14: Table showing sequencing results for Daphnia Magna from Figure 3.17.

SAMPLE	ID	PRIMER	MASTER MIX	PCR Protocol	REFERENCE	% IDENTITY	A. Number
Daphnia Magna	1A	18S	Q5	Touchdown	Bucklin et al., 2019	100	AM490278.1
Daphnia Magna	1B	18S	Q5	Touchdown	Djurhuus et al. 2018	X	
Daphnia Magna	1C	18S	Environmental	Touchdown	Bucklin et al., 2019	100	EU370423.1
Daphnia Magna	1D	18S	Environmental	Touchdown	Djurhuus et al. 2018	X	
Daphnia Magna	1E	18S	Q5	Bucklin et al., 2019	Bucklin et al., 2019	100	AM490278.1
Daphnia Magna	1F	18S	Q5	Bucklin et al., 2019	Djurhuus et al. 2018	X	
Daphnia Magna	1G	18S	Environmental	Bucklin et al., 2019	Bucklin et al., 2019	X	

Daphnia Magna	1H	18S	Environmental	Bucklin et al., 2019	Djurhuus et al. 2018	X	

When analysis was repeated on mesocosm samples using *Daphnia* Magna as a positive control, results show correct size amplicons with either both primer-sets having the Illumina tag or with a second round PCR to attach the Illumina primer tag (Figure 3.13). Sequenced samples are shown in Table 3.15.

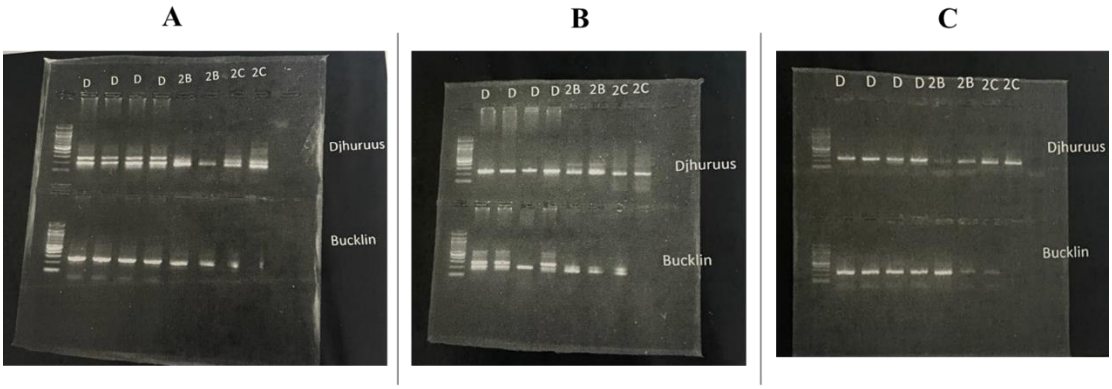


Figure 3. 13: PCR amplification of cultured Daphnia Magna provided by CEA used as positive control (D) and mesocosm samples from day 2 (2B and 2C).

All gel samples were analysed using two different primer sets: in the bottom row, Bucklin et al., (2019) and in the top row, Djurhuus et al., (2018). In figure "A" samples were analysed with the primer sets having the Illumina tag on, in figure "B", samples were analysed with the primer sets having no tag on and finally in figure "C", samples were analysed using a second round PCR to add the Illumina tag. 10 µL PCR product loaded per well; 3 µL 100 bp Ladder loaded. PCR negative controls (-) were negative for amplification. Samples successfully amplified at the right size (229 bp Djurhuus and 275 Bucklin with tag, 200 bp Djurhuus and 207 Bucklin with no tag).

Table 3. 15: Table showing sequencing results for Daphnia Magna from Figure 3.18.

SAMPLE	ID	PRIMER	TAG	2nd ROUND	REFERENCE	% IDENTITY	A.Number
Daphnia Magna	1	18S	YES	NO	Djurhuus et al. 2018	97.71	AM490278.1
Daphnia Magna	1	18S	YES	NO	Djurhuus et al. 2018	100	AM490278.1
Daphnia Magna	2	18S	YES	NO	Bucklin et al., 2019	95.89	AM490278.1

Daphnia Magna	2	18S	YES	NO	Bucklin et al., 2019	X	X
Daphnia Magna	3	18S	NO	NO	Djurhuus et al. 2018.	X	X
Daphnia Magna	3	18S	NO	NO	Djurhuus et al. 2018	100	EU370423.1
Daphnia Magna	4	18S	NO	NO	Bucklin et al., 2019	100	AM490279.1
Daphnia Magna	4	18S	NO	NO	Bucklin et al., 2019	X	X
Daphnia Magna	5	18S	YES	YES	Djurhuus et al. 2018	95.92	AM490278.1
Daphnia Magna	5	18S	YES	YES	Djurhuus et al. 2018	94.64	AM490278.1
Daphnia Magna	6	18S	YES	YES	Bucklin et al., 2019	99.45	AM490278.1
Daphnia Magna	6	18S	YES	YES	Bucklin et al., 2019	97.74	AM490278.1

When it was tested if better amplification was achieved using neat freshly extracted DNA or DNA diluted to 5 ng/ µl according to Bucklin et al., (2019), results showed a slightly better amplification in the diluted samples. A comparison between old extractions stored at -20°C and freshly made extractions resulted in a better-quality amplification when freshly extracted samples were used and washing steps were included as shown in Figure 3.14.

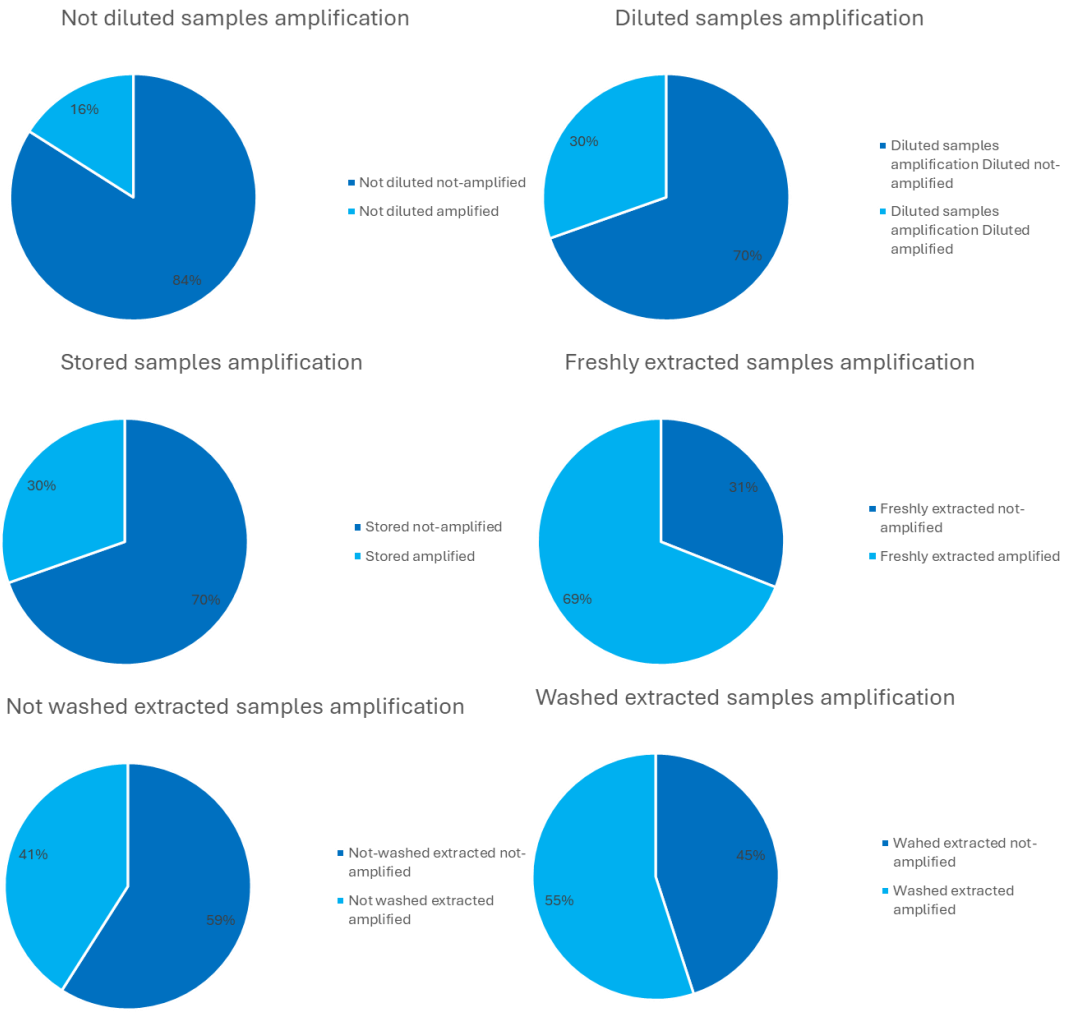


Figure 3. 14: Summary of successful and unsuccessful amplification using different techniques to optimise DNA-extractions of zooplankton mixed samples.

Results show a better amplification when samples are diluted (30.4%) compared to non-diluted samples (16%); a better amplification when samples are freshly extracted (69%) compared to frozen stored samples (30.4%); finally washed samples showed a slightly better amplification (55%) than unwashed samples (41%). The methodology was sequentially improved starting to test diluted and not diluted samples. As the diluted worked better, only diluted samples were used for the next optimisation step. This strategy was utilised for each optimisation step.

A final repeat was carried out to select the best primer set i.e. either Djurhuus et al. (2018) or Bucklin et al. (2019) (Appendix S5). Results of the step-by-step repeat of Bucklin et al., (2018) including diluted samples plus the wash step showed better amplification than that of Djurhuus et al (2018) and the previous protocol used.

Finally, the positive control (*Tigriopus Californicus*) was selected for use in metabarcoding amplifications and when tested the results showed a correctly sized amplicon. This control was used as it is a marine zooplankton species (Copepod) that we would not expect to see within our mesocosm samples.

After all the optimisation for 18S amplification, results for the metabarcoding preparation showed correctly sized amplicons in the mesocosm samples and positive control and no contamination (Figure 3.15).

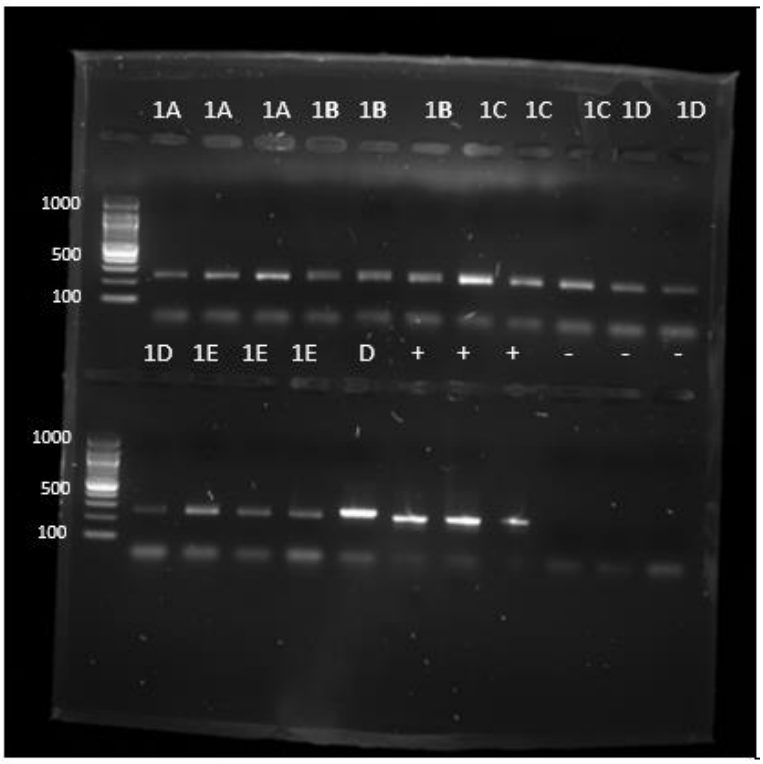


Figure 3. 15: PCR amplification of sampling day 1 with 18S primer set (Bucklin et al. 2019) using Q5 mastermix.

Successful amplification of all DNA extracts from 5 mesocosms (A-E plus *Daphnia* (D); amplicon size 275bp). 10 µL PCR product loaded per well in triplicates; 3 µL 100 bp Ladder loaded. PCR negative controls (-) were negative for amplification. Positive control DNA (*Tigriopus Californicus*) also successfully amplified.

Metabarcoding of zooplankton 18S using the primers described by Bucklin et al. (2019) was the final chosen method for molecular analysis as the 18S could be successfully amplified from a range for mesocosm samples, and a reference database of zooplankton genera could be compiled from recorded 18S sequences.

3.3 Discussion

Metabarcoding and the culturing of plankton genera are both essential tools in the field of aquatic ecology, offering distinct advantages and limitations. Metabarcoding has revolutionized plankton analysis by providing a high-throughput, DNA-based approach to taxonomic identification and community characterization. This method allows for the simultaneous analysis of multiple taxa from complex environmental samples, making it ideal for detecting rare or unculturable species that may be misidentified in traditional microscopy-based techniques (Lentendu et al., 2018; Pawlowski et al., 2014). While metabarcoding has transformed the field, it has its limitations, particularly concerning reference databases' quality and comprehensiveness. The reliability of taxonomic assignments is closely tied to the availability and accuracy of reference sequences (Amaral-Zettler et al., 2009; Egge et al., 2013). This challenge is heightened in the context of phytoplankton, which consists of a diverse and continuously evolving group of microorganisms (de Vargas et al., 2015). Metabarcoding may also struggle to differentiate between closely related species or strains due to the short DNA fragments typically sequenced (Zhan et al., 2013; Metfies et al., 2016). Such challenges are particularly relevant in phytoplankton research, where subtle taxonomic distinctions can have profound ecological implications (Lovejoy et al., 2002). Phytoplankton databases are still mostly incomplete, therefore it is still difficult to sequence and match genera or species in the blasting process. This has been shown when the databases for both genes, especially the 18s were created. There is a scarcity of phytoplankton sequences available for the selected CEA genera. This does not allow us to compare morphological and molecular analysis in phytoplankton genera. Because of this, the idea of the analysis of phytoplankton genera through metabarcoding was discarded. On the other hand, databases with zooplankton sequences are more complete and regularly updated making zooplankton identification through molecular techniques feasible. Two widely utilized genetic markers, the 18S ribosomal RNA (rRNA) gene and the Cytochrome c Oxidase subunit I (COI) gene, are commonly targeted for plankton identification and characterization. The 18S gene, due to its extensive taxonomic applicability, serves as a linchpin in cataloguing the diversity of planktonic microorganisms, spanning protists, algae, and metazoans (Amaral-Zettler et al., 2009).

Researchers can readily access sequences of the 18S gene in prominent databases like GenBank, thereby providing an indispensable resource for characterizing plankton communities at various taxonomic levels. Complementing this, the COI gene, favoured for zooplankton and metazoan identification, is accessible in databases including BOLD and GenBank. Its significance lies in its ability to provide species-level resolution and unveil cryptic diversity within planktonic communities (Leray et al., 2013). This extensive availability of both the 18S and COI genes in online databases facilitates the molecular analysis of zooplankton even though the 18S gene is still less available than the COI gene. All the selected genera from CEA mesocosms were retrieved apart from the juvenile form of Copepods for both genera. However, zooplankton sequences for the 18s gene were mostly not available from BOLD database reducing the sequences availability for this gene.

The culturing of zooplankton genera for molecular analysis offers both advantages and disadvantages that can significantly influence data accuracy and comprehensiveness. Culturing allows researchers to maintain controlled experimental conditions, ensuring that environmental factors are consistent and predictable (Lampert and Sommer, 2007). Moreover, it provides a relatively abundant source of target organisms, guaranteeing enough specimens for molecular analysis (Culver et al., 2000). This approach enables species-specific experiments, allowing for in-depth studies of particular zooplankton genera (Orsini et al., 2012). However, culturing zooplankton is time-consuming. It can be labour and resource-intensive, requiring specialised facilities, equipment, and maintenance (Rahman et a., 2023). Furthermore, the environmental conditions in laboratory cultures may lead to ecological and genetic drift, potentially altering the traits and genetic makeup of the cultured zooplankton (Lundsgaard-Hansen et al., 2012). These laboratory conditions may not fully represent the diversity and ecological interactions found in natural ecosystems, potentially leading to biased results (De Meester et al., 2002). Here, the culturing results of the four monocultures has shown the difficulty in keeping the zooplankton alive when appropriate SOPs are not available. In fact, the only successful culture was *Daphnia*. Zooplankton can seem easy to culture thanks to its fast reproduction cycle. However, zooplankton are also extremely sensitive to environmental changes and like all other animals, it is affected by parasites if not in the right conditions. By maintaining a clean container and water quality in zooplankton

cultivation, undesirable organisms can be eliminated more effectively, hence avoiding any issues that may contribute to low zooplankton output (Rahman et al., 2023).

In addition to these considerations, the technique of isolating zooplankton under a microscope using a needle was possible. This precise method minimizes contamination, reduces disruption to the environment, and enables the selective isolation of specific zooplankton specimens. However, it is time-consuming, has the potential to damage specimens, offers limited throughput, and requires a high level of skill. For unclear or turbid samples, isolating specific genera was extremely hard as other genera or contaminants were easily sucked into the syringe-needle causing impurities.

This study set out to investigate possible metabarcoding techniques to both identify and quantify relative abundance of indicator-genera in in-field ecotoxicity studies to replace or compliment traditional microscopy techniques.

The main aim of this chapter was to identify the most feasible, reliable, and time-effective method to sequence both phytoplankton and zooplankton genera. The approach was to develop and optimise metabarcoding or genera-specific qPCR protocols. The phytoplankton metabarcoding method was discarded as there is still low sequence availability on databases for the selected genes. The main focus of this study shifted towards zooplankton genera found at CEA only. Genus-specific analysis via qPCR assay were also discarded as results were inconsistent. Metabarcoding was selected as the best approach to analyse zooplankton genera as, once a methodology was developed and improved, it became the most consistent and reliable method to analyse zooplankton mixed samples. Its use will be further explored in the next few chapters to understand its application in mesocosm studies and pesticide trials.

Chapter 4. Assessment of zooplankton metabarcoding as a monitoring approach compared to taxonomic identification by microscopy

4.1 Introduction

Zooplankton play a fundamental role in aquatic ecosystem dynamics; therefore, understanding their composition and structure is essential in ecotoxicity studies. Zooplankton are traditionally identified through morphological analyses. This method is highly time-consuming, requires taxonomic expertise and, depending on the sample quality, can only sometimes identify the full range of genera present (Creer et al., 2010). However, with the use of molecular techniques, such as metabarcoding, the identification process of zooplankton samples has been proven to be facilitated through DNA analysis (Pawlowski et al., 2018). Traditional zooplankton identification methods involve morphological examination using microscopy. This method is prone to errors, especially in the case of rare or not-intact genera, and if species and not genera identification is required for ecotoxicological analyses, it is only sometimes effective (Pawlowski et al., 2018). Metabarcoding has revolutionised the field of biodiversity assessment thanks to advances in DNA-sequencing technology. The estimation of biodiversity with DNA metabarcoding using high-throughput sequencing (HTS) is becoming an important tool for surveying biodiversity thanks to the broad taxonomic coverage and the possibility of increased sample processing speed allowing to increase the sampling effort (frequency and spatial coverage) with sustainable costs (Schroeder et al., 2020). Mesocosm studies, designed to simulate a natural ecosystem with controlled environmental conditions, provide one of the best environments for investigating ecological processes in ecotoxicology (Liess et al., 2016). Studies designed in mesocosms allow variable manipulation, experiment replication, and observation and analysis of organisms' interactions in a controlled setting (Maltby and Hills, 2008).

This chapter aims to analyse the comparative advantages and limitations of both metabarcoding and molecular techniques in zooplankton identification and to determine if metabarcoding could replace or enhance traditional identification methods.

4.2 Results

4.2.1 Quantitative Analysis of Water Chemistry in Mesocosms

All sampling dates will be shown as weeks (Week 22=25th May 2020, Week 24=8th June 2020, Week 26=22nd of June 2020, Week 28=6th July 2020, Week 30=20th July 2020, Week 33=10th August 2020, Week 37=1st September 2020) for better clarity of the sampling points. Results of the six environmental variables (temperature, pH, dissolved oxygen, conductivity, chlorophyll, and turbidity) are summarised in Table 4.1, N.B. turbidity is measured in nephelometric turbidity units (NTU). The results were plotted to visually represent the variation between the collection days (Figure 4.1). The results from the one-way ANOVA tests are summarised in Table 4.2. A significant seasonal difference was found in oxygen levels, pH, conductivity, and turbidity. Tukey post-hoc test results are summarised in Table S7 (Appendix).

Table 4. 1: Environmental conditions (Temperature, pH, dissolved oxygen, conductivity, chlorophyll, and turbidity) for each mesocosm at each collection day.

Where Week 22=25th May, Week 24=8th June, Week 26=22nd of June, Week 28=6th July, Week 30=20th July, Week 33=10th August, Week 37=1st September.

Environmental Conditions						
Date	**Temp (°C)**	**Dissolved Oxygen (mg/L)**	**pH**	**Conductivity (uS)**	**Turbidity (NTU)**	**Chlorophyll (ug/L)**
162						
Week 22	14.72	9.4	8.03	753	2.8	10.4
Week 24	11.51	10.69	8.09	724	0.3	3.1
Week 26	12.19	11.88	8.09	616	1.3	3.5
Week 28	19.06	10.84	8.61	0.589	3.3	4.3
Week 30	16.93	10.45	9.18	561	2.4	4.4
Week 33	14.51	11.98	9.61	599	1.2	6.6
Week 37	17.07	10.13	9.62	488	0.8	2.4
164						
Week 22	11.51	10.69	8.09	724	0.3	3.1
Week 24	12.19	11.88	8.09	616	1.3	3.5
Week 26	19.06	10.84	8.61	0.589	3.3	4.3
Week 28	16.93	10.45	9.18	561	2.4	4.4
Week 30	17.16	10.37	9.29	500	4.5	4.2
Week 33	14.55	11.69	9.79	556	4.9	6.8
Week 37	17.34	10.04	9.58	439	4.9	5.8
168						
Week 22	15.05	10.68	8.12	579	0.9	3.5
Week 24	11.64	12.92	8.31	534	0	2.4
Week 26	12.19	12.9	8.6	563	0.1	3
Week 28	19.23	11.57	8.96	0.55	3.2	4.2
Week 30	16.65	10.28	9.13	532	3.3	3.5

Week 33	14.58	11.49	9.68	441	5.7	7.9
Week 37	17.27	9.84	9.27	466	1.1	2.7
173						
Week 22	14.81	11.29	8.66	606	3.8	12.7
Week 24	11.47	12.99	8.06	607	1.2	2.1
Week 26	12.43	12.07	8.2	583	0.8	3.4
Week 28	19.56	12.81	8.97	0.509	1.5	5.7
Week 30	17.01	11.89	9.77	472	3.9	4.7
Week 33	15.38	13.39	9.96	609	8.7	12.8
Week 37	17.3	10.25	9.49	395	0.7	4
183						
Week 22	14.75	8.35	7.9	626	0.6	3.4
Week 24	11.47	12.99	8.06	607	1.2	2.1
Week 26	11.72	12.56	7.96	602	0.2	1.7
Week 28	19.63	12.45	8.92	0.649	2.8	4.3
Week 30	16.96	11.36	9.25	634	4.3	3.5
Week 33	15.12	11.9	9.82	598	3.1	5.7
Week 37	17.28	10.46	9.45	539	0.4	3.6

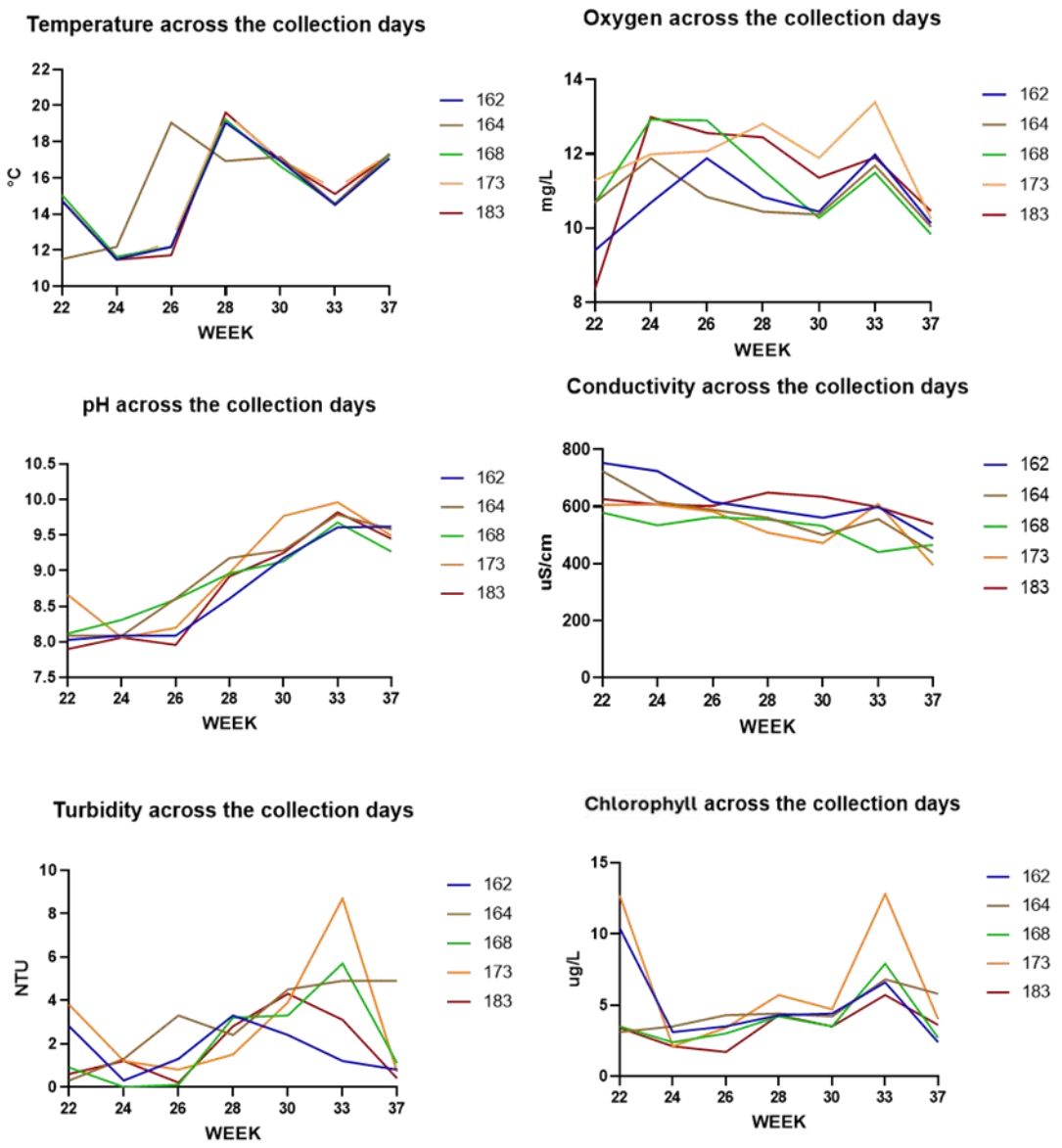


Figure 4. 1: Visual representation of the environmental variables within five mesocosms (162, 164, 168, 173 and 183) over time (from week 22 to week 37): temperature, pH, dissolved oxygen, specific conductance, turbidity, and chlorophyll.

Table 4. 2: Summary of the one-way ANOVA test on the six environmental variables, where "*" indicates a significant difference.

Environmental variable	P value
Temperature ℃	0.0632
Oxygen	0.0072 *
pH	<0.01 *
Conductivity	<0.01 *
Turbidity	<0.01 *
Chlorophyll	0.0632

4.2.2 Taxonomic Characterisation of Zooplankton Through Morphological Identification

The aim here was to identify the most prevalent genera of zooplankton by the 'gold standard' microscopy method. This morphological identification across all 35 mesocosm samples detected 20 genera (Figure 4.2) belonging to three zooplankton groups (Figure 4.3)

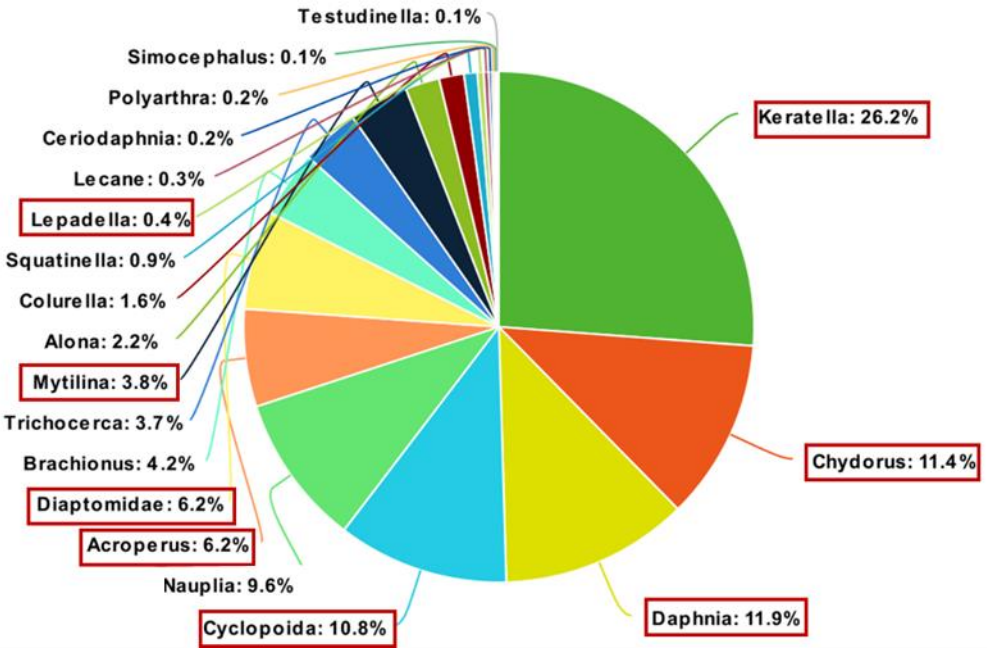


Figure 4.2: Morphology genera presence and abundance detected in morphological analyses of the 35 mesocosm samples.

The percentage was obtained by taking the average for each genus of all collection dates. Highlighted in red are the nine most abundant selected genera.

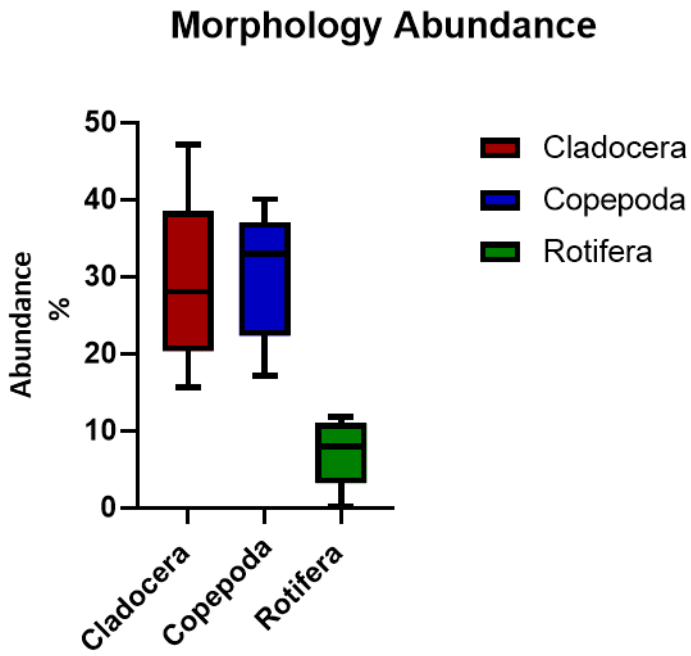


Figure 4. 3: Zooplankton families' abundance found in the 35 mesocosm samples analysed morphologically.

Out of the 20 genera identified, eight were selected for further analysis (Figure 4.4), the other genera were present in less than two mesocosms (replicate) per collection day and were not represented well across the season. By comparing data from twelve HCD mesocosm studies conducted at Cambridge Environmental Assessments' mesocosm facility between 2011 and 2017 with available data from the same site and mesocosm type (where zooplankton morphological data was determined) as presented in Brooks et al. (2019) (Figure 4.5) and juxtaposing these findings with the morphological results of the current study, a correlation in the abundance of certain genera becomes evident. In fact, five of the nine selected genera (*Chydorus, Daphnia, Mytillina, Nauplia* and *Cladopelma*) had the same pattern of abundance as the previous study. The *Cyclopoidae, Diaptomidae* and *Keratella* genera have a slightly different pattern from the previous study and, the *Acroperus* genera, one of the most abundant genera found in this study, was not found previously in sufficient abundance to have been included in the HCD paper.

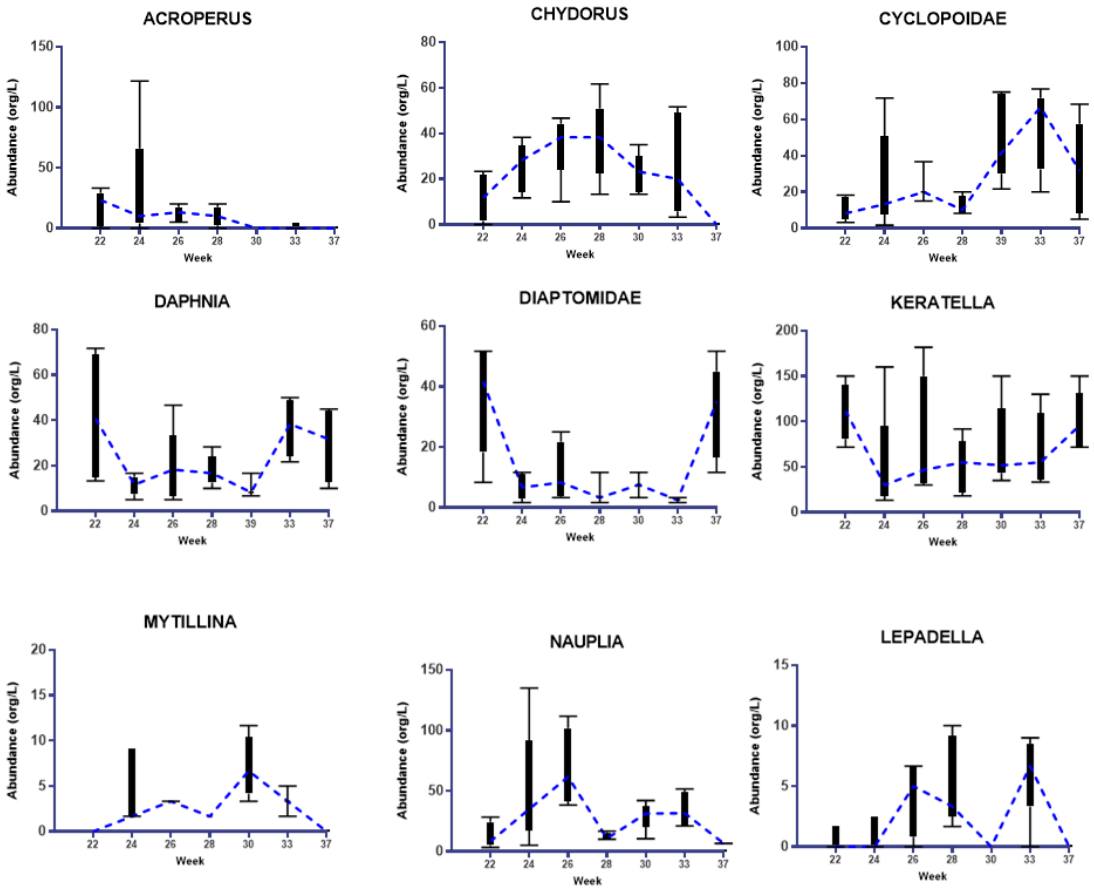


Figure 4. 4: Morphological identification and abundance of the nine selected zooplankton genera at each collection day.

Collection dates are shown on the X-axis while relative abundance is shown on the Y-axis. The blue dashed line represents the connection between each collection point at the median level.

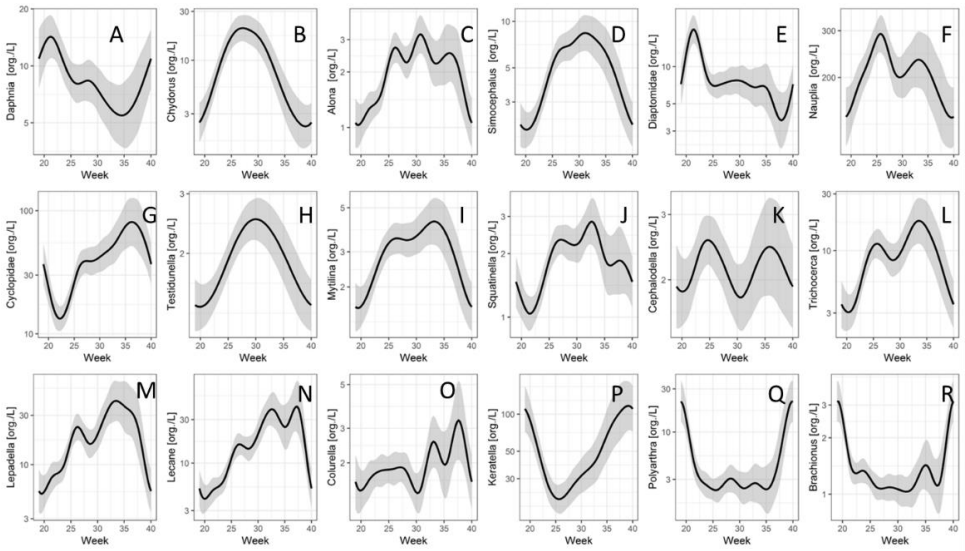


Figure 4. 5: Zooplankton abundances during a mesocosm season using data collated from control replicates in 12 mesocosm studies carried out by CEA between 2011 and 2017 Historical control Data, (HCD).

Where a–d different *Cladocerae*; e-g different *Copepoda* h–r different *Rotiferae*; species; names given on y-axes; Mesocosm season ends in Week 40 = circa 4th October. Solid black lines represent the fit from generalised additive model (GAN); grey shading indicates 95% confidence interval. Reproduced from (Brooks et al., 2019).

4.2.3 Metabarcoding Analysis of Zooplankton: Taxonomic Identification and Biodiversity Assessment

The aim here was to identify the most prevalent zooplankton genera by molecular analysis and determine whether these genera were also identified by morphological analysis.

To achieve this, high-throughput sequencing data underwent several stages of bioinformatics processing, each critical in refining the dataset to accurately assign taxonomy. The following table (Table 4.3) summarises the sequences at each stage, detailing the number of sequences retained or lost, and ultimately, how many sequences were successfully assigned taxonomy. Sequence matrices were examined to assess the average read quality both prior to and following quality control (QC). Additionally, comparing the average read number per sample was considered beneficial for determining the comparability of the samples, with an average of 416,050.79 reads per sample before QC and 241,271.59 reads per sample after QC.

Table 4. 3: Summary of Sequence Retention and Taxonomic Assignment at Each Stage of the Bioinformatics Workflow.

Software	Read Count	Description
RAW DATA	18.6 M	Represents the initial sequencing output from the Illumina platform, containing all raw reads generated before any processing or filtering.
FLASH	10.97 M	FLASH (Fast Length Adjustment of SHort reads) merges paired-end reads that overlap, producing longer, more accurate reads for downstream analysis.
TRIMMOMATIC	11,35 M	Involves removing low-quality bases and adapter sequences from the raw reads, ensuring that only high-quality, clean sequences are retained.
CUTADAPT	10,42 M	A tool that removes adapter sequences, primers, and other unwanted sequences from reads, refining them for more accurate alignment and analysis.
SEQTK	10,42 M	A lightweight tool used for processing sequences in FASTA or FASTQ format, typically for tasks like filtering, trimming, and subsampling.
Assigned taxonomy by BLAST	588,707	BLAST (Basic Local Alignment Search Tool) assigns taxonomy to sequences by comparing them against known sequences in a database, identifying species present.

During the metabarcoding preparation, two random samples (1A and 7A) were selected to check PCR reproducibility using 3, 6, 9 and 12 replicates. Results of the ANOVA test showed no significant difference ($p=0.99$ for sample 1A and $p=0.98$ for sample 7A).

Therefore, PCR in triplicates was performed in all subsequent analysis as no statistical difference was found using more replicates.

The two repeat metabarcoding data sets (R1 and R2) were tested against each other to check for reproducibility. However, when the average between the two repeats was taken, three genera out of 23 for each repeat were found only in one repeat and not in the other one (*Chydoridae, Colurella, Daphniidae* in R1 and *Asplanchna, Bravhionidae* and *Polyarthra* in R2). These genera were not any of the nine genera selected by morphological analysis, the NGS data for these genera were counted in the total genera detected using metabarcoding analyses. The *Nauplia* genera was not present in the barcoding database and could not be analysed by this method.

Results of the paired T-test showed no significant difference between the two metabarcoding repeats for the eight genera selected for NGS analysis by morphological examination (Table 4.4).

Table 4.4: Results of paired T-test between R1 and R2 of metabarcoding.

No significant differences were found for any of the eight selected zooplankton genera and thus one repeat will be considered explanatory enough in further analyses.

Genera	Metabarcoding R1 vs R2
Acroperus	0.83
Chydorus	0.1
Cyclopoidae	0.97
Diaptomidae	0.91
Daphnia	> 0.99
Keratella	0.99
Lepadella	0.92
Mytillina	0.85

From this point onwards, the average between the two repeats was taken for each genus for further analyses as no significant difference between the two was determined.

A total of 23 genera were detected by metabarcoding analysis (Figure 4.6) where genera with less than 0.1% abundance are not presented. The genera detected belonged to three different zooplankton families (Figure 4.7).

As the *Nauplia* genus was not identified in the metabarcoding database, i.e., there was no sequence information available in publicly available databases (known as a barcoding gap), the eight most abundant genera were analysed instead of nine (*Diaptomidae, Lepadella, Daphnia, Acroperus, Keratella, Mytillina, Cyclopoidae, Chydorus*).

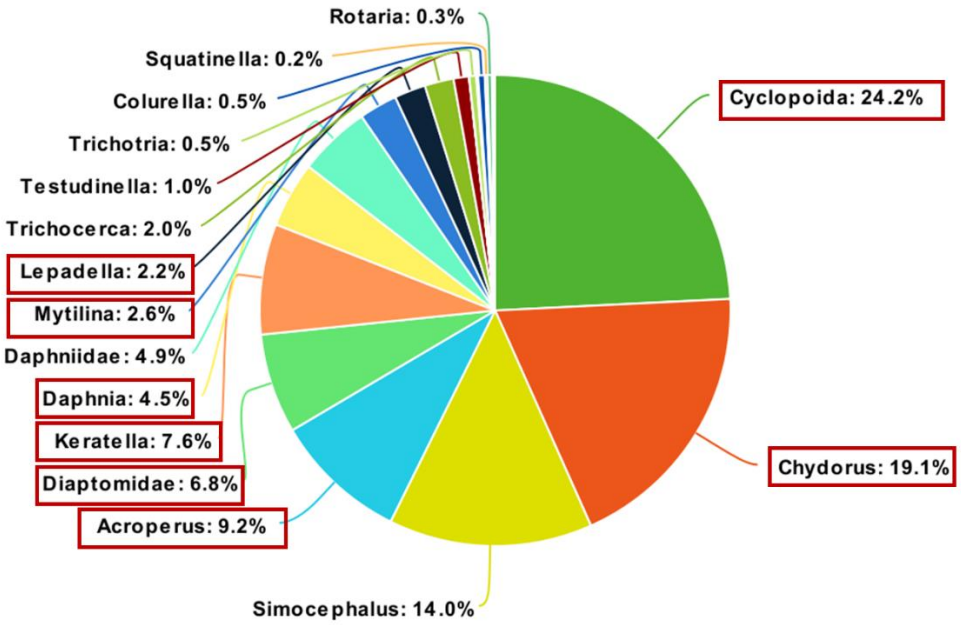


Figure 4. 6: Metabarcoding genera presence and abundance detected in morphological analyses of the 35 mesocosm samples.

The percentage was obtained by taking the average for each genus of all collection dates. Highlighted in red are the nine most abundant selected genera.

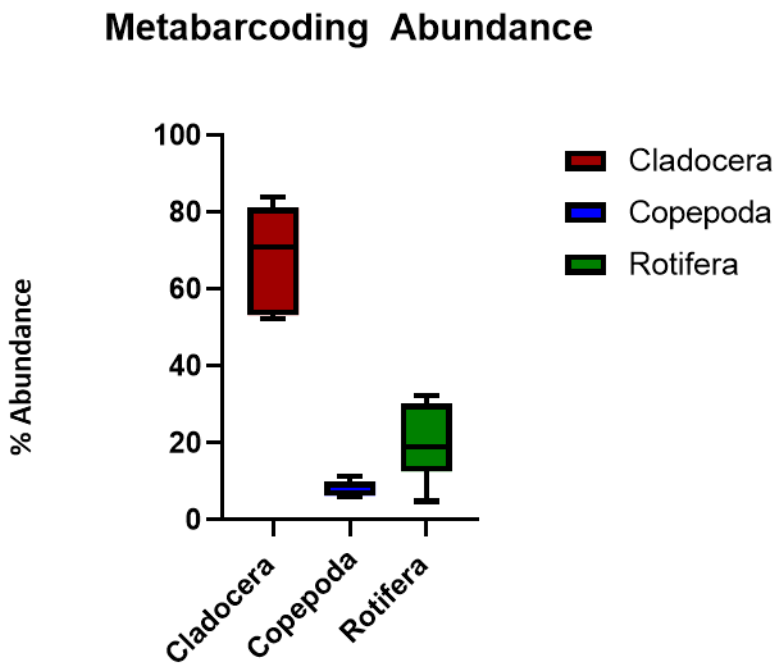


Figure 4. 7: Zooplankton groups' abundance found in the 35 mesocosm samples analysed with metabarcoding.

4.2.4 Comparative Analysis of Zooplankton Taxonomy: Morphological Assessment vs. Metabarcoding Approach

The aim here was to compare the most prevalent genera of zooplankton by both methodologies and to determine whether metabarcoding could replace or aid traditional identification techniques.

The bar charts representing abundance in both methodologies are shown in Figure 4.8. The comparison showed a visual difference in genera abundance between morphology and metabarcoding. This was then confirmed in the Venn diagram (Figure 4.9) showing a total of 26 different genera found, where 17 (65%) of them were identified by both methodologies, 3 only by morphology and 6 only by metabarcoding.

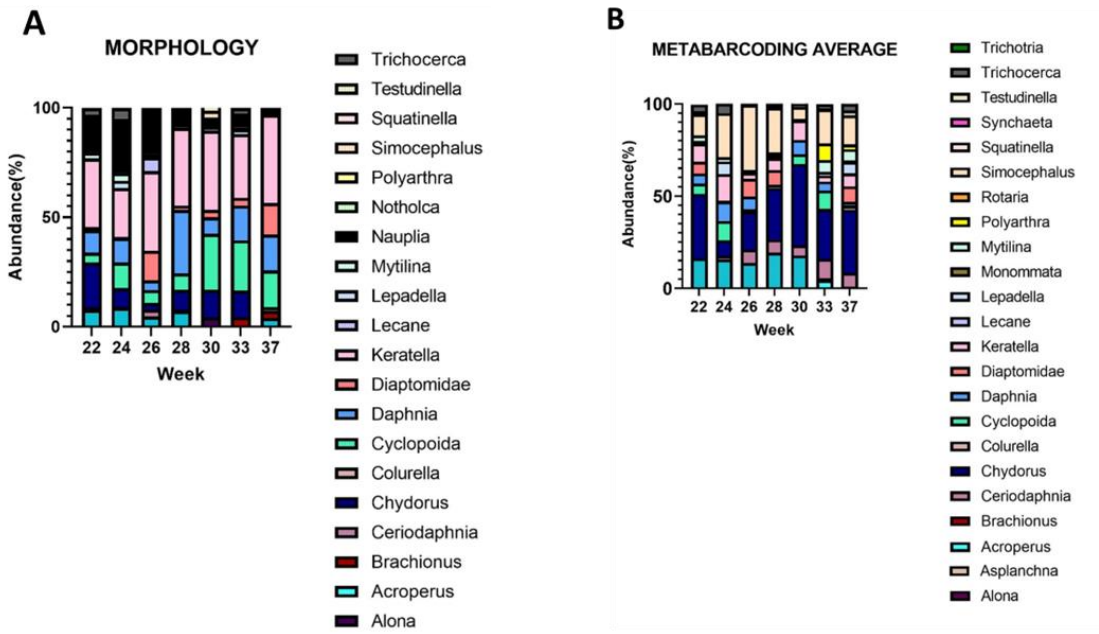


Figure 4. 8: Bar chart showing the genera abundance for both methodologies.

The morphology per week (A) throughout the zooplankton season and, the genera abundance for the average (between R1 and R2) metabarcoding (B). Morphology has detected 20 genera while metabarcoding has detected 23.

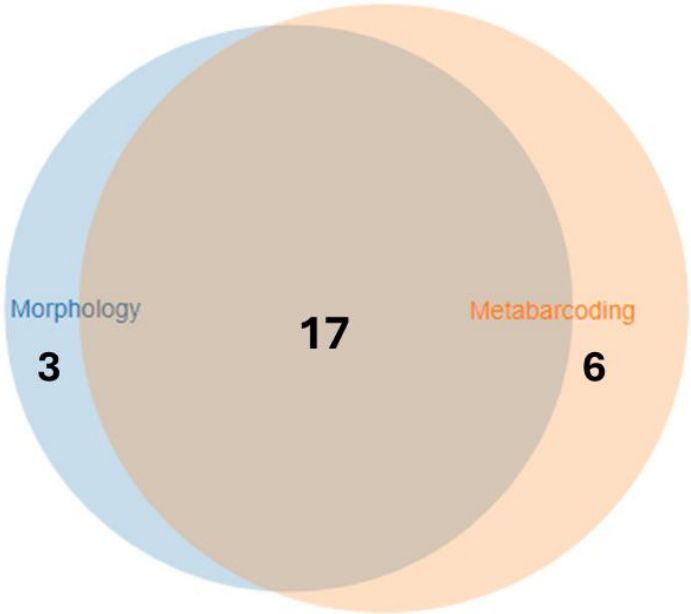


Figure 4. 9: Venn diagram of morphology and metabarcoding identification.

Morphology has detected 20 genera, while metabarcoding has detected 23. Three genera were detected by Morphology and not metabarcoding, six genera were detected

metabarcoding and not morphology, and 17 genera were detected by both methods, a 65% intersection.

As stated before, the statistical analyses were focused only on the eight most abundant genera detected by morphological analysis, over at least three sampling times. Results of the comparison scatter plot showed a similar pattern for seven out of the eight genera (*Diaptomidae, Lepadella, Acroperus, Keratella, Mytillina, Cyclopoidae* and *Chydorus*) but not for *Daphnia* (Figure 4.10). A final test was made to check for correlation between the two methodologies by making a Spearman Rank correlation test (Figure 4.11). The Spearman correlation test was conducted to assess the relationships between different genera, providing correlation coefficients (ρ) and associated p-values summarised in Table 4.4. *Acroperus* and *Chydorus* exhibited moderate positive correlations of 0.58 and 0.64, respectively, but the correlations were not statistically significant (p-values of 0.17 and 0.12). *Cyclopoidae* showed a weak positive correlation (0.35, p = 0.44), while *Daphnia* displayed a strong negative correlation (-0.69, p = 0.065), approaching statistical significance with a p-value of 0.085. *Diaptomidae* exhibited a relatively strong positive correlation (0.67, p = 0.1). *Keratella, Lepadella,* and *Simocephalus* showed weak to moderate correlations (0.28, 0.53, -0.39) with p-values higher than 0.05. *Mytilina* displayed a very weak negative correlation (-0.08, p = 0.86). In summary, while some genera showed moderate to strong correlations, many were not statistically significant at the 0.05 level.

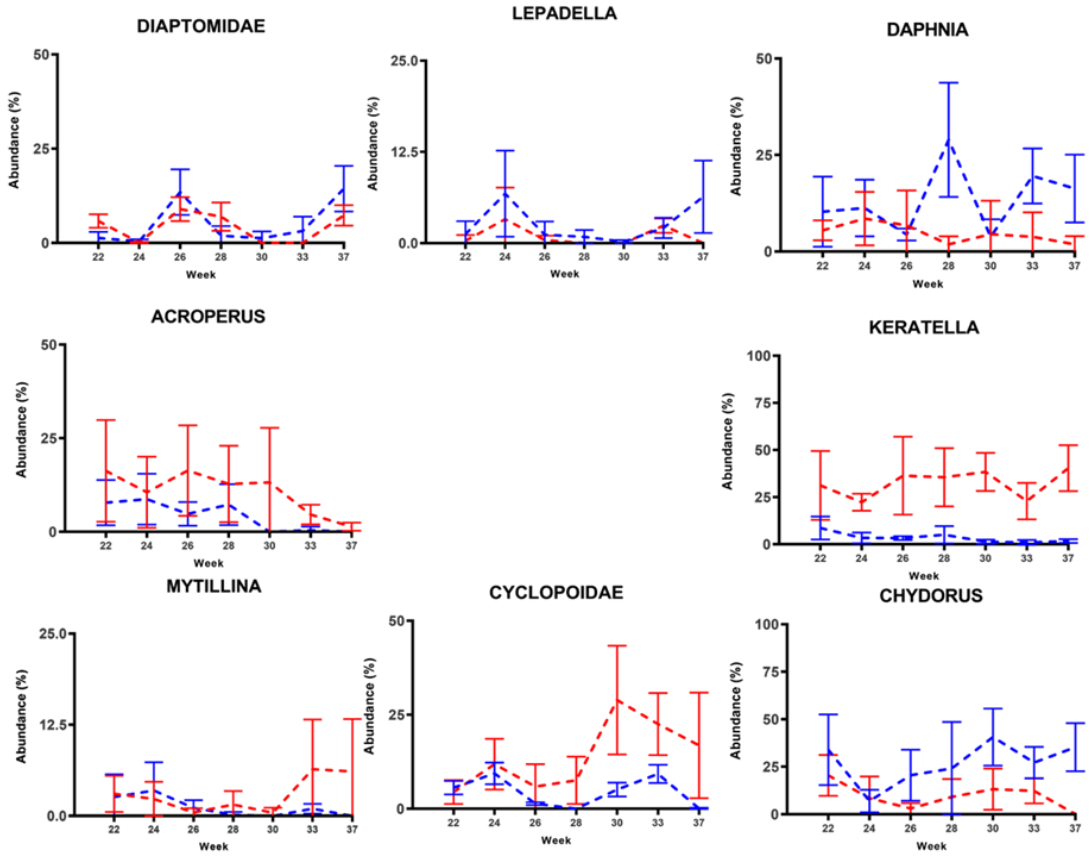


Figure 4. 10: Visual representation of the eight selected genera identification through morphology (blue) and average between R1 and R2 metabarcoding (red) compared throughout the season.

Error bars show the mean and error standard deviation.

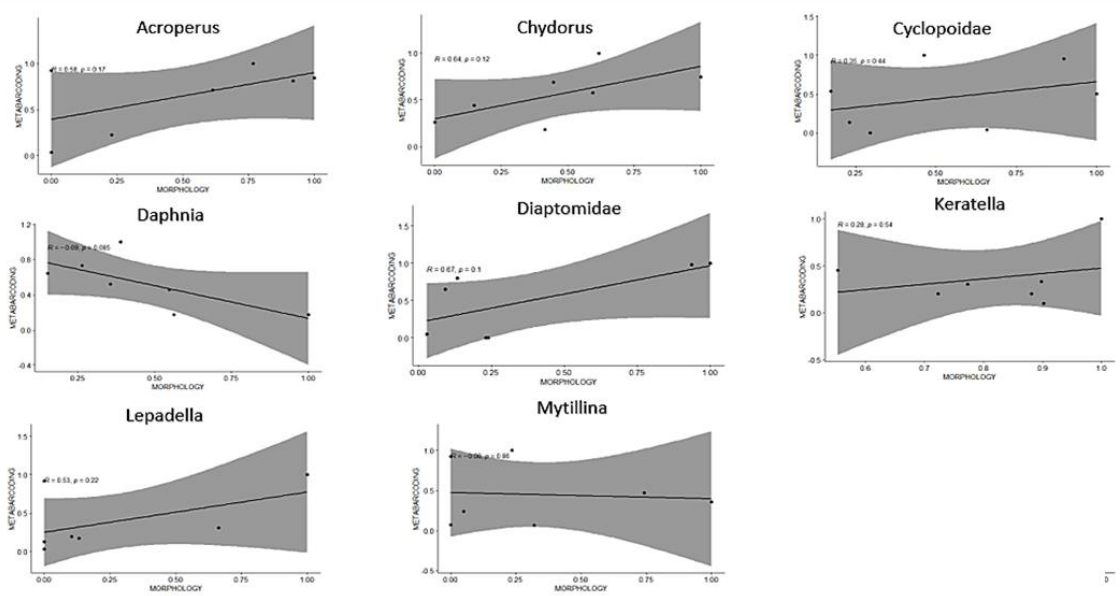


Figure 4. 11: Results of Spearman correlation test performed on the nine selected zooplankton genera.

The X-axis represents morphology, and the Y-axis represents metabarcoding. *Acroperus* and *Chydorus* showed a moderate positive correlation (0.58 and 0.64) but were not statistically significant (p-values: 0.17 and 0.12). *Cyclopoidae* showed a weak positive correlation (0.35, p: 0.44), while *Daphnia* showed a strong negative correlation (-0.69, p:0.085). *Diaptomidae* showed a relatively strong positive correlation (0.67, p: 0.1). *Keratella* and *Lepadella* showed a weak to moderate correlation (0.28,0.53, -0.39) with p-values above 0.05. Finally, *Mytillina* showed a very weak negative correlation (-0.08, p: 0.86).

Table 4. 4: Summary of the Spearman correlation test between morphology and metabarcoding on the nine selected genera.

Genera	R-value (ρ)	P-value
Acroperus	0.58	0.17
Chydorus	0.64	0.12
Cyclopoidae	0.35	0.44
Daphnia	-0.69	0.085
Diaptomidae	0.67	0.1
Keratella	0.28	0.54
Lepadella	0.53	0.21
Mytillina	-0.08	0.86
Simocephalus	-0.39	0.39

The similarity test showed differences in the internal community of zooplankton across sampling dates (Figure 4.12). the results of the ANOSIM revealed significant differences in community structure among the seven collection dates, for both morphology and metabarcoding. In the case of morphology, the R-value of 0.45 indicates a moderate level of dissimilarity among the collection dates, and the associated p-value of 0.001 indicates that these differences are statistically significant. Similarly, for metabarcoding, the R-value of 0.14 suggests a lower level of dissimilarity compared to morphology, but the p-value of 0.01 still indicates statistical significance. These findings suggest that, as expected, the composition of communities varied significantly over the different collection dates, highlighting temporal variations in both morphological and metabarcoding assessments of community structure. Further exploration and analysis may be necessary to understand the specific factors contributing to these temporal differences in community composition.

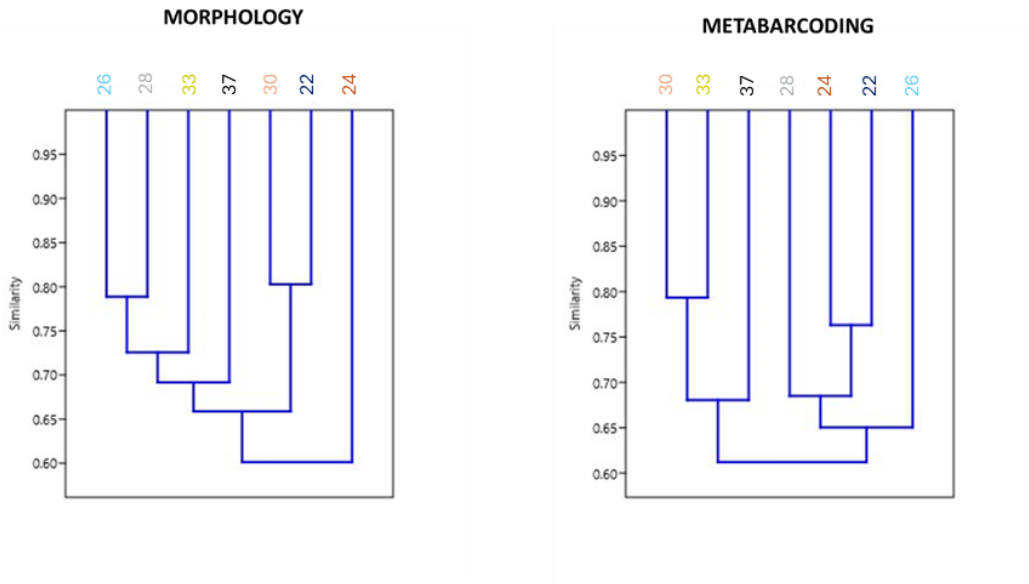


Figure 4. 12: Visual representation of the internal zooplankton community across sampling dates.

Each node shows a divergence. Sampling weeks on the same branch are more closely related to sampling weeks on different branches.

4.2.5 Relationship between water chemistry and zooplankton abundance

The aim here was to investigate whether environmental variables affected zooplankton abundance. In the analysis of the mesocosm study, a Mantel Test was employed to explore the relationships between six environmental variables and the species Bray-Curtis dissimilarity matrix within the eight abundant genera selected by morphological analysis (Figure 4.13). Temperature and turbidity exhibited no significant relationship with genera dissimilarity, as indicated by Mantel statistics (R values) of 0.29 and 0.08, with associated p-values of 0.12 and 0.32, respectively. Similarly, oxygen, conductivity, and chlorophyll demonstrated a non-significant negative correlation with species dissimilarity, reflected in Mantel statistics (R values) of -0.32, -0.07, and -0.1, and p-values of 0.93, 0.62, and 0.64, respectively. Lastly, pH displayed a non-significant positive correlation with species dissimilarity, with a Mantel statistic (R value) of 0.00 and a p-value of 0.45. These results collectively suggest that, within the studied mesocosms, the examined environmental variables do not exhibit statistically significant associations with the dissimilarity in species composition.

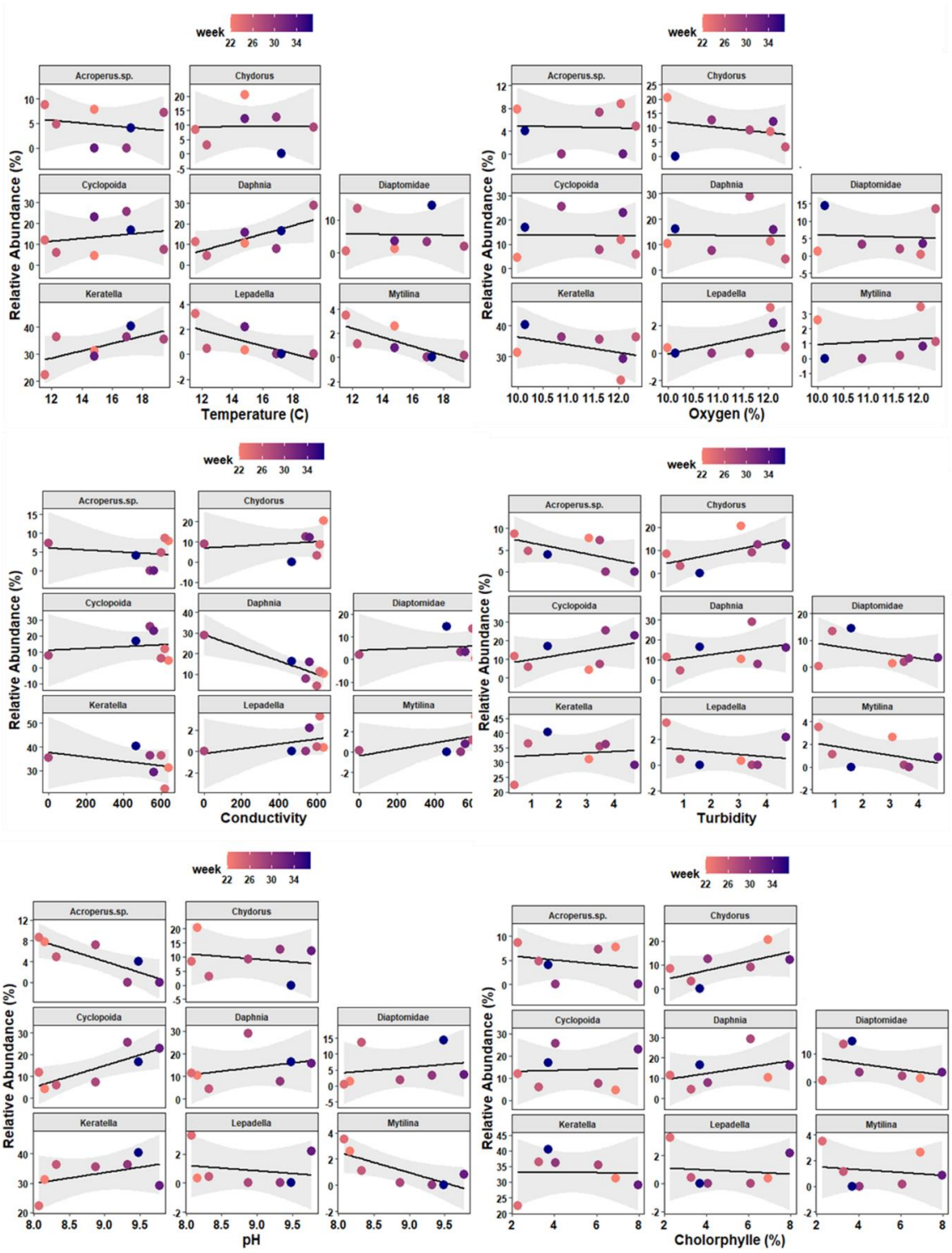


Figure 4. 13: Mantel test on the six environmental variables influence on selected zooplankton genera found across the season.

Temperature and turbidity showed no significant relationship with species dissimilarity (Mantel R values of 0.29 and 0.08, p-values of 0.12 and 0.32). Oxygen, conductivity and chlorophyll displayed non-significant negative correlations with species dissimilarity (Mantel R values of -0.32, -0.07 and -0.1, p-values of 0.93,0.62 and 0.64, respectively). Lastly, pH exhibited a non-significant positive correlation with species dissimilarity (Mantel R value of 0.00, p-value of 0.45).

4.3 Discussion

Environmental factors can affect the life cycles of zooplankton. Changes in physicochemical conditions directly, and indirectly, affect zooplankton distribution and occurrence (Dorak 2013). Water temperature and dissolved oxygen values are two of the most critical factors affecting the abundance of zooplankton (Park et al. 2000). Dissolved oxygen concentration reflects the dominating biological and physical processes in aquatic environments, and it is one of the most critical parameters to determine water quality. Water temperature is essential as it manages the chemical and biological activity of organisms in aquatic life. It has been proven that as the temperature increases, the abundance and diversity of zooplankton species are higher in aquatic ecosystems (Rasconi et al., 2015). Significant differences were found in four environmental variables out of six (Oxygen, pH, Conductivity and Turbidity), throughout the season. The differences in seasonality of these environmental variables could have caused a change, albeit not significant, in zooplankton abundance within the season. This should be considered when interpreting data on the effects of any toxicant.

The morphological and molecular data of the seasonal mesocosms samples showed high variation of the selected eight common genera across the sampling season. Compared with the previous study conducted on 12 mesocosms on the same site (Brooks et al. 2019), a similar pattern was found for most genera. However, differences were expected as the 12-mesocosm studies (Brooks et al. 2019) contain data from 12 seasons. While one might anticipate the season to remain relatively consistent each year, fluctuations in environmental factors could lead to variations in zooplankton abundance and impact different genera to varying degrees.

This study aimed to establish an efficient method to prepare zooplankton samples for metabarcoding. Creer et al. (2010), Pawlowski et al. (2018), and Zhan et al. (2013) have demonstrated the utility of metabarcoding in biodiversity assessment, particularly in the analysis of meiofaunal biospheres and the detection of rare species in aquatic communities. While the NCBI nt database allows for comparing multiple loci against a consistent database, more accurate taxonomic assignments are likely when using specialised databases like SILVA for 18S sequences (Lindeque et al. 2013). Regardless of the chosen database, assigning taxonomy to Taxonomic Units may be biased or

impeded due to a lack of reference sequences. This limitation was evident in the case of the copepod genus Nauplii in our study, where the identification of certain taxa observed via microscopy becomes impossible if their sequences are absent in the databases. Furthermore, some databases do not necessitate voucher specimens for species identification during sequence submission, potentially resulting in inaccurate or ambiguous annotations. The need for caution in interpreting sequence comparison results, particularly concerning rare or unexpected species, until databases are more comprehensively populated is essential (Djurhuus et al., 2018).

The comparison between morphological and metabarcoding identification showed no significant correlation between the selected 8 genera's abundance. However, a similar number of genera was found, even though both methods have detected genera do not present in the other. In fact, a total of 26 different genera were found, where 17 of them were identified by both methodologies, 3 only by morphology and 6 only by metabarcoding. The fewer genera detected by morphology could be due to misidentification during morphological analysis, especially in more turbid samples. A total of 65% of the genera were identified by both methodologies, showing that molecular results support the hypothesis that DNA metabarcoding can augment the morphological identification for estimating the diversity and composition of the zooplankton assemblage. As many studies have shown (Djurhuus et al., 2018; Schroeder et al., 2020), metabarcoding is often able to detect more genera than morphology. This suggests that the combined use of DNA metabarcoding and morphological analysis could complement each other to better understand zooplankton biodiversity. However, this lack of direct correlation for each genera indicates that whilst both methods can measure relative zooplankton population abundance, and seasonal changes in those populations, what the methods are measuring are distinct. The next chapter will investigate whether analysis by both methods will provide the same or distinct outcomes from a pesticide trial.

Chapter 5. Ecotoxicity study using Roundup™

5.1 Introduction

Ecotoxicity studies are crucial for testing, and safely assessing, new pesticides or pesticide formulations. Most ecotoxicological studies are still performed in laboratories, and these studies usually provide information on the direct effects on the test organism. This makes realistic extrapolation to the field situation uncertain (Svendsen et al., 1997). In ecotoxicology, interactions between a whole community and the environment are often more informative (Sylwestrzak et al., 2021) and are achievable using mesocosm experiments. Mesocosm experiments are designed to isolate a part of an ecosystem, giving some control of the physio-chemical conditions and composition of the biota (Svendsen et al., 1997). These environmental conditions are often affected using pesticides of which there are several kinds, including herbicides; insecticides; fungicides; and nematicides (Pathak et al., 2022).

Herbicides are chemical compounds primarily used to control weeds, unwanted plants and often algae. This chapter will focus on a glyphosate-based herbicide (GBH), Roundup™ from Monsanto. Glyphosate is the most frequently used herbicide worldwide, including in the EU, and it has been used for several decades (Tarazona et al., 2017; Goncalves et al., 2019; Vera et al., 2012). It acts by inhibiting the enzyme 5-Enolpyruvyl Shikimicacid 3-phosphate synthase (EPSPS) which disrupts the amino acids' aromatic biosynthesis, reducing protein synthesis and growth and eventually causing cellular disruption and death (Vera et al., 2012). However, this herbicide often affects non-target organisms, including aquatic species. Furthermore, other compounds, such as surfactants, are often added to the commercial formulation to increase adhesion and, thus, penetration to the leaf surface and absorbance (Goncalves et al., 2019) which can again affect non-target organisms. Glyphosate is highly hydrophilic, its half-life in aquatic environments is less than seven days, and it dissipates by 50 % in a maximum of two weeks (Sylwestrzak et al., 2021; Goncalves et al., 2019). Whilst glyphosate does not directly affect zooplankton, in aquatic environments it causes the death of plants used by zooplankton communities as refuge and feeding systems, leaving them more exposed to predators and triggering a chain effect (Goncalves et al., 2019). Zooplankton are commonly used as a bioindicator of water quality for environmental conditions since they have a short generation time and a fast reproduction cycle (Islam et al., 2023).

Identification of the zooplankton community and diversity is therefore essential in ecotoxicology studies.

Metabarcoding is revolutionising how we assess biological communities, and its application to ecotoxicity studies could lead to the quicker and more economical identification of organisms compared to traditional microscopy methods. Whilst the application of metabarcoding techniques to identify indicator species has been broadly studied, to the best of my knowledge, the application of metabarcoding to identify zooplankton in ecotoxicity studies remains poorly investigated. However, some emergent studies have investigated the analysis of zooplankton using metabarcoding (Bucklin et al., 2019; Zhao et al., 2021; Djurhuus et al., 2018 and Ershova et al., 2021). Therefore, in this study, the hypothesis is that metabarcoding techniques can be applied to identify and quantify the relative abundance of zooplankton genera within in-field ecotoxicity studies, to replace, or complement, traditional microscopy techniques.

The specific objectives for this chapter were to analyse the effect of glyphosate on zooplankton genera as individuals and as a community, and to look for possible correlations between morphological identification and metabarcoding methods during a pilot-scale ecotoxicology trial.

5.2 Results

5.2.2 Quantitative Analysis of Water Chemistry in Mesocosms

Pictures were taken on sampling day 1 pre-application and on sampling day 9 post-application (20 weeks later; Figure 5.1). Pictures on sampling day 1 pre-application look empty as the macrophytes have only just been introduced to the mesocosms and were yet to grow/establish. Pictures on sampling day 9 post-application clearly show the difference in macrophyte growth between the control and high-dose samples, with the control samples showing substantially more growth than the dosed mesocosms. All graphs throughout this study are colour coded as such: black for control, green for low-dose (5.5 mg a.i./L) and red for high-dose (22.0 mg a.i./L). Mesocosm 149LD (low-dose) had less plant growth compared to the other two low-dose mesocosm. Also, mesocosm 147HD (high dose) showed higher plant growth compared to the other two high-dose mesocosms. In addition, low and high glyphosate dose mesocosms presented higher floating filamentous algae levels than the control mesocosms.

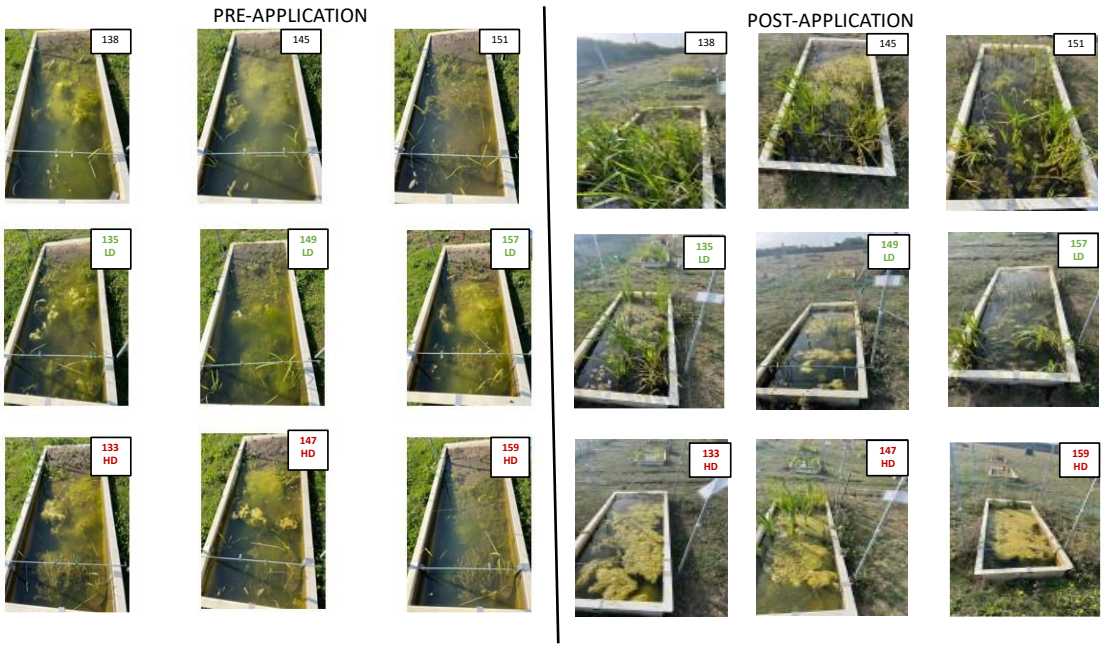


Figure 5. 1: The nine selected mesocosms pictures comparison on the first day of trial- week 18 (pre-application on the left) and on the last day of trial -week 38 (post-application on the right).

Where LD is the low-dose (5.5 mg a.i./L) and HD is the high-dose (22.0 mg a.i./L).

The scatter plot graph created to visually represent the five environmental variables (temperature, oxygen, pH, conductivity, and turbidity) showed the difference between the three herbicide concentrations throughout the season (Figure 5.2). No data was collected for Chlorophyll due to equipment failure. The vertical dotted line in the graphs represents the application day. When observing the graphs, it can be noted that the plots at the three different concentrations of glyphosate followed the same pattern for both temperature and conductivity. Sudden decreases in conductivity, such as that seen in Week 26, can be due to heavy rainfall (Franceschini et al., 2019). However, when the precipitation patterns were looked at for the season, no significant precipitation was found on that, or previous, days so this remains unexplained. Additionally, the high-concentration glyphosate plot showed higher values compared to the control plot for pH, oxygen, and turbidity. A pH over 7 was expected in the control as per the previous sampling year (2020, Section 4.2.1, Chapter 4). When the changes throughout the season were analysed using the ANOVA test, the only significant differences were found in the low concentration (5.5 mg a.i./L) of glyphosate for oxygen (p-value= 0.03) and in the high concentration of glyphosate for pH (p-value= 0.03) (Table 5.1). Furthermore, when a comparison was made between the environmental variables using

the RM-ANOVA and Dunnett's post hoc test pre-application (Table 5.2A), a significant difference was globally found for the temperature (p=0.04). Post-application (Table 5.2B), a significant difference was globally found for the temperature variable (0.03) and more specifically between the control and the high-concentration of glyphosate (p= 0.03): a significant difference was globally found for the oxygen variable (p= 0.01) and more specifically between the control and the high-concentration of glyphosate (p= 0.01); and finally a significant difference was globally found for the pH (p < 0.01) and more specifically between the control and the high-concentration of glyphosate for turbidity (p-value = 0.01).

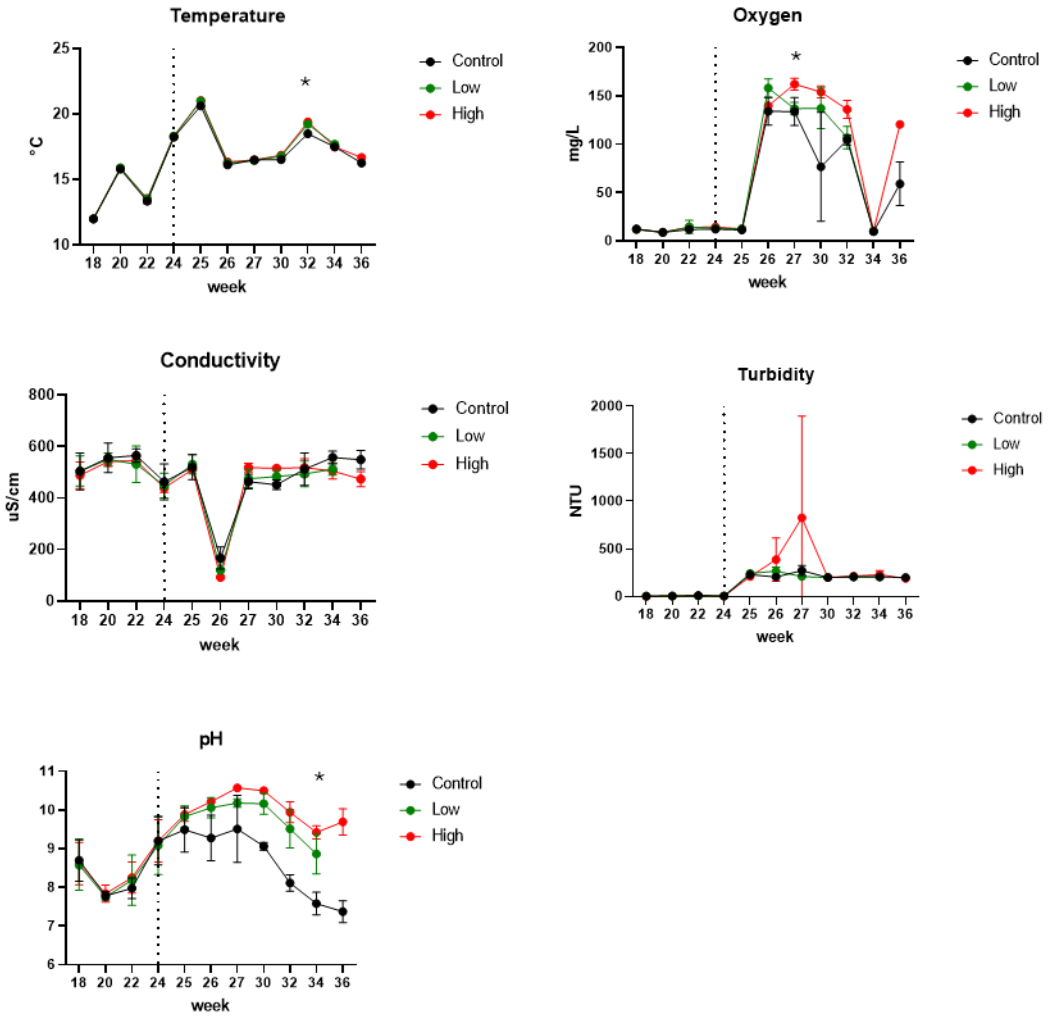


Figure 5. 2: Comparison of the five environmental variables analysed in the study throughout the season (starting on week 18 and ending in week 36) for low and high glyphosate addition compared to the control.

Each line represents a different dose where the blue line= control, the orange line=low-dose(5.5mg a.i./L) and the red line =high-dose(22mg a.i./L). The vertical line projected

in week 25 represents the first sampling point post-application and divides pre- and post-application for easier comparison. Significant differences ($p < 0.05$) were shown by:*.

Table 5. 1: Environmental variables change throughout the season at each concentration.

Where C= control, L=low-dose (5.5mg a.i./L) and H=high-dose (22mg a.i./L). Significant differences were shown by:*. Where The F-Value is a ratio of the variance between groups to the variance within groups and the Pr(>F) is the p-value associated with the F-Value. It indicates the probability of obtaining the observed F-Value (or a more extreme one) if the null hypothesis is true.

ENVIRONMENTAL VARIABLE	Concentration	F-Value	Pr(>F)
Temperature	C	2.18	0.174
	L	3.58	0.095
	H	2.55	0.14
Oxygen	C	1.64	0.23
	L	6.16	0.03*
	H	3.2	0.1
pH	C	0.97	0.34
	L	3.81	0.08
	H	6.13	0.03*
Conductivity	C	0	0.98
	L	0.13	0.72
	H	0.02	0.89
Turbidity	C	12.48	0.21
	L	11.83	< 0.05
	H	1.76	< 0.05

5.2.3 Concentration-Dependent Impacts of Roundup™ Herbicide on Zooplankton: Insights into Genera, Group and Community-Level Responses

The analysis performed to assess Roundup™ effects on zooplankton community, group and genera are summarised in Figure 5.3. The statistical analyses were carried out in identical order for both methodologies.

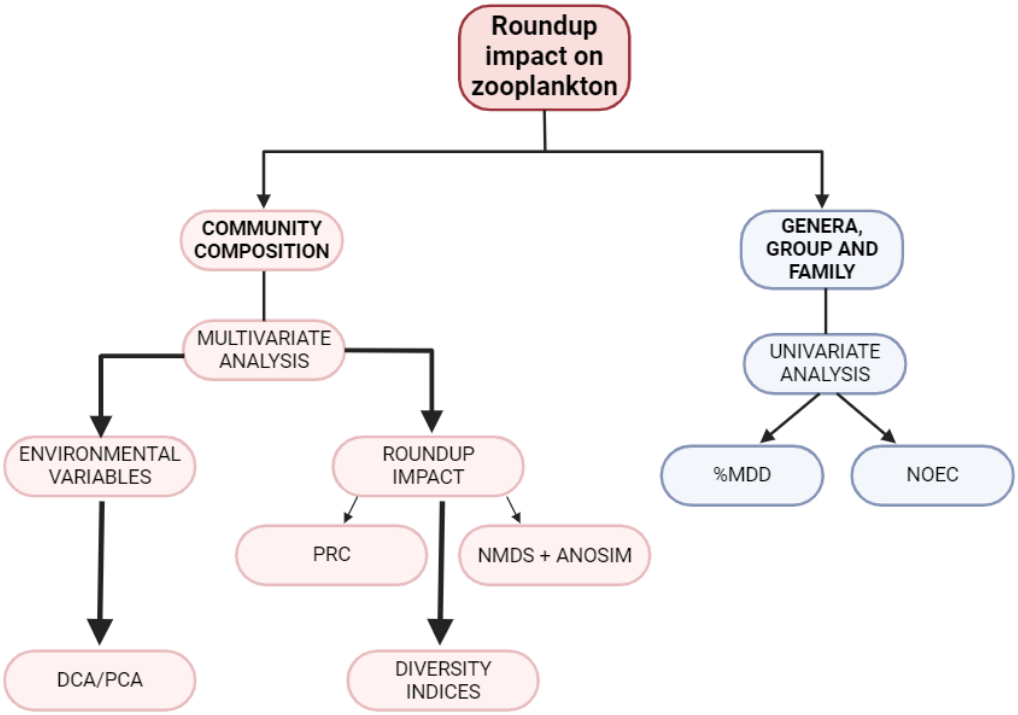


Figure 5. 3: Summary of statistical analyses performed on zooplankton genera, group, family, and community composition to assess environmental variables and Roundup™ impacts.

5.2.3.1. Impacts in community composition of environmental variables when analysing zooplankton by microscopy

DCA was employed to assess the correlation between community distribution, environmental variables, and glyphosate doses (Low 5.5 mg a.i./L and high 22.0 mg a.i./L) across the study. Results showed an adjusted explain variation of 26.69%. The primary axis aims to identify the correlation between species distribution and the scoring, representing the optimal correlation achievable for a total variation of 22.77% (S8, Appendix). Doses were treated as factors (red) and environmental variables were treated as continuous variables and thus shown as arrows (green) as shown in Figure 5.4. As the doses clustered together, the community was not significantly affected by them. For the environmental variables, the longer the arrow the more impact it had on

the community. In this case, Days and, thus, the season had the biggest impact on community changes. pH and dissolved oxygen had a bigger impact than conductivity, temperature, and turbidity. The correlation between the temperature and the day, shown by both arrows going in the same direction, was expected as it became warmer throughout the season. Temperature, turbidity, and dissolved oxygen seemed to correlate with higher temperatures. Dissolved oxygen usually results in higher algae presence, and with higher pH, more algae are found in a system (Liu et al., 2016). The Day factor resulted almost parallel to the first axis, while pH and conductivity resulted in opposite directions. Again, this was expected as they are directly related and always go in different directions. From a chemical perspective, dissolved oxygen and turbidity were expected between temperature and pH as stated before, with higher pH, more algae formed and thus more oxygenation in the system and turbidity caused by the algae.

The beginning of the season is shown on the left side of the graph while the end of the season is shown on the right side of the graph. Analysing species distribution patterns revealed noteworthy shifts in community composition over the study period. *Lecane* demonstrated proliferation while *Brachionus* diminished, suggesting an inverse relationship between these species. The clustering tendencies of *Trichocerca, Trichotria, Chydorus,* and *Lepadella*, as well as *Cyclopoidae, Diaptomidae, Nauplia Simocephalus,* and *Notholca,* indicated their propensity for co-occurrence. The clustering of *Copepods*, particularly *Nauplia,* mirrored their larval stage prevalence.

Moreover, the outer regions of the graph highlighted rapid fluctuations in *Rotifera* populations, indicative of their characteristic propensity for rapid population dynamics. Notably, the absence of discernible clustering patterns suggested minimal impact of glyphosate doses on community composition.

In summary, the DCA analysis provided intricate insights into the complex interplay between environmental variables and species distribution, offering valuable scientific insights into community dynamics over the study duration. These insights need further exploring by using the univariate analysis to inspect the doses effects on each zooplankton genera, family, and group.

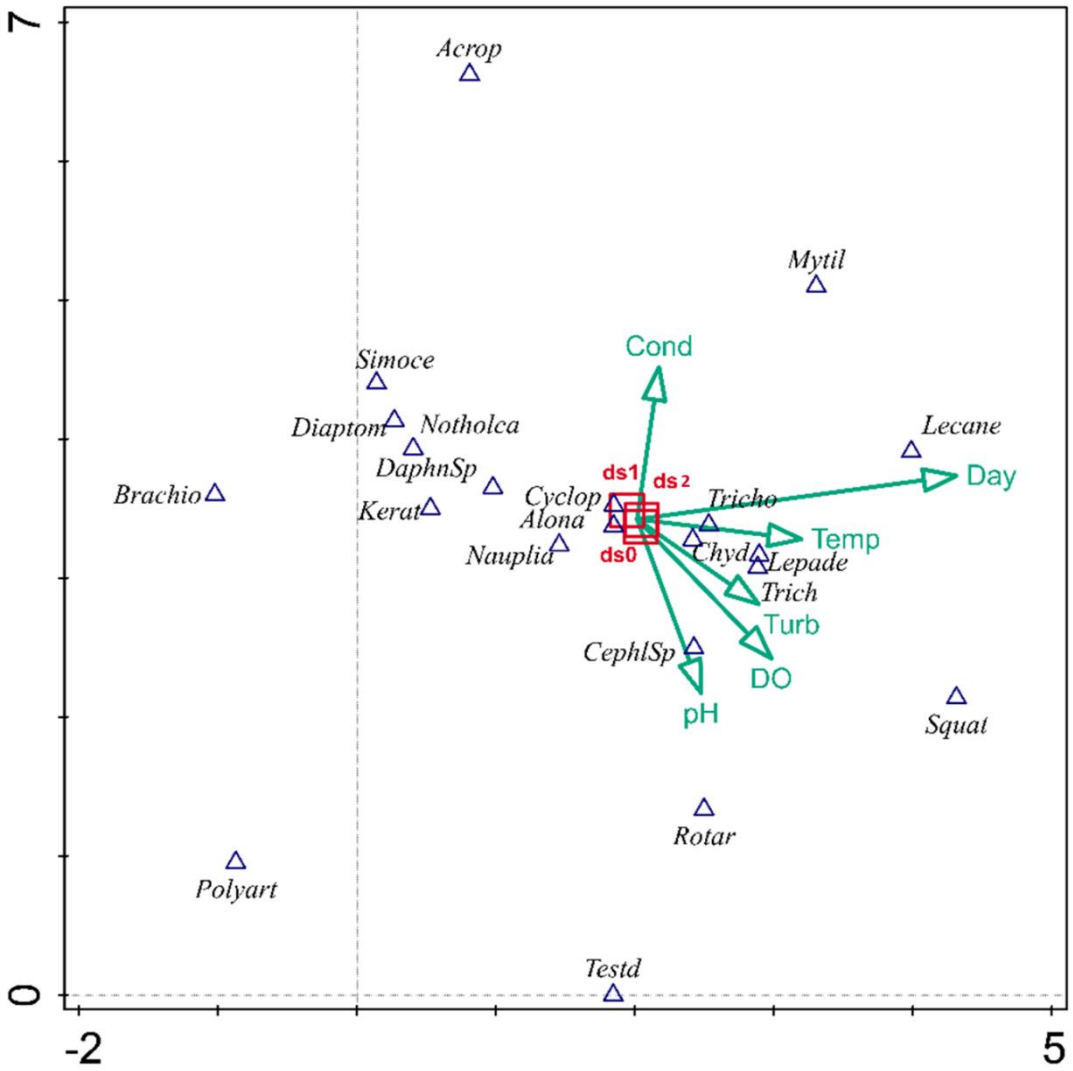


Figure 5. 4: DCA ordination of the species abundance and environmental variables.

Variables are represented by arrows, species are depicted by triangles and the three glyphosate doses are represented in red squares where ds0 = control, ds1= low dose (5.5 mg a.i./L) and ds2= high dose (22.0 mg a.i./L).

5.2.3.2. Data analysis for glyphosate effects on community composition in morphological analysis

All genera were used in the community composition analysis even the one later excluded from the %MDDS and NOEC univariate analyses. The Principal Response Curve (PRC) analysis produced three fundamental results for ecotoxicological studies: Test of significance, site scores and species scores (Szöcs et al., 2014). The statistical

significance of the first PRC axis can be assessed through permutations, determining the extent of variation attributable to treatment effects.

Site scores signify deviations of treated communities from the control, depicted as the mean difference of site scores between treatment and control on the first axis over time. This graphical representation effectively portrays the temporal dynamics of treatment effects on communities (Figure 5.5A). Finally, species scores on the first axis indicate the taxa responsible for the observed treatment-related differences. Taxa with higher absolute scores wield a more pronounced influence on the principal response curve, elucidating their contribution to the observed pattern (Figure 5.5B).

The plotted lines in Figure 5.5A represented the control (black line), low-dose (blue line, 5.5 mg a.i./L), and high-dose (red line, 22.0 mg a.i./L) treatments. Prior to application, the community composition within the mesocosm, assigned to the respective doses, exhibited notable divergence, evidenced by peaks around Day -20 from the application. After application, the community compositions for both doses, when compared to the control, converged, displaying remarkably similar patterns. Notably, a dose-response trend is evident for both doses, although not statistically significant (pseudo-F=0.3, P=0.204).

The species scores, as detailed in Table S9 in the appendix, correspond to the numbers on the axes in Figure 5.5B. Proximity to the axis denotes similarity in response to the chemical; hence, genera situated at the extremes had the strongest impact. When interpreting species scores, it is important to consult the univariate analysis as a low taxon weight does not necessarily equate to a negative response. A taxon may exhibit a robust response divergent from the global pattern, resulting in a low species score (van den Brink and ter Braak, 1999). In the morphological analysis, four genera (*Chydorus, Alona, Testudinella* and *Cephalodella*) resulted in the strongest reaction from glyphosate. *Chydorus* and *Alona* exhibit a negative impact, whereas *Testudinella* and *Cephalodella* demonstrate a positive impact.

In essence, the community was found to be quite different, but not significantly so. While the PRC analysis provides valuable insights into community dynamics in response to glyphosate doses, a comprehensive understanding necessitates careful consideration of the nuanced interplay between taxa and doses, ideally supplemented by univariate analyses to elucidate complex relationships.

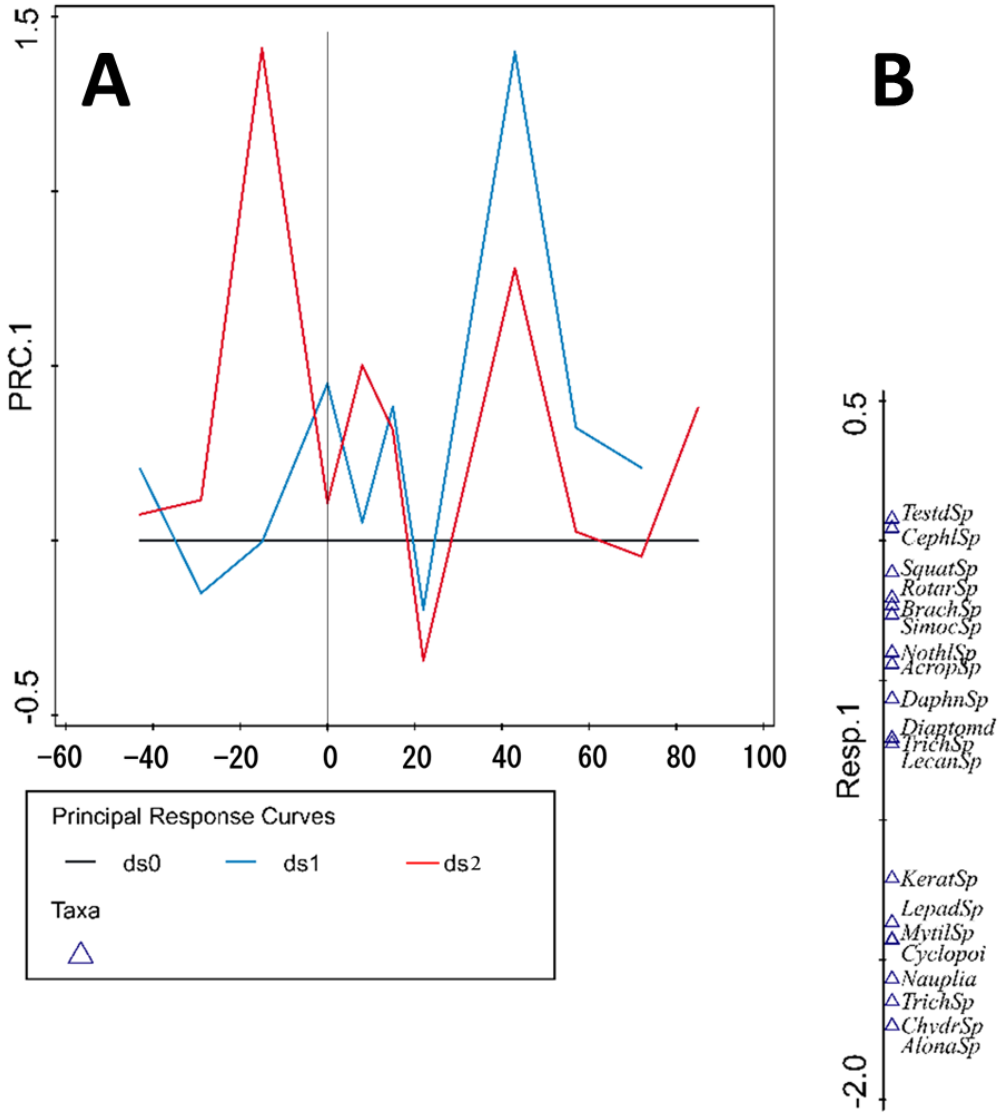


Figure 5. 5: Principal Response Curve for morphology data.

As a special form of partial RDA on the zooplankton community to assess the effect of treatment along time (pseudo-F=0.3, P=0.204). Where ds0 = control, ds1= low dose (5.5 mg a.i./L) and ds2= high dose (22.0 mg a.i./L). The first axis displays the highest fraction of variation that can be explained by the explanatory variables. The mean difference of site scores between treatment and control on first axis displays the deviation of communities from control.

Figure 5.5B: Species scores on first axis responsible for PRC.

The community composition was also assessed by analysing its changes at each collection points for the control and both doses.

Low-stress value for all concentrations was found in the NMDS plot indicating a reliable representation of the dissimilarities in the reduced-dimensional space (Figure 5.6 A). For all analysis: Each sampling point is positioned in the two-dimensional space according to its dissimilarity to others. Points closer in the configuration are more ecologically similar, whereas those farther apart are more dissimilar. The stress levels denoted as 0.05 for the control group, 0.16 for the low-dose group, and 0.11 for the high-dose group, serve as indicators measuring the effectiveness in capturing dissimilarities among data points within the reduced-dimensional space. These stress levels provide insights into the degree of distortion in representing the original data structure, with lower values suggesting a more faithful representation.

The stress metric gauges the extent to which the proximity of samples in a reduced-dimensional space, typically two-dimensional, aligns with their actual multivariate distances. Reduced stress values signify higher alignment, making them preferable (Zhu and Yu., 2009).

The NMDS analysis for the control was conducted to explore dissimilarity patterns among the time points (week 18, 20, 22, 24, 25, 26, 27, 30, 32, 34 and 36). The coordinates along Axis 1 and Axis 2 represent the ecological distances between the samples. Axis 1 explains the largest proportion of dissimilarity (81.16%), while Axis 2 contributes 13.67%. Each sampling point was positioned in a two-dimensional space according to its dissimilarity to others. Points closer in the configuration are more ecologically similar, whereas those farther apart are more dissimilar. Notably, the samples exhibit a clear separation along both Axis 1 and Axis 2, indicating distinct ecological compositions. For instance, time points 18, 20, and 22 clustered together, showing similar ecological characteristics, while samples 30, 32, 34, and 36 form a distinct cluster.

The NMDS analysis for the low dosage showed that Axis 1 explains 66% of the dissimilarity, while Axis 2 contributed 1.9%. The stress value of 0.16 suggested a moderately good fit of the data to the reduced dimensional space, indicating that the NMDS adequately represents the dissimilarity in the dataset. The samples exhibited distinct spatial separation along both axes, reflecting dissimilar ecological

compositions. Samples 18, 20, and 22 are clustered closely together, indicating ecological similarities, while samples 34 and 36 are positioned farther apart, suggesting greater dissimilarity.

The NMDS analysis for the high dose showed that Axis 1 explains 76% of the dissimilarity, while Axis 2 contributes 29%. The stress value of 0.11 indicated a good fit of the data to the reduced dimensional space, suggesting that the NMDS adequately represents the dissimilarity in the dataset. Samples are positioned distinctly along both axes, reflecting dissimilar ecological compositions. For instance, collection points 18, 20, and 22 exhibit similar ecological profiles, while collection points 30 and 32 are positioned farther apart, indicating greater dissimilarity.

This grouping is reinforced by the hierarchical clustering tree based on the Bray-Curtis similarity index (Figure 5.6B).

NMDS was used as a simple method of visual interpretation to compare the overall structure of the zooplankton community at the different sampling points in the two glyphosate doses and a control. The ANOSIM statistically tests the significant differences in these groups. In ANOSIM, the R-value ranges from -1 to 1, where 1 indicates the complete dissimilarity between groups, 0 indicates no dissimilarity), and -1 suggests that dissimilarities are greater within groups than between groups (Somerfield et al., 2021). For the control and the two glyphosate doses, the ANOSIM results were similar where a significant difference was found in all data sets ($R=0.39$, $p = < 0.01$, $R=0.36$, $p= <0.01$ and $R=0.33$, $p= <0.01$ respectively) (Figure 5.6A).

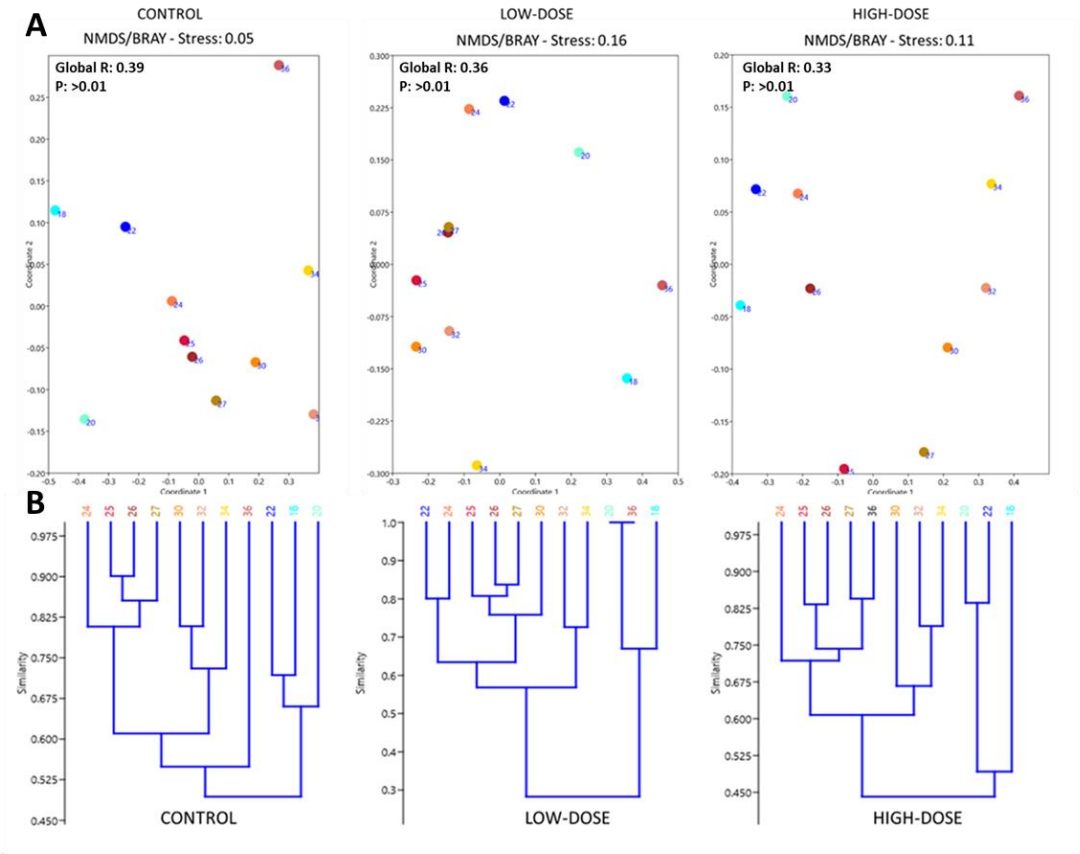


Figure 5. 6: Non-Metric multidimensional scaling (NMDS) analysis based on Bray Curtis dissimilarity matrix.

The coordinates along Axis 1 and Axis 2 represent the ecological distances between the samples. Each sampling point is positioned in the two-dimensional space according to its dissimilarity to others. Points closer in the configuration are more ecologically similar, whereas those farther apart are more dissimilar. The stress levels, denoted as 0.05 for the control group, 0.16 for the low-dose group, and 0.11 for the high-dose group, serve as indicators measuring the effectiveness in capturing dissimilarities among data points within the reduced-dimensional space. These stress levels provide insights into the degree of distortion in representing the original data structure, with lower values suggesting a more faithful representation. ANOSIM resulted in: R=0.39, p = < 0.01 for control, R=0.36, p= <0.01 for low glyphosate dose and R=0.33, p= <0.01 for high glyphosate dose.

5.6 B: Hierarchical clustering tree showing the relationships among sampled points using the Bray-Curtis similarity index. The tree illustrates the hierarchical arrangement of ecological similarities and dissimilarities among the samples. Points that cluster together share higher ecological similarity, while those in separate branches exhibit greater dissimilarity. The Bray-Curtis similarity index serves as a measure of community composition, providing insights into the grouping of samples based on their species composition and abundance patterns.

The Shannon diversity index (H) and Chao-1 showed similarity in control, low-Dose (5.5 mg a.i./L), and high-Dose (22 mg a.i./L) treatments (Figure 5.7). The average (H) for morphology (control) pre-application was 1.82, and for post-application was 2.07; for morphology (low-dose for glyphosate) pre-application was 1.88 and for post-application was 2.01; and finally, for morphology (high-dose for glyphosate) pre-application was 1.90 and for post-application was 1.76. When the two glyphosate treatments plus control were analysed using RM ANOVA, no significant differences were observed in pre-application. However, the RM ANOVA global post-application revealed a significant global value of $p<0.01$. The post-hoc Dunnett's tests post-application was found to be not significant between the control and the low-dose of glyphosate but was significantly different between the control and high-dose of glyphosate ($p<0.01$).

The Chao-1 index assumes that the number of organisms identified for a taxon has a Poisson distribution and corrects for variance. It is useful for data sets skewed toward low abundance. The average Chao-1 for morphology (control) pre-application was 14, and post-application was 16.71; for morphology (low-dose for glyphosate) pre-application was 14.5 and post-application was 15.57; and finally, for morphology (high-dose for glyphosate) pre-application was 7.55 and post-application was 12.85. When the three treatments were analysed using RM ANOVA, no significant differences were observed in pre-application (Table 5.3). However, the global post-application revealed a significant global value of $p<0.01$. The post-hoc Dunnett's tests post-application was not significant between the control and the low dose of glyphosate but was significantly different between the control and high-dose of glyphosate ($p=0.02$).

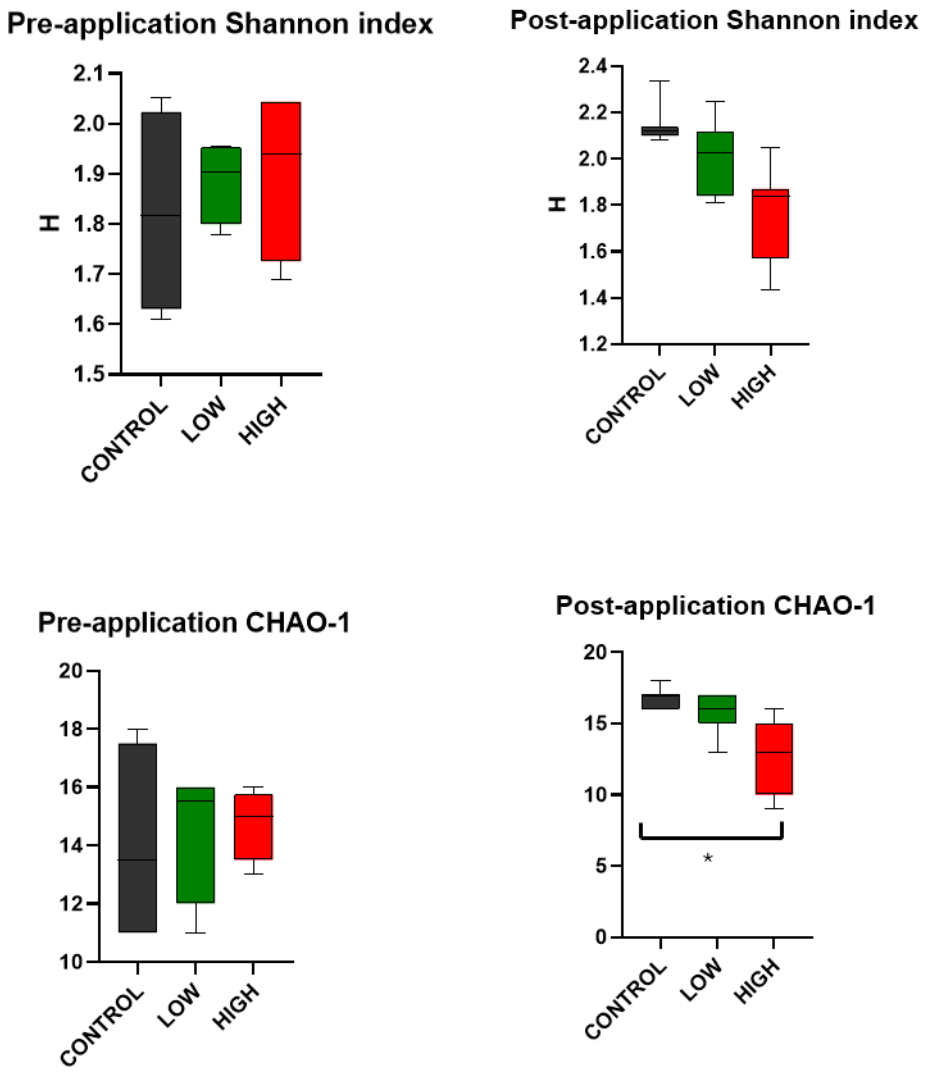


Figure 5. 7: Comparison of Shannon Diversity Index and Chao-1 Index in Response to Control, Low-Dose, and High-Dose Treatments pre- and post-application. Significant differences ($p < 0.05$) were shown by:*.

5.2.3.3 Morphological Analysis of Glyphosate Impact on Zooplankton at genera, family and group level

In the previous section, the analyses were performed to assess the effect of Glyphosate on a community composition level. In this section, analyses were performed to assess the impact of Glyphosate on each genera group and family of zooplankton.

In the study, 23 discrete zooplankton taxa were identified. All identified zooplankton taxa were in a higher taxonomic group; *Cladocera*, *Copepoda* and *Rotifera*. The zooplankton taxa, five grouped families (e.g. *Daphniidae*) and the three groups,

Cladocera, *Copepoda* and *Rotifera*, were used in the univariate analysis, including determining %MDD (Table 5.5) and the $NOEC_{population}$ and overall $NOEAEC_{population}$. Additionally, the total number of zooplankton and the number of zooplankton taxa were analysed similarly. However, the following taxa results were considered not reliable for univariate analysis due to being low and sporadic: *Brachionus*, *Polyarthra* and *Synchaeta*. Data from these taxa were analysed at a higher taxonomic level and thus included in the corresponding grouped family and/or *Cladocera*, *Copepoda* or *Rotifera*, as per the recommendation of EFSA (2013).

For each zooplankton genera, family and group, a category was assigned (Table 5.2) as the first step following the Aquatic Guidance Document (EFSA 2013), especially using the summary table in Figure 5.8. Taxa that were not possible to be analysed because they were too sporadic or absent were excluded from the table. These MDD classes can be used to categorise taxa sampled in the microcosm/mesocosm experiment based on their MDDs. Then, the %MDD values were clustered into five classes based of the EFSA 2013 and readapted by Brock et al., (2015) (Fig 5.8).

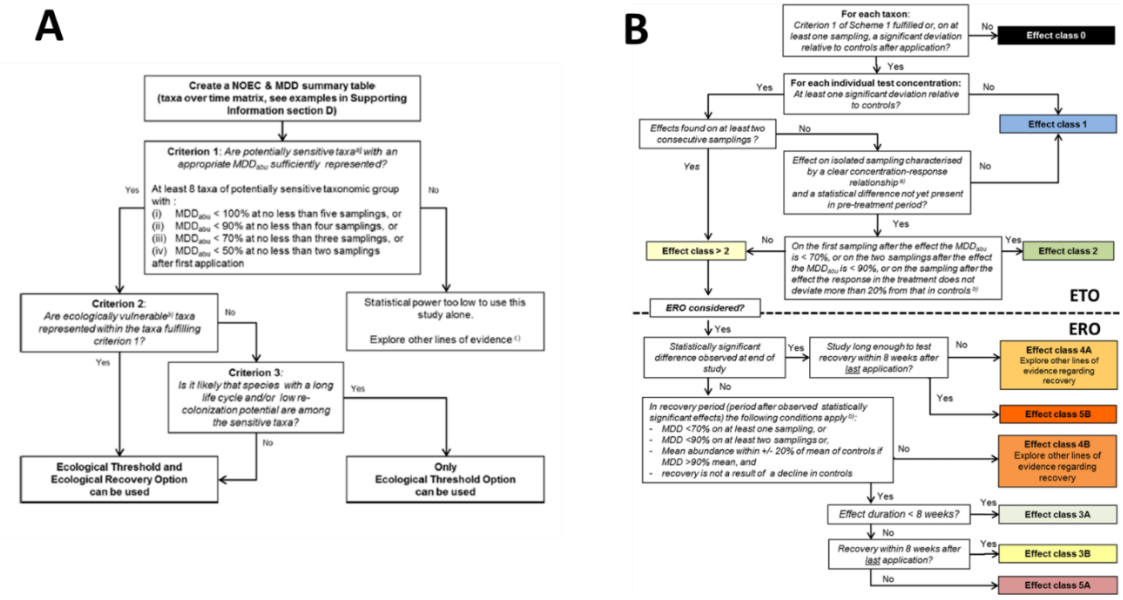


Figure 5. 8: A Decision scheme 1 to assess the reliability of a microcosm/mesocosm study to derive regulatory acceptable concentrations (RACs) based on treatment-related effects of pesticide exposure.

Informed by e.g. available single species and semi-field tests and other read-across information (a)). Ecologically vulnerable due to potential intrinsic sensitivity to the test item, likelihood of exposure, long life cycle (e.g. bi-, uni- or semi-voltine) and/or low immigration potential (b)). For example, focussed population level and

microcosm/mesocosm studies addressing additional sensitive species or population modelling (Taken from Brock et al., 2015).

5.8B: Decision scheme 2 for the derivation of effect classes for treatment-related effects (focus on treatment-related declines) on population abundance from results of microcosm/mesocosm studies. The MDDabu values mentioned in the decision scheme do not apply to indirect effects in the form of increases in population abundance if the NOECs of these treatment-related increases are associated with MDDabu values >100 % or if no MDDabu can be calculated due to the absence of the taxon in control test systems (n.c.). A clear concentration–response relationship for direct effects is characterized by a monotonous treatment-related decrease in abundance while in addition, the statistical difference coincides with a high enough mean abundance of the taxon in controls (a)). When selecting a certain minimum abundance for a taxon in controls, the argumentation for this should be provided. If a significant effect is observed in the application period, the next sampling should occur within a week. If the high %MDDabu in the post-effect period can be explained ecologically (e.g. emergence of insects) and a justification is given that this phenomenon will also occur under realistic field conditions, some flexibility of the MDD criterion is recommended (b)) (Taken from Brock et al., 2015).

As stated in Brock et al., (2015), category 1 encompasses taxa demonstrating sufficient statistical power to potentially exhibit treatment-related responses, including the establishment of a no adverse effect concentration. This category employs specific MDD criteria, such as thresholds at multiple sampling occasions post-initial test item application, to identify taxa exhibiting consistent treatment-related declines or statistically significant increases in population abundance. Examples provided illustrate species falling within this category. Category 2 comprises taxa failing to meet the stringent MDD criteria of Category 1 but displaying a Lowest Observed Effect Concentration (LOEC) at least on one sampling occasion. These taxa manifest statistically significant decreases or increases in population abundance under specific conditions, as delineated, and examples highlight treatment-related changes in population abundance. Category 3 encompasses taxa that neither meet the MDD criterion for Category 1 nor demonstrate significant differences with controls across samplings. Such taxa are excluded from the evaluation of treatment-related responses or effect class derivation. Examples include taxa exhibiting no significant differences compared to controls throughout the study. The statistical findings for each taxa group and sampling method are summarized into tables based on these categories, facilitating the assessment of treatment effects and study reliability in demonstrating such effects.

These categories, alongside the MDD classes, provide a structured framework for evaluating treatment effects on sampled taxa in aquatic experiments. The subsequent phase involves evaluating microcosm/mesocosm studies based on effect classes integrating the Minimum Detectable Difference (MDD) concept. In this adaptation, effect classes, as previously proposed by De Jong et al. (2008) and the EFSA Aquatic Guidance Document (EFSA 2013), are slightly modified to incorporate MDD requirements effectively. **Effect Class 0** denotes instances where treatment-related effects cannot be statistically evaluated. This class is consistently assigned to Category 3 taxa, indicating questionable reliability of the study if relevant endpoints exhibit this class consistently. For Category 1 and Category 2 taxa, the following effect classes apply: **Effect Class 1** denotes the absence of treatment-related effects, supported both statistically and ecologically. **Effect Class 2** signifies slight effects, typically observed at individual samplings and meeting specific criteria for recovery consideration. **Effect Class 3A** signifies pronounced short-term effects followed by recovery within 8 weeks, while **Effect Class 3B** indicates prolonged effects with recovery within 8 weeks post the last application. **Effect Class 4A** denotes significant short-term effects in a study too brief to demonstrate complete recovery within 8 weeks. **Effect Class 4B** indicates significant short-term effects with recovery evaluation hindered due to high MDD values. **Effect Class 5A** represents pronounced long-term effects followed by eventual recovery within the year of application, while **Effect Class 5B** signifies prolonged effects without demonstrated recovery before the termination of the experiment or winter onset.

Table 5. 2: Summary endpoints for zooplankton abundance in univariate analysis for morphological analysis.

Group	Taxa/group	Effect classes according to EFSA (2013) and Brock *et al* (2014), at given nominal concentrations (mg./L)			
		5.5	22	%MDD Cat.	NOEC
Cladocera	*Acroperus*	-	-	3	
	Alona/Alonella sp.	-	-	3	
	Chydorus sp.	1	1	1	≥ 22
	Chydoridae (grouped)	1	1	1	≥ 22

	Simocephalus sp.	-	-	2	
	Daphnia	1	2-4B	2	5.5
	Daphniidae (grouped)	1	2-4B	2	5.5
	Total Cladocera (grouped)	1	1	2	≥ 22
Copepoda	Cyclopoida	1	1	1	≥ 22
	Diaptomidae	2-4B	1	2	5.5
	Nauplii	-	-	3	
	Total Copepoda (grouped)	1	1	1	≥ 22
Rotifera	*Keratella* sp.	2-4B	1	2	5.5
	Total Brachionidae (grouped)	2-4B	1	2	5.5
	Lecane sp.	-	-	3	
	Colurella sp.	-	-	3	
	Lepadella sp.	-	-	3	
	Total Lepadellidae (grouped)	-	-	3	
	Cephalodella sp.	-	-	3	
	Mytilina sp.	-	-	3	
	Total Notommatidae (grouped)	-	-	3	
	Rotaria sp.	-	-	3	
	Trichocerca sp.	-	-	3	
	Trichotria sp.	-	-	3	
	Total Rotifera (grouped)	1	1	1	
Total No. Zooplankton		1	1	1	≥ 22
No. Taxa		1	1	1	≥ 22
No. Families		1	1	1	≥ 22
Overall Lowest NOEC$_{population}$ (Class 2)		<5.5			
Overall NOEAEC$_{population}$		X			

Effect classes (0, 1, 2, 2-4A, 2-4B, 3A, 3A-4A, 3A-4B, 5A and 5B) are based upon statistically significant NOECs occurring according to the principles outlined by (EFSA, 2013) and adapted by Brock *et al.* (2015).

Note on effect class 2: This is only assigned for effects on isolated occasions where characterised by a clear causal response relationship. Further, recovery from declines has to be reliably confirmed on the sampling occasion following the decrease in abundance to assign effect class 2.

Note on effect class 2-4A: If a delayed response is observed on the last sampling only, this is indicated as effect class 2-4A.

+ Increase in abundance in comparison to the control (indirect effect)

Abundance data was analysed, the log10 for all the y axis was used when graphs were made for all the genera, families and groups (*Cladocera, Copepoda* and *Rotifera*). This was done to better show the differences in abundance even at low abundance levels. Within the Cladocera group (Figure 5.9), five genera were identified: *Acroperus, Alona/Alonella, Chydorus, Daphnia* and *Simocephalus*. These were grouped into the two families *Chydoridae* and *Daphniidae*. Out of these five genera, *Acroperus* and *Alona* were classified as category 3 and thus not suitable for analysis. *Simocephalus* was categorised as category 2 as a significant difference was found between the control and both low and high doses. However, these differences were then not considered reliable as the %MDD was > 100%. Both the *Chydorus* genera and *Chydoridae* family were assigned a Category 1 and an effect class of one. This is because taxa were characterised by sufficient statistical power to demonstrate treatment-related responses (category 1). However, no statistically significant effects were observed as a result of treatment.

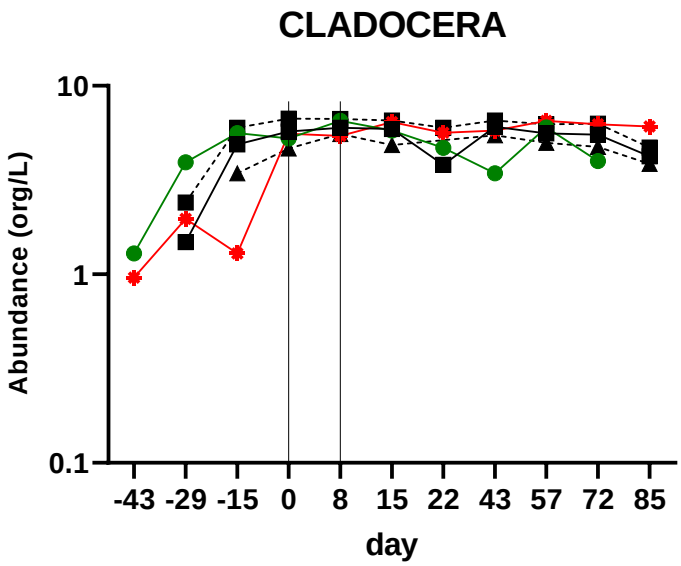


Figure 5. 9: Zooplankton abundance *Cladocera.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

Both the *Daphnia* genera (Figure 5.10) and the *Daphniidae* family (Figure 5.10) were assigned a category 2 and an effect class of 1 for the low-dose (5.5 mg a.i./L). However, for the high dose (22.0 mg a.i./L), an effect class of 2-4B was assigned as Significant short-term effects demonstrated but recovery could not be properly evaluated due to high %MDDabu values in recovery period.

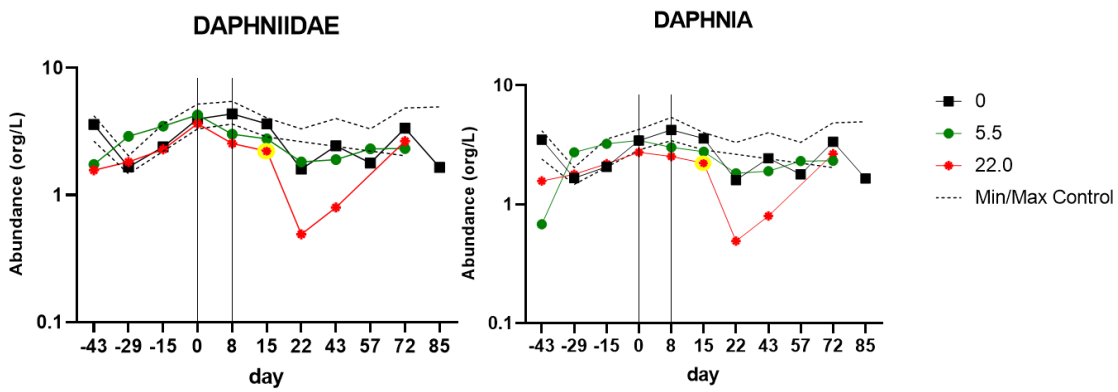


Figure 5. 10: Zooplankton abundance *Daphniidae* and *Daphnia*.

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

Significant differences were highlighted in yellow at the detected time.

Within the *Copepoda* group (Figure 5.11), three genera were identified: *Cyclopoida, Diaptomidae* and *Nauplia*. Out of these three genera, one of them (*Nauplia*), was classified as category 3 and thus not suitable for analysis.

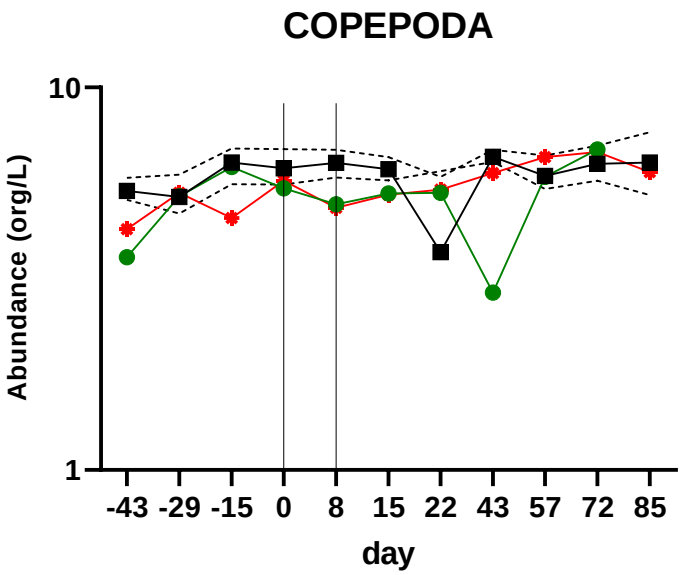


Figure 5. 11: Zooplankton abundance *Copepoda.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

Out of the remaining genera, *Cyclopoida* was assigned a category one and an effect class of one. This is because taxa were characterised by sufficient statistical power to potentially demonstrate treatment-related responses (category 1). The *Diaptomidae* genera (Figure 5.12) was assigned a category 2 as taxa did not meet the %MDDabu criterion for category one, an effect class of 2-4B as significant short-term effects demonstrated but recovery cannot be properly evaluated due to high %MDDabu values in recovery period and an effect class of 1 as no (statistically and/or ecologically significant) effects observed because of the high treatment.

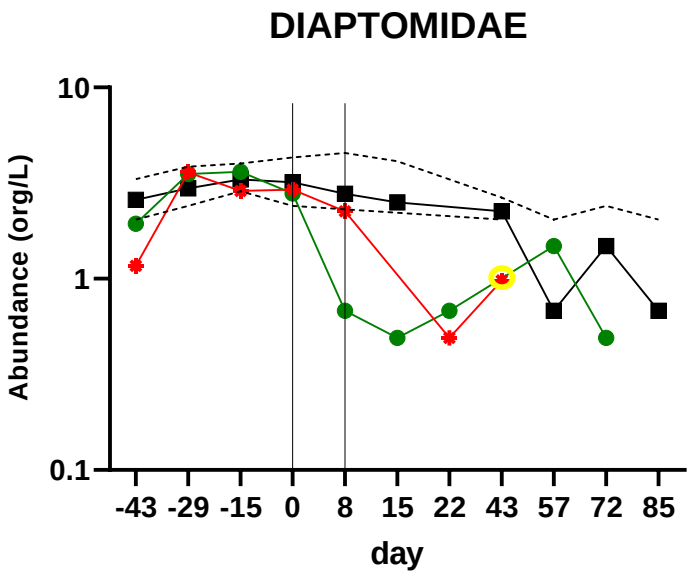


Figure 5. 12: Zooplankton abundance *Diaptomidae.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found. Significant differences were highlighted in yellow at the detected time.

Within the *Rotifera* group (Figure 5.13), nine genera were identified: *Keratella, Brachionus, Lecane, Colurella, Lepadella, Mytillina, Rotaria, Trichocerca* and *Trichotria*. These were grouped into the three families *Brachionidae, Lepadellidae* and *Notomatidae*. Out of these nine genera, eight of them were classified as category 3 and thus not suitable for analysis. Also, the *Lepadellidae* and *Notommatidae* family was assigned to category 3.

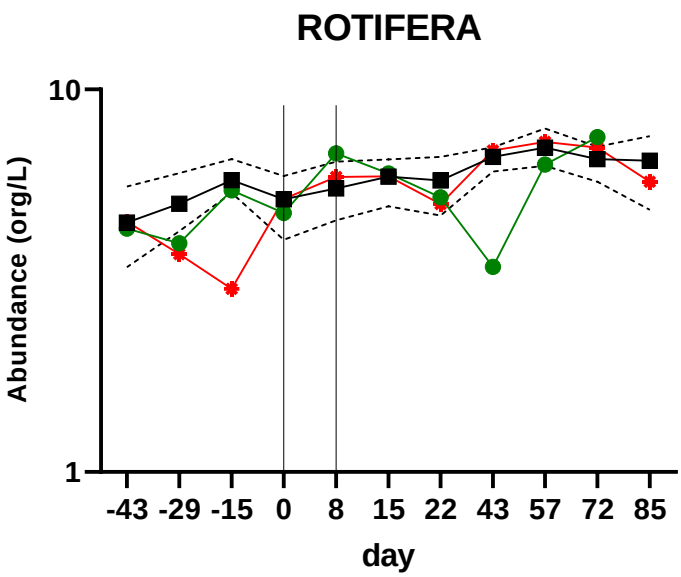


Figure 5. 13: Zooplankton abundance *Rotifera.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

Category 2 was assigned to the *Keratella* genera (Figure 5.14) and the *Brachionidae* family (Figure 5.14). This category comprises taxa characterised by statistically significant decreases in population abundance on samplings when the MDDabu values are <100 %.

As stated before, the *Keratella* genera was the only genera assigned to a different Effect Class. Before the treatment with the test item Glyphosate (Days -43, -29 and -15), no statistically significant differences were detected for the 5.5 mg a.i./L assigned mesocosms in comparison with the control. However, a statistically significant lower abundance was observed on Day -15 at 22.0 mg a.i./L compared to the control. Post-application, a significant decrease in the 5.5 mg a.i./L was observed on Day 43, however the MDD% on the following day was >100% and thus not reliable, *Keratella* was assigned an Effect class of 4B for 5.5 mg a.i./L and 1 for 22.0 mg a.i./L.

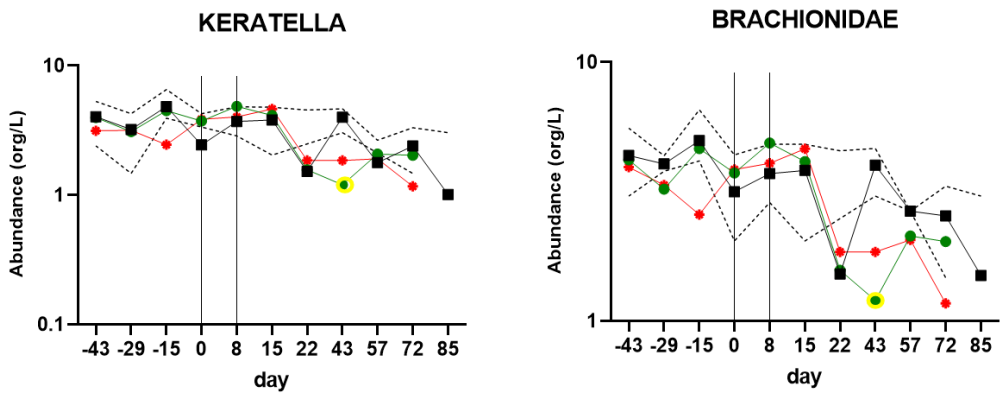


Figure 5. 14: Zooplankton abundance *Keratella* and *Brachionidae.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found. Significant differences were highlighted in yellow at the detected time.

The %MDD for total zooplankton abundance (Figure 5.15) ranged from 19.98 to 81.54, following the application of the test item. This indicated that a significant effect could be detected from small differences between the controls and treatment concentrations throughout the study. Therefore, the data for this taxon was Category 1. Before the initial treatment with the test item (Days -43 and -29), no statistically significant differences were detected for the assigned 5.5 mg a.i./L and 22.0 mg a.i./L mesocosms in comparison with the control mesocosms. However, an increase in abundance was observed in the control between Day -15 and 0 and statistically significant difference was found at 22 mg a.i./L. Following the test item application (Day 0) until the end of the study (Day 85), no treatment-related statistically significant effects or visual trends were observed at 5.5 and 22.0 mg a.i./L in comparison to the control and therefore an Effect Class 1 was assigned for these treatments.

Based on the findings described above the overall NOEC and NOEAEC (community and population) of this mesocosm study for Glyphosate is recommended to be less than 5.5 mg a.i./L.

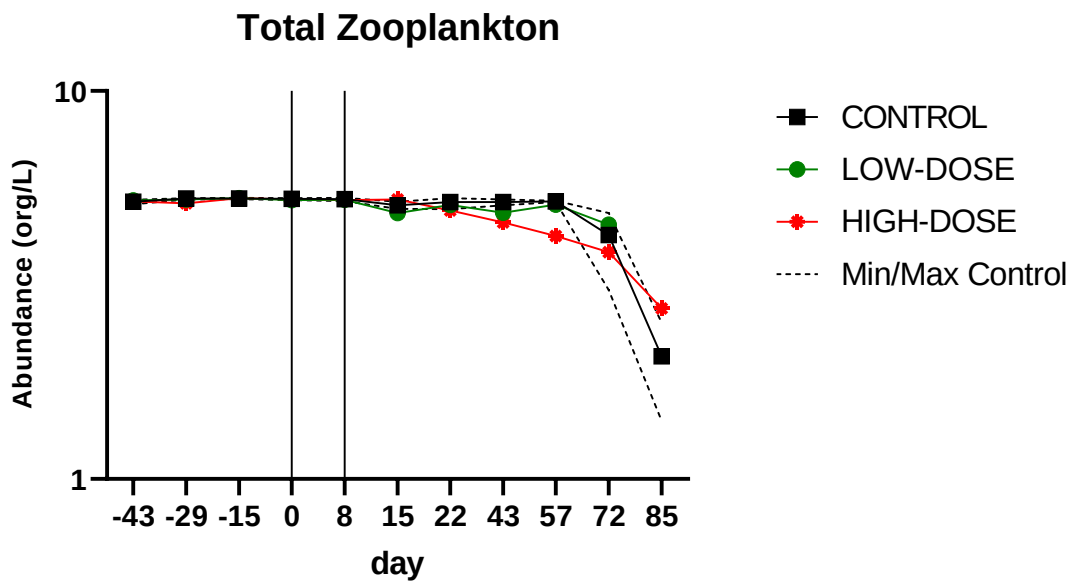


Figure 5. 15: Total Zooplankton abundance.

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

5.2.4. Generation of an updated database for zooplankton 18s region sequences

A total of 1,596 zooplankton sequences were retrieved from GenBank to analyse the zooplankton samples collected in 2022 based on the genera found in the list provided by CEA including the positive control (*Tigriopus Californicus*). The predominant zooplankton sequences are those of the *Rotifera* group (67.92%) followed by the *Copepoda* group (16.77%), and finally, the smallest portion was found for the *Cladocera* group (15.31%) (Figure 5.16).

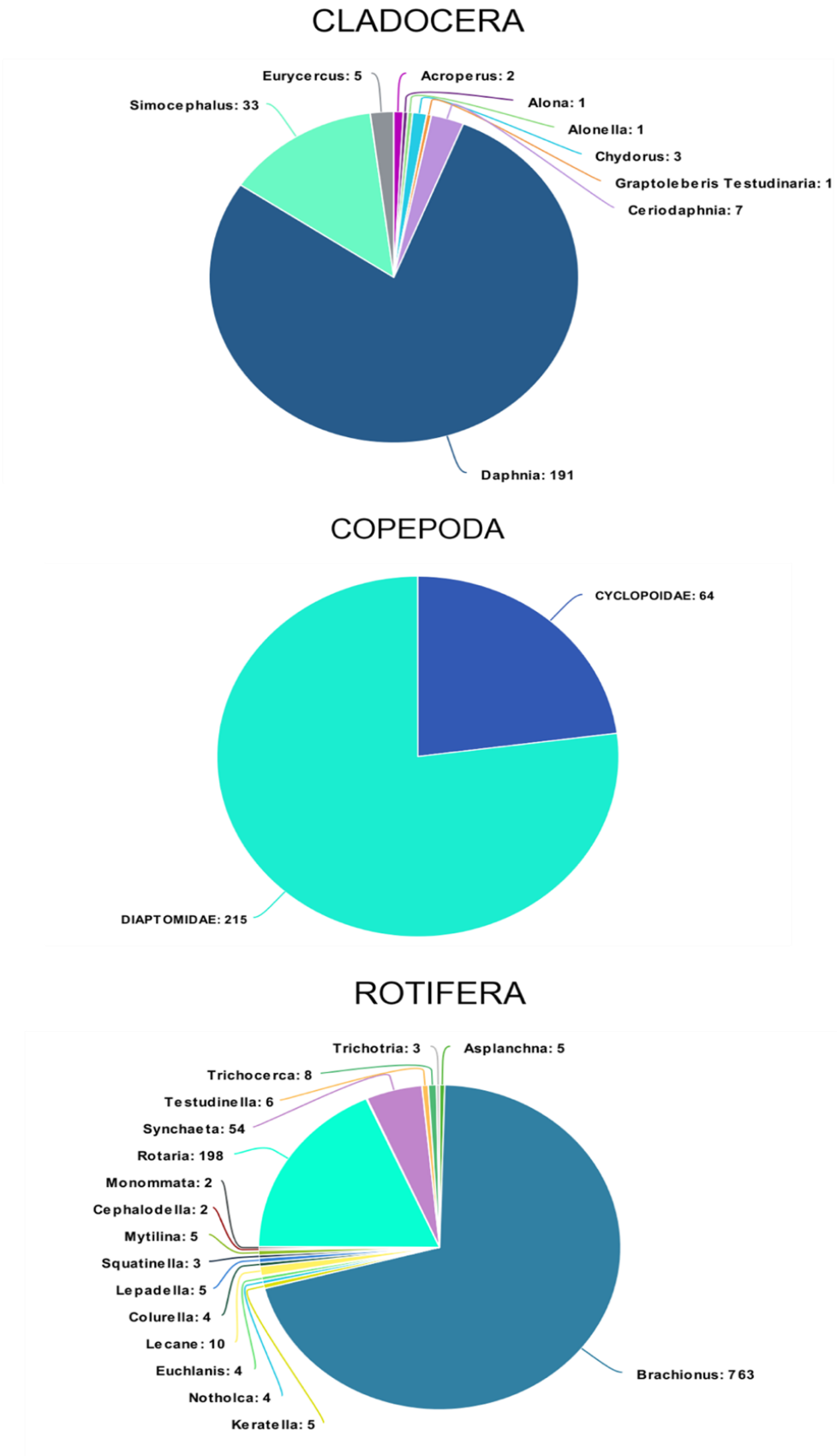


Figure 5. 16: Sequences availability in NCBI GenBank for the zooplankton genera present in the list provided by CEA for the 18s gene.

5.2.4.1 18s gene fragment sequencing of mesocosm samples

A total of 117 samples were collected and analysed. Out of these, six did not result in successful PCR amplification so were discarded: Sample 9 (PO29) from post-application Day 2, samples 2, 5, 7, 8 (PO42, 5, 7, 8 respectively) from post-application Day 4, and sample 4 (PO54) from post-application day 5. (Table 5.3). In addition, due to cost constraints only 96 samples, including positive and negative controls, could be sent to DeepSeq for Illumina sequencing therefore further samples were discarded as follows: the remainder of post-application days 4 and 9; and the three low doses of post-application day 8 (Table 5.4). It was decided to exclude all post-application day 4 as several samples did not amplify by PCR. All post-application day 9 were excluded as this was the latest time point in the season and thus the one with a higher probability to be the least informative (assuming mesocosm recovery had already occurred by this point). Finally, the low doses of post-application Day 8 were excluded as the most information were likely to be provided by the control and high-dose mesocosm comparisons.

Table 5. 3: Total samples obtained after PCR where samples shaded in black represent unsuccessful PCR amplification.

ID	ALL SAMPLES OBTAINED INCLUDING NON-WORKING SAMPLES								
	CONTROL			LOW DOSE			HIGH DOSE		
	1	2	3	4	5	6	7	8	9
PRE-APP1									
PRE-APP2									
PRE-APP3									
APP-DAY									
POST-APP1									
POST-APP2									
POST-APP3									
POST-APP4									
POST-APP5									
POST-APP6									
POST-APP7									
POST-APP8									

POST-APP9									

Table 5. 4: Excluded samples after PCR.

where samples coloured in black represent unsuccessful PCR amplification and samples coloured in red represent the other samples excluded from further analysis.

ID	EXCLUDED SAMPLES								
	CONTROL			LOW DOSE			HIGH DOSE		
	1	2	3	4	5	6	7	8	9
PRE-APP1									
PRE-APP2									
PRE-APP3									
APP-DAY									
POST-APP1									
POST-APP2									
POST-APP3									
POST-APP4									
POST-APP5									
POST-APP6									
POST-APP7									
POST-APP8									
POST-APP9									

The DeepSeq results from Illumina sequencing were analysed with the previously used script (Chapter 4, Section 4.8.1).

To achieve this, high-throughput sequencing data underwent several stages of bioinformatics processing, each critical in refining the dataset to accurately assign taxonomy. The following table (Table 5.5) summarises the sequences at each stage, detailing the number of sequences retained or lost, and ultimately, how many sequences were successfully assigned taxonomy.

Table 5. 5: Summary of Sequence Retention and Taxonomic Assignment at Each Stage of the Bioinformatics Workflow.

Software	Read Count	Description
RAW DATA	20.6 M	Represents the initial sequencing output from the Illumina platform, containing all raw reads generated before any processing or filtering.
FLASH	12.15 M	FLASH (Fast Length Adjustment of SHort reads) merges paired-end reads that overlap, producing longer, more accurate reads for downstream analysis.
TRIMMOMATIC	12,57 M	Involves removing low-quality bases and adapter sequences from the raw reads, ensuring that only high-quality, clean sequences are retained.
CUTADAPT	11,56 M	A tool that removes adapter sequences, primers, and other unwanted sequences from reads, refining them for more accurate alignment and analysis.
SEQTK	11,56 M	A lightweight tool used for processing sequences in FASTA or FASTQ format, typically for tasks like filtering, trimming, and subsampling.
Assigned taxonomy by BLAST	651.179	BLAST (Basic Local Alignment Search Tool) assigns taxonomy to sequences by comparing them against known sequences in a database, identifying species present.

However, another script (BANZAI) was tested to compare results. A summary of the script pipeline can be found in Figure 5.17. This script was designed for analysing Illumina-generated environmental DNA sequence data (O'Donnell et al., 2016).

The results of the BANZAI script were found to be inconclusive. The script underwent extensive modifications and updates to address issues as the original script was no longer compatible with the current NGS data, likely due to the use of outdated and deprecated packages or libraries. Outdated packages are software components that have been superceeded by newer versions, and deprecated packages are those that are no longer recommended for use and might not receive updates or support. As a result, attempts to run the script led to conflicts and unreliable results.

To rectify these issues and ensure the script's functionality in a contemporary context, substantial modifications are necessary. This involves updating or replacing deprecated packages with more current alternatives, resolving conflicts between different versions, and adapting the code to meet the requirements of the current software ecosystem which is outside the scope of this study. Therefore, this script was not used further.

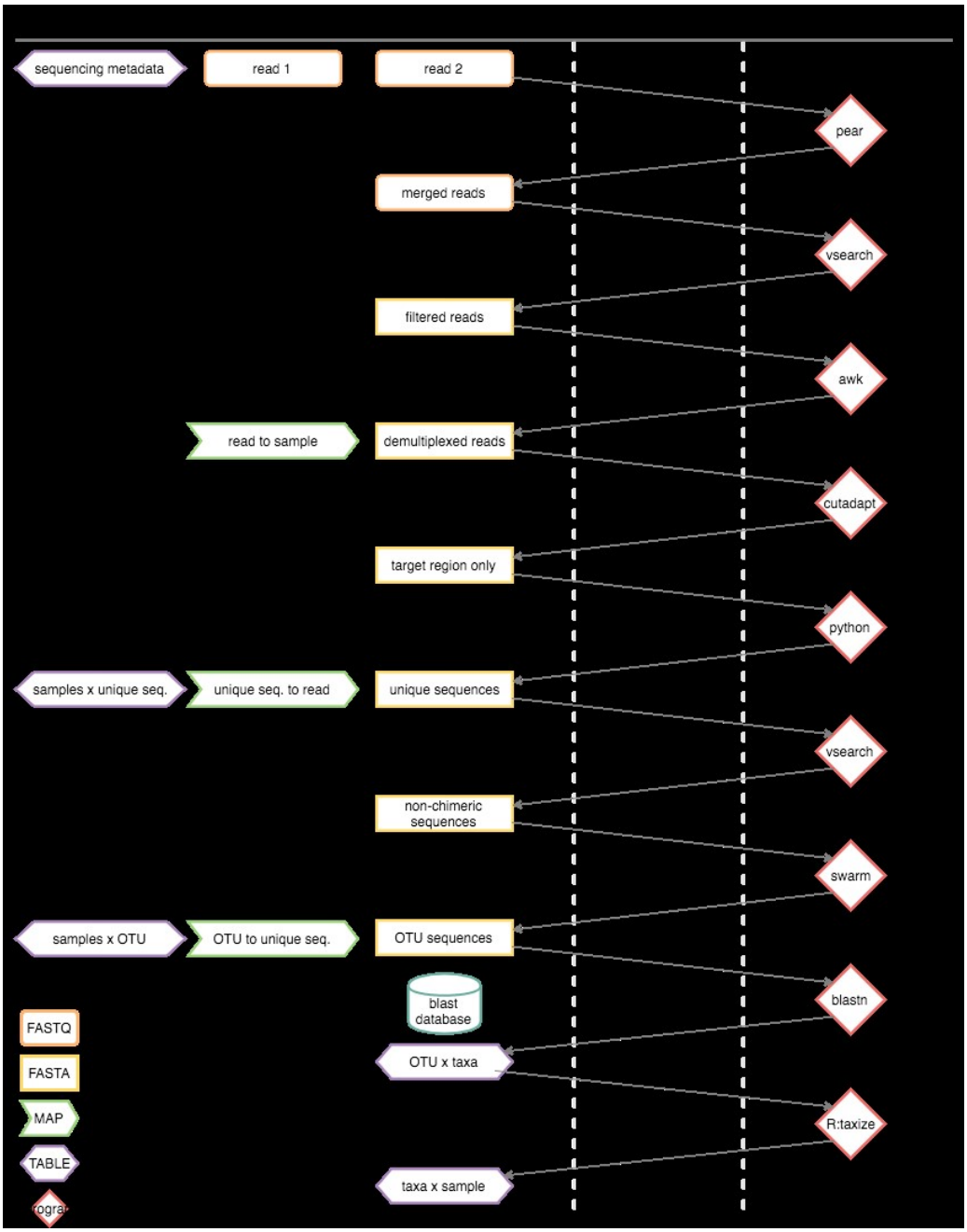


Figure 5. 17: Summary of the BANZAI pipeline taken from (O'Donnell et al., 2016).

5.2.4.2 Impacts in community composition of environmental variables when analysing zooplankton by metabarcoding.

For the metabarcoding analysis, PCA was utilised to assess the influence of Glyphosate and environmental doses on community dynamics (Figure 5.18). Results showed an adjusted explain variation of 5.83% (S10, Appendix). As this is a linear model unlike the DCA used to analyse morphological data, the genera are shown as arrows instead

of points. The closer the arrows, the more similar the genera behaved in the community. Exactly like the DCA, the environmental variables are shown as factors and the longer is the arrow, the more impact they have on the community. The control and the two doses were found very distant from each other indicating a difference between them. High dose and day are going in the same direction showing that the high dose effect is increasing with time. The same pattern was seen as for morphology analysis where Day, pH and conductivity are going in different directions. Again, temperature, oxygen and turbidity were located between pH and Day for the same reasons previously explained in the DCA analysis. The *Notholca, Keratella* and *Daphnia* genera located opposite to the day factor indicating a higher presence at the beginning of the study. On the other hand, *Lecane, Lepadella* and *Keratella* located more towards the right side of the graph, and so were found more towards the end of the study.

Diaptomidae, Cyclopoidae and *Chydorus* results showed they have a very strong impact on the community as their arrows were longer than the other genera. *Diaptomidae* and *Copepoda* are both Copepods, and they can be in competition with each other for food so if *Diaptomidae* dies in the high dose, there is a higher chance to find *Cyclopoidae* in that dose.

In summary, the PCA analysis provided intricate insights into the complex interplay between environmental variables and species distribution, offering valuable scientific insights into community dynamics over the study duration. These insights need further exploring by using the univariate analysis to inspect the dose effects on each zooplankton genera, family, and group.

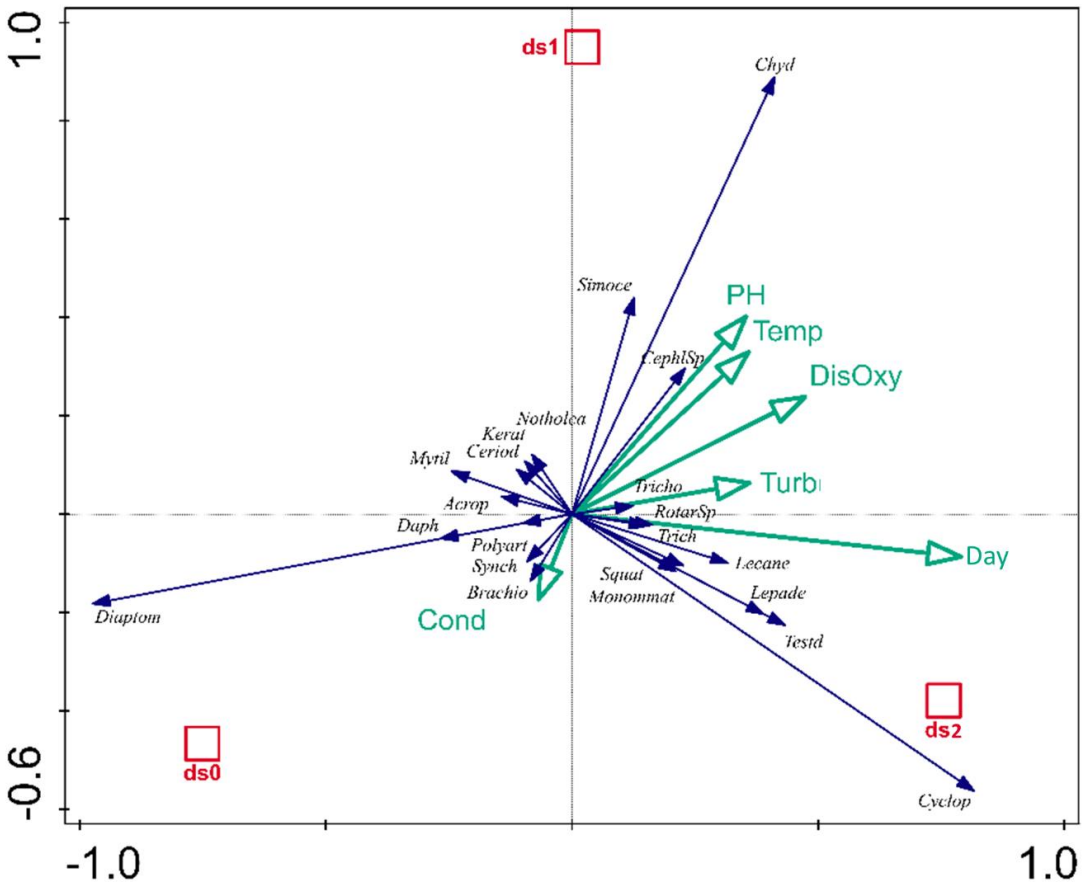


Figure 5. 18: PCA ordination of the species abundance and environmental variables.

Where variables are represented by green arrows, species are depicted by blue arrows and the three Glyphosate doses are represented in red squares where ds0 = control, ds1= low dose (5.5 mg a.i./L) and ds2= high dose (22.0 mg a.i./L).

5.2.4.3. Data analysis for Glyphosate effects on community composition in molecular analysis

The following analyses were performed to assess the effect of Glyphosate on a community composition level on metabarcoding data. This is the same analysis utilised to assess the effects of Glyphosate doses on community morphology in the previous section.

The plotted lines in Figure 5.19A represented the control (black line), low-dose (blue line, 5.5 mg a.i./L), and high-dose (red line, 22.0 mg a.i./L) treatments. Prior to application, the community composition within the mesocosm, assigned to the respective doses, exhibited notable divergence. After application, however, the

community compositions for both doses, when compared to the control, converged, displaying remarkably similar patterns. Notably, a significant dose-response trend was found for both doses (pseudo-F=1.2, P=0.004).

The species scores, as detailed in Table S11 in the appendix, correspond to the numbers on the axes in Figure 5.19B. Once more, in this analysis, *Diaptomidae*, *Cyclopoidae*, and *Chydorus* occupied the extremities, indicating their prominent influence on community dynamics, while other genera clustered together towards the centre point.

In essence, the community was found significantly impacted by the Glyphosate doses. While the PRC analysis provides valuable insights into community dynamics in response to Glyphosate doses, a comprehensive understanding necessitates careful consideration of the nuanced interplay between taxa and doses, ideally supplemented by univariate analyses to elucidate complex relationships.

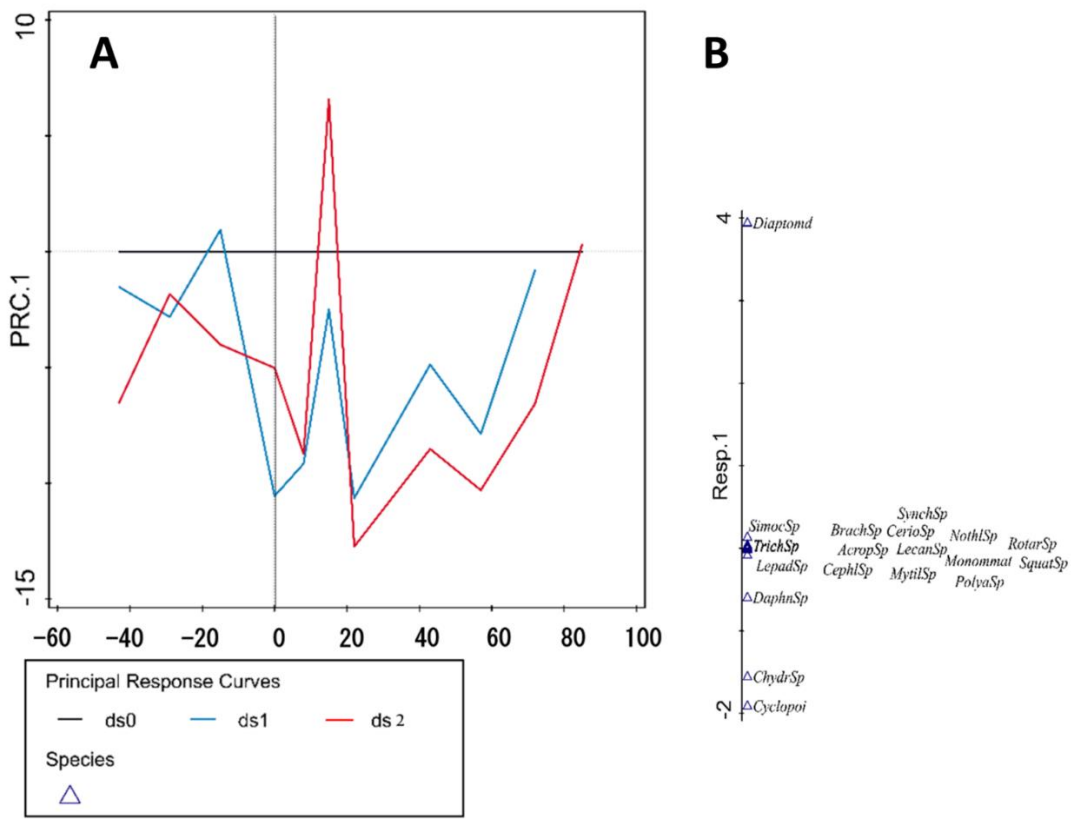


Figure 5. 19:A Principal Response Curves for metabarcoding data to assess the effect of treatment over time (pseudo-F=1.2, P=0.004).

Where ds0 = control, ds1= low dose (5.5 mg a.i./L) and ds2= high dose (22.0 mg a.i./L). The first axis displays the highest fraction of variation that can be explained by the explanatory variables. The mean difference of site scores between treatment and control on first axis displays the deviation of communities from control.

Figure 5.19B: Species scores on first axis responsible for PRC.

A NMDS plot showed low-stress values for all concentrations indicating a reliable representation of the dissimilarities in the reduced-dimensional space (Figure 5.20A). Each sampling point is positioned in the two-dimensional space according to its dissimilarity to others. Points closer in the configuration are more ecologically similar, whereas those farther apart are more dissimilar. Stress level (control: 0.05, low-dose=0.10 and high dose=0.07) measure that reflects how well the dissimilarities among the points in the reduced-dimensional space. The NMDS plot for the control displayed the ecological dissimilarities among samples based on their community compositions. The stress value of 0.04 suggested that the NMDS configuration accurately represents the original dissimilarities in the ecological data. Axis 1, explained 87% of the variation,

highlighting the primary gradient in the community composition, while Axis 2, contributed 6% of the variation. Outlier datapoint 36 stands out as further apart from the others on Axis 1.

The NMDS analysis for the low dose was conducted on ecological community data, generating two-dimensional scores for each sample along Axis 1 (explaining 64% of the variation) and Axis 2 (contributing 12% of the variation). The stress value was low at 0.10, suggesting a reliable representation of ecological dissimilarities. Along Axis 1, the 11 time points (weeks 18, 20, 22, 24, 25, 26, 27, 30, 32, 34 and 36) exhibited a gradient from negative to positive values, signifying a directional shift in community composition. Axis 2 further contributed to the separation of samples, emphasizing additional ecological distinctions. Outlier datapoint 36 stands out as it is positioned farther from the others on both axes.

The NMDS analysis for the high dose generated two-dimensional scores for each sample, allowing for the visualization of ecological dissimilarities along Axis 1 and Axis 2. The stress value of 0.07 indicates a good fit of the ordination, suggesting that the resulting NMDS plot reliably represents the community structure. Axis 1 explaining 91% of the variation, exhibited a directional shift from negative to positive values, indicating a gradient in community composition. Axis 2, with a value close to zero, suggested less influence on the separation of samples. Sampling dates in control appeared more scattered randomly compared to the low and high dose of Glyphosate where the sampling dates appeared to be more cluster within pre- and post-application.

NMDS was used as a simple method of visual interpretations to compare the overall structure of zooplankton community at the different sampling points in the three Glyphosate doses. This grouping is reinforced by the hierarchical clustering tree based on the Bray Curtis similarity index, indicating some differences in clustering between pre- and post-application as shown in the two different clustering in high Glyphosate dose (Figure 5.20B). However, In the second cluster where all the pre-application samples are located, a post-application sampling day (week 26) can be found. The ANOSIM tests the significant differences in these groups. For the control and the two Glyphosate doses, the ANOSIM results were similar where a significant difference was found in all three of them ($R=0.42$, $p = < 0.01$, $R=0.51$, $p= < 0.01$ and $R=0.66$, $p= <$

0.01 respectively) (Figure 5.20A). This illustrated how the sampling points varied between each other throughout the season.

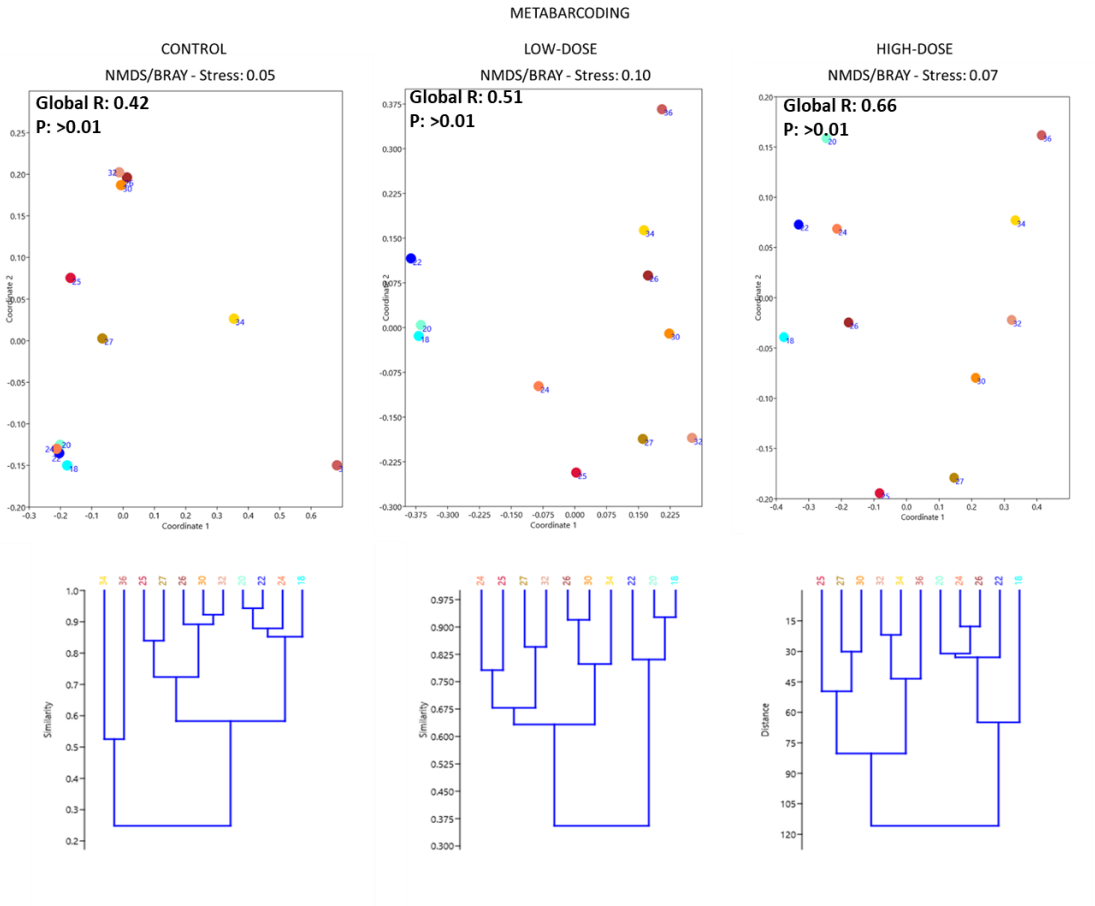


Figure 5. 20: (A) NMDS analysis based on Bray Curtis dissimilarity matrix.

The coordinates along Axis 1 and Axis 2 represent the ecological distances between the samples. Each sampling point is positioned in the two-dimensional space according to its dissimilarity to others. Points closer in the configuration are more ecologically similar, whereas those farther apart are more dissimilar. Stress level (Control: 0.05, low-dose=0.10 and high dose=0.07) measure that reflects how well the dissimilarities among the points in the reduced-dimensional space. ANOSIM resulted in R=0.42, p = < 0.01 for control, R=0.51, p= < 0.01 for the low Glyphosate dose and R=0.66, p= < 0.01 for the high Glyphosate dose, y)

5.20B: Hierarchical clustering tree depicting the relationships among sampled points using the Bray-Curtis similarity index. The tree illustrates the hierarchical arrangement of ecological similarities and dissimilarities among the samples. Points that cluster together share a higher ecological similarity, while those in separate branches exhibit greater dissimilarity. The Bray-Curtis similarity index serves as a measure of community composition, providing insights into the grouping of samples based on their species composition and abundance patterns. The low-dose (week 36) is not present as it was excluded from analysis.

The Shannon diversity index (H) and Chao-1 showed similarity in control, Glyphosate low-Dose (5.5 mg a.i./L), and high-Dose (22 mg a.i./L) treatments (Figure 5.21). The average (H) for metabarcoding (control) pre-application was 0.71, and post-application was 1.24; for metabarcoding (low-dose for Glyphosate) pre-application was 1.02 and post-application was 1.35; and finally, for metabarcoding (high-dose for Glyphosate) pre-application was 1.07 and post-application was 1.04. When the three treatments were analysed using RM ANOVA, no significant differences were observed in both pre- and post-application.

The Chao-1 index assumes that the number of organisms identified for a taxon has a Poisson distribution and corrects for variance. It is useful for data sets skewed toward low-abundance. The average Chao-1 for morphology (control) pre-application was 16.75, and post-application was 17; for morphology (low-dose for Glyphosate) pre-application was 17 and post-application was 17.33; and finally, for morphology (high-dose for Glyphosate) pre-application was 15.5 and post-application was 15.85. When the three treatments were analysed using RM ANOVA, no significant differences were observed in both pre- and post-application.

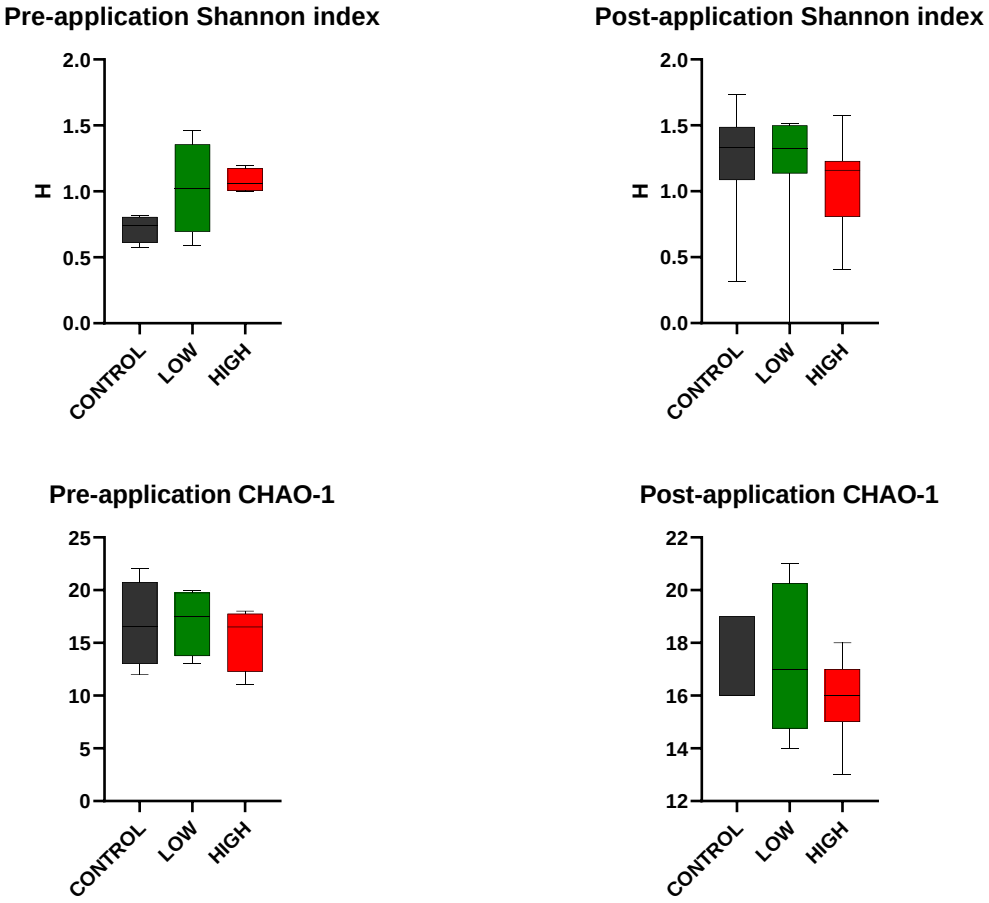


Figure 5. 21: Comparison of Shannon Diversity Index and Chao-1 Index in Response to Control, Low-Dose, and High-Dose Treatments pre- and post-application.

5.2.4.4 Molecular Analysis of Glyphosate Impact on Zooplankton

In the previous section, the analyses were performed to assess the effect of Glyphosate on a community composition level. The following analyses were performed to assess the impact of Glyphosate on each genera group and family of zooplankton.

In the study, 28 discrete zooplankton taxa were identified. All identified zooplankton taxa were in a higher taxonomic group; *Cladocera*, *Copepoda* or *Rotifera*. The zooplankton taxa, five grouped families (e.g. *Daphniidae*) and the three groups, *Cladocera*, *Copepoda* and *Rotifera*, were used in the univariate analysis, including determining %MDD (Table 5.6) and the NOEC using the same criteria as 5.2.3.1.1. Taxa that were not possible to analyses because they were too sporadic or absent were excluded from the table.

Table 5. 6: Summary endpoints for zooplankton abundance in univariate analysis for metabarcoding analysis.

Group	Taxa/group	Effect classes according to EFSA (2013) and Brock et al (2014), at given nominal concentrations (mg./L)			
		5.5	22	%MDD Cat.	NOEC
Cladocera	*Alona/Alonella*	-	-	3	
	Acroperus	-	-	3	
	Chydorus	1	1	1	≥ 22
	Chydoridae (grouped)	1	1	1	≥ 22
	Simocephalus	1	2-4B	2	5.5
	Daphnia	-	-	3	
	Ceriodaphnia	-	-	3	
	Daphniidae (grouped)	1	2-4B	2	5.5
	Total Cladocera (grouped)	1	1	1	≥ 22
Copepoda	Cyclopoida	2	2	1	<5.5
	Diaptomidae	2-4B	2-4B	1	<5.5
	Total Copepoda (grouped)	1	1	1	≥ 22
Rotifera	*Keratella*	-	-	3	
	Brachionus	-	-	2	
	Total Brachionidae (grouped)	-	-	3	
	Notholca	-	-	3	
	Lecane .	-	-	3	
	Colurella	-	-	3	
	Lepadella	-	-	2	
	Total Lepadellidae (grouped)	-	-	2	
	Squatinella	-	-	3	
	Cephalodella	-	-	3	
	Mytilina	-	-	2	
	Total Notommatidae (grouped)	-	-	3	
	Rotaria	-	-	2	
	Polyarthra	-	-	3	
	Synchaeta	-	-	3	
	Trichocerca	-	-	3	
	Trichotria	-	-	3	
	Total Rotifera (grouped)	-	-	3	

Total No. Zooplankton	2	1	1	<5.5
No. Taxa	2	1	1	<5.5
No. Families	1	1	1	≥ 22
Overall Lowest NOEC$_{population}$ (Class 2)	<5.5			
Overall NOEAEC$_{population}$	X			

Effect classes (0, 1, 2, 2-4A, 2-4B, 3A, 3A-4A, 3A-4B, 5A and 5B) are based upon statistically significant NOECs occurring according to the principles outlined by (EFSA, 2013) and adapted by Brock *et al.* (2015).

Note on effect class 2: This is only assigned for effects on isolated occasions where characterised by a clear causal response relationship. Further, recovery from declines has to be reliably confirmed on the sampling occasion following the decrease in abundance to assign effect class 2.

Note on effect class 2-4A: If a delayed response is observed on the last sampling only, this is indicated as effect class 2-4A.

+ Increase in abundance in comparison to the control (indirect effect)

Within the Cladocera group (Figure 5.22), seven genera were identified: *Acroperus, Alona/Alonella, Chydorus, Daphnia, Ceriodaphnia* and *Simocephalus*. These were grouped into two families, *Chydoridae* and *Daphniidae*. Out of these seven genera, four of them were classified as category 3 and thus not suitable for analysis.

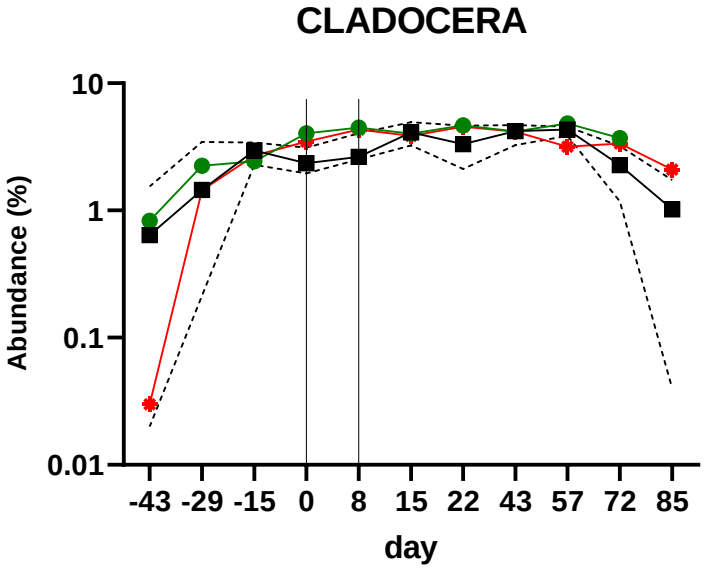


Figure 5. 22: Zooplankton abundance *Cladocera.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

Both the *Chydorus* genera and the *Chydoridae* family were assigned a category of 1 and an effect class of one. This is because taxa were characterised by sufficient statistical power to potentially demonstrate treatment-related responses (category 1). However, no statistically significant effects were observed because of treatment.

The *Simocephalus* (Figure 5.23) genera and the *Daphniidae* family (Figure 5.24) were the only Cladoceran assigned to a group 2 and an effect class of 2-4B as significant short-term effects demonstrated but recovery cannot be properly evaluated due to high %MDDabu values in recovery period.

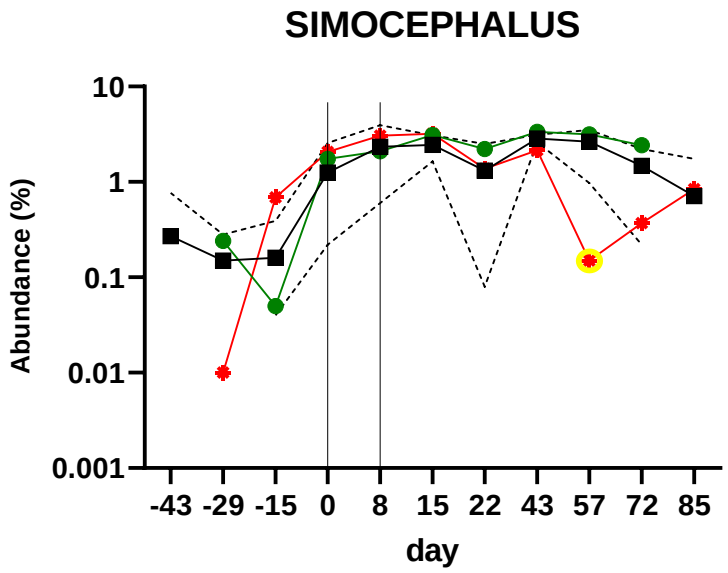


Figure 5. 23: Zooplankton abundance *Simocephalus.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found. Significant differences were highlighted in yellow at the detected time.

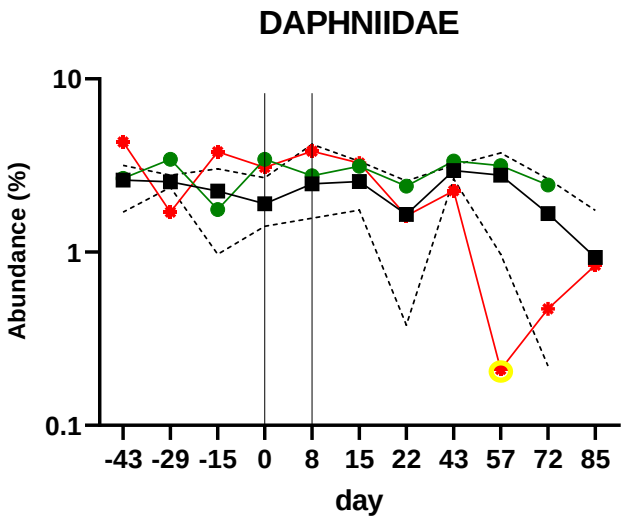


Figure 5. 24: Zooplankton abundance *Daphniidae.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found. Significant differences were highlighted in yellow at the detected time.

Within the *Copepoda* group (Figure 5.25), two genera were identified: *Cyclopoidae* and *Diaptomidae*

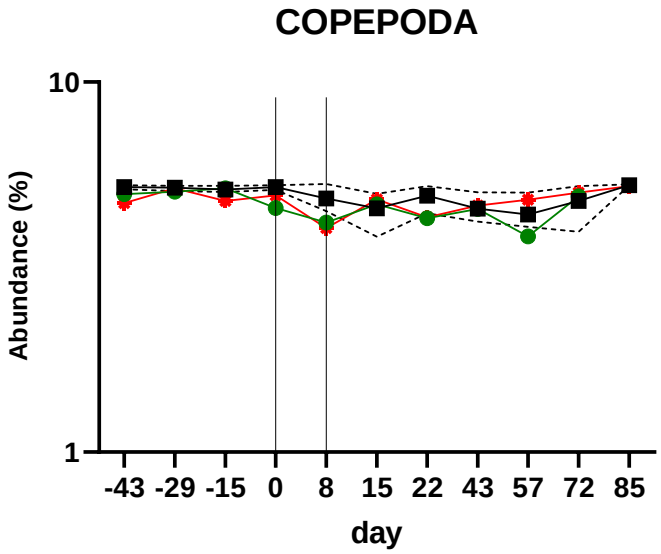


Figure 5. 25: Zooplankton abundance *Copepoda.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

Both these genera were assigned a category of 1 as Taxa characterised by a sufficient statistical power to potentially demonstrate treatment-related responses and consequently also a no adverse effect concentration. The *Cyclopoidae* genera (Figure 5.26), was assigned an effect class of two at both doses as statistically significant effects concern short-term and/or quantitatively restricted responses usually observed at individual samplings only.

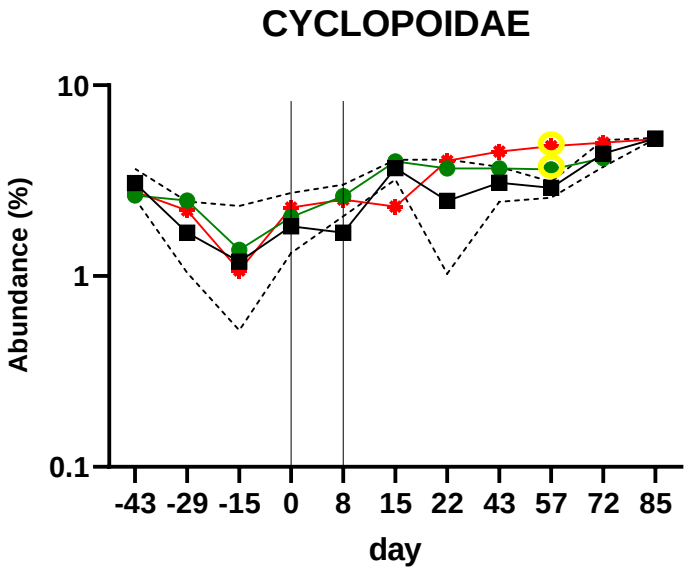


Figure 5. 26: Zooplankton abundance *Cyclopoidae.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found. Significant differences were highlighted in yellow at the detected time.

The *Diaptomidae* genera (Figure 5.27), was assigned an effect class of 2-4B at both doses as significant short-term effects demonstrated, but recovery cannot be properly evaluated due to high %MDDabu values in the recovery period.

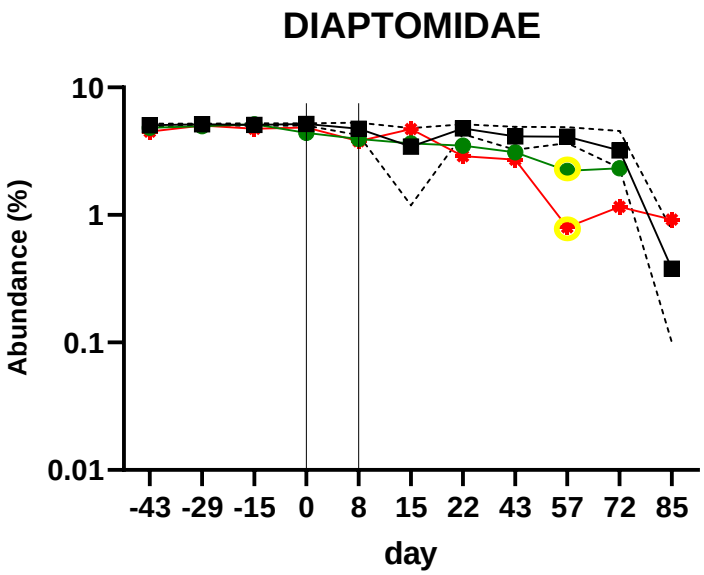


Figure 5. 27: Zooplankton abundance *Diaptomidae*.

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found. Significant differences were highlighted in yellow at the detected time.

Within the *Rotifera* group (Figure 5.28), fourteen genera were identified: *Keratella, Brachionus, Notholca, Lecane, Colurella, Lepadella, Squatinella, Cephalodella, Mytillina, Rotaria, Polyarthra, Synchaeta, Trichocerca* and *Trichotria*. These were grouped into the three families *Brachionidae* and *Lepadellidae* and *Notommatidae*. All genera and families except *Brachionus, Lepadella, Lepadellidae, Mytillina* and *Rotaria* were classified as category 3 and thus not suitable for analysis. The genera previously stated that were not classified as Category 3, were assigned a Category 2 as a significant difference was found between the control and or two doses. However, the %MDD was too high to be determined reliable.

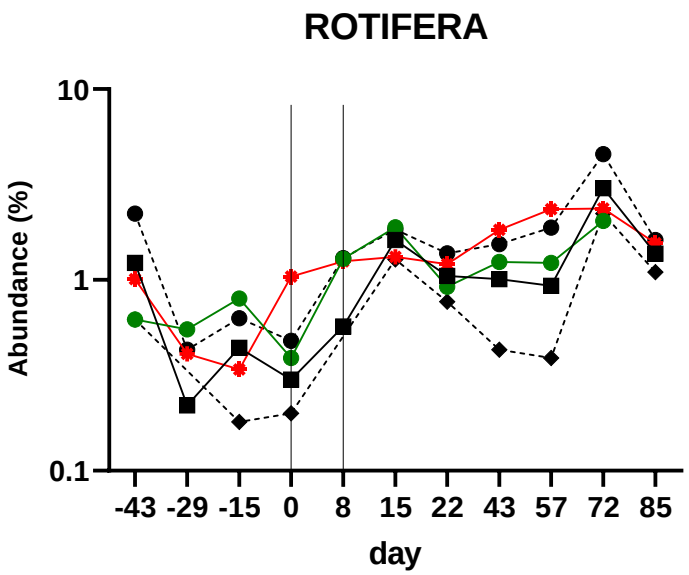


Figure 5. 28: Zooplankton abundance *Rotifera.*

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

The %MDD for total zooplankton abundance (Figure 5.29) ranged from 12 to 34, following the application of the test item. This indicated that a significant effect could potentially be detected from small differences between the controls and treatment concentrations throughout the study. Therefore, the data for this taxon were considered to be Category 1. Prior to the initial treatment with the test item (Days -43 and -29), no statistically significant differences were detected for the 5.5 mg a.i./L and at 22.0 mg a.i./L assigned mesocosms in comparison with the control. Following the test item application (Day 0) until the end of the study (Day 85), statistically significant effects or visual trends were observed at 5.5 but not at 22.0 mg a.i./L in comparison to the control and therefore an Effect Class 1 was assigned for these treatments.

Based on the findings described above the overall NOEC and NOEAEC (community and population) of this mesocosm study for Glyphosate is recommended to be less than 5.5 mg a.i./L.

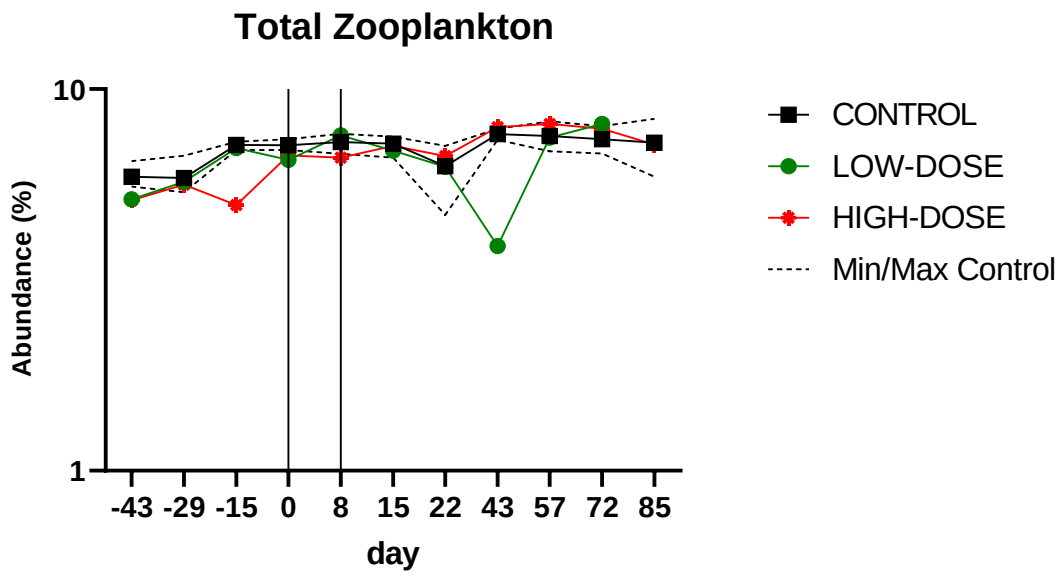


Figure 5. 29: Total zooplankton abundance.

Where control = 0 mg a.i/L (black), 5.5 (mg a.i./L) = low-dose (green), 22.0 (mg a.i./L) = high-dose (red) and finally the black dotted line is the range score of the control where min is the lowest value found and max is the highest value found.

5.2.5 Comparative Analysis of Zooplankton Taxonomy: Morphological Assessment vs. Metabarcoding Approach

5.2.5.1 Morphology and Metabarcoding Analysis Comparison

While the morphological analysis presented a more balanced distribution of genera, the results from the percentage abundance bar charts revealed that metabarcoding identified a higher number of genera (28) compared to morphological analysis (23) (Figure 5.30). A total of 22 genera were identified by both methodologies with metabarcoding identifying an additional 6 genera and morphological analysis identifying a single additional genus (Figure 5.30).

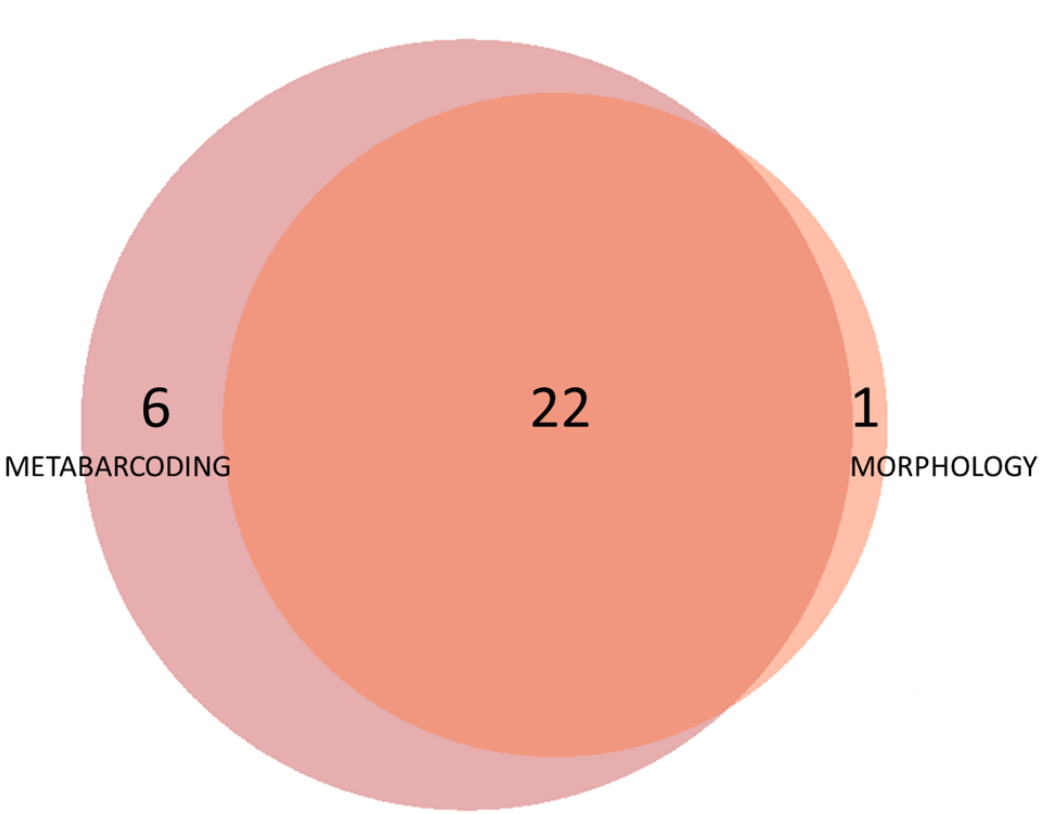


Figure 5. 30: Comparison of genera detected by both methodologies.

A total of 22 genera were identified by both methodologies (75.86%) with metabarcoding identifying an additional 6 genera and morphological analysis identifying a single additional genus.

The morphological approach exhibited a more even distribution in its percentage abundance bar chart. (5.31). A visual comparison of both methodologies was made for each concentration of glyphosate used (Figure 5.32). In the control, a total of 23 and 24 genera were detected in morphological analysis and metabarcoding identification, respectively; in both the low-dose and high -dose, a total of 22 and 23 genera were detected in morphological and metabarcoding identification, respectively. It was also observed that the morphological outcomes appeared similar visually across all three concentrations. However, the metabarcoding results exhibited similarity only between the control and low-dose, with the high-dose showing discernible differences toward the end of the season.

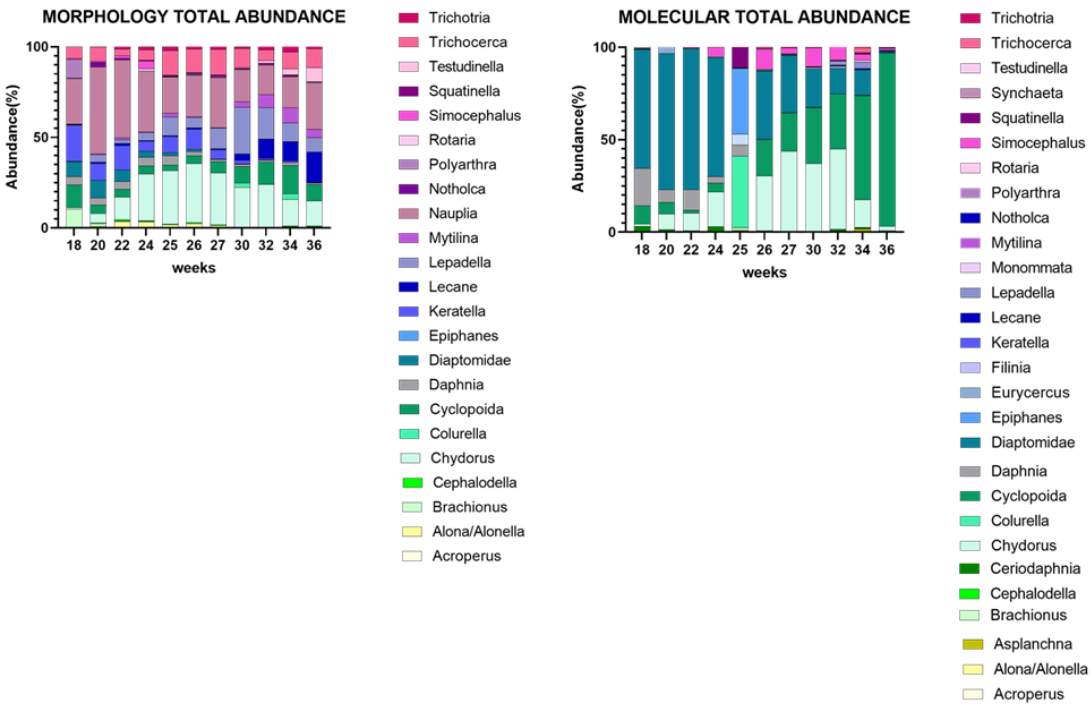


Figure 5. 31: Comparison of zooplankton genera abundance (%) at each sampling day (combined for all mesocosms) between the two methodologies.

On the left, the bar chart of morphology shows 23 genera detected; on the right, the bar chart of molecular analysis shows 28 genera detected.

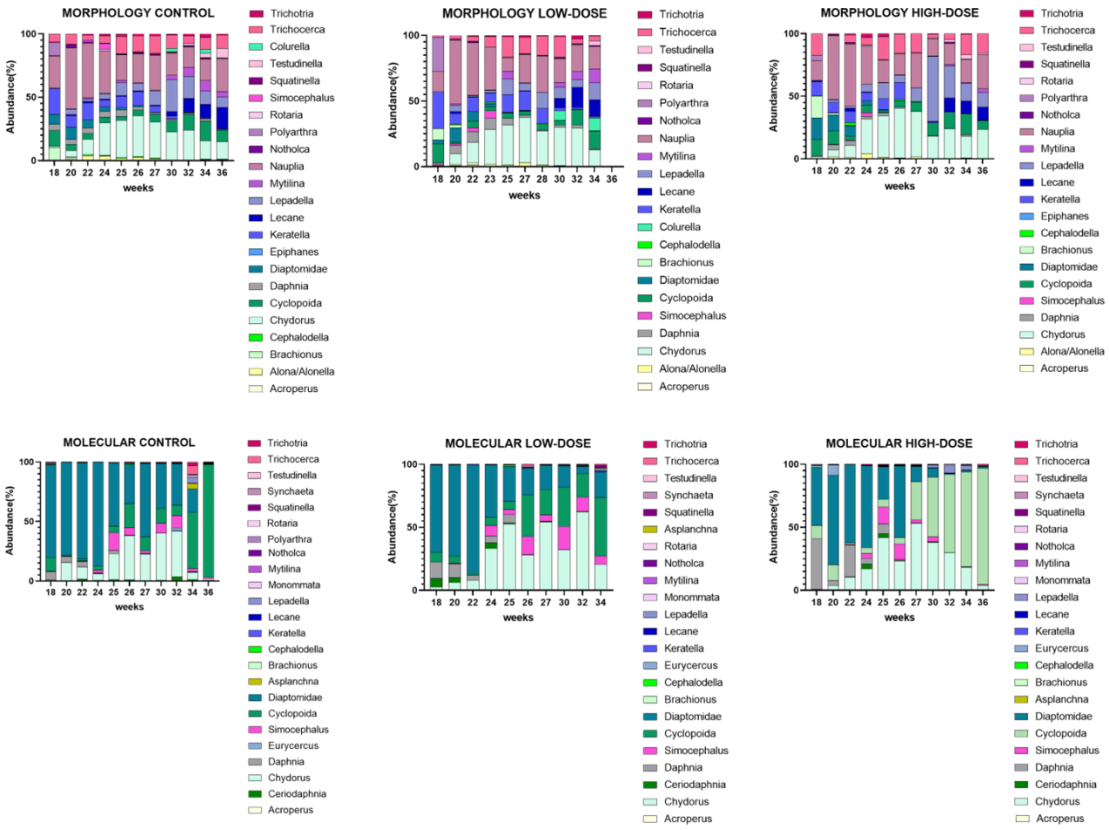


Figure 5. 32: Comparison of zooplankton genera abundance (%) at each sampling day at the selected Glyphosate doses for both methodologies.

In the top row control, low-dose and high-dose in morphological analyses are shown. In the bottom row, control, low-dose and high-dose in molecular analyses are shown.

Finally, the Spearman correlation coefficient was calculated between morphological and molecular data for various genera to perform a correlation plot on the eight most abundant genera detected in the two methodologies (Figure 5.33). The values ranged from -1 to 1, where -1 indicates a perfect negative correlation, 1 indicates a perfect positive correlation, and 0 indicates no correlation. The correlation coefficients provided showed the relationship between morphology and metabarcoding at both the genus and group levels. The conducted analysis, correlations between morphological and molecular variables for different genera were examined (Figure 5.20). For *Chydorus*, a robust and significant positive correlation (R = 0.85, P = 0.001) was observed, indicating a strong association between the two sets of variables. Conversely,

Cyclopoidae exhibited a moderate positive correlation (R = 0.51), but the result was not statistically significant (P = 0.11), suggesting that the observed correlation could potentially be attributed to random chance. *Keratella* displayed a strong positive correlation (R = 0.68, P = 0.022), with statistical significance, emphasizing a reliable association between morphological and molecular features. *Daphnia* exhibited a strong positive correlation (R = 0.72, P = 0.013), indicating a significant relationship between the variables. In the case of *Lepadella*, a moderate positive correlation (R = 0.48) was noted, but the correlation lacked statistical significance (P = 0.14). *Trichotria* demonstrated a strong positive correlation (R = 0.74, P = 0.008), with a statistically significant association between morphological and molecular variables. *Trichocerca*, on the other hand, displayed a weak negative correlation (R = -0.15), and the result was not statistically significant (P = 0.65), suggesting a lack of a significant relationship. Finally, *Diaptomidae* exhibited a very strong positive correlation (R = 0.88, P <0.001), indicating a highly significant and robust association between morphological and molecular variables. In the analysis conducted at the group level (Figure 5.34), correlations between morphological and molecular variables for *Cladocera*, *Copepoda*, and *Rotifera* were explored. For *Cladocera*, a modest positive correlation (R = 0.18) was observed, but the result lacked statistical significance (P = 0.6). *Copepoda* exhibited a moderate positive correlation (R = 0.47), but the result was not statistically significant (P = 0.14). In the case of *Rotifera*, a weak negative correlation (R = -0.04) was identified, and the result was not statistically significant (P = 0.22).

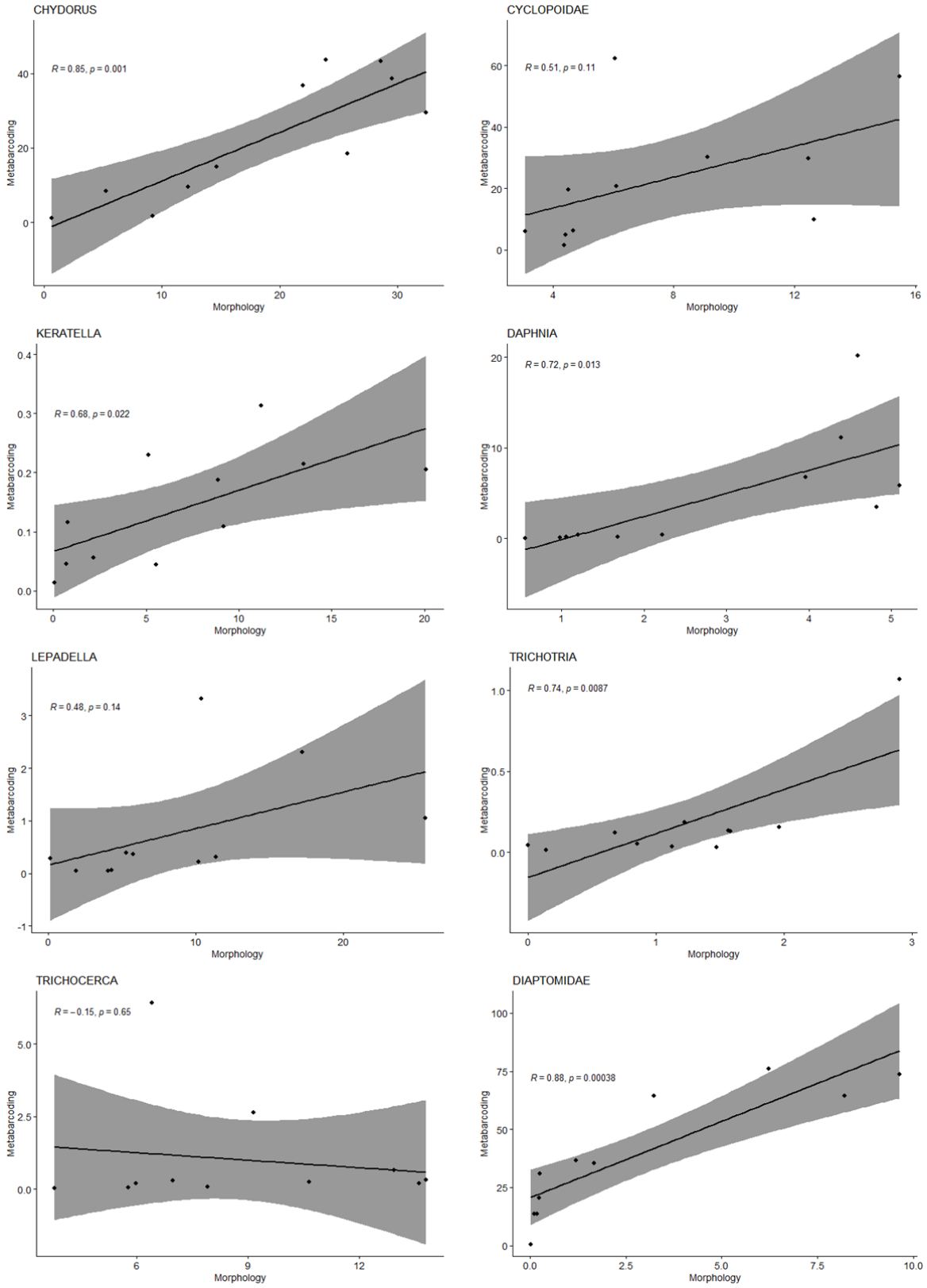


Figure 5. 33: Pearson Correlation test between morphology and metabarcoding data for the eight selected genera.

The values ranged from -1 to 1, where -1 indicates a perfect negative correlation, 1 indicates a perfect positive correlation, and 0 indicates no correlation.

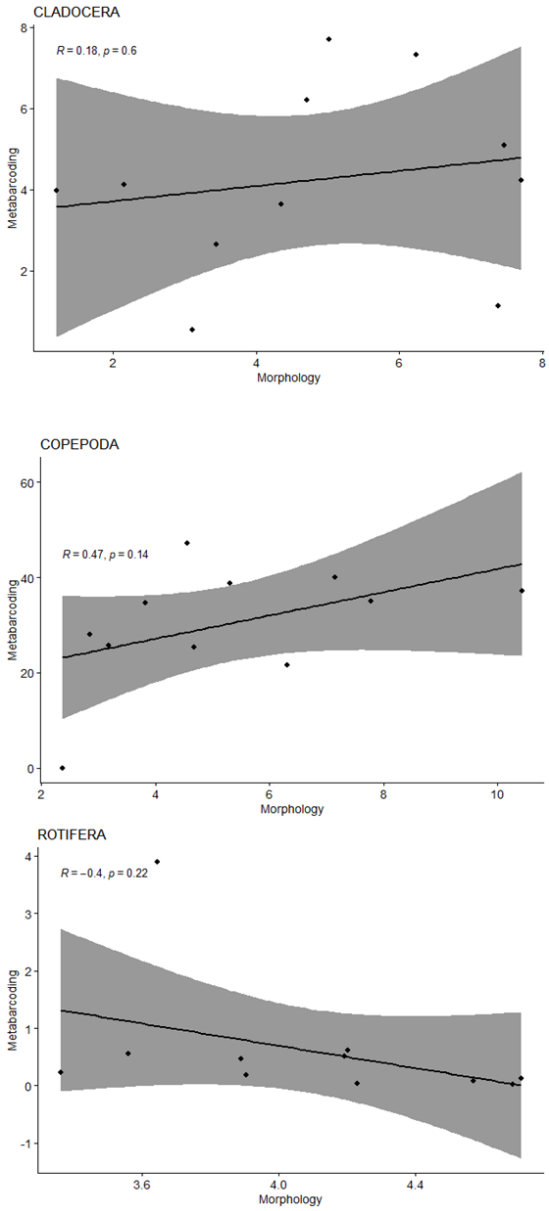


Figure 5. 34: Pearson Correlation test between morphology and metabarcoding data for the three zooplankton groups (*Cladocera, Copepoda* and *Rotifera*).

The values ranged from -1 to 1, where -1 indicates a perfect negative correlation, 1 indicates a perfect positive correlation, and 0 indicates no correlation.

5.3 Discussion

This study looked at how Glyphosate affects zooplankton presence and abundance in mesocosm test systems. In this investigation, two distinct methodological approaches were employed: metabarcoding, involving DNA analysis, and traditional

morphological analysis by microscopy. The primary objective was not only to assess the impact of Glyphosate at different concentrations on zooplankton but also to analyse the concordance between the outcomes from molecular and morphological methodologies at genera, group and trial level. Additionally, the study considers the implications of the results for the broader understanding of the ecological consequences of pesticide exposure on zooplankton communities.

The following concentrations were set to target zooplankton as individuals and as a community. Effects for both glyphosate doses (5.5 and 22.0 mg a.i/L) were expected to be seen within the first 24 hours after the application. The different dosing concentrations were expected to have different effects for each zooplankton genera (Hebert et al, 2020). Glyphosate was expected to cause substantial biomass decreases in *Rotifers* at around = 0.3 mg a.i./L, in *Cladocerans* at around 0.7 mg a.i./L and in *Copepods* at around 5.5 mg a.i./L (Hebert et al, 2020). A recovery was expected for both doses within the timeframe of the study.

5.3.1 Water quality and the effects of Glyphosate application

For the water quality environmental variables measured, significant differences were found in temperature, dissolved oxygen, and pH post-application of the higher dose of the pesticide. Although temperature also varied pre-application, the overall data indicate that the high dose of Glyphosate altered the physio-chemical properties of water. Glyphosate is a weak acid, so could directly lower the pH (Lopez-Valcarcel et al., 2023), however, this was not the case in the current study, as the high dose of glyphosate showed a higher pH than the control.

5.3.2 Glyphosate effect on Zooplankton and community composition

Counter to expectations, the community composition in morphological analyses was not impacted by the Glyphosate doses when looked at in The Detrended Correspondence Analysis as the doses clustered together. This was also reflected in the Principal Response Curve Analysis where doses showed several peaks compared to the control but no significant differences were found ($p = > 0.05$). The Day factor seems to have the most impact on the community as the zooplankton community composition changes throughout the season. In fact, the beginning of the season was found to be dominated by the *Cladocera* and *Copepoda* groups, compared to the end of the season, where mostly *Rotifers* were found. Seasonality plays a crucial role in the life cycles of zooplankton groups such as copepods, C*ladocerans*, and Rotifers. Copepods, exhibit

seasonal fluctuations influenced by factors like temperature, light availability, and food abundance, leading to variations in their abundance and species composition throughout the year (Choi et al., 2020). *Cladocerans*, on the other hand, show a peak in abundance during spring and summer months, driven by favourable conditions for reproduction and growth, followed by a decline during colder seasons (Jeppesen et al., 2011). Rotifers, with their rapid reproductive rates and ability to enter dormant stages, display a diverse array of seasonal adaptations, with some species thriving in summer blooms while others persisting through harsher conditions by forming resting eggs (Gilbert, 2016). These contrasting seasonal dynamics highlight the varied ecological strategies employed by these microcrustaceans to cope with changing environmental conditions. The PRC also showed that *Testudinella*, and *Cephalodella* were positively impacted by the Glyphosate herbicide and *Chydorus* and *Alona* were negatively impacted. This was further investigated using univariate analysis, but no significant differences were found between the control and any of the Glyphosate doses for these four genera.

The NMDS for the morphology data, combined with the hierarchical clustering, showed a difference in clustering among collection dates throughout the season where, in the control are randomly scattered and in the low and high Glyphosate doses pre- and post-application are clustered separately. This clustering was tested through the ANOSIM. A significant difference ($p = <0.01$) between sampling points was found in both control and two doses indicating the pesticide-affected zooplankton communities. The community composition was also assessed by calculating and comparing two different diversity indices (Shannon and Chao-1) between pre- and post-application. Shannon diversity index (H) is a commonly used diversity index that takes into account both the abundance and evenness of genera present in the community where higher values indicate higher diversity (Rajagopal et al., 2009). The value found in the Shannon diversity index for the high dose of Glyphosate during post-application was lower compared to the one of the control and low dose. Therefore, a higher diversity was found in post-application for control and low-dose while a higher diversity was found in pre-application in the high dose. Furthermore, no significant difference was found pre-application when tested with RM-ANOVA. A significant difference was found in the global post-application and between control and high-dose ($p = < 0.01$). The Chao-1 indicator showed that for higher values, there was higher diversity (Stephan et al., 2017). The post-application for control and two doses had a higher number and thus

greater diversity. The results of the RM-ANOVA on Chao-1 pre-application were not significantly different. However, they were significantly different during post-application ($p = <0.01$). There was also a significant difference between the control and high dose in the post hoc Dunnett's test.

Assessments of microcosm/mesocosm studies heavily rely on determining concentrations where no observable adverse effects occur (NOECs) across various population-level parameters. Ideally, these studies should include a power analysis for the concentration-response relationships that establish these NOECs, along with measurement endpoints where significant effects may not be readily apparent. Post-analysis, an indication of the statistical power can be inferred through calculated minimum detectable differences (MDDs). The MDD essentially signifies the magnitude of difference between treatment and control means required to identify a statistically significant effect (Brock et al., 2015)

The overall NOEAEC value obtained from morphological analyses indicated a recommended dose of Glyphosate lower than 5.5 mg a.i./L. Not many genera were affected by the Glyphosate at both doses. However, it appears that the low-dose (5.5 mg a.i./L) had a bigger effect than the high dose (22.0 mg a.i./L), as the majority of significance for this dose was non-monotonous. Non-monotonous dose-response refers to a situation in toxicity testing where the dose-response relationship between the concentration of a substance and its effect on organisms does not follow a consistent trend. In other words, there may be fluctuations or irregularities in the response of the organisms to increasing concentrations of the substance. This can make it challenging to determine a clear threshold concentration at which no adverse effects are observed, as the response of the organisms may vary unpredictably across different concentration levels. As a result, special consideration and interpretation may be needed when assessing the toxicity of substances with non-monotonous dose-response relationships (Vanderberg et al., 2013).

Unlike the morphological analyses, metabarcoding showed a concise difference between doses in the Principal Component Analysis and the Principal Response Curve. In fact, the control and the two Glyphosate doses in the PCA resulted in extremely scattered and the PRC analysis showed a significant difference between the two doses and the control ($p=0.004$). Similarly to the morphological analyses, the season had the

major impact on the community composition. However, unlike morphology, the two *Copepoda (Cyclopoidae* and *Diaptomidae*) and Chydorus were the genera reacting the most from exposure to Glyphosate. When this effect was looked at in the univariate analysis, a significant difference for the two *Copepoda* was found at both doses confirming the impact from the multivariate analyses.

The NMDS combined with the hierarchical clustering showed a difference in clustering among collection dates throughout the season where, in the control are randomly scattered and in the low and high Glyphosate doses pre- and post-application are clustered separately especially in the low-dose. This clustering was tested through the ANOSIM. A significant difference ($p = <0.01$) between sampling points was found in both control and two doses. Finally, the community composition was assessed by calculating and comparing two different diversity indices (Shannon and Chao-1) between pre- and post-application. The value found in the Shannon diversity index for the high dose of glyphosate during post-application had lower variation compared to the one of the control and low-dose. No significant difference was found pre-application and post-application when tested with RM-ANOVA. For the Chao-1 analysis, the post-application for control and two doses had a higher value and thus greater diversity. Results of the RM-ANOVA on Chao-1 pre-application showed no significant difference between conditions.

The overall NOEAEC value obtained from metabarcoding analyses indicated a recommended dose of Glyphosate lower than 5.5 mg a.i./L. Not many genera were affected by the Glyphosate at both doses. However, it seems that the low-dose (5.5 mg a.i./L) is having a bigger effect than the high dose (22.0 mg a.i./L) as the majority of significance for this dose was non-monotonous. These results agreed with the morphological analysis showing that both methodologies have determined the same NOEC of > 5.5 mg a.i./L.

In this subsection of the study, Glyphosate was expected to cause substantial biomass decreases in *Rotifers*, *Cladocerans* and *Copepoda* at the doses applied (Hebert et al, 2020). This was indeed realised, in both methodologies, at genera level for Cladocera and Copepoda. The Rotifera group showed a decrease at genera level for two genera. However, no decrease for this group was detected in the metabarcoding analyses.

Overall, the morphological analyses showed no significant differences in the community composition. However, some genera had a strong response to both doses. The difference at the community structure at pre- and post-application was shown in the clustering in low and high Glyphosate doses, indicating its effect at the different sampling points. The ANOSIM also confirmed a significant variation between the global sampling points. However, this variation was also found in the control. Finally, The Shannon index indicated a higher diversity post-application in both control and low-doses probably due to the zooplankton seasonality compared to the pre-application where most of the zooplankton are still dormant (Jensen, 2019). However, the Shannon index clearly showed that the high dose of Glyphosate post-application had lower variation confirming again its effect on zooplankton.

In contrast, the metabarcoding analyses showed a significant difference in the community composition. The difference in the community structure at pre- and post-application was shown in the clustering in low and high Glyphosate doses, confirming its effect at the different sampling points. The ANOSIM also confirmed a significant variation between the global sampling points. However, again, this variation was also found in the control. Finally, The Shannon index indicated a higher diversity post-application in both control and low-doses probably due to the zooplankton seasonality compared to the pre-application. The Shannon index clearly showed that the high dose of glyphosate, post-application had lower variation indicating an effect on zooplankton.

Given the results of both morphological and metabarcoding analyses, which indicate minimal significant differences in community composition following Glyphosate application, it is important to explore potential reasons behind the limited observed impact on zooplankton assemblages. One possibility is that zooplankton possess adaptive mechanisms enabling them to mitigate the effects of environmental stressors such as herbicides. For instance, zooplankton may alter their feeding strategies, perhaps by exploiting alternative food sources or enhancing their efficiency in nutrient uptake, thereby compensating for any reduction in phytoplankton biomass caused by Glyphosate (Vanderberg et al., 2013). Furthermore, abiotic factors and seasonal dynamics likely play a crucial role in shaping zooplankton communities, potentially overshadowing any herbicide-induced effects (Brock et al., 2015). The significant influence of the 'Day' factor in the analyses suggests that seasonal variations such as changes in temperature, light, and nutrient availability could be driving community

changes more than Glyphosate exposure (Choi et al., 2020). Additionally, the interaction between these seasonal factors and zooplankton's life cycle stages, which include periods of dormancy and population peaks, might buffer these organisms against environmental changes, including herbicide exposure (Jeppesen et al., 2011; Gilbert, 2016). Thus, the lack of substantial impact observed in this study may result from a complex interplay of biological resilience, seasonal dynamics, and abiotic conditions, rather than an absence of herbicide effects.

5.3.3 Comparison between morphology and metabarcoding analyses

The major point of this study was to identify if metabarcoding could mimic morphological analysis during a pesticide trial. Even though the metabarcoding did not mimic the same significant differences found for each genus and family, both methodologies established that a dose lower than the selected low-dose (5.5 mg a.i./L) would needed for the Glyphosate in a true risk assessment. This was a pilot study with lower number of doses compared to a normal full ecotoxicological study and a specific recommended NOEC dose is required other than a dose smaller than 5.5 mg a.i./L.

A positive correlation between five (*Chydorus, Keratella, Daphnia, Trichotria* and *Diaptomidae*) out of the eight selected genera was found between the two methodologies when analysed through Pearson's correlation indicating that 62.5 % of the genera's abundance was similar throughout the season. However, no significant correlation was found for any of the groups (*Cladocera, Copepoda* and *Rotifera*) between the two methodologies. It is not clear why the correlation seen here was not consistent between seasons (data reported in Chapter 4).

5.3.4 Conclusions

The comparison of the two detection methodologies in this chapter indicated that detection of the majority of genera demonstrated a strong correlation between the two methods (5 out of 8 genera tested), unlike in Chapter 4. Metabarcoding detected more genera than morphological analysis. This result is comparable with data reported in Chapter 4.

This chapter looked at a small-scale pilot study for the herbicide active ingredient glyphosate present within Glyphosate. The selected doses of Glyphosate were expected to reduce zooplankton populations Hebert et al., (2020). However, by comparison, the effects seen by both morphological and metabarcoding analysis methods were relatively subtle. The effects of Glyphosate were found to be different between the two methodologies in community composition as significant differences were found only in metabarcoding analyses. However, both methodologies established a NOEC lower than 5.5 mg a.i./L. Whilst the patterns of genera vary between methods the pilot study data indicates the methods may produce similar outcomes for ecotoxicity trials.

There is a possibility that a more effective analysis could have been performed by sampling after 24, 48, 72 and 96 h to exclude an immediate effect and possible quick recovery rather than from sampling one-week post-application, as some studies have sampled one day post-application (Hebert et al., 2021 and Baker et al., 2016). This study attempted to show the relationships between zooplankton communities, environmental variables, and pesticide exposure.

The integrated approach of morphology and metabarcoding provided an understanding of community dynamics. These findings hold significance in the context of environmental monitoring and contribute to the broader discourse on the ecological consequences of pesticide exposure in aquatic ecosystems.

Chapter 6. Oxford Nanopore sequencing

6.1 Introduction

Oxford Nanopore Technologies (ONT) remains a revolutionary method in DNA sequencing, offering a unique approach compared to traditional methods. ONT employs nanopores, small protein-based channels, to directly read DNA strands in real-time as they pass through these pores (Lu et al., 2016). This approach provides long read lengths, advantageous for applications such as de novo genome assembly and structural variant detection. In contrast, Illumina sequencing utilises a short-read technology

based on fluorescently labelled nucleotides and has been the cornerstone of high-throughput sequencing for years. While Illumina excels in accuracy, ONT's technology offers longer reads, capturing complex genomic regions more comprehensively.

The comparison between ONT and Illumina sequencing involves factors such as read length, accuracy, cost, and applicability to specific research objectives. Although hybrid approaches combining the strengths of both technologies have demonstrated the complementarity of short and long-read sequencing strategies in genomics research (Goodwin et al., 2016; Logsdon et al., 2020; Nurk et al., 2020), recent advancements and the rapidly changing landscape of sequencing technology have shifted the focus towards optimising the use of individual platforms rather than hybrid methods.

Recent Innovations include adaptive sampling, ultra-long reads, high-fidelity chemistry and scalability and throughput. One of the latest features introduced by ONT is adaptive sampling, which allows for the real-time selection of target DNA sequences during sequencing. This technology can enrich specific regions of interest, reducing sequencing costs and improving efficiency (Payne et al., 2021). ONT has made significant strides in generating ultra-long reads, with some sequencing runs achieving read lengths exceeding 4 Mbp. These ultra-long reads are particularly useful for resolving complex genomic regions and structural variations (Jain et al., 2020). ONT continues to improve its sequencing accuracy with new chemistries and flow cells. The Q20+ chemistry, for instance, has been reported to significantly reduce error rates, bringing them closer to those of short-read technologies (Vaser et al., 2021). Furthermore, the release of PromethION, a high-throughput sequencer by ONT, has expanded the scalability of nanopore sequencing, making it suitable for large-scale genomic projects (Shafin et al., 2021).

The ONT MinION can identify multiple species across various samples in a single sequencing run. The potential advantages of utilising the MinION include the ability to process samples independently, handle small sample batches for quicker results (as sequencing runs can be paused and the flow cell reused), conduct on-site sequencing (as the MinION is a portable device connectable to a laptop), and generate longer reads with an average length of 8 kbp for optimal sequencing. However, a notable drawback is the higher DNA sequencing error rates, with Tyler et al. (2018) reporting an average error rate of 6% (using R9.4 flow cells), in contrast to the significantly lower average

error rate of 0.24% observed with Illumina platforms (Pfeiffer et al., 2018). This higher error rate may compromise the reliability of species detection results, especially with the short read lengths usually associated with community analyses. Additionally, there might be increased per-sample costs and potential limitations in scalability with the MinION, although these aspects need further exploration in the context of species detection. Overcoming these limitations could lead to a substantial reduction in sample turnaround time, bridging the gap between sample collection and obtaining results, thereby enhancing the effectiveness of environmental DNA (eDNA) approaches (Fitzpatrick et al., 2021; Pomerantz et al., 2023).

Integrating Oxford Nanopore Sequencing (ONT) into ecotoxicological studies opens new frontiers in decoding the genomic responses of zooplankton to environmental stressors. By providing real-time, long-read sequencing capabilities, ONT enables a comprehensive exploration of the intricate genetic makeup of these key aquatic organisms. The use of ONT in ecotoxicology aligns with the broader shift towards single-molecule sequencing methodologies (Goodwin et al., 2016). However, this approach is relatively new and constantly updating, with limited references on MinION barcoding of zooplankton short fragment reads.

In the previous chapter, metabarcoding was assessed to identify zooplankton genera and relative abundance using Illumina sequencing. In this chapter, the study aims to analyse how the ONT platform performs. The primary objectives involve a comprehensive investigation into the effectiveness of ONT sequencing specifically applied to zooplankton samples within ecotoxicity trials. This involves exploring the ability of Nanopore sequencing to generate high-quality, short-read data that can capture the intricate genetic diversity present in these environmental samples. The second aim is to conduct a detailed comparative analysis between Oxford Nanopore and Illumina sequencing methodologies, focusing on sequencing accuracy, read length distribution, and cost-effectiveness within the specific framework of zooplankton ecotoxicity trials. The goal of comparing the strengths and limitations of both sequencing platforms is to clarify the most suitable and efficient approach for obtaining comprehensive genomic insights from zooplankton communities under ecotoxicological stress. The outcomes of this comparative analysis will contribute valuable insights to the broader scientific community engaged in environmental genomics, shedding light on the optimal sequencing strategy for ecotoxicity studies involving zooplankton samples.

6.2 Results

6.2.1 Comparative Analysis of Illumina and Oxford Nanopore Sequencing Results for Zooplankton identification

A summary of the process is illustrated in Figure 6.1, and a comparison of the costs involved in Table 6.1.

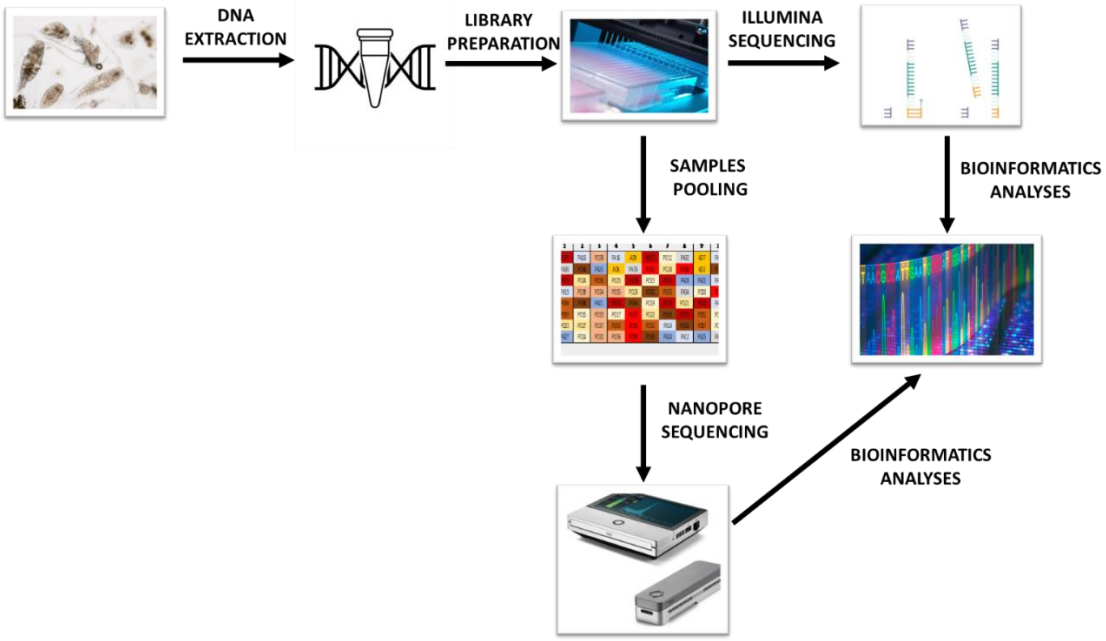


Figure 6. 1: Flowchart depicting the analysis process of zooplankton samples, from DNA extraction to NGS sequencing.

Table 6. 1: Comparison of Illumina and Nanopore sequencing costs (for a 96 samples run).

In this table, communal costs such as DNA-extraction kits, Qubit assay kit, primers and tubes were excluded.

Illumina	Price (£)	Nanopore	Price (£)
96 amplicon library preps after first stage PCR (£19.42/each)	1864.32	Native Barcoding kit (96 V14)	720

One 250PE MiSeq run (MiSeq V2 500 cycle kit)	1568.55	NEB Blunt/TA Ligase Master Mix (NEB, cat # M0367	176
		NEBNext Ultra II End repair/dA-tailing Module (NEB, cat # E7546)	752
		NEBNext Quick Ligation Module (NEB, cat # E6056)	1,200
		Bovine Serum Albumin (BSA) (50 mg/ml)	223
		Flow-cell (R10.4.1)	630
		Flow-cell wash kit	90
		Flow-cell Priming kit	35
		MinIon	8,100.00
Total for one run	**3516.37**	**Total without Minion sequencing device**	**2627.2**

Results from the custom-made script created to analyse the zooplankton samples from Nanopore sequencing were inconclusive as no ranking was created at the end of the PERL script. Because of this, no direct comparison could be made for each sample between Illumina and Nanopore.

The MinION sequencing run generated 8.9 M raw reads in total, while Illumina MiSeq sequencing generated 20.6 M paired end reads (Figure 6.2). The Illumina sequences were already indexed by DeepSeq so no access to the original raw reads was obtained. A step-by-step table (Table 6.2) was made to show the sequence loss for both Nanopore with no mismatches and one mismatch at each bioinformatic step.

Table 6.2: Nanopore Sequencing: Sequence Processing and Taxonomy Assignment Summary.

SOFTWARE	Software Description	Read Count (no mismatches)	Read Count (one mismatch)
RAW DATA	The initial sequence data generated by the Nanopore MinION sequencer, representing the total number of raw reads obtained from the sequencing run.	8,9 M	8,9 M

Chopper	Software used for trimming and filtering low-quality reads and adapter sequences from the raw data, improving the overall quality of the sequence data.	8,48 M	8,48 M
SEQTK	A toolkit for processing sequences, SEQTK is used here for further quality control and sequence filtering, reducing the data to high-confidence reads.	8,07 M	8,07 M
INDEXING	The process of adding index sequences to reads, allowing for the identification of samples. Fewer reads may remain due to stricter matching criteria (no mismatches).	1.4M	5.7M
CUTADAPT	A tool used to remove unwanted sequences from the reads, such as adapters or primers, resulting in high-quality, final sequences for downstream analysis.	489,987	1,496,610
Assigned Taxonomy	The final number of reads that were successfully assigned a taxonomic classification after passing through the entire bioinformatics pipeline.	43,533	133,063

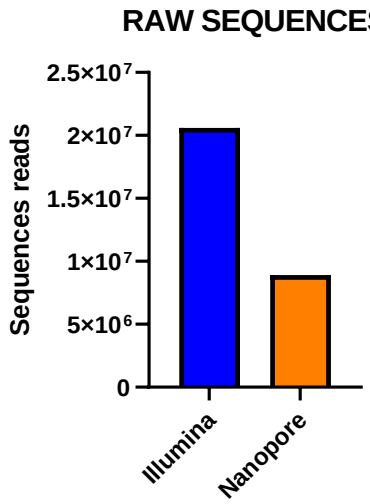


Figure 6. 2: Total sequence reads comparison between Illumina and Nanopore before bioinformatics analyses (Illumina sequences were already indexed by DeepDeq).

The demultiplexing process of 8.9 M raw Nanopore sequences was conducted both with and without allowing one mismatch using a wildcard. Without considering any mismatches, the demultiplexing procedure successfully identified and assigned 1.4 M sequences (Figure 6.4) to their respective indexes, indicating a demultiplexing success rate of just 16%. This method therefore led to a substantial sequence loss, with approximately 84% of the original sequences being discarded. Conversely, when allowing for one mismatch using a wildcard, the number of retained sequences increased to 5.7 M (Figure 6.3), representing approximately 65% of the total raw sequences. Despite the improvement in demultiplexing efficiency with the inclusion of mismatches, there remained a sequence loss of approximately 35%. These results suggest that while permitting slight variations in indexes can enhance demultiplexing success, further optimization is necessary to minimize sequence loss and improve overall efficiency. These 1.4M or 5.7M indexed reads with Naopore analysis are in comparison to 20.6M indexed reads produced with the Illumina MiSeq analysis (Figure 6.4).

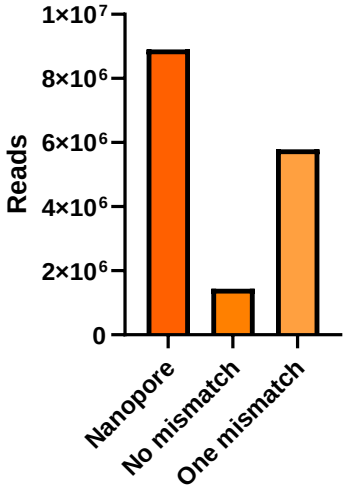


Figure 6. 3: Differences in sequence reads before binning (Nanopore) and after binning the indexes allowing no and then one mismatch.

These findings underscore the distinct characteristics and challenges associated with each sequencing technology and highlight the importance of optimizing demultiplexing strategies to maximize sequencing output and accuracy.

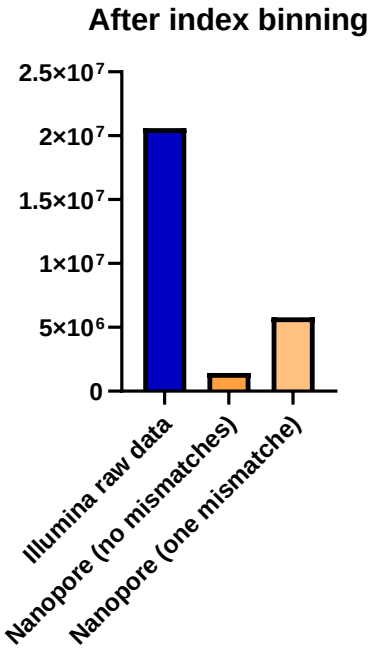


Figure 6. 4: Comparison of Illumina sequences and Nanopore after indexes binning.

The 80 samples run using the Illumina platform were successfully blasted against the custom-made zooplankton database. For the Nanopore samples when percentage

similarity was set to 95% both with no mismatch and one mismatch index datasets this resulted in 50 samples and 70 samples blasting, respectively. BLAST assigned taxonomic identities to 100% of Illumina samples, 62.5% for Nanopore no mismatches, and 87.5% for Nanopore one mismatch (Figure 6.5). When the percentage similarity was set to 80% for the Nanopore samples both with no mismatch and one mismatch resulted in 55 samples blasting and 75 samples blasting respectively.

In summary, the overall data indicates that for Illumina-derived sequences, 20.6 M indexed reads and 80 samples matched to a sequence during the BLAST at 95% match. For ONT-derived no-mismatches in the index sequences, 8.9 M raw reads led to 1.4 M indexed reads and 50 samples matched to a sequence during the BLAST at 95% match (and 55 samples at 80% match). Finally, for ONT-derived one mismatch allowed within index sequences, 8.9 M raw reads led to 5.7 M indexed reads and 70 samples matched to a sequence during the BLAST at 95% match (and 75 sequences at 80% match).

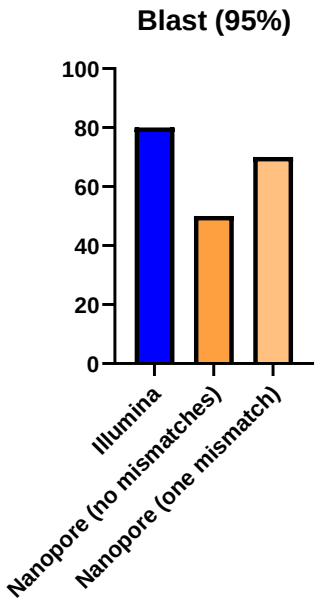


Figure 6. 5: Number of samples that provided sequence that matched to a database sequence at 95% identity (NGS platform indicated, index sequences matched at 100% or allowed a single mismatch as indicated).

6.2.1.1 Total genera detected by Nanopore and Illumina sequencing

A total of 28 genera were detected by Illumina sequencing, whereas eleven genera were detected by Nanopore with no mismatches. and 15 genera by Nanopore one mismatch (Figure 6.6). The eleven genera identified by Nanopore no-mismatches were also detected by Illumina sequencing. Similarly, the 15 genera detected by Nanopore one mismatch were detected by Illumina sequencing (Figure 6.7).

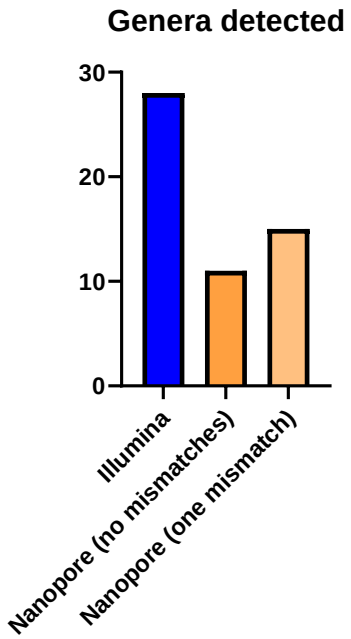


Figure 6. 6: Comparison of the number of genera detected by both methodologies.

Where NGS platform indicated, index sequences matched at 100% or allowed a single mismatch.

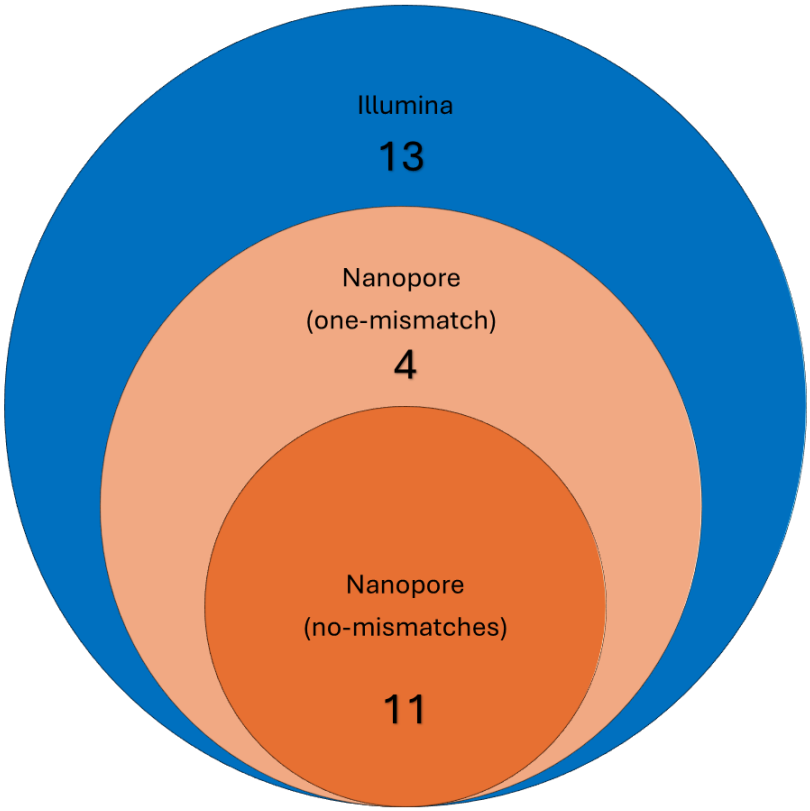


Figure 6. 7: Communal number detected by the two methodologies.

A total of 28 genera were detected by Illumina sequencing, 11 genera were detected by Nanopore sequencing with no mismatches and finally, 15 genera were detected by Nanopore sequencing with one mismatch in index sequences. All genera detected were communal within each method.

6.7 Discussion

The introduction of the MinION by Oxford Nanopore Technologies (ONT) in 2014 marked a significant advancement in nucleic acid sequencing for several reasons. Firstly, its lower initial and per-base sequencing costs, priced at 1,000 USD for the entry starter pack, made it more accessible. Secondly, its capability to produce long-read sequencing, now extending up to approximately 4 Mb, addressed the limitation of short read-lengths associated with Illumina sequencing, which typically reaches about 500 bp. Additionally, its compact size, portability, and real-time data generation capabilities addressed common criticisms of existing sequencing technologies (Chang et al., 2023). Since its introduction, Nanopore sequencing has found widespread application in numerous whole-genome sequencing and metagenomic studies. However, the use of nanopore sequencing for metabarcoding purposes remains relatively uncommon, as evidenced by the limited number of published papers, particularly in biodiversity-

related fields. Early studies primarily focused on microbial taxa, with limited attention given to non-microbial species until more recently. Recent studies have utilised Nanopore metabarcoding for biodiversity assessment, community characterisation, species-specific detection, and even diet analysis (Chang et al., 2023). Despite the promise shown by nanopore sequencing in metabarcoding, its adoption has been slowed down by two main factors. Firstly, the relatively high error rates in Nanopore reads, ranging from around 20% to as high as 38% (Chang et al., 2023), compared to the low error rate of Illumina sequencing at 0.24%, raises concerns about the accuracy of species identification in DNA metabarcoding.

In this study, following Nanopore sequencing, quality filtering of the sequenced reads was conducted, resulting in an average retention of 16% of the initial reads. This figure may appear notably lower when compared to conventional sequencing methods. For example, other studies have also found a lower proportion when comparing Nanopore and Illumina sequencing reads, such as 37% reported by Menegon et al. (2017) and 50% by Semmouri et al. (2020). Moreover, a stricter quality threshold was implemented in this study to minimise potential bias stemming from base-calling errors, thus ensuring accurate species identification. Errors detected in the resulting sequences can be generated in three ways: DNA template degradation; errors during amplification; and errors during sequencing, with the latter's magnitude varying amongst the different sequencing technologies (Semmouri et al., 2021). Tyler et al. (2018) reported an average error rate of 6% with R9.4 flow cells, contrasting with Illumina platforms' average error rate of only 0.24% (Pfeiffer et al., 2018), potentially impacting the reliability of species detection outcomes.

Moreover, Nanopore turnaround times were significantly faster; in fact, completing the one run of 80 samples, inclusive of sample preparation took only three days, compared to the approximately 2–4 weeks required for Illumina sequencing. A final point to consider was the cost between the two methodologies. Various studies have compared sequencing costs between nanopore and Illumina for metabarcoding, and it is generally agreed upon that nanopore metabarcoding with the MinION is generally cheaper than Illumina MiSeq (Van der Reis et al., 2022; Stevens et al., 2023). We reduced reagent costs further by pooling the already tagged library from the Illumina barcodes. This enabled us to pool multiple samples into just 10 pools for nanopore library preparation without further need to barcode them.

In this study, it was demonstrated that Nanopore metabarcoding did not generate metabarcodes with Illumina-like quality. The most notable difference was in taxonomic units richness which are assigned taxonomy using reference databases of known sequences, where we obtained 28 Illumina genera, compared to 11 nanopore genera for no mismatches and 15 for one mismatch, with 11 genera shared across both platforms (28% congruence).

Chapter 7: General Discussion

The primary objectives of this study were to analyse the comparative advantages and limitations of both metabarcoding and molecular techniques in identifying zooplankton. Additionally, this research aimed to assess whether metabarcoding could effectively replace or improve traditional morphological identification methods. As part of the investigation, the impact of the herbicide Glyphosate on zooplankton genera both at the individual and community levels was assessed. This part of the study aimed to examine potential correlations between morphological identification and metabarcoding methods within the context of an ecotoxicology trial.

These objectives were met by comparing the zooplankton genera identified and the relative abundance of zooplankton obtained through both methodologies (morphology and metabarcoding) during two zooplankton collection seasons, one of which included a pilot ecotoxicity trial involving a herbicide (Glyphosate).

7.1 Outcomes of the Study

7.1.1 Optimisation of molecular detection of phytoplankton and zooplankton

Plankton are often used as a bio-indicator of aquatic pollution since they have a short generation time and respond rapidly to environmental changes and anthropogenic disturbances (Barka et al., 2020). In this study, the most sustainable, reliable, and time-efficient methodologies for the molecular detection of both phytoplankton and zooplankton genera were evaluated. This work involved the development and

refinement of metabarcoding or genus-specific qPCR protocols. However, the phytoplankton metabarcoding approach was not taken forward due to insufficient sequence availability in databases for the designated genes. Consequently, the study's focus was on zooplankton genera detection. Furthermore, genus-specific analysis of zooplankton through qPCR assays produced inconsistent findings and was not further developed. Metabarcoding emerged as the preferred approach for zooplankton genus analysis, demonstrating superior consistency and reliability once a robust methodology was established. This was established by testing the most commonly used primers in zooplankton metabarcoding (Folmer et al., 1994, Bucklin et al., 2019 and Djurhuus et al., 2018) and with the final optimised PCR method for metabarcoding based on the protocol used in Bucklin et al., (2019).

7.1.2 Comparison of molecular analysis of zooplankton populations with morphological examination

The application of metabarcoding to zooplankton analysis is still a novel process. Although the use of metabarcoding methods to identify indicator species has been extensively examined, to my knowledge, there is a limited exploration into the application of metabarcoding for identifying zooplankton (Bucklin et al., 2019; Zhao et al., 2021; Djurhuus et al., 2018; and Ershova et al., 2021).

In this project, a correlation between morphology and metabarcoding techniques was observed in zooplankton analyses. In the seasonal study, conducted between May and September 2020, metabarcoding and traditional gold standard microscopy methodologies were compared across five different mesocosm samples over a season of sampling. Despite each method detecting genera absent in the other, the results revealed a similar number of genera. Specifically, a total of 26 different genera were identified, with 17 detected by both methodologies (65% correlation), 3 exclusively by morphology, and 6 exclusively by metabarcoding.

Results were consistent in the next sampling year (2022) where the ecotoxicology pilot study was assessed, as again, they revealed that metabarcoding identified a higher number of genera (28) compared to morphological analysis (23). Both methodologies identified the same 22 genera with metabarcoding identifying an additional 5 genera and morphological analysis identifying a single additional genus (Nauplii – the larval

stage of *Copepoda*). However, the gene sequence does not change between larvae and adult stages, thus the metabarcoding only assigns sequences to the adult stage. The fewer genera detected via morphology in both experiments could be due to potential misidentifications, particularly in the most turbid samples as due to their low-clarity identification of similar genera can be complicated (Yang et al., 2017).

Furthermore, no significant correlation was found in the first sampling year when the most abundant genera (*Diaptomidae, Lepadella, Daphnia, Acroperus, Keratella, Mytillina, Cyclopoidae, Chydorus*) for both morphology and metabarcoding were looked at. The next sampling year showed a positive correlation between five (*Chydorus, Keratella, Daphnia, Trichotria and Diaptomidae*) out of the eight selected genera when comparing the two methodologies and analysis through Pearson's correlation, indicating that 62.5 % of the genera's abundance was similar throughout the season. However, no significant correlation was found for any of the groups (*Cladocera, Copepoda* and *Rotifera*) between the two methodologies.

As previously mentioned, zooplankton are well-suited for metabarcoding analysis due to their global presence in water bodies and the easiness of sampling. While the collection of zooplankton does not demand specialised skills or expensive equipment, the processing of samples continues to be limited by the few successful methods available in most zooplankton monitoring surveys (Ershova et al., 2021). Identifying zooplankton samples through morphological analysis currently remains the gold standard method for quantitative assessment of zooplankton at the species level. This time-consuming approach requires a trained taxonomist, and the processing of a single sample can range from several hours to a few days depending on the quality and the abundance of organisms in the sample. Additionally, this method is inadequate for quantifying organisms that deform or disintegrate in fixatives or those that lack sufficient morphological features for visual identification such as larval stage (Ershova et al., 2021). This method is also prone to human bias. As a result of the high time and expertise requirements, many months/years generally pass between sample collection and data acquisition, which may cause significant delays in ecotoxicological trials.

On the other hand, in recent years, DNA metabarcoding has revolutionized zooplankton identification by leveraging differences in DNA sequences among taxa (Hebert et al., 2003). This technique significantly reduces the time and increases the availability of

expertise required for identification. However, most molecular studies have traditionally focused on individual species or specific groups and only recently has metabarcoding been widely applied for biodiversity and community analyses in zooplankton research (Bucklin et al., 2016, 2019; Zhang et al., 2018). Metabarcoding has the potential to detect all species in a community, regardless of their developmental stage or the preservation of distinguishing features. Therefore, it can rapidly identify shifts in biodiversity and community composition, monitor rare species that may not be visually identifiable in sufficient numbers, and detect invasive species.

The outcomes of metabarcoding studies primarily depend on the selection of markers, which must offer ample taxonomic coverage and resolution tailored to the specific studied taxa. One limitation of metabarcoding is identifying an appropriate target gene region and the corresponding primers. Preferably, the selected primers should be universally applicable and effective across all the chosen taxa. Moreover, the amplified DNA region must exhibit sufficient genetic variation between species while remaining stable enabling a detailed species-level identification (Ershova et al., 2021). One of the most used genes for plankton barcoding is the mitochondrial cytochrome oxidase I (COI). However, this gene has been proven to be too long for most next-generation sequencing methods. Some studies have been proven to be successful when using a smaller fragment of this gene (Folmer et al., 1994). In this study, a more consistent amplification was achieved using the 18s gene (Bucklin et al., 2019) compared to the COI gene even when using the Folmer primer set. In fact, the majority of metabarcoding studies on planktonic organisms have been performed using different regions of the 18s gene (Bucklin et al., 2016) as this region is well-conserved across most taxa. However, the 18s gene also has poor taxonomic resolution for the metazoan division which includes zooplankton (Mohrbeck et al., 2015). Finally, when considering which target region is optimal for the identification of the selected taxa, the reference sequence completeness of the database is essential (Ershova et al., 2021).

Reference databases are growing for several of the gene regions most frequently used for zooplankton metabarcoding. The most notable are the GenBank database and the Barcode of Life Database, which provide a valuable reference library for species identification (Bucklin et al., 2010). Recent implementations of metabarcoding have shown valuable insights into the genetic diversity of both freshwater and marine planktonic communities (Zhang et al., 2023). However, the functional evaluation of

communities and biodiversity through metabarcoding is not highly accurate due to the scarcity of comprehensive reference barcode databases. In fact, in certain investigations, over 40% of identified taxonomic units could not be definitively assigned to a taxonomic category (Yang et al., 2017). In the current study, the juvenile stage of *Diaptomidae* and *Cyclopoidae* (*Nauplii*), were not found by metabarcoding as directly assigned to the adult stage. This is not the case in the morphological data as *Nauplii* are separately identified as too similar to distinguish to which adult genera belong. This could be one of the reasons why the two methodologies do not correlate with the *Copepoda* abundance. In fact, the *Nauplii* genera is one of the most abundant in morphological analyses. Similarly, *Copepoda* predominates the metabarcoding abundance probably because the numbers found in morphology for Nauplii are seen in the adult stage in metabarcoding. Furthermore, other three genera found in the list of species, genera and groups present within mesocosms (data provided by CEA; Chapter 4, Table 4.1) were not available for the 18s gene in the databases. This could significantly influence the metabarcoding outcome in terms of relative abundance especially when comparing it to the morphological analyses.

In this study, the metabarcoding analysis identified three additional genera in the first season and 5 in the second one compared to the morphological analysis. This suggests that DNA metabarcoding enhances the accuracy/sensitivity of morphological identification in estimating zooplankton diversity and composition. This is consistent with previous studies (Djurhuus et al., 2018; Schroeder et al., 2020), where metabarcoding detected more genera than morphology alone. Although metabarcoding has become widely used for detecting community diversity, its quantitative reliability remains a subject of debate as previous studies have often detected a poor correlation between sequence reads and abundance (Bucklin et al., 2016; Lamb et al., 2019; Santoferrara, 2019; van der Loos and Nijland, 2020). One of the causes of this discrepancy is the differences in size between organisms and their life-stage causing the variation of the amount of DNA present in each organism. Additionally, biases introduced during DNA extraction and polymerase chain reaction (PCR) procedures exacerbate this issue by causing uneven recovery and amplification of genetic material among organisms (Ershova et al., 2021). However, if such affects are organism specific (and consistent for metabarcoding analysis) then the method should be able to detect changes in presence and relative abundance of different taxa between samples and over

time. But again, this may contribute to a lack of correlation for relative abundance of taxa measured by morphological analysis.

The additional detections observed through metabarcoding could plausibly be attributed to cryptic variation within zooplankton communities. Cryptic species, which are genetically distinct but morphologically similar, often elude traditional morphological identification methods. Metabarcoding, which employs high-throughput sequencing of specific genetic markers, enables the detection of these subtle genetic differences, thereby revealing a greater extent of biodiversity than might be discerned through conventional approaches (Hebert et al., 2004). Cryptic variation is particularly prevalent among zooplankton, a group characterised by vast diversity and significant challenges in morphological identification. Zooplankton species frequently exhibit phenotypic plasticity, wherein environmental factors induce morphological changes, further complicating species identification using traditional methods (Adamowicz et al., 2009). Metabarcoding circumvents these limitations by targeting genetic sequences, which can uncover hidden diversity within what might otherwise be classified as a single species based solely on morphological traits. This enhanced detection capability is crucial for a comprehensive understanding of zooplankton biodiversity and the ecological roles these organisms play. Cryptic species may exhibit differing ecological functions, including varied responses to environmental stressors or distinct roles within the aquatic food web (Pfenninger & Schwenk, 2007). Thus, the additional detections facilitated by metabarcoding not only broaden the understanding of zooplankton diversity but also provide deeper insights into the intricate ecological interactions within aquatic ecosystems.

7.1.3 Validation of analysis methods by application to a pilot-ecotoxicity study

7.1.3.1 Glyphosate effects on Zooplankton at genera, group and community level

This study reported for the first time the use of Roundup™ on zooplankton genera, groups, and communities in a mesocosm study. Glyphosate is the most frequently used herbicide both worldwide and, in the EU, and it has been used for several decades (Tarazona et al., 2017; Goncalves et al., 2019; Vera et al., 2012). However, there are still very few studies on its effect on zooplankton (Geyer et al., 2016) and in the majority of cases is always combined with another pesticide, insecticide or herbicide (Hebert et

al.,2020; Andrade et al., 2021). Out of these studies, most have been performed in laboratory facilities (Portinho et al., 2018; Lopez-Valcarcel et al., 2024). This is because setting up ecotoxicity studies in the laboratory, which is a simplified and controlled setting, facilitates the identification of modes of action and threshold (Hebert et al., 2020). Usually, laboratory testing only focuses on a single species overlooking the direct and indirect effects of the chemical at the community level (Rohr et al., 2006). According to Hebert et al., (2020), it is important to study the effects of the chemical by acknowledging the wider influence of species and trophic interactions on the reactions to pollutants; comprehending the interplay among concurrent agrochemicals; integrating temporal fluctuations to evaluate time-related impacts of agrochemicals or community dynamics, including recuperation intervals; and ultimately, factoring in physicochemical variables that may influence the movement, toxicity, and breakdown of agrochemicals, especially pesticides. In-field experiments performed in mesocosms can replace and complement experiments performed in laboratory facilities by addressing the previously stated points (Gessner and Tlili, 2016).

The application of pesticides in the natural environment is affected by environmental factors such as temperature, pH, conductivity, nutrients, and organic supply. The present study analysed the effects of Glyphosate on zooplankton genera, groups, and communities in mesocosms. Mesocosm experiments provide one of the most accurate representations of real conditions in ecotoxicity testing. They facilitate the study of interactions in aquatic bodies in the long term by minimising environmental variability (Kohler and Triebskorn, 2013). However, it is important to design them with the right system complexity and conditions. Several reports have demonstrated the efficacy of mesocosm use in aquatic ecotoxicology studies (Carpenter and Lodge, 1986; Loerracher et al., 2023). In this study, the glyphosate target doses in Glyphosate were set at 5.5 mg a.i./L (low dose), and 22.0 mg a.i./L (high dose). These doses were chosen to ensure minimal impact on zooplankton at the low dose, while the high dose aimed for a significant reduction or decline in the zooplankton community. The set concentrations aimed to target zooplankton as individuals and as a community, with effects expected within 24 hours of application as selected doses exceeded the EC_{50} on *Daphnia Magna* as indicated in the safety data sheet of the chemical (12.5 mg a.i./L) and were selected based on previous studies (Hébert et al., 2020). Dosing concentrations were anticipated to affect each zooplankton genus differently, with Glyphosate

expected to cause significant biomass decreases in certain genera at specific concentrations. Most studies on zooplankton genera have focused on *Daphnia* in laboratory experiments, therefore a direct comparison between data found in this study and the literature can be made by knowing that as stated before, laboratory and in-field analyses are affected differently by the environment. For the water quality environmental variables measured, significant differences were found in temperature, dissolved oxygen, and pH post-application of the higher dose of the pesticide. Although temperature also varied pre-application, the overall data indicate that the high dose of Glyphosate altered the physio-chemical properties of the water.

7.1.3.1.1 Glyphosate impacts in taxa

Morphological analyses showed a significant difference in *Daphnia* abundance between the control and the high Glyphosate concentration (22.0 mg a.i./L). However, no significant difference was found between the control and either of the two Glyphosate doses for *Daphnia* in the metabarcoding analyses. Some studies have found higher chronic toxicity throughout the life cycle of *Daphnia magna* at very low test concentrations of glyphosate (as an active ingredient in Roundup^{TM}) (between 0.05 mg L^{-1} and 0.45 mg L^{-1}) with negative outcomes at the population level, including diminished size, decreased growth rate, infertility, and even instances of miscarriage, can ensue (Lopez-Vakcarcel et al., 2023). In other studies, using glyphosate (as an active ingredient in Roundup^{TM}), a significant decrease in *Daphnia* was found at LC_{50} 11 mg a.i./L (Baker et al., 2016), and LC_{50} after 96h exposure time in *Daphnia magna* at 14 mg a.i./L (Lares et al., 2022).

At equivalent concentrations of the active ingredient, Roundup^{TM} exhibits greater toxicity compared to glyphosate, resulting in adverse effects on survival, behaviour, and the majority of physiological characteristics (Janssens and Stoks, 2017), which suggests that in this case, other substances present in the commercial formula contribute through synergistic effects to increase the toxicity towards non-target organisms (Rico-Martinez et al., 2012).

7.1.3.1.2 Glyphosate impacts on community composition

In this study, the morphological analysis showed that Glyphosate doses did not significantly alter zooplankton community composition according to Detrended

Correspondence Analysis and Principal Response Curve Analysis. However, there were noticeable differences in clustering patterns among sampling points, indicating some impact of Glyphosate. Seasonal changes were highlighted as a major influence, with shifts observed in dominant zooplankton groups throughout the year. Although Glyphosate had varied effects on different genera, significant differences were not always detected in univariate analyses. Contrastingly, the metabarcoding analyses revealed a significant distinction in community composition. Clustering patterns at different sampling points indicated a clear effect of Glyphosate, evident in low and high doses. The ANOSIM confirmed significant variations between sampling points globally, although similar variations were observed in the control group. Furthermore, the Shannon index highlighted higher diversity post-application, especially in control and low doses, which is likely influenced by zooplankton seasonality. Notably, the Shannon index indicated lower diversity of zooplankton post-application with a high dose of Glyphosate, indicating its impact on zooplankton.

In summary, while morphological analyses did not show significant community composition changes, metabarcoding provided evidence of Glyphosate's effect on zooplankton communities. The seasonality played a crucial role, the study indicated a NOEC of less than5.5 mg a.i./L and the metabarcoding analysis offered detailed insights into the impact of Glyphosate on specific genera.

7.1.3.2 Efficiency of Metabarcoding in mimicking Morphological Analyses in an ecotoxicological trial

The primary objective of this study was to compare the efficacy of metabarcoding against morphological analysis in a pesticide trial. While metabarcoding did not mirror the exact significant differences observed for each genus and family of morphological identification, both approaches concurred that a dosage lower than the NOEC (lower than 5.5 mg a.i./L) is preferable for Glyphosate application. In summary, the NOEC represents the concentration of a substance at which no statistically significant effects are observed, and it is a key parameter in assessing the environmental impact of pesticides in microcosm and mesocosm studies (Brook et al., 2015). The NOEC is still commonly used in ecotoxicology studies; however, many recent studies claim it is not a reliable toxicity metric, especially in cases where the concentration-response relationship is shallow, control responses are highly variable, or there is no discernible

pattern in the response except at high concentrations (Fox and Landis, 2016; Jager, 2012; Landis and Chapman 201; Crane and Newman 2000). The assertion that the NOEC can overcome these challenges is deemed unconvincing, as it essentially responds to noise in the data and lacks meaningful statistical support. Regardless of these criticisms, the NOEC is still extensively used and prioritized by EFSA. The NOEC can be calculated using both the Dunnett and William tests. The main difference between them is that the Williams test assumes a monotonic dose–response relationship, while such an assumption is not made for the Dunnett test (Jaki and Hothorn, 2013). The NOEC in this study was calculated using Dunnett's test as no direct effect of the herbicide Glyphosate was assumed on zooplankton and so a monotonic dose-response may not occur (Jaki and Hothorn, 2013).

In this study, the results obtained underscored the various responses of zooplankton genera to the different doses of Glyphosate. Some taxa showed resilience to the two Glyphosate doses while others showed higher sensitivity. This reinforced the fact that studying the effects of community composition instead of focusing only on one genus is extremely important.

7.1.4 In field molecular analysis: "potential and limitations"

The Oxford Nanopore Technology (ONT) NGS platform was applied to short-read fragments on Zooplankton mixed DNA samples. This represents a recent scientific approach as not many studies have applied this sequencing method to zooplankton samples in ecotoxicity studies. The integration of ONT into ecotoxicological studies has opened new frontiers in decoding the genomic responses of zooplankton to environmental stressors. By providing real-time, long-read sequencing capabilities, ONT enables a comprehensive exploration of the intricate genetic makeup of these key aquatic organisms. The use of ONT in ecotoxicology aligns with the broader shift towards single-molecule sequencing methodologies (Goodwin et al., 2016). However, this approach is still new and constantly updating with little to no references on MinIon barcoding on zooplankton short fragment reads. Despite the promise shown by Nanopore sequencing in metabarcoding, its adoption has been slowed down by two main factors. Firstly, a high error rate (Tyler et al. 2018; Pfeiffer et al., 2018), potentially impacting the reliability of species detection outcomes. Secondly, the scarcity of

programs tailored to process nanopore reads for metabarcoding (Chang et al., 2023). In the current study, quality filtering of the sequenced reads was performed following Nanopore sequencing, resulting in the retention of an average of just 16% of the initial raw reads. This is broadly similar to other studies, such as 37% retention reported by Menegon et al. (2017) and 50% by Semmouri et al. (2020). The low conversion of raw reads to indexed sequences is likely due to a high error rate. Errors can arise from DNA template degradation, amplification errors, and sequencing errors, with the magnitude of sequencing errors varying across different sequencing technologies (Semmouri et al., 2021).

Differing from conventional sequencing technologies, the MinION offers portability, enabling its usage beyond the confines of laboratories and facilitating field applications (Semmouri et al., 2020). As highlighted by Dumschott et al., (2020), the MinION can be easily connected to a standard laptop via USB 3.0 and weighs a mere 103 g, allowing sequencing activities to be conducted at various locations with power access. Examples of its versatile deployment include field studies for identifying closely related plant species in Snowdonia National Park (Parker et al., 2017), on-site analysis of Ebolavirus samples in West Africa (Quick et al., 2016), and identification of Cassava virus strains on farms in East Africa (Boykin et al., 2018).

Similarly to Illumina platforms, MinION offers the capability to detect multiple species from multiple samples in a single sequencing operation. The potential advantages of employing the MinION include the ability to process samples independently without relying on external services, handle small batches of samples to shorten turnaround time, conduct on-site sequencing using its portable nature, and generate longer reads averaging 8 kbp for optimal sequencing (Egeter et al., 2022). However, a primary potential drawback of MinION usage is its tendency to yield higher DNA sequencing error rates as discussed above. Additionally, there might be increased costs per sample and reduced scalability with MinION usage, although this aspect requires further exploration in its application to species detection. Overcoming existing limitations could enable the MinION to significantly reduce sample turnaround time, thereby minimizing the delay between sample acquisition and obtaining results.

Nanopore technology can sequence considerably longer reads compared to Illumina technology. Some eDNA studies have successfully sequenced fragments longer than

the typical 50–250 bp fragments, suggesting the potential utility of nanopore technology for sequencing such amplicons in future studies (Egeter et al., 2022 and Doorenspleet et al., 2021). As mentioned, another potential advantage of the MinION sequencer is its portability, enabling DNA sequencing in a portable laboratory setting. While this approach was not attempted in our current study, the laboratory protocols and bioinformatic software provided could be adapted to such conditions, as they do not necessarily rely on bulky equipment or an active internet connection. This would facilitate obtaining rapid on-site or in-field results, with the added benefit of minimizing potential laboratory-induced contamination. This could potentially be useful in ecotoxicity studies that are conducted in remote areas (Parker et al., 2017). Overall, the potential of Nanopore sequencing to perform in-field analyses will have an extremely important impact on some remote ecotoxicity studies.

In terms of bioinformatics analysis of metabarcoding data generated by Nanopore, there are fewer available tools compared to those for Illumina data. Furthermore, there are few comprehensive tools for metabarcoding utilising MinION workflows. At the time of the study, an R10.4.1 flow cell was utilised, but newer flow cell versions with improved sequencing quality are now available, which are likely to streamline bioinformatic processing and subsequent taxonomic assignment, thus facilitating the integration of MinION as a routine tool for biomonitoring (Egeter et al., 2022).

One significant advantage of employing MinION for metabarcoding compared to the Illumina platform is the flexibility to end sequencing runs before exhausting a flow cell, enabling multiple runs from a single flow cell. This feature is particularly advantageous when dealing with small sample sizes, as a small pool of samples can be processed, loaded, and sequenced independently without the need to wait for additional samples to fill a run (Chang et al., 2023). However, it is important to note that when reusing a flow cell, despite employing a DNase wash, there may still be some carry-over contamination from previous runs. Therefore, it is recommended to utilise different barcodes to enhance the detection and removal of carry-over contamination. In contrast, with Illumina-based approaches, smaller projects often remain on hold until a sufficient number of libraries are prepared from other projects to justify utilizing a full MiSeq/HiSeq flow cell (Egeter et al., 2022). Moreover, Nanopore turnaround times were significantly faster; in fact, completing the one run of 80 samples, inclusive of sample preparation took only three days, compared to the approximately 2–4 weeks

required for Illumina sequencing. Even in the context of zooplankton biomonitoring, where sampling intervals may occur as frequently as every two weeks, a nanopore-based metabarcoding approach would facilitate the rapid generation of results, thereby enhancing the operational feasibility of proposed routine biomonitoring strategies (Chang et al., 2023).

A final point to consider was the cost between the two methodologies as nanopore metabarcoding with the MinION is generally cheaper than Illumina MiSeq (Van der Reis et al., 2022; Stevens et al., 2023). Here, reagent costs were reduced further by pooling the already tagged library from the Illumina barcodes. However, the current study revealed that Nanopore metabarcoding did not produce metabarcodes of Illumina-like quality. The most significant distinction was observed in taxonomic units richness, where 28 Illumina genera were obtained compared to 11 nanopore generas with no mismatches, and a further 4 genera with one mismatch.

Moving forward, there is a clear trajectory for improvement. Recent advancements in nanopore read accuracy, coupled with new bioinformatic pipelines, suggest that nanopore sequencing could yield more accurate metabarcoding outcomes that are approaching conventional Illumina sequencing accuracy, without the need for extensive sequence polishing as previously required, especially for long-read fragments. Further evaluation studies will be necessary to assess the impact of these enhancements on metabarcoding. While MinION's lower per-base accuracy presents bioinformatic challenges, ongoing technological advancements, and the development of bioinformatics tools are progressively mitigating this issue.

7.1.4.1 Comparative Analysis of Nanopore and Illumina Sequencing Technologies: Balancing Cost, Time Efficiency, and Accuracy in Metabarcoding

When evaluating the comparative advantages of nanopore and Illumina sequencing technologies in the context of metabarcoding, it is essential to consider the dimensions of time efficiency, cost-effectiveness, and accuracy. Cost considerations are a significant factor, with nanopore sequencing, particularly using the MinION platform, emerging as a more economical option compared to Illumina MiSeq. This cost advantage has been substantiated by several studies (Van der Reis et al., 2022; Stevens et al., 2023), which highlight that the overall expenditure for a nanopore sequencing run

is substantially lower, particularly when reagent costs are minimised by pooling pre-tagged libraries from Illumina barcodes. Specifically, excluding the initial investment in the MinION device, the total cost of a nanopore run for 96 samples is £2627.20, compared to £3516.37 for an equivalent Illumina run.

In terms of time efficiency, nanopore sequencing provides a marked advantage, with a complete run, including sample preparation, being accomplished within three days. This is significantly faster than the two to four weeks typically required for Illumina sequencing. However, these benefits in cost and speed must be carefully weighed against the issue of accuracy. Nanopore sequencing is characterised by a higher error rate, ranging from 20% to 38% (Chang et al., 2023), which is considerably higher than the 0.24% error rate associated with Illumina sequencing. This disparity in error rates raises concerns regarding the reliability of species identification and the overall precision of nanopore-based metabarcoding. Consequently, while nanopore sequencing offers substantial advantages in terms of reduced costs and expedited processing times, the significantly higher error rate necessitates careful consideration and potentially supplementary validation steps to ensure the accuracy and integrity of the results obtained.

7.2 Limitations of the study

One of the significant limitations of this study was the use of an herbicide (Glyphosate) instead of an insecticide. This decision was made due to the affordability of the product, as 4,650 L (775 L per six mesocosms) of water had to be covered with the test solution. As Glyphosate is a herbicide, a non-direct effect was expected on zooplankton compared to a direct effect that an insecticide could have had. Indirect effects of herbicides are defined as observed effects on consumer populations in freshwater invertebrates that are not caused by direct toxicity but due to adverse effects on primary producers such as algae and macrophytes (Rico-Martinez et al., 2012). Specifically, glyphosate acts by inhibiting the enzyme 5-Enolpyruvyl Shikimicacid 3-phosphate synthase (EPSPS) which disrupts the amino acids' aromatic biosynthesis, reducing protein synthesis and growth and eventually causing cellular disruption and death in phytoplankton and algae (Vera et al., 2012). On the other hand, insecticides target insects' nervous systems or insect-specific processes such as moulting, inducing the

development of diverse ecotoxicological tests to evaluate their effects (Duke et al., 2023). In addition, accurately assessing the effects of herbicides and insecticides is challenging due to the complexity of natural ecosystems, which encompass diverse interactions and factors. (Hanazato, 2001).

Also, this was a pilot study and not a complete ecotoxicity trial. The limitations of this study were the number of mesocosms used and thus doses to be tested. In a complete ecotoxicity study usually, multiple doses of toxicant are applied to ensure a more detailed dose-response analysis in community composition. For example, Lopez-Mancisidor et al., (2008), tested four Chlorphyrifos nominal concentration using three replicates per treatment. The controls used were five replicates to reduce background noise in statistical analyses. Finnegan et al., (2018), tested five doses of Thiamethoxan with four controls. Also, Hebert et al., (2020) applied seven nominal concentrations of glyphosate and three controls. In fact, in this study, the suggested NOEC concentration resulted to be < 5.5 mg a.i./L (lowest dose). If more doses were tested, a more specific NOEC could have been found. Furthermore, as a herbicide was applied, phytoplankton sampling and identification would have contributed to a better understanding of the efficacy of the Glyphosate in the mesocosms.

In the current study, a further limitation was clearly seen when using Nanopore to analyse short reads (270 bp). The Nanopore platform is mostly suited for long reads, reaching up to 4 M reads (© 2008 - 2024 Oxford Nanopore Technologies plc). Results showed poor-quality reads compared to Illumina sequencing. However, recent studies have shown good-quality reads when sequencing long-read zooplankton genomes (Semmouri et al., 2021).

Definitively, the completeness of reference databases still plays a major role in the reliability of DNA metabarcoding studies, regardless of the sequencing technology in use, (Nistal-Garcia et al., 2021). Zooplankton sequences on databases for the 18s gene are still relatively scarce and not all species are present. This led to the use of an incomplete reference database which was likely to contribute to discrepancies when compared with the traditional microscopy data. Finally, a further limitation was the quality of the mixed zooplankton samples as the turbidity has led to difficulties in identifying the zooplankton morphologically, possibly leading to genus misidentification.

7.3 Future prospects

Future work should include the repetition of an ecotoxicity study with all the variants usually used to assess the effect of the pesticide. This includes the use of an insecticide to assess the direct effects of the toxicant on the zooplankton. By doing so, a possible wipe-out of the zooplankton community for the high dose could be obtained and a comparison of the recovery between the two methodologies could be studied. Appropriate examples would include the originally proposed insecticides that were discarded due to their affordability: Imidacloprid (Merga and Brink, 2021) and Acetamipirid (Gacem et al., 2022). Also, a higher number of nominal concentrations of the toxicant and more repeats for each one of them (including the control) would be recommended to enhance precision, increase signal-to-noise ratio, and improve the probability of detecting active effects. The current aquatic guidance document (EFSA, 2013), recommends the use of at least five test concentrations with at least two replicates per treatment. Moreover, a higher number of control replicates should be looked at compared to the nominal concentrations (Brock et al., 2015). Such a future mesocosm trial could be used to effectively demonstrate the ability of metabarcoding to enhance and/or replace traditional morphological studies.

Furthermore, new assessments using Nanopore sequencing could be tested when the technology becomes more suitable for accurate short-fragment regions, or by applying long-reads to overcome relatively high sequence error rates.

7.4 Concluding remarks

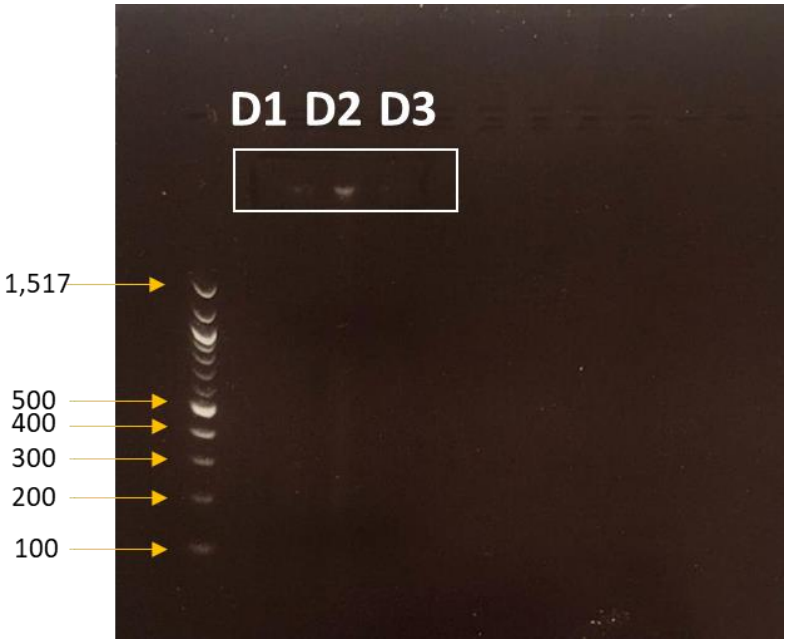
The metabarcoding analysis using the 18s gene resulted in an efficient way to study zooplankton diversity, as metabarcoding identified more taxa compared to the morphological analyses and at a higher taxonomic level. Simultaneously, various constraints were encountered, for example, some taxa present in the given list from CEA were not identifiable through metabarcoding, and the relative abundance did not consistently match the observed densities. Analyses applied to the seasonal mesocosm study showed no correlation between the selected eight genera abundance, however, a similar number of genera were detected. A total of 26 genera were found, 17 of them were identified by both methodologies, 3 only by morphology and 6 only by metabarcoding. The analyses applied in the pilot ecotoxicity study showed that the

metabarcoding did not mimic the same significant differences found for each genus and family by morphological analysis, but both methodologies established that a dose lower than the selected low-dose (5.5 mg a.i./L) is needed to observe 'no effect' for glyphosate. The metabarcoding analysis identified a higher number of genera (28) compared to morphological analysis (23). Both methodologies identified a total of 22 genera with metabarcoding identifying an additional 6 genera and morphological analysis identifying a single additional genus. The comparison of the two detection methodologies in the exotoxicity trial indicated that the detection of most genera demonstrated a strong correlation between the two methods (5 out of 8 genera tested).

Overall, it was deduced that by combining both morphology and metabarcoding techniques, more extensive results can be obtained. However, metabarcoding detected more genera than morphological analyses and an effect was found as expected for the glyphosate doses, indicating that metabarcoding could potentially replace traditional morphological analysis. Benefits would include: less time taken for analysis; and also, removal of the misidentification problems that occur in morphological analysis caused by poor quality samples.

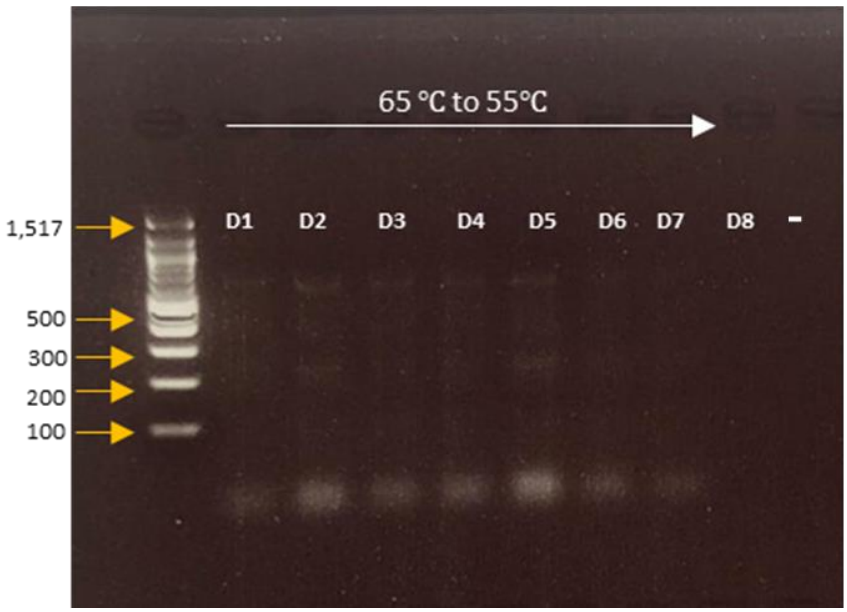
Despite the potential of metabarcoding to enhance our understanding of global plankton diversity patterns, it remains imperative to uphold expertise in morphological taxonomic identification of zooplankton. This ensures the ongoing validation and enhancement of metabarcoding methodologies via continued improvement of sequence databases. An integrative approach, combining both morphological and molecular methods, is advocated for a thorough evaluation and science-driven management of diversity within marine ecosystems.

Appendix



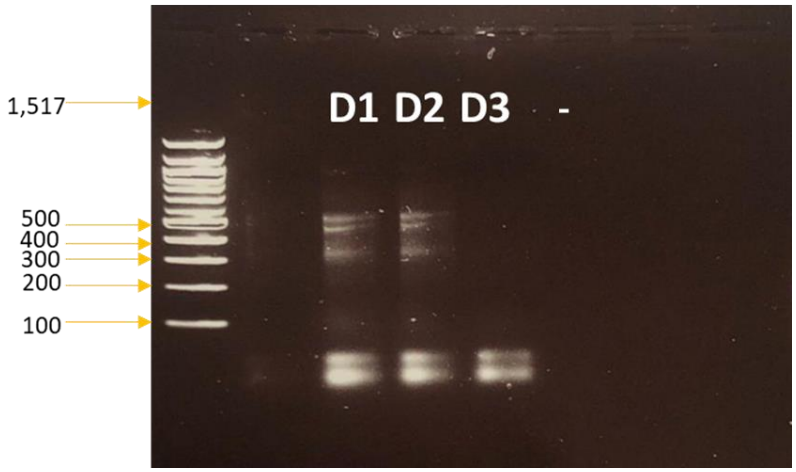
S1: *Daphnia* DNA-Genomics electrophoresis.

DNA-extraction of *Daphnia* samples (three replicates) on gel electrophoresis without qPCR in order to check DNA integrity. The boxed area represents the loaded DNA in triplicates.



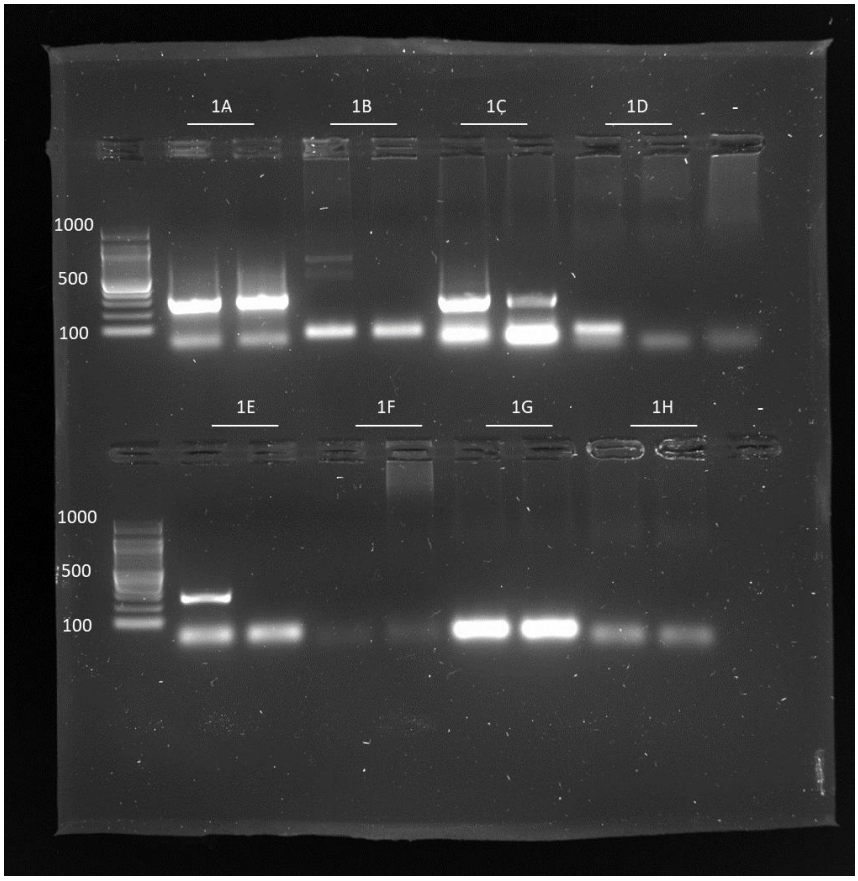
S2: Amplification of *Daphnia* using COI Primer 2 genus-specifc primer.

Gel electrophoresis of the gradient qPCR (from 65°C to 55°C) on 8 *Daphnia* replicates (expected fragment size 170 bp) using the second COI genus-specific primer-set (COI Primer_2) and 1Kb ladder.



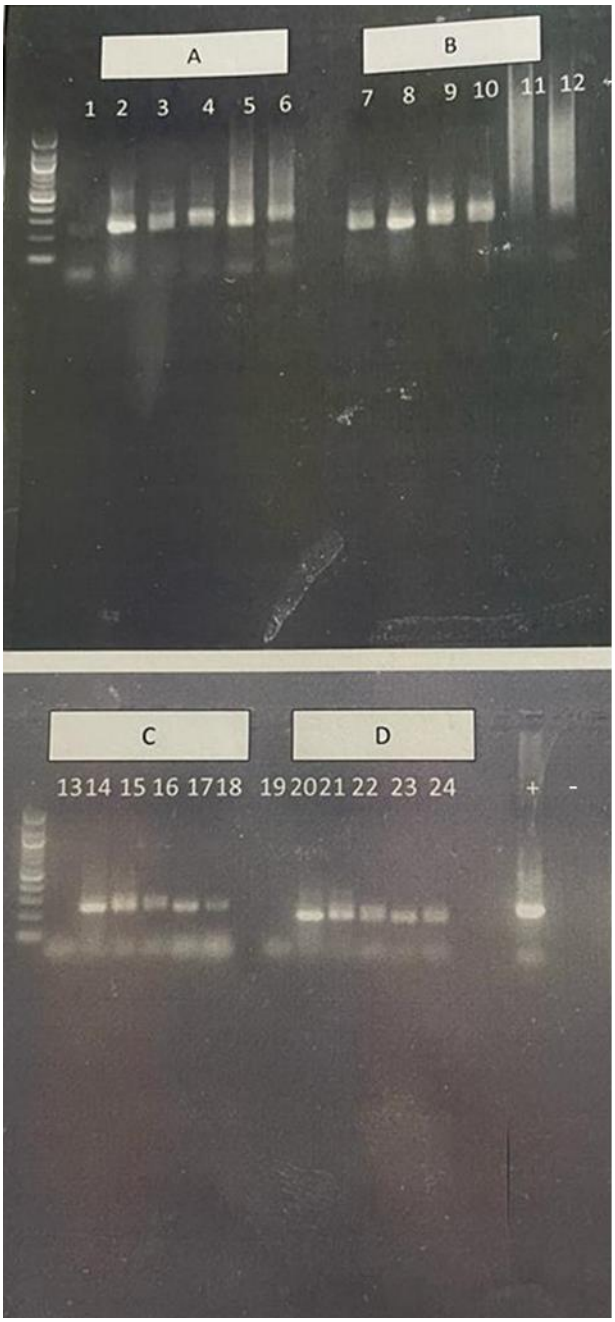
S3: Amplification of three *Daphnia* replicates using the third genus-specific primer-set "COI Primer_3.

A 1Kb ladder was used and the expected fragment size was 204 bp.



S4: PCR amplification of cultured Daphnia Magna provided by CEA.

Samples 1A-1D were analysed using a touchdown PCR while samples 1E-H were analysed using the PCR protocol from Bucklin et al., (2019). Also, samples 1A,1B,1E and 1F were analysed using the Q5 mastermix while samples 1C,1D,1G and 1H were analysed using the environmental mastermix. Finally, two different primer sets were tested: Bucklin et al., (2019) for samples (1A,1C,1E and 1G) while Djurhuus et al., 2018 (1B,1D,1F and 1H). 10 µL PCR product loaded per well; 3 µL 100 bp Ladder loaded. PCR negative controls (-) were negative for amplification. Samples 1A, 1C and 1E successfully amplified at the right size (229 bp Djurhuus and 275bp Bucklin).



S5: PCR amplification of randomly selected mesocosm samples from day 1 (ID:1,4,7,10,13,16,19 and 22), from day 4 (ID:3,5,8,11,14,17,20 and 23) and finally from day 7 (ID:4,6,9,13,15,18,21 and 24).

Samples were analysed using two different primer sets: “A” and “C” using Bucklin et al., (2019); and “B” and “D” using Djurhuus et al., (2018). Samples “A” and “B” were tested using a touchdown PCR program (Don 1991) while amplification in “C” and “D” was carried out using the Bucklin protocol.Also, samples 1-3, 7-9, 13-15 and 19-21 were samples before the dilutions while samples 2-6, 10-12, 16-18 and 22-24 were diluted at 5ng/ul according to Bucklin et al.,(2019). 10 µL PCR product loaded per well; 3 µL 100 bp Ladder loaded. PCR negative controls (-) were negative for amplification. Samples successfully amplified at the right size (229 bp Djurhuus and 275 bp Bucklin).

S6 Metagenomics

DEEP Seq ID	Sample Name	Concentration post Amp PCR	Volume	PCR Cycles	Index Kit	I7_Index_ID	index	I5_Index_ID	index 2	Index Plate Position	Final Lib Qubit (ng/ul)
ds1231_1	PA11	5.68	20	12	IDT for Illumina UD Index set B	UDP0097	TGCCGGTCAG	UDP0097	CCTGATACAA	A01	0.39
ds1231_2	PA12	14.45	20	12	IDT for Illumina UD Index set B	UDP0098	CACTCAATTC	UDP0098	TTAAGTTGTG	B01	7.31
ds1231_3	PA13	14.18	20	12	IDT for Illumina UD Index set B	UDP0099	TCTCACACGC	UDP0099	CGGACAGTGA	C01	4.66
ds1231_4	PA14	14.27	20	12	IDT for Illumina UD Index set B	UDP0100	TCAATGGAGA	UDP0100	GCACTACAAC	D01	1.21
ds1231_5	PA15	9.59	20	12	IDT for Illumina UD Index set B	UDP0101	ATATGCATGT	UDP0101	TGGTGCCTGG	E01	3.02
ds1231_6	PA16	18.28	20	12	IDT for Illumina UD Index set B	UDP0102	ATGGCGCCTG	UDP0102	TCCACGGCCT	F01	1.06
ds1231_7	PA17	9.18	20	12	IDT for Illumina UD Index set B	UDP0103	TCCGTTATGT	UDP0103	TTGTAGTGTA	G01	5.29
ds1231_8	PA18	11.83	20	12	IDT for Illumina UD Index set B	UDP0104	GGTCTATTAA	UDP0104	CCACGACACG	H01	0.62
ds1231_9	PA19	12.62	20	12	IDT for Illumina UD Index set B	UDP0105	CAGCAATCGT	UDP0105	TGTGATGTAT	A02	7.75
ds1231_10	PA21	10.91	20	12	IDT for Illumina UD Index set B	UDP0106	TTCTGTAGAA	UDP0106	GAGCGCAATA	B02	7.02
ds1231_11	PA22	77.13	20	12	IDT for Illumina UD Index set B	UDP0107	GAACGCAATA	UDP0107	ATCTTACTGT	C02	0.008
ds1231_12	PA23	9.98	20	12	IDT for Illumina UD Index set B	UDP0108	AGTACTCATG	UDP0108	ATGTCGTGGT	D02	11.9
ds1231_13	PA24	15.08	20	12	IDT for Illumina UD Index set B	UDP0109	GGTAGAATTA	UDP0109	GTAGCCATCA	E02	14.4
ds1231_14	PA25	10.76	20	12	IDT for Illumina UD Index set B	UDP0110	TAATTAGCGT	UDP0110	TGGTTAAGAA	F02	10.2
ds1231_15	PA26	22.07	20	12	IDT for Illumina UD Index set B	UDP0111	ATTAACAAGG	UDP0111	TGTTGTTCGT	G02	9.89
ds1231_16	PA27	12.66	20	12	IDT for Illumina UD Index set B	UDP0112	TGATGGCTAC	UDP0112	CCAACAACAT	H02	1.2
ds1231_17	PA28	20.91	20	12	IDT for Illumina UD Index set B	UDP0113	GAATTACAAG	UDP0113	ACCGGCTCAG	A03	11.2
ds1231_18	PA29	34.7	20	12	IDT for Illumina UD Index set B	UDP0114	TAGAATTGGA	UDP0114	GTTAATCTGA	B03	0.1

ds1231_19	PA31	7.01	20	12	IDT for Illumina UD Index set B	UDP0115	AGGCAGCTCT	UDP0115	CGGCTAACGT	C03	12
ds1231_20	PA32	3.88	20	12	IDT for Illumina UD Index set B	UDP0116	ATCGGCGAAG	UDP0116	TCCAAGAATT	D03	8.74
ds1231_21	PA33	36.88	20	12	IDT for Illumina UD Index set B	UDP0117	CCGTGACCGA	UDP0117	CCGAACGTTG	E03	7.58
ds1231_22	PA34	9.94	20	12	IDT for Illumina UD Index set B	UDP0118	ATACTTGTTC	UDP0118	TAACCGCCGA	F03	7.08
ds1231_23	PA35	6.62	20	12	IDT for Illumina UD Index set B	UDP0119	TCCGCCAATT	UDP0119	CTCCGTGCTG	G03	2.74
ds1231_24	PA36	3.11	20	12	IDT for Illumina UD Index set B	UDP0120	AGGACAGGCC	UDP0120	CATTCCAGCT	H03	6.39
ds1231_25	PA37	9.75	20	12	IDT for Illumina UD Index set B	UDP0121	AGAGAACCTA	UDP0121	GGTTATGCTA	A04	7.3
ds1231_26	PA38	4.21	20	12	IDT for Illumina UD Index set B	UDP0122	GATATTGTGT	UDP0122	ACCACACGGT	B04	7.43
ds1231_27	PA39	1.2	20	12	IDT for Illumina UD Index set B	UDP0123	CGTACAGGAA	UDP0123	TAGGTTCTCT	C04	5.6
ds1231_28	AD1	2.22	20	12	IDT for Illumina UD Index set B	UDP0124	CTGCGTTACC	UDP0124	TATGGCTCGA	D04	5.25
ds1231_29	AD2	2.26	20	12	IDT for Illumina UD Index set B	UDP0125	AGGCCGTGGA	UDP0125	CTCGTGCGTT	E04	5.12
ds1231_30	AD3	5.64	20	12	IDT for Illumina UD Index set B	UDP0126	AGGAGGTATC	UDP0126	CCAGTTGGCA	F04	4.2
ds1231_31	AD4	9.54	20	12	IDT for Illumina UD Index set B	UDP0127	GCTGACGTTG	UDP0127	TGTTCGCATT	G04	7.14
ds1231_32	AD5	7.84	20	12	IDT for Illumina UD Index set B	UDP0128	CTAATAACCG	UDP0128	AACCGCATCG	H04	3.12
ds1231_33	AD6	9.71	20	12	IDT for Illumina UD Index set B	UDP0129	TCTAGGCGCG	UDP0129	CGAAGGTTAA	A05	3.62
ds1231_34	AD7	12.2	20	12	IDT for Illumina UD Index set B	UDP0130	ATAGCCAAGA	UDP0130	AGTGCCACTG	B05	12.4
ds1231_35	AD8	29.97	20	12	IDT for Illumina UD Index set B	UDP0131	TTCGGTGTGA	UDP0131	GAACAAGTAT	C05	12.4
ds1231_36	AD9	26.36	20	12	IDT for Illumina UD Index set B	UDP0132	ATGTAACGTT	UDP0132	ACGATTGCTG	D05	22.2
ds1231_37	PO11	6.5	20	12	IDT for Illumina UD Index set B	UDP0133	AACGAGGCCG	UDP0133	ATACCTGGAT	E05	22.3
ds1231_38	PO12	6.26	20	12	IDT for Illumina UD Index set B	UDP0134	TGGTGTTATG	UDP0134	TCCAATTCTA	F05	18
ds1231_39	PO13	80.24	20	12	IDT for Illumina UD Index set B	UDP0135	TGGCCTCTGT	UDP0135	TGAGACAGCG	G05	0.009
ds1231_40	PO14	5.5	20	12	IDT for Illumina UD Index set B	UDP0136	CCAGGCACCA	UDP0136	ACGCTAATTA	H05	1.93
ds1231_41	PO15	17.73	20	12	IDT for Illumina UD Index set B	UDP0137	CCGGTTCCTA	UDP0137	TATATTCGAG	A06	1.07
ds1231_42	PO16	3.21	20	12	IDT for Illumina UD Index set B	UDP0138	GGCCAATATT	UDP0138	CGGTCCGATA	B06	13.8

ds1231_43	PO17	0.61	20	12	IDT for Illumina UD Index set B	UDP0139	GAATACCTAT	UDP0139	ACAATAGAGT	C06	8.49
ds1231_44	PO18	2.86	20	12	IDT for Illumina UD Index set B	UDP0140	TACGTGAAGG	UDP0140	CGGTTATTAG	D06	12.7
ds1231_45	PO19	6.94	20	12	IDT for Illumina UD Index set B	UDP0141	CTTATTGGCC	UDP0141	GATAACAAGT	E06	7.24
ds1231_46	PO21	10.28	20	12	IDT for Illumina UD Index set B	UDP0142	ACAACTACTG	UDP0142	AGTTATCACA	F06	22.2
ds1231_47	PO22	5.43	20	12	IDT for Illumina UD Index set B	UDP0143	GTTGGATGAA	UDP0143	TTCCAGGTAA	G06	15.5
ds1231_48	PO23	4.97	20	12	IDT for Illumina UD Index set B	UDP0144	AATCCAATTG	UDP0144	CATGTAGAGG	H06	15.1
ds1231_49	PO24	3.86	20	12	IDT for Illumina UD Index set B	UDP0145	TATGATGGCC	UDP0145	GATTGTCATA	A07	13.2
ds1231_50	PO25	7.47	20	12	IDT for Illumina UD Index set B	UDP0146	CGCAGCAATT	UDP0146	ATTCCGCTAT	B07	26.5
ds1231_51	PO26	14.98	20	12	IDT for Illumina UD Index set B	UDP0147	ACGTTCCTTA	UDP0147	GACCGCTGTG	C07	14.7
ds1231_52	PO27	15.09	20	12	IDT for Illumina UD Index set B	UDP0148	CCGCGTATAG	UDP0148	TAGGAACCGG	D07	15.4
ds1231_53	PO28	6.85	20	12	IDT for Illumina UD Index set B	UDP0149	GATTCTGAAT	UDP0149	AGCGGTGGAC	E07	11
ds1231_54	PO31	4.73	20	12	IDT for Illumina UD Index set B	UDP0150	TAGAGAATAC	UDP0150	TATAGATTCG	F07	11.2
ds1231_55	PO32	11.06	20	12	IDT for Illumina UD Index set B	UDP0151	TTGTATCAGG	UDP0151	ACAGAGGCCA	G07	8.21
ds1231_56	PO33	14.54	20	12	IDT for Illumina UD Index set B	UDP0152	CACAGCGGTC	UDP0152	ATTCCTATTG	H07	22.7
ds1231_57	PO34	2.8	20	12	IDT for Illumina UD Index set B	UDP0153	CCACGCTGAA	UDP0153	TATTCCTCAG	A08	0.6
ds1231_58	PO35	1.11	20	12	IDT for Illumina UD Index set B	UDP0154	GTTCGGAGTT	UDP0154	CGCCTTCTGA	B08	1.68
ds1231_59	PO36	10.2	20	12	IDT for Illumina UD Index set B	UDP0155	ATAGCGGAAT	UDP0155	GCGCAGAGTA	C08	8.82
ds1231_60	PO37	6.43	20	12	IDT for Illumina UD Index set B	UDP0156	GCAATATTCA	UDP0156	GGCGCCAATT	D08	12.3
ds1231_61	PO38	12.5	20	12	IDT for Illumina UD Index set B	UDP0157	CTAGATTGCG	UDP0157	AGATATGGCG	E08	9.9
ds1231_62	PO39	9.46	20	12	IDT for Illumina UD Index set B	UDP0158	CGATGCGGTT	UDP0158	CCTGCTTGGT	F08	4.13
ds1231_63	PO51	3.31	20	12	IDT for Illumina UD Index set B	UDP0159	TCCGGACTAG	UDP0159	GACGAACAAT	G08	2.92
ds1231_64	PO52	1.35	20	12	IDT for Illumina UD Index set B	UDP0160	GTGACGGAGC	UDP0160	TGGCGGTCCA	H08	3.56
ds1231_65	PO53	6.76	20	12	IDT for Illumina UD Index set B	UDP0161	AATTCCATCT	UDP0161	CTTCAGTTAC	A09	11.9
ds1231_66	PO55	29.39	20	12	IDT for Illumina UD Index set B	UDP0162	TTAACGGTGT	UDP0162	TCCTGACCGT	B09	3.96

ds1231_67	PO56	6.03	20	12	IDT for Illumina UD Index set B	UDP0163	ACTTGTTATC	UDP0163	CGCGCCTAGA	C09	18.3
ds1231_68	PO57	7.76	20	12	IDT for Illumina UD Index set B	UDP0164	CGTGTACCAG	UDP0164	AGGATAAGTT	D09	3.24
ds1231_69	PO58	5.53	20	12	IDT for Illumina UD Index set B	UDP0165	TTAACCTTCG	UDP0165	AGGCCAGACA	E09	3.31
ds1231_70	PO59	2.09	20	12	IDT for Illumina UD Index set B	UDP0166	CATATGCGAT	UDP0166	CCTTGAACGG	F09	4.91
ds1231_71	PO61	10.33	20	12	IDT for Illumina UD Index set B	UDP0167	AGCCTATGAT	UDP0167	CACCACCTAC	G09	24.8
ds1231_72	PO62	2.2	20	12	IDT for Illumina UD Index set B	UDP0168	TATGACAATC	UDP0168	TTGCTTGTAT	H09	11
ds1231_73	PO63	2.85	20	12	IDT for Illumina UD Index set B	UDP0169	ATGTTGTTGG	UDP0169	CAATCTATGA	A10	9.4
ds1231_74	PO64	6.92	20	12	IDT for Illumina UD Index set B	UDP0170	GCACCACCAA	UDP0170	TGGTACTGAT	B10	18.7
ds1231_75	PO65	5.1	20	12	IDT for Illumina UD Index set B	UDP0171	AGGCGTTCGC	UDP0171	TTCATCCAAC	C10	16.4
ds1231_76	PO66	26	20	12	IDT for Illumina UD Index set B	UDP0172	CCTCCGGTTG	UDP0172	CATAACACCA	D10	10.8
ds1231_77	PO67	16.49	20	12	IDT for Illumina UD Index set B	UDP0173	GTCCACCGCT	UDP0173	TCCTATTAGC	E10	17.5
ds1231_78	PO68	10.88	20	12	IDT for Illumina UD Index set B	UDP0174	ATTGTTCGTC	UDP0174	TCTCTAGATT	F10	24
ds1231_79	PO69	15.63	20	12	IDT for Illumina UD Index set B	UDP0175	GGACCAGTGG	UDP0175	CGCGAGCCTA	G10	18.2
ds1231_80	PO71	1.15	20	12	IDT for Illumina UD Index set B	UDP0176	CCTTCTAACA	UDP0176	GATAAGCTCT	H10	6.2
ds1231_81	PO72	6.88	20	12	IDT for Illumina UD Index set B	UDP0177	CTCGAATATA	UDP0177	GAGATGTCGA	A11	15.6
ds1231_82	PO73	2.38	20	12	IDT for Illumina UD Index set B	UDP0178	GATCGTCGCG	UDP0178	CTGGATATGT	B11	11.3
ds1231_83	PO74	4.55	20	12	IDT for Illumina UD Index set B	UDP0179	TATCCGAGGC	UDP0179	GGCCAATAAG	C11	25.3
ds1231_84	PO75	3.55	20	12	IDT for Illumina UD Index set B	UDP0180	CGCTGTCTCA	UDP0180	ATTACTCACC	D11	22.3
ds1231_85	PO76	6.57	20	12	IDT for Illumina UD Index set B	UDP0181	AATGCGAACA	UDP0181	AATTGGCGGA	E11	35
ds1231_86	PO77	9.58	20	12	IDT for Illumina UD Index set B	UDP0182	AATTCTTGGA	UDP0182	TTGTCAACTT	F11	33.6
ds1231_87	PO78	2.67	20	12	IDT for Illumina UD Index set B	UDP0183	TTCCTACAGC	UDP0183	GGCGAATTCT	G11	12.9
ds1231_88	PO79	1.78	20	12	IDT for Illumina UD Index set B	UDP0184	ATCCAGGTAT	UDP0184	CAACGTCAGC	H11	9.68
ds1231_89	PO81	6.83	20	12	IDT for Illumina UD Index set B	UDP0185	ACGGTCCAAC	UDP0185	TCTTACATCA	A12	1.2
ds1231_90	PO82	9.31	20	12	IDT for Illumina UD Index set B	UDP0186	GTAACTTGGT	UDP0186	CGCCATACCT	B12	9.95

ds1231_91	PO83	3.31	20	12	IDT for Illumina UD Index set B	UDP0187	AGCGCCACCAC	UDP0187	CTAATGTCTT	C12	6.31
ds1231_92	PO87	6.88	20	12	IDT for Illumina UD Index set B	UDP0188	TGCTACTGCC	UDP0188	CAACCGGAGG	D12	22.4
ds1231_93	PO88	10.62	20	12	IDT for Illumina UD Index set B	UDP0189	CAACACCGCA	UDP0189	GGCAGTAGCA	E12	13.6
ds1231_94	PO89	9.68	20	12	IDT for Illumina UD Index set B	UDP0190	CACCTTAATC	UDP0190	TTAGGATAGA	F12	16.8
ds1231_95	Positive	12	20	12	IDT for Illumina UD Index set B	UDP0191	TTGAATGTTG	UDP0191	CGCAATCTAG	G12	12.1
ds1231_96	Negative	8.72	20	12	IDT for Illumina UD Index set B	UDP0192	CCGGTAACAC	UDP0192	GAGTTGTACT	H12	0.006

Table S 7: Tukey post-hoc test on environmental variables season 1.

Ph						
Tukey's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Below threshold?	Summary	Adjusted P Value	
1 vs. 2	0.038	-0.6830 to 0.7590	No	ns	>0.9999	A-B
1 vs. 3	-0.132	-1.019 to 0.7548	No	ns	0.9804	A-C
1 vs. 4	-0.768	-1.489 to -0.04750	Yes	*	0.0403	A-D
1 vs. 5	-1.164	-1.443 to -0.8847	Yes	***	0.0002	A-E
1 vs. 6	-1.612	-2.115 to -1.109	Yes	***	0.0006	A-F
1 vs. 7	-1.322	-2.048 to -0.5962	Yes	**	0.0058	A-G
2 vs. 3	-0.17	-0.7158 to 0.3758	No	ns	0.7156	B-C
2 vs. 4	-0.806	-1.305 to -0.3067	Yes	**	0.0091	B-D
2 vs. 5	-1.202	-1.922 to -0.4822	Yes	**	0.008	B-E
2 vs. 6	-1.65	-2.113 to -1.187	Yes	***	0.0004	B-F
2 vs. 7	-1.36	-1.873 to -0.8471	Yes	**	0.0014	B-G
3 vs. 4	-0.636	-1.155 to -0.1166	Yes	*	0.0249	C-D
3 vs. 5	-1.032	-1.988 to -0.07581	Yes	*	0.0386	C-E
3 vs. 6	-1.48	-2.248 to -0.7120	Yes	**	0.0047	C-F

3 vs. 7	-1.19	-2.005 to -0.3748	Yes	*	0.0132	C-G
4 vs. 5	-0.396	-1.037 to 0.2453	No	ns	0.2113	D-E
4 vs. 6	-0.844	-1.229 to -0.4595	Yes	**	0.0028	D-F
4 vs. 7	-0.554	-1.158 to 0.04952	No	ns	0.0663	D-G
5 vs. 6	-0.448	-0.7915 to -0.1045	Yes	*	0.0199	E-F
5 vs. 7	-0.158	-0.7594 to 0.4434	No	ns	0.8221	E-G
6 vs. 7	0.29	-0.1413 to 0.7213	No	ns	0.1686	F-G
Conductivity						
Tukey's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Below threshold?	Summary	Adjusted P Value	
1 vs. 2	40	-52.57 to 132.6	No	ns	0.4586	A-B
1 vs. 3	67	-73.65 to 207.7	No	ns	0.3822	A-C
1 vs. 4	85	-101.1 to 271.1	No	ns	0.4146	A-D
1 vs. 5	117.8	-99.24 to 334.8	No	ns	0.2892	A-E
1 vs. 6	97	-78.37 to 272.4	No	ns	0.2768	A-F
1 vs. 7	192.2	-6.041 to 390.4	No	ns	0.0554	A-G
2 vs. 3	27	-85.57 to 139.6	No	ns	0.8668	B-C
2 vs. 4	45	-123.9 to 213.9	No	ns	0.8143	B-D
2 vs. 5	77.8	-111.2 to 266.8	No	ns	0.4977	B-E
2 vs. 6	57	-63.69 to 177.7	No	ns	0.3887	B-F
2 vs. 7	152.2	-25.49 to 329.9	No	ns	0.0826	B-G
3 vs. 4	18	-79.48 to 115.5	No	ns	0.9497	C-D
3 vs. 5	50.8	-73.12 to 174.7	No	ns	0.5011	C-E
3 vs. 6	30	-94.44 to 154.4	No	ns	0.8645	C-F
3 vs. 7	125.2	18.06 to 232.3	Yes	*	0.0293	C-G
4 vs. 5	32.8	-6.606 to 72.21	No	ns	0.0903	D-E

4 vs. 6	12	-164.3 to 188.3	No	ns	0.9997	D-F
4 vs. 7	107.2	78.93 to 135.5	Yes	***	0.0003	D-G
5 vs. 6	-20.8	-216.4 to 174.8	No	ns	0.9963	E-F
5 vs. 7	74.4	45.25 to 103.5	Yes	**	0.0016	E-G
6 vs. 7	95.2	-99.85 to 290.3	No	ns	0.364	F-G
Turbidity						
Tukey's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Below threshold?	Summary	Adjusted P Value	
1 vs. 2	0.88	-2.870 to 4.630	No	ns	0.8762	A-B
1 vs. 3	0.54	-4.396 to 5.476	No	ns	0.9957	A-C
1 vs. 4	-0.96	-5.346 to 3.426	No	ns	0.903	A-D
1 vs. 5	-2	-6.633 to 2.633	No	ns	0.4593	A-E
1 vs. 6	-3.04	-9.230 to 3.150	No	ns	0.3593	A-F
1 vs. 7	0.1	-6.478 to 6.678	No	ns	>0.9999	A-G
2 vs. 3	-0.34	-2.977 to 2.297	No	ns	0.9901	B-C
2 vs. 4	-1.84	-4.610 to 0.9303	No	ns	0.1743	B-D
2 vs. 5	-2.88	-3.977 to -1.783	Yes	**	0.0014	B-E
2 vs. 6	-3.92	-9.955 to 2.115	No	ns	0.1851	B-F
2 vs. 7	-0.78	-4.686 to 3.126	No	ns	0.9312	B-G
3 vs. 4	-1.5	-5.100 to 2.100	No	ns	0.4878	C-D
3 vs. 5	-2.54	-5.501 to 0.4212	No	ns	0.0822	C-E
3 vs. 6	-3.58	-10.69 to 3.531	No	ns	0.3413	C-F
3 vs. 7	-0.44	-2.336 to 1.456	No	ns	0.8809	C-G
4 vs. 5	-1.04	-4.161 to 2.081	No	ns	0.6661	D-E
4 vs. 6	-2.08	-9.742 to 5.582	No	ns	0.8038	D-F
4 vs. 7	1.06	-3.630 to 5.750	No	ns	0.8913	D-G

5 vs. 6	-1.04	-6.773 to 4.693	No	ns	0.9532	E-F
5 vs. 7	2.1	-1.592 to 5.792	No	ns	0.259	E-G
6 vs. 7	3.14	-4.199 to 10.48	No	ns	0.4665	F-G

Table S8: Summary table produced by CANOCO for DCA unconstrained-supplementary variables analysis. Total variation is 0.81160, supplementary variables account for 33.06% (adjusted explained variation is 26.69%).

Statistic	Axis 1	Axis 2	Axis 3	Axis 4
Eigenvalues	0.1848	0.1161	0.0562	0.0395
Explained variation (cumulative)	22.77	37.07	44	48.87
Gradient length	2.86	3.8	1.3	1.46
Pseudo-canonical correlation (suppl.)	0.9066	0.6222	0.5033	0.449

Table S9: Species scores for PRC in morphological analysis.

	Resp.1
Alona/Alonella sp.	-1.7361
Chydorus sp.	-1.7349
Trichocerca sp.	-1.6486
Nauplia	-1.5673
Cyclopoida	-1.4313
Mytilina sp.	-1.4267
Lepadella sp.	-1.3655
Keratella sp.	-1.2086
Trichotria sp.	-0.7264
Lecane sp.	-0.7082
Diaptomidae	-0.7069
Daphnia sp.	-0.5661
Acroperus sp.	-0.4421
Notholca sp.	-0.4002
Simocephalus sp.	-0.2657
Brachionus sp.	-0.2322
Rotaria sp.	-0.204
Squatinella sp.	-0.1137
Polyarthra sp.	0.0097
Cephalodella sp.	0.0444
Testudinella sp.	0.0783

Table S10: Summary table produced by CANOCO for PCA unconstrained-supplementary variables analysis. Total variation is 239421.525, supplementary variables account for 58.72% (adjusted explained variation is 54.83%).

Statistic	Axis 1	Axis 2	Axis 3	Axis 4
Eigenvalues	0.6497	0.2726	0.0494	0.0187
Explained variation (cumulative)	64.97	92.23	97.17	99.04
Pseudo-canonical correlation (suppl.)	0.8454	0.6279	0.4747	0.4348

Table S11: Species scores for PRC in metabarcoding analysis.

	Resp.1
Cyclopoida	-1.9141
Chydorus sp.	-1.5597
Daphnia sp.	-0.5995
Lepadella sp.	-0.0867
Cephalodella sp.	-0.0211
Ceriodaphnia sp.	-0.0146
Brachionus sp.	-0.003
Mytilina sp.	-0.0002
Notholca sp.	-0.0001
Polyarthra sp.	0
Testudinella sp.	0.0002
Acroperus sp.	0.0003
Keratella sp.	0.0004
Rotaria sp.	0.0012
Synchaeta sp.	0.0012
Monommata	0.0023
Squatinella sp.	0.0062
Lecane sp.	0.0064
Trichotria sp.	0.0224
Trichocerca sp.	0.0338
Simocephalus sp.	0.131
Diaptomidae	3.9392

Bibliography

Agrigem.(n.d.) HomePage Retrieved from:

https://www.agrigem.co.uk/product/Roundup™-pro-vantage-480-5l/

Adamowicz, S. J., Thielsch, A., & Hebert, P. D. N. (2009). Cryptic diversity in a widespread freshwater zooplankter revealed using DNA barcoding. Molecular Ecology, 18(17), 3787-3795.

Allan, J. D. (1976). Life history patterns in zooplankton. The American Naturalist, 110(971), 165-180.

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., & Huse, S. M. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS ONE, 4(7), e6372.

Anderson, D. M., Glibert, P. M., & Burkholder, J. M. (2002). Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. Estuaries, 25(4), 704-726.

Andrade, V. S., Gutierrez, M. F., Reno, U., Popielarz, A., Gervasio, S., & Gagneten, A. M. (2021). Synergy between glyphosate and cypermethrin formulations on zooplankton: evidences from a single-specie test and a community mesocosm experiment. Environmental Science and Pollution Research, 28(2021), 26885–26894

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Araújo, M. F., Castanheira, E. M. S., & Sousa, S. F. (2023). The Buzz on Insecticides: A Review of Uses, Molecular Structures, Targets, Adverse Effects, and Alternatives. Molecules, 28(8), 3641.

Baker, L. F., Mudge, J. F., Thompson, D. G., Houlahan, J. E., & Kid, K. A. (2016). The combined influence of two agricultural contaminants on natural communities of phytoplankton and zooplankton. Ecotoxicology, 25, 1021–1032.

Barka, S., Gdara, I., Ouanes-Ben Othmen, Z., Mouelhi, S., El Bour, M., & Hamza-Chaffai, A. (2020). Seasonal ecotoxicological monitoring of freshwater zooplankton in Bir Mcherga dam (Tunisia). Environmental Science and Pollution Research, 27(2020), 5670–5680.

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental Conditions Influence eDNA Persistence in Aquatic Systems. Environmental Science & Technology.

Bartram, J. (1996). Water Quality Monitoring: A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes (ISBN 0-419-21730-4).

Battaglin, W.A., Meyer, M.T., Kuivila, K.M., Dietze, J.E., 2014. Glyphosate and its degradation product AMPA occur frequently and widely in U.S. soils, surface water, groundwater, and precipitation. J. Am. Water Resour. Assoc. 50, 275–290

Bayat A. (2002). Science, medicine, and the future: Bioinformatics. BMJ (Clinical research ed.), 324(7344), 1018–1022.

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., ... & Boss, E. S. (2006). Climate-driven trends in contemporary ocean productivity. Nature, 444(7120), 752-755.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114-2120.

Bonacorsi, S., Visseaux, B., Bouzid, D., Pareja, J., Rao, S. N., Manissero, D., Hansen, G., & Vila, J. (2021). Systematic Review on the Correlation of Quantitative PCR Cycle Threshold Values of Gastrointestinal Pathogens With Patient Clinical Presentation and Outcomes. Frontiers in medicine, 8, 711809. https://doi.org/10.3389/fmed.2021.711809

Borggaard, O. K., & Gimsing, A. L. (2008). Fate of glyphosate in soil and the possibility of leaching to ground and surface waters: a review. Pest Management Science: Formerly Pesticide Science, 64(4), 441-456.

Borics, G., Abonyi, A., Salmaso, N., & Ptacnik, R. (2021). Freshwater phytoplankton diversity: models, drivers and implications for ecosystem properties. Hydrobiologia, 848, 53–75.

Boudou, A., & Ribeyre, F. (1997). Aquatic Ecotoxicology: From the Ecosystem to the Cellular and Molecular Levels. Environmental Health Perspectives, 105(Supplement 1), 21-35.

Boykin LM, Ghalab A, De Marchi BR, et al.. 2018. Real time portable genome sequencing for global food security

Branton, D., et al. (2008). The potential and challenges of nanopore sequencing. Nat. Biotechnol., 26(10), 1146–1153.

Brock, T. C. M., Hammers-Wirtz, M., Hommen, U., Preuss, T. G., Ratte, H.-T., Roessink, I., Strauss, T., & Van den Brink, P. J. (2015). The minimum detectable difference (MDD) and the interpretation of treatment-related effects of pesticides in experimental ecosystems. Environ Sci Pollut Res, 22, 1160–1174.

Brooks, A. C., Foudoulakis, M., Schuster, H. S., & Wheeler, J. R. (2019). Historical control data for the interpretation of ecotoxicity data: are we missing a trick? Ecotoxicology, 28, 1198–1209.

Bucklin, A., Batta-Lona, P. G., Questel, J. M., Wiebe, P. H., Richardson, D. E., Copley, N. J., & O'Brien, T. D. (2022). COI Metabarcoding of Zooplankton Species Diversity for Time-Series Monitoring of the NW Atlantic Continental Shelf.Frontiers in Marine Science, Volume(9).

Bucklin, A., Hopcroft, R. R., Kosobokova, K. N., Nigro, L. M.,Ortman, B. D., Jennings, R. M., and Sweetman, C. J. (2010). DNA barcoding of Arctic Ocean holozooplankton for species identification and recognition. Deep-Sea Research II, 57: 40–48.

Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A., & Lehtiniemi, M. (2016). Metabarcoding of marine zooplankton: prospects, progress and pitfalls. Journal of Plankton Research, 38(3), 393–400.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., ... & Knight, R. (2012). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences, 108(Supplement 1), 4516-4522.

Carpenter, S. R. (1996). Microcosm experiments have limited relevance for community and ecosystem ecology. Ecology, 77(3), 677-680.

Chang, G. K., Chain, F. J., Abbott, C. L., & Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. Evolutionary Applications, 11, 1901–1914.

Chang, J. J. M., Ip, Y. C. A., Neo, W. L., Mowe, M. A. D., Jaafar, Z., & Huang, D. (2023). Primed and ready: Nanopore metabarcoding can now recover highly accurate consensus barcodes that are generally indel-free. bioRxiv.

Choi, J. K., Lee, Y. J., Park, J. S., & Kang, J. H. (2020). Seasonal variations in the abundance and composition of zooplankton community in a temperate shallow estuary (Gamak Bay, Korea). Ocean Science Journal, 55(1), 63-73.

Cloern, J. E., & Jassby, A. D. (2010). Patterns and scales of phytoplankton variability in estuarine–coastal ecosystems. Estuaries and Coasts, 33(2), 230-241.

Corell, J., & Rodríguez-Ezpeleta, N. (2014). Tuning of protocols and marker selection to evaluate the diversity of zooplankton using metabarcoding. Revista de Investigación Marina AZTI-Tecnalia.

Crane M, Newman MC. What level of effect is a no observed effect? Environ Toxicol Chem. 2000 Feb. 19:516–519.

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., ... & Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. Methods in Ecology and Evolution, 7(9), 1008-1018.

Cristescu, M. E. (2014). From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends in Ecology & Evolution, 29(10), 566-571.

Culver, D. A., Beisner, B. E., & Taylor, B. E. (2000). Evaluation of the precision and accuracy of an ANOVA model to estimate plankton diversity. Limnology and Oceanography, 45(2), 543-550.

De Jong FMW, Brock TCM, Foekema EM, Leeuwangh P. (2008). Guidance for summarizing and evaluating aquatic micro- and mesocosm studies. RIVM Report 601506009/2008. RIVM, Bilthoven.

De Meester, L., Gómez, A., Okamura, B., & Schwenk, K. (2002). The monopolization hypothesis and the dispersal–gene flow paradox in aquatic organisms. Acta Oecologica, 23(3), 121-135.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., et al. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science, 348*, 1261605.

Diamanti-Kandarakis, E., Bourguignon, J.-P., Giudice, L. C., Hauser, R., Prins, G. S., Soto, A. M., Zoeller, R. T., & Gore, A. C. (2009). Endocrine-Disrupting Chemicals: An Endocrine Society Scientific Statement. Endocrine Reviews, 30(4), 293–342.

Diaz, R. J., & Rosenberg, R. (2008). Spreading dead zones and consequences for marine ecosystems. Science, 321(5891), 926-929.

Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., Muller-Karger, F., & Breitbart, M. (2018). Evaluation of marine zooplankton community structure through environmental DNA metabarcoding. Limnology and Oceanography: Methods. Advance online publication.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15-21.

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., & Mattick, J. S. (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Research, 19(14), 4008. DOI: 10.1093/nar/19.14.4008.

Doorenspleet, K., Jansen, L., Oosterbroek, S., Bos, O., Kamermans, P., Janse, M., Wurz, E., Murk, A., & Nijland, R. (2021). High resolution species detection: accurate long read eDNA metabarcoding of North Sea fish using Oxford Nanopore sequencing. bioRxiv.

Dorak, Z., Gaygusuz, Ö., Tarkan, A.S., Aydın, H. (2013). Diurnal vertical distribution of zooplankton in a newly formed reservoir (Tahtalı Reservoir, Kocaeli): the role of abiotic factors and chlorophyll a. Turkish Journal of Zoology, 37, 218-227.

Duke, S. O., Pan, Z., Chittiboyina, A. G., Swale, D. R., & Sparks, T. C. (2023). Molecular targets of insecticides and herbicides – Are there useful overlaps? Pesticide Biochemistry and Physiology, 191, 105340.

Duke, S.O., Powles, S.B. (2008). Glyphosate: A once-in-a-century herbicide. Pest Management Science, 64, 319–325. doi: 10.1002/ps.1518.

Dumschott, K., Schmidt, M. H.-W., Chawla, H. S., Snowdon, R., & Usadel, B. (2020). Oxford Nanopore sequencing: new opportunities for plant genomics? Journal of Experimental Botany, 71(18), 5313–5322.

Dutilleul, P., Stockwell, J. D., Frigon, D., & Legendre, P. (2000). The Mantel Test versus Pearson's Correlation Analysis: Assessment of the Differences for Biological and Environmental Studies. Journal of Agricultural, Biological, and Environmental Statistics, 5(2), 131-150.

Ebert, D. (2005). Ecology, epidemiology, and evolution of parasitism in *Daphnia*. National Library of Medicine.

Edwards, M., & Richardson, A. J. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. Nature, 430(7002), 881-884.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues). (2013). Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Journal, 11(7), 3290. doi:10.2903/j.efsa.2013.3290.

Egeter, B., Verissimo, J. S., Lopes-Lima, M., Fonseca, N. A., et al. (2022). Speeding up the detection of invasive bivalve species using environmental DNA: A Nanopore and Illumina sequencing comparison. Molecular Ecology Resources, 22(6).

Egge, E. S., Bittner, L., Andersen, T., Audic, S., de Vargas, C., Edvardsen, & Vaulot, D. (2013). 454 pyrosequencing to describe microbial eukaryotic community composition, diversity and relative abundance: a test for marine haptophytes. PLoS ONE, 8(9), e74371.

Eggen, R., Behra, R., Burkhardt-Holm, P., Escher, P., & Schweigert, B. I. (2004). Challenges in ecotoxicology. Peer Reviewed.

Elías-Gutiérrez, M., Valdez-Moreno, M., Topan, J., Young, M. R., & Cohuo-Colli, J. A. (2018). Improved protocols to accelerate the assembly of DNA barcode reference libraries for freshwater zooplankton. Ecology and Evolution, 8(5), 3002-3018.

Emraj, D. A., Hossain, M. A., Ye, Q., Qin, J. G., & Leterme, S. C. (2017). Plankton bioindicators of environmental conditions in coastal lagoons. Estuarine, Coastal and Shelf Science, 184, 102-114.

Endo, N., Sato, K., Matsumura, K., Yoshimura, E., Odaka, Y., & Nogata, Y. (2010). Species-specific detection and quantification of common barnacle larvae from the Japanese coast using quantitative real-time PCR. Biofouling, 26, 901-911.

Ershova, E. A., Wangensteen, O. S., Descoteaux, R., Barth-Jensen, C., & Præbel, K. (2021). Metabarcoding as a quantitative tool for estimating biodiversity and relative biomass of marine zooplankton. ICES Journal of Marine Science, 78(9), 3342–3355.

Fabry, V. J., Seibel, B. A., Feely, R. A., & Orr, J. C. (2008). Impacts of ocean acidification on marine fauna and ecosystem processes. ICES Journal of Marine Science, 65(3), 414-432.

Feio, M. J., Filipe, A. F., Garcia-Raventos, A., Ardura Gutiérrez, A., Calapez, A. R., Pujante, A. M., Mortágua, A., Murria, C., Díaz de Quijano, D., & Martins, F. (2020). Advances in the use of molecular tools in ecological and biodiversity assessment of aquatic ecosystems. Limnetica.

Fernández de Puelles, M.L., Gazá, M., Cabanellas-Reboredo, M., Santandreu, M.d.M., Irigoien, X., González-Gordillo, J.I., Duarte, C.M., & Hernández-León, S. (2019). Zooplankton Abundance and Diversity in the Tropical and Subtropical Ocean. Instituto Español de Oceanografí.

Finnegan, M., Emburey, S., Hommen, U., Hamer, M., & et al. (2018). A freshwater mesocosm study into the effects of the neonicotinoid insecticide thiamethoxam at multiple trophic levels. Environmental Pollution, 242(Pt B), 1444-1457.

Fitzpatrick, M. C., et al. (2021). Environmental DNA: what's behind the scene. Molecular Ecology.

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular marine biology and biotechnology, 3(5), 294-299.

Forró, L., Korovchinsky, N. M., Kotov, A. A., & Petrusek, A. (2007). Global diversity of Cladocerans (Cladocera; Crustacea) in freshwater. Freshwater animal diversity assessment. Springer.

Fossi, M. C. (1994). Nondestructive biomarkers in ecotoxicology. Environmental Health Perspectives, 102(Suppl 12), 49-54.

Fox, D. R., & Landis, W. G. (2016). Don't be fooled—A no-observed-effect concentration is no substitute for a poor concentration–response experiment. Environmental Toxicology and Chemistry. DOI: 10.1002/etc.3459.

Franceschini, S., Tancioni, L., Lorenzoni, M., Mattei, F., & Scardi, M. (2019). An ecologically constrained procedure for sensitivity analysis of Artificial Neural Networks and other empirical models. PLOS ONE, 14(1), e0211445. DOI: 10.1371/journal.pone.0211445.

Franz, J. E., Mao, M. K., & Sikorski, J. A. (1997). Glyphosate: A unique global herbicide. ACS Monograph 189, American Chemical Society, Washington, DC.

Friedheim, S. (2016). Comparison of species identification methods: DNA barcoding versus morphological taxonomy. Horizons, 1, 13.

Gacem, H., Bendali-Saoudi, F., Serradj, N., Houmani, M., & Soltani, N. (2022). Risk assessment of the neonicotinoid insecticide acetamiprid on two non-target species, *Daphnia* magna Straus, 1820 (Crustacea, Cladocera) and Plea minutissima Leach, 1817 (Insecta, Heteroptera). Ecole Normal Supérieure d'Enseignement Technologique Azzaba.

Galan, M., Pages, M., & Cosson, J.-F. (2012). Next-generation sequencing for rodent barcoding: species identification from fresh, degraded and environmental samples. PloS one, 7, e48374.

Gaygusuz, Ö., & Dorak, Z. (2013). Species composition and diversity of the zooplankton fauna of Darlik Stream (İstanbul - Turkey) and its tributaries. Journal of Aquatic Ecology, 7(4).

Gessner, M. O., & Tlili, A. (2016). Fostering integration of freshwater ecology with ecotoxicology. Freshwater Biology, 61, 1991–2001.

Geyer, R. L., Smith, G. R., & Rettig, J. E. (2016). Effects of Roundup™ formulations, nutrient addition, and Western mosquitofish (Gambusia affinis) on aquatic communities. Environmental Science and Pollution Research, 23, 11729–11739.

Giesy, J. P., Dobson, S., & Solomon, K. R. (2000). Ecotoxicological risk assessment for Roundup herbicide. Reviews of Environmental Contamination and Toxicology, 167, 35-120.

Gilbert, J. J. (2016). Rotifers as models for the biology of aging. In Rotifers (pp. 147-160). Springer, Cham.

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., Spear, S. F., McKee, A., Oyler-McCance, S. J., Cornman, R. S., Laramie, M. B., Mahon, A. R., Lance, R. F., Pilliod, D. S., Strickler, K. M., Waits, L. P., Fremier, A. K., Takahara, T., Herder, J. E., & Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in Ecology and Evolution, 7, 1299-1307. DOI: 10.1111/2041-210X.12595

Gonçalves, B. B., Giaquinto, P. C., Silva, D. D. S., Neto, C. D. M. E. S., Lima, A. A. D., Darosci, A. A. B., Portinho, J. L., Carvalho, W. F., & Rocha, T. L. (Year). Chapter 2: Ecotoxicology of Glyphosate-Based Herbicides on Aquatic Environment. In Biochemical Toxicology: Heavy Metals and Nanomaterials

Gonzalez, A., & Bradley, R. L. (1998). How predictable are the ecosystem responses to global change? A comparison of a tropical forest and a boreal forest. Science, 281(5375), 1292-1293.

Goodall-Copestake, W., Tarling, G., & Murphy, E. (2012). On the comparison of population-level estimates of haplotype and nucleotide diversity: a case study using the gene cox1 in animals. Heredity, 109, 50-56.

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics, 17(6), 333–351.

Gourmelon, A., & Ahtiainen, J. (2007). Developing Test Guidelines on invertebrate development and reproduction for the assessment of chemicals, including potential endocrine active substances—The OECD perspective. Ecotoxicology, 16, 161-167.

Gupta, N., & Verma, V. K. (2019). Next-Generation Sequencing and Its Application: Empowering in Public Health Beyond Reality. Microbial Technology for the Welfare of Society, 17, 313–341. doi: 10.1007/978-981-13-8844-6_15.

Hairston, N. G., & Kearns, C. M. (2002). Temporal dispersal: Ecological and evolutionary aspects of zooplankton egg banks and the role of sediment mixing. Integrative and Comparative Biology, 42(3), 481-491.

Hammer, Ø., Harper, D.A.T., Ryan, P.D. 2001. PAST: Paleontological Statistics software package for education and data analysis. Paleontologia Electronica 4(1): 9pp.

Hanazato, T. (1998). Response of a zooplankton community to insecticide application in experimental ponds: A review and the implications of the effects of chemicals on the structure and functioning of freshwater communities. Environmental Pollution, 101(3), 361-373.

Hanazato, T. (2001). Pesticide effects on freshwater zooplankton: An ecological perspective. Environmental Pollution, 112(1), 1-10.

Hansen, B., & Bech, G. (1996). Bacteria associated with a marine planktonic copepod in culture. II. Bacterial colonization of fecal pellets. Aquatic Microbial Ecology, 10(1), 39-51.

Hanson, M., Wolff, B., Green, J., Kivi, M., Panter, G., Warne, M. S. J., Ågerstrand, M., & Sumpter, J. (2017). How we can make ecotoxicology more valuable to environmental protection. Science of The Total Environment, 578, 228-235.

Hatakeyama, S., Shiraishi, H., & Sugaya, Y. (1991). Monitoring of the overall pesticide toxicity of river water to aquatic organisms using a freshwater shrimp, Paratya compressa improvisa. Chemosphere, 22, 229-235.

Hays, G. C., Richardson, A. J., & Robinson, C. (2005). Climate change and marine plankton. Trends in Ecology & Evolution, 20(6), 337-344.

Hébert, M. P., Fugère, V., Beisner, B. E., da Costa, N. B., Barrett, R. D. H., Bell, G., Shapiro, B. J., Yargeau, V., Gonzalez, A., & Fussmann, G. F. (2020). Widespread agrochemicals differentially affect zooplankton biomass and community structure. bioRxiv

Hébert, M.-P., Fugère, V., Beisner, B. E., da Costa, N. B., Barrett, R. D. H., Bell, G., Shapiro, B. J., Yargeau, V., Gonzalez, A., & Fussmann, G. F. (2021). Widespread agrochemicals differentially affect zooplankton biomass and community structure. Ecological Applications.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society B: Biological Sciences, 270(1512), 313–321. doi:10.1098/rspb.2002.2218

Helgen, J. C., Larson, N., & Anderson, R. (1988). Responses of zooplankton and Chaoborus to temephos in a natural pond and in the laboratory. Archives of Environmental Contamination and Toxicology, 17, 459-471.

Heuschele, J., & Selander, E. (2014). The chemical ecology of copepods. Journal of Plankton Research, 36, 895-913.

Hoffman, D. J., Rattner, B. A., Burton Jr, G. A., & Cairns Jr, J. (2002). Handbook of ecotoxicology.

Honggang, Z., Baoshan, C., Zhiming, Z., & Xiaoyun, F. (2012). Species diversity and distribution for zooplankton in the intertidal wetlands of the Pearl River estuary, China. Procedia Environmental Sciences, 13, 2383-2393.

Hutchinson, G. E. (1984). Limnological methods (2nd ed.). Wiley.

Islam, S. K. S., & Midya, S. (2023). Growth regulatory pattern of zooplankton in herbicide and antibiotic contaminated aquatic ecosystem: An overview. Watershed Ecology and the Environment, 5, 153-160.

Jager, T. (2012). Bad habits die hard: The NOEC's persistence reflects poorly on ecotoxicology. Environmental Toxicology and Chemistry, 31(2), 228-229. DOI: 10.1002/etc.746

Jain, M., et al. (2020). Nanopore sequencing and assembly of a human genome with ultra-long reads. Nature Biotechnology.

Jain, M., Olsen, H. E., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biology, 17(239), 1-2.

Jain, M., Olsen, H. E., Turner, D. J., Stoddart, D., Bulazel, K. V., Paten, B. Akeson, M. (2018). Linear assembly of a human centromere on the Y chromosome. Nature Biotechnology, 36(4), 321–323.

Jaki, T., & Hothorn, L. A. (2013). Statistical evaluation of toxicological assays: Dunnett or Williams test—take both. *Regulatory Toxicology*, 87, 1901–1910.

Janssens, L., & Stoks, R. (2017). Stronger effects of Roundup than its active ingredient glyphosate in damselfly larvae. Aquatic Toxicology, 189, 88-94.

Jensen, T. C. (2019). Winter decrease of zooplankton abundance and biomass in subalpine oligotrophic Lake Atnsjøen (SE Norway). Journal of Limnology. DOI: 10.4081/jlimnol.2019.1877.

Jeppesen, E., Nõges, T., Davidson, T. A., Haberman, J., Nõges, P., Blank, K., ... & Lauridsen, T. L. (2011). Zooplankton as indicators in lakes: A scientific-based plea for including zooplankton in the ecological quality assessment of lakes according to the European Water Framework Directive (WFD). Hydrobiologia, 676(1), 279-297.

Jobling, S., Nolan, M., Tyler, C. R., Brighty, G., & Sumpter, J. P. (1998). Widespread sexual disruption in wild fish. Environmental Science & Technology, 32(17), 2498–2506.

Juan-García, A., Pakkanen, H., Juan, C., & Vehniäinen, E. (2023). Alterations in *Daphnia* magna exposed to enniatin B and beauvericin provide additional value as environmental indicators. Ecotoxicology and Environmental Safety, 249, 114427.

Kasianowicz, J et. al.. (1996) Characterization of individual polynucleotide molecules using a membrane channel. Proceedings of the National Academy of Sciences. www.pnas.org/cgi/doi/10.1073/pnas.93.24.13770.

Kazi, T. G., Arain, M. B., Jamali, M. K., Jalbani, N., Afridi, H. I., Sarfraz, R. A., Baig, J. A., & Shah, A. Q. (2009). Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. Ecotoxicology and Environmental Safety, 72, 301–309.

Khan, N., Muller, J., Khan, S. H., Amjad, S., Nizamani, S., & Bhanger, M. I. (2010). Organochlorine pesticides (OCPs) contaminants in sediments from Karachi harbour, Pakistan. Journal of the Chemical Society of Pakistan, 32(4), 542-549.

Kiørboe, T. (2008). A Mechanistic Approach to Plankton Ecology. Princeton University Press.

Kirk, H., & Freeland, J. R. (2011). Applications and implications of neutral versus non-neutral markers in molecular ecology. International Journal of Molecular Sciences, 12, 3966-3988.

Kitamura, R., Ishii, K., Maeda, I., Kozaki, T., Iwabuchi, K., & Saito, T. (2016). Evaluation of bacterial communities by bacteriome analysis targeting 16S rRNA genes and quantitative analysis of ammonia monooxygenase gene in different types of compost. Journal of Bioscience and Bioengineering, 121(1), 57-65.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic acids research, 41(1), e1-e1.

Knapik, K. (2013). Genetic analysis of bacteriophages from clinical and environmental samples. July 2013. DOI: 10.13140/RG.2.2.29564.08327.

Köhler, H. R., & Triebskorn, R. (2013). Wildlife ecotoxicology of pesticides: can we track effects to the population level and beyond? Science, 341(6147), 759-765.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and environmental microbiology, 79(17), 5112-5120.

Kreutzweiser, D. P., Back, R. C., Sutton, T. M., Pangle, K. L., & Thompson, D. G. (2004). Aquatic mesocosm assessments of a neem (azadirachtin) insecticide at environmentally realistic concentrations—2: Zooplankton community responses and recovery. Ecotoxicology and Environmental Safety, 59, 194-204.

Kreutzweiser, D. P., Back, R. C., Sutton, T. M., Thompson, D. G., & Scarr, T. A. (2002). Community-level disruptions among zooplankton of pond enclosures treated with a neem (azadirachtin) insecticide. Aquatic Toxicology, 56, 257-273.

Kuczynski, J., Stombaugh, J., Walters, W. A., González, A., Caporaso, J. G., & Knight, R. (2012). Using QIIME to analyze 16S rRNA gene sequences from microbial communities. Current protocols in microbiology, 27(1), 1E-5.

Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: a meta-analytical approach. Molecular Ecology, 28(2), 420–430. DOI: 10.1111/mec.14920.

Lampert, W. (1989). The adaptive significance of diel vertical migration of zooplankton. Functional Ecology, 3(1), 21-27.

Lampert, W., & Sommer, U. (2007). Limnoecology: The Ecology of Lakes and Streams. Oxford University Press.

Landis W, Chapman PM. Well past time to stop using NOELs and LOELs. Integrated Environmental Assessment and Management. 2011; 7:vi–viii. [PubMed: 21932339]

Lares, B. A., Vignatti, A. M., Echaniz, S. A., & Gutiérrez, M. F. (2022). Effects of glyphosate on cladocera: A synthetic review. Aquatic Toxicology, 249, 106232.

Lentendu, G., Mahé, F., Bass, D., Rueckert, S., Stoeck, T., & Dunthorn, M. (2018). Consistency of metabarcoding and morphological approaches for water quality assessment. Environmental Science & Technology, 52(22), 12994-13005.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., & Boehm, J. T. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Frontiers in Zoology, 10(1), 34.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 34(18), 3094-3100.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics, 25(14), 1754-1760.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics, 30(7), 923-930.

Liess, M., Foit, K., Becker, A., Hassold, E., Dolciotti, I., Kattwinkel, M., Rastall, A. (2016). Ethical issues in environmental science: The case of mesocosm research. Environmental Science & Technology, 50(14), 7347–7353.

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., Atkinson, A., & Ianora, A. (2013). Next generation sequencing reveals the hidden diversity of zooplankton assemblages. PLoS One, 8(10), e81327

Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2020). A practical guide to DNA metabarcoding for entomological ecologists. Ecological Entomology, 45, 373-385.

Liu, N., Yang, Y., Li, F., Ge, F., & Kuang, Y. (2016). Importance of controlling pH-dependent dissolved inorganic carbon to prevent algal bloom outbreaks. Bioresource Technology, 220, 246-252.

Liu, Y., Wang, L., Pan, B., Wang, C., Bao, S., & Nie, X. (2017). Toxic effects of diclofenac on life history parameters and the expression of detoxification-related genes in *Daphnia* magna. Aquatic Toxicology, 183, 104-113.

Loerracher, A.-K., Schmidt, J., Ebke, P., Schmolke, A., Abi-Akar, F., Galic, N., & Ashauer, R. (2023). Characterization of patterns and variability in the dynamics of outdoor aquatic mesocosms: exploring the capabilities and challenges in data supporting aquatic system models. Ecotoxicology, 32, 782–801. DOI: 10.1007/s10646-023-02685-0.

Logsdon, G. A., et al. (2020). High-throughput sequencing: technology and applications. Trends in Biotechnology.

López-Mancisidor, P., Carbonell, G., Fernández, C., & Tarazona, J. V. (2008). Ecological impact of repeated applications of chlorpyrifos on zooplankton community in mesocosms under Mediterranean conditions. Ecotoxicology, 17, 811–825.

López-Valcárcel, M. E., Del Arco, A., & Parra, G. (2023). Sublethal exposure to agrochemicals impairs zooplankton ability to face future global change challenges. Science of The Total Environment, 873, Article 162020.

López-Valcárcel, M. E., del Arco, A., & Parra, G. (2024). Zooplankton vulnerability to glyphosate exacerbated by global change. Science of The Total Environment, 913, 169806.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology, 15(12), 550.

Lovejoy, C., Legendre, L., Martineau, M.-J., & von Quillfeldt, C. H. (2002). Distribution of phytoplankton and other protists in the North Water. Deep-Sea Research Part II: Topical Studies in Oceanography, 49, 5027–5047.

Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics*, 14(5), 265-279.

Lundsgaard-Hansen, B., Matthews, B., von Rueden, J., Seehausen, O., & Wagner, C. E. (2012). Ecological speciation in postglacial European whitefish: rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. Ecology and Evolution, 2(2), 477-498.

Magoc, T., & Salzberg, S. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics, 27(21), 2957–2963.

Maltby, L., & Hills, L. (2008). Spray drift of pesticides and stream macroinvertebrates: Experimental evidence of impacts and effectiveness of mitigation measures. Environmental Pollution, 156, 1112–1120.

Maltby, L., Blake, N., Brock, T. C. M., Van den Brink, P. J., Acreman, M., & Dunbar, M. J. (2018). Introducing model ecosystems in ecotoxicology. Environmental Science & Technology, 52(5), 2574–2580.

Mardis, E. R. (2008). Next-Generation DNA Sequencing Methods. Annual Review of Genomics and Human Genetics, 9(1), 387–402. DOI: 10.1146/annurev.genom.9.081307.164359.

Martin (2011) Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet Journal. 2011;17:10–12.

Martins-Gomes, C., Silva, T. L., Andreani, T., & Silva, A. M. (Year). Glyphosate vs. Glyphosate-Based Herbicides Exposure: A Review on Their Toxicity. Journal of Environmental Toxicology, 35(4), 567-582.

Matozzo, V., Fabrello, J., Masiero, L., Ferraccioli, F., Finos, L., Pastore, P., Di Gangi, I. M., & Bogialli, S. (2018). Ecotoxicological risk assessment for the herbicide glyphosate to non-target aquatic species: A case study with the mussel Mytilus galloprovincialis. Environmental Pollution, 233, 623-632.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research, 20(9), 1297-1303.

McManus, G. B., & Katz, L. A. (2009). Molecular and morphological methods for identifying plankton: what makes a successful marriage? Journal of Plankton Research, 31, 1119-1129.

Meier, R., Shiyang, K., Vaidya, G., & Ng, P. K. L. (2006). DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. Systematic Biology, 55(5), 715–728. doi:10.1080/10635150600969864

Meiklejohn, K. A., Damaso, N., & Robertson, J. M. (2019). Assessment of BOLD and GenBank–Their accuracy and reliability for the identification of biological materials. PloS One, 14, e0217084.

Menegon M., Cantaloni C., Rodriguez- Prieto A., Centomo C., Abdelfattah A., Rossato M., Bernardi M. et al. 2017. On site DNA barcoding by nanopore sequencing. PLoS One, 12: e0184741.

Merga, L. B., & Van den Brink, P. J. (2021). Ecological effects of imidacloprid on a tropical freshwater ecosystem and subsequent recovery dynamics. Science of The Total Environment, 784, 147167.

Mertens, M., Höss, S., Neumann, G., Afzal, J., & Reichenbecher, W. (2018). Glyphosate, a chelating agent-relevant for ecological risk assessment? Environmental Science and Pollution Research International, 25(6), 5298-5317. DOI: 10.1007/s11356-017-1080-1.

Metfies, K., Hwan, H. K., Yun, Y. S., & Wilson, W. H. (2016). Development of a quantitative PCR assay for monitoring Pseudo-nitzschia spp. (Bacillariophyceae) and associated domoic acid production in cultured and natural samples. Journal of Plankton Research, 38(6), 1423-1432.

Metzker, M. L. (2010). Sequencing technologies - the next generation. Nature Reviews Genetics, 11(1), 31-46. DOI: 10.1038/nrg2626.

Min, M. A., Needham, D. M., Sudek, S., Truelove, N. K., Pitz, K. J., Chavez, G. M., Poirier, C., Gardeler, B., von der Esch, E., Ludwig, A., Riebesell, U., Worden, A. Z., & Chavez, F. P. (2022). Ecological divergence of a mesocosm in an eastern boundary upwelling system assessed with multi-marker environmental DNA metabarcoding.

Mohrbeck, I., Raupach, M. J., Martínez Arbizu, P., Knebelsberger, T., & Laakmann, S. (2015). High-Throughput Sequencing-The Key to Rapid Biodiversity Assessment of Marine Metazoa? PLoS ONE, 10(10), e0140342. DOI: 10.1371/journal.pone.0140342.

Moiseenko, T. I. (2008). Aquatic ecotoxicology: Theoretical principles and practical application. Water Resources, 35, 530–541. DOI: 10.1134/S0097807808050047.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., & Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symposia on Quantitative Biology, 51(Pt 1), 263-273.

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. Molecular Ecology, 22, 2841.

Nistal-García, A., García-García, P., García-Girón, J., Borrego-Ramos, M., Blanco, S., & Bécares, E. (2021). DNA metabarcoding and morphological methods show complementary patterns in the metacommunity organization of lentic epiphytic diatoms. Science of The Total Environment, 786, 147410.

Nurk, S., et al. (2020). The complete sequence of a human genome. Science.

O'donnell, J. L., Kelly, R. P., Lowell, N. C., and Port, J. A.: Indexed PCR primers induce template- Specific bias in Large-Scale DNA sequencing studies, PLoS One, 11, 1–11,

Orsini, L., Mergeay, J., Vanoverbeke, J., De Meester, L., & Pfrender, M. E. (2013). Causes of hybridization and the evolution of hybrid zones in the *Daphnia* longispina species complex. Ecology and Evolution, 3(11), 3701-3714.

Oxford Nanopore Technologies. Medaka. https://nanoporetech.github.io/medaka.

Oxford Nanopore Technologies. MinION Analysis Pipeline (MAP) protocol. https://community.nanoporetech.com/protocols/MAPprotocol/v/map_9095_v2_rev_j_14dec2020

Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. Science, 276(5313), 734-740.

Parker J, Helmstetter AJ, Devey D, Wilkinson T, Papadopulos AST. 2017. Field-based species identification of closely-related plants using real-time nanopore sequencing. *Scientific Reports* 7, 8345.

Parmar, T. K., Rawtani, D., & Agrawal, Y. (2016). Bioindicators: the natural indicator of environmental pollution. Frontiers in life science, 9, 110-118.

Pathak, V. M., Verma, V. K., Rawat, B. S., Kaur, B., Babu, N., Sharma, A., Dewali, S., Yadav, M., Kumari, R., Singh, S., Mohapatra, A., Pandey, V., Rana, N., & Cunill, J. M. (2022). Current status of pesticide effects on environment, human health and it's eco-friendly management as bioremediation: A comprehensive review. Frontiers in Microbiology.

Pawlowski, J., Esling, P., Lejzerowicz, F., Cedhagen, T., & Wilding, T. A. (2014). Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. Molecular Ecology Resources, 14(6), 1129-1140.

Pawlowski, M. B., Branstrator, D. K., & Hrabik, T. R. (2018). Major shift in the phenology of crustacean biomass in western Lake Superior associated with temperature anomaly. Journal of Great Lakes Research, 44, 788–797.

Payne, A., Holmes, N., Clarke, T., Munro, R., Debebe, B. J., & Loose, M. (2021). Readfish enables targeted nanopore sequencing of gigabase-sized genomes. Nature Biotechnology.

Pennak, Robert W. (1946). The dynamics of fresh-water plankton populations. Ecological Monographs, 16(4), 339-355.

Persson, A., Hansson, L. A., & Bronmark, C. (2007). Effects of fish on nutritional ecology of zooplankton in a humic lake. Freshwater Biology, 52(5), 957-966.

Petersen, J. E., Cornwell, J. C., & Kemp, W. M. (2009). Implicit scaling in the design of experimental aquatic ecosystems. Oikos, 118(1), 1-9.

Pfeiffer, F., Gröber, C., Blank, M., Händler, K., Beyer, M., Schultze, J. L., et al. (2018). Systematic evaluation of error rates and causes in short samples in next-generation sequencing. Scientific Reports, 8, 10950

Pfenninger, M., & Schwenk, K. (2007). Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evolutionary Biology, 7, 121.

Pitz, K. J., Guo, J., Johnson, S. B., Campbell, T. L., Zhang, H., Vrijenhoek, R. C., Chavez, F. P., Geller, J. (2020). Zooplankton biogeographic boundaries in the California Current System as determined from metabarcoding. PLoS One, 15(6), e0235159. doi: 10.1371/journal.pone.0235159

Pomerantz, A., et al. (2023). Field-based portable sequencing for rapid biodiversity assessments. Nature Biotechnology.

Poonam, T., Bhattacharya, T., & Chakraborty, S. (2013). Water quality indices—important tools for water quality assessment: a review. International Journal of Advances in Chemistry, 1, 15-28.

Portinho, J. L., Nielsen, D. L., Daré, L., Henry, R., Oliveira, R. C., & Branco, C. C. Z. (2018). Mixture of commercial herbicides based on 2,4-D and glyphosate mixture can suppress the emergence of zooplankton from sediments. Chemosphere, 203, 151-159. doi: 10.1016/j.chemosphere.2018.03.156.

Questel, J. M., Hopcroft, R. R., DeHart, H. M., Kosobokova, K., Smoot, C., & Bucklin, A. (2021). Metabarcoding Analysis of Zooplankton Biodiversity of the Pacific-Arctic Chukchi Borderlands Region. Marine Biodiversity, 51(4). DOI: 10.1007/s12526-020-01136-x.

Quick J, Loman NJ, Duraffour S, et al.. 2016. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530, 228–232.

Rahman, H., Azani, N., Suhaimi, H., Yatim, S. R., Yuslan, A., & Ra, N. W. (2023). A Review on Different Zooplankton Culturing Techniques and Common Problems Associated with Declining Density. In International Conference on Sustainable Fisheries and Marine 2022 (INFISMA-2022), IOP Conf. Series: Earth and Environmental Science, 1147, 012012.

Rajagopal, T., Thangamani, A., Sevarkodiyone, S. P., Sekar, M., & Archunan, G. (2010). Zooplankton diversity and physico-chemical conditions in three perennial ponds of Virudhunagar district, Tamilnadu. Journal of Environmental Biology, 31, 265-272.

Rasconi, S., Gall, A., Winter, K., & Kainz, M. J. (2015). Increasing water temperature triggers dominance of small freshwater plankton. PLOS ONE.

Ratnasingham, S. & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes 7, 355-364. DOI: 10.1111/j.1471-8286.2006.01678.

Relyea, R., & Hoverman, J. (2006). Assessing the ecology in ecotoxicology: a review and synthesis in freshwater systems. *Ecology Letters, 9*(10), 1157-1171.

Reynolds, C. S. (2006). The Ecology of Phytoplankton. Cambridge University Press.

Rhoads, A., & Au, K. F. (2015). PacBio Sequencing and Its Applications. Genomics, Proteomics & Bioinformatics, 13(5), 278–289.

Richardson, A. J. (2008). In hot water: zooplankton and climate change. ICES Journal of Marine Science, 65, 279-295.

Rico-Martínez, R., Arias-Almeida, J. C., Pérez-Legaspi, I. A. (2012). Adverse Effects of Herbicides on Freshwater Zooplankton. In Herbicides - Properties, Synthesis and Control of Weeds. InTech. DOI: 10.5772/33558.

Ringelberg, J. (2010). Diel Vertical Migration of Zooplankton in Lakes and Oceans: Causal Explanations and Adaptive Significances. Springer.

Rohr, J. R., Kerby, J. L., & Sih, A. (2006). Community ecology as a framework for predicting contaminant effects. *Trends in Ecology and Evolution*, 21, 606–613.

Sánchez-Bayo, F. (2011). Impacts of agricultural pesticides on terrestrial ecosystems. In Ecological Impacts of Toxic Chemicals (pp. 63-87). Bentham Science Publishers Ltd.

Santoferrara, L. F. (2019). Current practice in plankton metabarcoding: optimization and error management. Journal of Plankton Research, 41, 571–582.

Sarkar, R. R., Pal, S., & Chattopadhyay, J. (2005). Role of two toxin-producing plankton and their effect on phytoplankton–zooplankton system – a mathematical study supported by experimental findings. Biosystems, 80(1), 11-23.

Sarma S, Nandini S, Gulati R (2005) Aquatic Biodiversity II. Dordrecht: Springer; 2005. Life history strategies of cladocerans: comparisons of tropical and temperate taxa; pp. 315–333. [Google Scholar]

Savin, M., Martin, J., Legresley, M., Giewat, M., & Rooney-Varga, J. (2004). Plankton diversity in the Bay of Fundy as measured by morphological and molecular methods. Microbial ecology, 48, 51-65.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., ... & Weber, C. F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and environmental microbiology, 75(23), 7537-7541.

Schmitt-Jansen, M., Veit, U., Dudel, G., & Altenburger, R. (2008). An ecological perspective in aquatic ecotoxicology: Approaches and challenges. Basic and Applied Ecology, 9(4), 337-345.

Schroeder, A., Pallavicini, A., Edomi, P., Pansera, M., & Camatti, E. (2021). Suitability of a dual COI marker for marine zooplankton DNA metabarcoding. Marine Environmental Research, 170, 105444. DOI: 10.1016/j.marenvres.2021.105444.

Schroeder, A., Stanković, D., Pallavicini, A., Gionechetti, F., Pansera, M., & Camatti, E. (2020). DNA metabarcoding and morphological analysis - Assessment of zooplankton biodiversity in transitional waters. Marine Environmental Research, 160, 104946.

Semmouri, I., De Schamphelaere, K. A. C., Willemse, S., Vandegehuchte, M. B., Janssen, C. R., & Asselman, J. (2021). Metabarcoding reveals hidden species and improves identification of marine zooplankton communities in the North Sea. ICES Journal of Marine Science, 78(9), 3411–3427.

Semmouri, I., De Schamphelaere, K., Mees, J., Janssen, C., & Asselman, J. (2020). Evaluating the potential of direct RNA nanopore sequencing: Metatranscriptomics highlights possible seasonal differences in a marine pelagic crustacean zooplankton community. Marine Environmental Research, 153, 104836.

Shafin, K., et al. (2021). Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. Nature Biotechnology.

Shannon, C. E., and Wiener, W. (1963). The Mathematic Theory of Communication. Urbana: University of Illinois Press.

Shokralla et al. (2012) Shokralla S, Spall JL, Gibson JF, Hajibabaei M. Next-generation sequencing technologies for environmental DNA research. Molecular Ecology. 2012;21:1794–1805. doi: 10.1111/j.1365-294X.2012.05538.x.

Singh, U. B., Ahluwalia, A., Sharma, C., Jindal, R., & Thakur, R. (2013). Planktonic indicators: A promising tool for monitoring water quality (early-warning signals). Ecology, Environment and Conservation, 19, 793-800.wn

Sládeček , V. (1983) Rotifers as indicators of water quality. Hydrobiologia. 1983;100:169–201. doi: 10.1007/BF00027429.

Ślósarczyk, K., Jakóbczyk-Karpierz, S., Różkowski, J., & Witkowski, A. J. (2021). Occurrence of Pharmaceuticals and Personal Care Products in the Water Environment of Poland: A Review. Water, 13(16), 2283.

Smayda, T. J. (1997). Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. Limnology and Oceanography, 42(5), 1137-1153.

Somerfield, P. J., Clarke, K. R., & Gorley, R. N. (2021). Analysis of similarities (ANOSIM) for 3-way designs. Austral Ecology, 46, 927–941.

Sommer, U., Aberle, N., Lengfellner, K., & Lewandowska, A. (2012). The Baltic Sea spring phytoplankton bloom in a changing climate: an experimental approach. Marine Biology, 159, 2479–2490.

Sommer, U., Adrian, R., De Senerpont Domis, L., Elser, J. J., Gaedke, U., Ibelings, B., Jeppesen, E., Lürling, M., Molinero, J. C., & Mooij, W. M. (2012). Beyond the Plankton Ecology Group (PEG) model: mechanisms driving plankton succession. Annual Review of Ecology, Evolution, and Systematics, 43, 429-448.

Sommer, U., Gliwicz, Z. M., Lampert, W., & Duncan, A. (1986). The PEG-model of seasonal succession of planktonic events in fresh waters. Archiv für Hydrobiologie, 106, 433-471.

Song, Cu., Choi, H., Jeon, MS., et al. (2021). Zooplankton diversity monitoring strategy for the urban coastal region using metabarcoding analysis. Sci Rep, 11, 24339. DOI: 10.1038/s41598-021-03656-3.

Stelzer, C. P. (2011). The cost of sex and competition between cyclical and obligate parthenogens in a rotifer. The American Naturalist, 177(2), E43-E53.

Stephan, L. R., Castilho-Noll, M. S. M., & Henry, R. (2017). Comparison among zooplankton communities in hydrologically different lentic ecosystems. Limnetica, 36(1), 99-112. DOI: 10.23818/limn.36.08

Stevens, B. M., Creed, T. B., Reardon, C. L., & Manter, D. K. (2023). Comparison of Oxford Nanopore Technologies and Illumina MiSeq sequencing with mock communities and agricultural soil. Scientific Reports, 13(1), 9323. doi: 10.1038/s41598-023-36101-8.

Stewart, R. I., Dossena, M., Bohan, D. A., Jeppesen, E., Kordas, R. L., Ledger, M. E., ... & Woodward, G. (2013). Mesocosm experiments as a tool for ecological climate-change research. Advances in Ecological Research, 48, 71-181.

Stollewerk, A. (2010). The water flea *Daphnia*-a 'new' model system for ecology and evolution? Journal of biology, 9, 1-4.

Suter II, G. W. (1983). Multispecies tests for environmental toxicology. Environment International, 9, 157-160.

Svendsen, C., & Weeks, J. M. (1997). A simple low-cost field mesocosm for ecotoxicological studies on earthworms. Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology, 117(1), 31-40.

Sylwestrzak, Z., Zgrundo, A., & Pniewski, F. (2021). Ecotoxicological studies on the effect of Roundup® (Glyphosate formulation) on marine benthic microalgae. International Journal of Environmental Research and Public Health, 18, 884. DOI: 10.3390/ijerph18030884.

Szöcs, E., Van den Brink, P. J., Lagadic, L., Caquet, T., Roucaute, M., Auber, A., Bayona, Y., Liess, M., Ebke, P., Ippolito, A., Ter Braak, C. J. F., Brock, T. C. M., & Schäfer, R. B. (2015). Analysing chemical-induced changes in macroinvertebrate communities in aquatic mesocosm experiments: A comparison of methods. Ecotoxicology. Accepted 17 January 2015.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution, 30, 2725-2729.

Tarazona, J. V., & Ramos-Peralonso, M. J. (2014). Ecotoxicology. In Encyclopedia of Toxicology (Third Edition).

Tarazona, J. V., Court-Marques, D., Tiramani, M., Reich, H., Pfeil, R., Istace, F., & Crivellente, F. (2017). Glyphosate toxicity and carcinogenicity: a review of the scientific basis of the European Union assessment and its differences with IARC. Archives of Toxicology, 91(8), 2723-2743. DOI: 10.1007/s00204-017-1962-5.

Torstensson, L., Börjesson, E., & Stenström, J. (1989). Efficacy and environmental fate of glyphosate on different types of weeds. Swedish Journal of Agricultural Research, 19, 169-174.

Traas, T.P., Van Leeuwen, K. (2007). Ecotoxicological Effects. In: Risk Assessment of Chemicals. DOI: 10.1007/978-1-4020-6102-8_7.

Tresnakova, N., Stara, A., & Velisek, J. (2021). Effects of Glyphosate and Its Metabolite AMPA on Aquatic Organisms. Applied Sciences, 11, 9004. DOI: 10.3390/app11199004.

Tyler, A. D., Mataseje, L., Urfano, C. J., Schmidt, L., Antonation, K. S., Mulvey, M. R., & Corbett, C. R. (2018). Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications. Scientific Reports, 8(1), 1–12.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. Nucleic Acids Research, 40, e115-e115.

Valenzuela-González, F., Martínez-Porchas, M., Villalpando-Canchola, E., & Vargas-Albores, F. (2016). Studying long 16S rDNA sequences with ultrafast-metagenomic sequence classification using exact alignments (Kraken). Journal of Microbiological Methods, 122, 38-42.

Van den Brink, P. J., & Ter Braak, C. J. F. (1999). Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress. Environmental Toxicology and Chemistry, 18(2), 138–148.

van der Loos, L. M., & Nijland, R. (2020). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. Molecular Ecology. DOI: 10.1111/mec.15592.

van der Reis, A. L., Beckley, L. E., Olivar, M. P., & Jeffs, A. G. (2022). Nanopore shortCread sequencing: A quick, costCeffective and accurate method for DNA metabarcoding. Environmental DNA, 5, 282–96.

Vandenberg, L. N., Colborn, T., Hayes, T. B., Heindel, J. J., Jacobs, D. R. Jr., Lee, D.-H., Shioda, T., Soto, A. M., vom Saal, F. S., Welshons, W. V., Zoeller, R. T., & Myers, J. P. (2012). Hormones and endocrine-disrupting chemicals: Low-dose effects and nonmonotonic dose responses. Endocrine Reviews, 33(3), 378–455. doi: 10.1210/er.2011-1050.

VanGuilder, H. D., Vrana, K. E., & Freeman, W. M. (2008). Twenty-five years of quantitative PCR for gene expression analysis. Biotechniques, 44, 619-626.

Vaser, R., & Šikić, M. (2021). Yet another de novo genome assembler. Nature Biotechnology.

Vera, M. S., Di Fiori, E., Lagomarsino, L., Sinistro, R., Escaray, R., Iummato, M. M., Juárez, A., Ríos de Molina, M. C., Tell, G., & Pizarro, H. (2012). Direct and indirect

effects of the glyphosate formulation Glifosato Atanor® on freshwater microbial communities. Ecotoxicology, 21(7), 1805–1816.

Wallace, R. L., Snell, T. W., Ricci, C., & Nogrady, T. (2006). Rotifera, vol 1: Biology, ecology and systematics, 2nd edn. Backhuys, Leiden.

Wang, C., Lin, X., Li, L., & Lin, S. (2016). Differential Growth Responses of Marine Phytoplankton to Herbicide Glyphosate. PloS One, 11(3), e0151633. DOI: 10.1371/journal.pone.0151633.

Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. Nature Biotechnology, 39(12), 1348–1365.

Woese, C. R. (1987). Bacterial evolution. Microbiological reviews, 51(2), 221-271.

Xiao, Y. M., Wang, J. A., Wang, M. A., Liu, J. P., Yuan, H. Z., & Qin, Z. H. (2010). Study on the inclusion complexes of flumorph and dimethomorph with β-cyclodextrin to improve fungicide formulation. Journal of the Chemical Society of Pakistan, 32, 363.

Yang, J., Yang, J., Zhang, X., Xie, Y., Song, C., Sun, J., Zhang, Y., Giesy, J. P., & Yu, H. (2017a). Ecogenomics of zooplankton community reveals ecological threshold of ammonia nitrogen. Environmental Science & Technology, 51, 3057–3064. doi: 10.1021/acs.est.6b05606.

Yoon, T.-H., Kang, H.-E., Kang, C.-K., Lee, S. H., Ahn, D.-H., Park, H., & Kim, H.-W. (2016). Development of a cost-effective metabarcoding strategy for analysis of the marine phytoplankton community. PeerJ, 4, e2115.

Zabaloy, M. C., Garland, J. L., & Gómez, M. A. (2008). Assessment of the impact of glyphosate application on soil microbial activity using multi-substrate enzyme assays. Environmental Pollution, 156(3), 1081-1088.

Zaidi, B. R., & Imam, S. H. (2008). Biodegradability. Encyclopedia of Ecology, 357-366. Reference Module in Earth Systems and Environmental Sciences.

Zhan, A., Hulak, M., Sylvester, F., Huang, X., & Adebayo, A. A. (2013). Invasive round goby Neogobius melanostomus and tubenose goby Proterorhinus semilunaris (Teleostei: Gobiidae) in the Danube: first record and invasion history. Journal of Applied Ichthyology, 29(2), 407-410.

Zhang, Y., Hu, M., & Li, A. (2024). Review of the Occurrence of Herbicides in Environmental Waters of Taihu Lake Basin and Its Potential Impact on Submerged Plants. Water, 16, 726. DOI: 10.3390/w16050726.

Zhang, Z., Li, J., Li, H., Wang, L., Zhou, Y., Li, S., Zhang, Z., Feng, K., & Deng, Y. (2023). Environmental DNA metabarcoding reveals the influence of human activities on microeukaryotic plankton along the Chinese coastline. Water Research, 233, 119730.

Zhao, L., Zhang, X., Xu, M., Mao, Y., & Huang, Y. (2021). DNA metabarcoding of zooplankton communities: species diversity and seasonal variation revealed by 18S rRNA and COI. PeerJ, 9, e11057.

Zhu, C., & Yu, J. (2009). Nonmetric multidimensional scaling corrects for population structure in association mapping with different sample types. Genetics, 182(3), 875–888. doi: 10.1534/genetics.108.098863.