DTU Library

Viral Metagenomics to Assess the Microbial Quality of Reused Wastewater

Hellmér, Maria

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Viral Metagenomics to Assess the Microbial Quality of Reused Wastewater

PhD Thesis
Maria Hellmér

March 2018

DTU Food
National Food Institute

Preface

This PhD thesis is submitted to meet the requirements for obtaining a PhD degree at the National Food Institute, Technical University of Denmark.

The work was carried out in the Division of Microbiology and Production, National Food Institute, Technical University of Denmark, in the period from December 2014 to March 2018. Part of the work was carried out at the department of Microbiology at University of Barcelona, Spain, under the supervision of Professor Rosina Girones (October-November 2015).

Abstract

Reused wastewater has been suggested as a stable freshwater source that answers the increased demand linked to the global population growth. Raw wastewater harbors a diverse population of virus and bacteria, including pathogens, of which several can cause disease in humans even at low concentrations. Among these the enteric viruses, including norovirus (NoV), are one of the most frequent causes of food and waterborne diseases worldwide. Moreover, the transmission of pathogens to human via food or water is often due to the reuse of insufficiently cleaned wastewater in food production. To ensure food safety it is therefore important to assure that the reclaimed wastewater is free from viral pathogens before it is used during the production of ready-to-eat food products.

Traditionally the microbial water quality has been accessed by use of fecal indicator bacteria. However, a number of studies agree in a lack of correlation between levels of fecal indicator bacteria and the presence of e.g. human pathogenic enteric viruses. To decrease the risk of foodborne outbreaks due to reused wastewater it is thus of great importance to investigate and acknowledge the whole viral content within the water to enable the use of proper treatment measures. To address this question, metagenomics sequencing could be a promising tool.

The aim of this PhD thesis was to investigate if viral metagenomics sequencing can been used to assess the microbial quality of water. Further goals were to investigate the change in viral community throughout the wastewater treatment procedures and if possible determinate the treatment needed before reuse of wastewater.

Initially the effect of different combinations of viral concentration and extraction methods on the detected viral metagenome in wastewater samples was investigated. The results showed that the particular selected method had a strong effect on the detected viral metagenome (**Hellmér I**). However, none of the 16 tested method combinations appeared to be superior in all tested parameters. Each method and method combination showed strengths and weaknesses. Based on the results obtained and two specific criteria, favoring the possibility to concentrate viruses from large volumes of water and the detection of genomes from RNA viruses, the monolithic adsorption filtration (MAF) combined with the extraction kit Nucleospin RNA XS were chosen in subsequent studies.

Following this an investigation of the usefulness of viral metagenomcis to assess the micobiological water quality was performed. In **Hellmér II** the viral population was investigated before and after different treatment steps in one conventional urban wastewater treatment plant (WWTP) and in one highly advanced hospital WWTP. In addition, quantitate PCR analysis was applied to investigate the levels of the broadly proposed human viral indicators adenovirus (HAdV) and JC polyomavirus (JCPyV) as well as the pathogens, NoV, *Salmonella,* and *Campylobacter*. Results showed that analysis of the total viral community structure

could differentiate microbial contaminated water from clean water without looking at specific pathogens. We observed a change in the viral community structure in effluent from the membrane bioreactor (MBR) at the hospital WWTP. The similarity of the viral composition in the MBR effluent to the community composition of sterile molecular grade water indicated a good microbial quality, although NoV and *Salmonella* were both detected in one of nine MBR effluent samples. At the hospital WWTP, the MBR effluent is further treated by ozone, granular activated carbon, and UV irradiation before being discharged. No microorganisms were detected in these samples. In the raw wastewater at the urban WWTP a seasonal pattern of the levls of NoV and HAdV was observed with a decrease in concentration during the summer months. This result contradicts previous observations of a stable concentration of HAdV throughout the year in wastewater in other countries, and may therefore question the usefulness of HAdV as a fecal indicator in waters.

During the PhD project we received drinking water suspected to be the cause of a NoV outbreak (**Hellmér III**). The received volume of water enabled a small comparison between viral concentration methods on naturally NoV contaminated drinking water. NoV was detected with RT-qPCR and characterized with Sanger sequencing in the water sample concentrated with MAF. It was however not possible to further verify this finding of NoV using viral metagenomcis, although the bioinformatic analysis did detect the two viral process controls initially spiked in the water sample. Failure to detect NoV could be due to method dependent biases observed in **Hellmér I** or that the initial concentration of NoV in the water sample may have been below the limit of detection.

In conclusion this PhD thesis has extended the knowledge of viral metagenomics for microbial water quality assessment and outbreak analysis. The effect of laboratory methods to prepare water samples for viral metagenomics sequencing has been investigated. This has increased our understanding of method dependent biases and stress the importance of carefully choosing the laboratory method best suited for the project in question. The PhD study demonstrated for the first time successful application of MAF to detect NoV in contaminated water implicated in a disease outbreak. The three studies within this PhD thesis also highlight the bioinformatic challenges in analyzing viral metagenomics data. Finally the presented work showed that viral metagenomcis could distinguish between the microbial water qualities throughout different stages of wastewater treatment. In combination with existing control measures metagenomics could be a powerful tool to analyze the effectiveness of the process in a wastewater treatment plant. Future studies are needed to validate these results before the method can be used in water management.

Dansk Resumé

Genbrug af spildevand har været foreslået som en mulig løsning på det vandbehov som den globale befolkningsvækst medfører. Råt spildevand indeholder en varieret population af virus og bakterier, også patogener, som kan medføre sygdom selv i lave koncentrationer. Blandt disse er enteriske virus, herunder norovirus (NoV), som er en af de hyppigste årsager til fødevare- og vandbårne sygdomme verden over. Smittevejen for patogener via fødevarer og vand er ofte via genbrug af utilstrækkeligt renset spildevand i fødevareproduktionen. For at sikre fødevaresikkerheden er det derfor vigtigt at det genbrugte vand er frit for patogene virus før vandet anvendes i produktionen af spiseklare fødevarer.

Den mikrobiologiske vandkvalitet er traditionelt blevet testet ved analyse for fækale indikatorbakterier. Der er imidlertid flere studier der viser, at der ikke er sammenhæng mellem niveauer af fækale indikatorbakterier og forekomsten af enteriske virus. For at reducere risiko for fødevarebårne udbrud er det vigtigt at kende det totale indhold af virus i vandet, for at kunne identificere og anvende en effektiv vandrensningsmetode. Til dette formål virker metagenom sekventering til at være en lovende metode.

Formålet med denne PhD afhandling var at undersøge om viral metagenomsekventering kan anvendes til at vurdere den mikrobiologiske vandkvalitet. Det var desuden et mål at undersøge ændringer i den virale population igennem rensningsprocessen af spildevand og om muligt bestemme den nødvendige vandrensningsteknik før det rensede spildevand kan genanvendes i fødevareproduktionen.

Projektet indledtes med at undersøg hvilken effekt forskellige kombinationer af opkoncentrerings og oprensnings metoder af virus havde på det påviste virale metagenom i spildevandsprøver. Det viste sig at den valgte kombination af metode havde en betydelig effekt på det påviste virale metagenom (**Hellmér I**). Det viste sig tillige at ingen af de 16 anvendte metodekombinationer var optimal på alle parametre. Hver metode havde sine styrker og svagheder. Baseret på resultater af to specifikke parametre, hvor muligheden for at koncentrere virus fra store vand voluminer samt muligheden for påvisning af genomer fra RNA virus blev prioriteret, udvalgtes den bedste metodekombination til anvendelse i de efterfølgende studier. Den udvalgte kombination bestod af koncentrering ved monolithisk adsorbtionsfiltrering (MAF) og oprensning med Nucleospin RNA XS.

Efterfølgende, blev anvendeligheden af metagenomanalyse til test af vandkvalitet undersøgt. I **Hellmér II** blev den virale population analyseret efter forskellige procetrin i et konventionelt spildevandsanlæg, samt i et avanceret hospitals-spildevandsanlæg. Udover anvendelsen af metagenomanalyser, blev der ved kvantitativ PCR analyse analyseret for de ofte foreslåede human virale indikatore, adenovirus (HAdV) og JC polymyxavirus (JCPyV) samt for patogenerne NoV, *Salmonella,* og *Campylobacter*.

Resultaterne viste at metagenomanalyse af den totale virale populationsstruktur, uden at fokusere på specifikke patogener, var i stand til at skelne imellem mikrobiologisk forurenet og rent vand. På hospitals-

spildevandsanlægget blev der observeret en ændring i den virale population efter behandling i en membran bioreaktor (MBR). Den virale sammensætning af MBR behandlet spildevand matchede overordnede sterilt vand, hvilket tydede på en god vandkvalitet, trods at der i en enkelt ud af ni MBR behandlede vandprøver kunne påvises både NoV og *Salmonella*. I spildevandsanlægget bliver det MBR behandlede vand efterfølgende udsat for ozon, filtreret gennem aktivt kul og UV bestrålet før det udledes. I disse prøver fandtes ingen mikroorganismer. I det rå spildevand fra det konventionelle spildevandsanlæg påvistes en sæsonvariation i niveauet af NoV og HAdV, med lavere koncentrationer observeret i sommermånederne. Dette i modsætning til tidligere observationer i udlandet, hvor der året rundt er påvist en stabil koncentration af HAdP i spildevand. En mulig sæsonvariation kan sætte spørgsmål ved anvendeligheden af HAdV som fækal indikator.

I forbindelse med projektet modtog vi en prøve af drikkevand der var mistænkt som årsag til et udbrud af NoV (**Hellmér III**). Dette tillod et mindre sammenligningsstudie af forskellige metoder til opkoncentrering af virus i drikkevand. NoV blev påvist ved RT-qPCR og karakteriseret ved Sanger sekventering i vandprøven opkoncentreret med MAF. Det var dog ikke muligt at verificere dette fund af NoV i vandet yderligere med viral metagenomanalyse, på trods af de bioinformatiske analyser resulterede i genfindelse af de to lignende virale proceskontroller der blev tilsat vandprøverne som positive kontroller. Dette kan skyldes metodebias der blev observeret i **Hellmér I** eller at koncentrationen af den tilsatte NoV har været under påvisningsgrænsen.

Denne PhD afhandling har udvidet kendskabet til brug af viral metagenomics ved vurdering af vandkvalitet og udbrudsopsporing. Forskellige laboratoriemetoder til forberedelse af vandprøver forinden metagenomanalyse for virus er blevet undersøgt, hvilket har øget vores forståelse af de enkelte metoders bias, og understreget vigtigheden af at vælge den bedst egnede metode til det givne formål. PhD projektet demonstrerede desuden for første gang succesfuld anvendelse af MAF til påvisning af NoV i forurenet drikkevand, der førte til udbrud af Roskildesyge. De tre studier i PhD projektet har også belyst udfordringerne ved brugen af bioinformatics ved analyse af virale metagenomdata. Endelig er det vist at viral metagenomisc kan skelne mellem den mikrobiologiske vandkvalitet gennem forskellige procestrin i et spildevandsrensningsanlæg. Kombineret med klassiske kontrolmetoder kan metagenomics derfor være et værdifuldt redskab til vurdering af effektiviteten i et rensningsanlæg. Yderligere studier er dog nødvendige for at validere disse resultater før metoden vil kunne tages i anvendelse.

List of papers

Three original papers, one published and two manuscripts, are included in this thesis.

I. Hjelmsø MH*, **Hellmér M***, Fernandez-Cassi X, Timoneda N, Lukjancenko O, Seidel M, Elsässer D, Aarestrup FM, Löfström C, Bofill-Mas S, Abril JF, Girones R, Schultz AC. (2017) 'Evaluation of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic Sequencing', Plos One, 12(1), p. e0170199. doi: 10.1371/journal.pone.0170199.
    * These authors contributed equally to this work.

II. **Hellmér M,** Lukjancenko O, Nielsen U, Stranddorf K, Kohle S, Seidel M, Aarestrup FM, Schultz AC. Viral metagenomics used to assess the microbial quality of water in two Danish wastewater treatment plants. *Manuscript in preparation*

III. **Hellmér M*,** Ahlmann-Nielsen A*, Rasmussen LD, Villif A, Jensen T, Lukjancenko O, Seidel M, Schultz AC. Application of Monolithic Absorption Filtration in the Investigation of a Waterborne Outbreak of Norovirus in Denmark, 2015. *Manuscript in preparation*
    * These authors contributed equally to this work.

Abbreviations

AS	Activated sludge
BOD	Biological oxygen demand
cDNA	Complementary DNA
COD	Chemical oxygen demand
ds	Double stranded genome
FIB	Fecal indicator bacteria
GE	Gastroenteritis
HAdV	Human adenovirus
HAV	Hepatitis A virus
HEV	Hepatitis E virus
MBR	Membrane bioreactors
MPN	Most probable number
MST	Microbial source tracking
NA	Nucleic acid
NGS	Next generation sequencing
NoV	Norovirus
PCR	Polymerase chain reaction
PFU	Plaque-forming units
ss	Single stranded genome
qPCR	quantitative PCR
RT-qPCR	Reverse transcriptase qPCR
WWTP	Waste water treatment plant

Table of Contents

1. Introduction

1.1 Motivation for the study

The global demand for freshwater for agricultural and industrial use is estimated to increase by 55% within the next 30 years (UN-WWAP, 2015), driven mainly by an estimated global population growth from7.5 billion today to 9.8 billion in 2050 (United Nations Department of Economic and Social Affairs, 2017). A growing population will also produce a larger volume of wastewater. In the United Nations world water development report from 2015, reused wastewater is suggested as a stable freshwater source (UN-WWAP, 2015), but before reusing wastewater it needs to be cleaned from harmful chemicals, toxins, and microorganisms (WHO, 2006).

Raw wastewater harbors a diverse viral and bacterial population, including human pathogens (Cantalupo *et al.*, 2011). To enable reuse of wastewater, antimicrobial treatment needs to be thorough since many of the pathogens can cause disease in humans even at low concentrations (Leclerc, Schwartzbrod and Dei-Cas, 2002; Koopmans and Duizer, 2004). Pathogens may lose infectivity during heat treatment, e.g. cooking (Araud *et al.*, 2016; Ceylan, McMahon and Garren, 2017), but if not sufficiently inactivated or if eaten raw, contamination of food products will pose a risk of illness to consumers. Examples of such products are irrigated vegetables and fruits (Ethelberg *et al.*, 2010), shellfish grown in contaminated waters (Rasmussen *et al.*, 2016), or berries rinsed and frozen with contaminated water (Bernard *et al.*, 2014).

The public health promotion of the nutritional value of vegetables and greens (FAO/WHO, 2004) has increased the demand for fresh produce in industrialized countries. However, during the last few years a high number of outbreaks related to fresh produce have been reported (Callejón *et al.*, 2015). To ensure food safety it is important to assure that reclaimed wastewater is free from pathogens before it is used during the production of ready-to-eat food products where no microbial inactivation steps are included in the food production chain.

Today there are no regulations within the EU stipulating the reduction of microorganisms during the treatment of wastewater (EEC Council, 1991) and no EC Directive regarding irrigation. There is a newly published European guideline regarding irrigation water (European Union, 2017), which relies on the use of the indicator organism *Escherichia coli (E. coli).* However, a number of studies show a lack of correlation between *E. coli* levels and presence of human pathogenic viruses and bacteria (Gerba *et al.*, 1979; Wéry *et al.*, 2008; Ferguson *et al.*, 2012; Zhou *et al.*, 2015; McMinn, Ashbolt and Korajkic, 2017). To decrease the risk of foodborne outbreaks due to reused wastewater it is of great importance to investigate the whole

microbial content within the water to enable the use of proper treatment measures. Metagenomics sequencing is therefore a promising tool to answer these questions (Nieuwenhuijse and Koopmans, 2017).

1.2 Objectives

The main aim of this thesis was to assess the risk of using wastewater for food production by determining the effect of different treatment steps on the total microbial content of the water. This was done by sequencing the total pool of nucleic acids, the metagenome, in water samples collected at different stages during the treatment of wastewater. Further goals were to investigate and assess the elimination of selected viruses and pathogenic bacteria throughout the wastewater treatment procedures. To achieve the goals four specific objectives were identified.

1. Develop a methodological pipeline that covers all steps, from sampling of water to the detection of microbial communities by the use of metagenomics sequencing.
2. Analyze the viral population after different steps along the treatment process in one conventional municipal wastewater treatment plant (WWTP) and one highly advanced hospital WWTP.
3. Investigate the correlation between the presence of pathogenic viruses and the current indicator organisms *E. coli* with the broadly proposed viral indicators human adenovirus (HAdV) and JC polyoma virus (JCPyV).
4. Validate the developed method pipeline during the investigation of water incriminated in a disease outbreak.

Objective 1 was addressed by a method study in the laboratory in which we compared the efficiency of viral concentration and genome extraction methods for studying viral metagenomics from raw wastewater (**Hellmér I**). The results obtained in **Hellmér I** were used as a guidance in planning the investigation of microbial elimination in two Danish WWTPs, one conventional and one pilot highly advance hospital WWTP (**Hellmér II**) to reach objectives 2 and 3. In addition, the methods developed in **Hellmér I** and the further experience on bioinformatic analysis of metagenomics data gained in **Hellmér II** were used to investigate water that may have caused a viral disease outbreak due to wastewater contamination (**Hellmér III**).

1.3 Outline

Chapter 2 provides an overview of the theoretical background related to the research in this PhD study. The chapter introduces the waterborne microbial pathogens typically present in the influent raw wastewater and how they are affected by the construction of WWTPs. Chapter 2 continues with a discussion on the risks associated with the reuse of treated wastewater backed up with an introduction of the current regulations. In chapter 3 the methodological processes used for analyzing viruses and bacteria in water are discussed, including the methods used in this thesis. In chapter 4 the findings from **Hellmér I, II, and III** are discussed,

and chapter 5 presents conclusions, impacts and further perspectives learned from this PhD thesis. Finally chapter 6 includes the three original papers and manuscripts produced during the PhD thesis.

2. Background

2.1 Waterborne pathogens

Infectious diseases transmitted through food are a huge public health burden worldwide, resulting in considerable morbidity and mortality (Kirk *et al.*, 2015; WHO, 2015). Foodborne infectious agents may contaminate foods through a number of possible transmission routes e.g. through contact with contaminated hands or surfaces during processing, or from contaminated water.

Waterborne pathogens are microorganisms that can be transmitted and cause disease through ingestion of contaminated water (Leclerc, Schwartzbrod and Dei-Cas, 2002). Most food- and waterborne pathogens are spread through the fecal-oral route and will reach the WWTP through excreted feces. High concentrations of food- and waterborne pathogens are thus found in raw wastewater as well as throughout the wastewater treatment process. Examples of selected viruses and bacteria and their levels present after different treatment steps are listed in Table 1. If poorly treated wastewater is directly used or contaminates the water used in food production pathogens could be spread to consumers and result in illness, Figure 1.

Table 1: Viral and bacterial concentrations in, log genome copies / L, throughout wastewater treatment.

	Raw wastewater influent	Primary effluent	Activated sludge effluent	MBR effluent	Ozone disinfection effluent	Reference
Norovirus	4.7	4.4	2.1	2.3		**Hellmér II**
		3.6	1.9	0.6		Francy et., al 2012
HAdV	4.7	4.0	1.6			**Hellmér II**
	8.0	6.0			6.9	Wang et.. al 2018
		4.0	1.8	0.5		Francy et., al 2012
JCPyV	4.1	3.6	1.9			**Hellmér II**
Salmonella	3.0	2.3	1.3	1.8		**Hellmér II**
Campylobacter	2.7	2.2	0.9			**Hellmér II**

Figure 1: Transmission route for waterborne pathogens spreading through contaminated foods.

This section will describe some of the most common viruses and bacteria causing illness through the consumption of fresh produce processed with contaminated water.

2.1.1 Viruses

Viruses are obligate parasites and cannot replicate outside a host. They are consists of a genome enclosed in a protein capsid, which can be surrounded by a lipid membrane or envelope, Figure 2. The viral genome consists of either RNA or DNA and could be single- or double-stranded (ss/ds). The ss genomes are either positive- or negative-sensed. The genome structure is highly variable and can be circular, linear or segmented (Acheson, 2007).

.
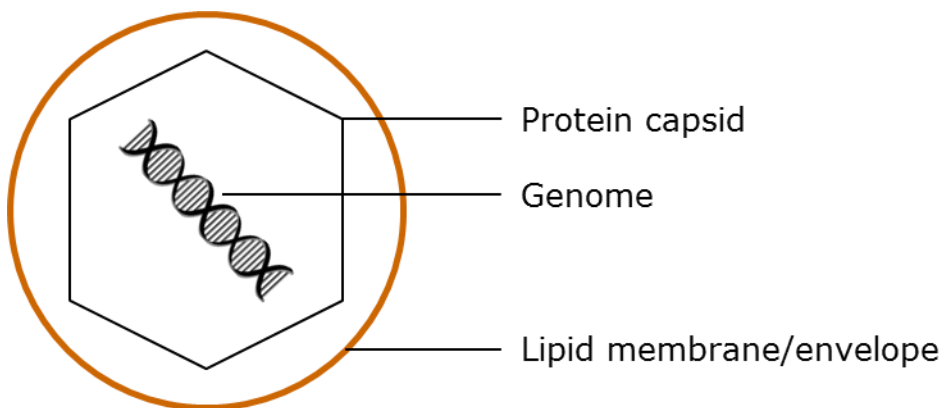
Figure 2. Schematic picture of a virus with a genome, protein capsid and lipid membrane/envelope

Naked viruses that do not have a membrane envelope are often more persistent in the environment. xamples like poliovirus have been found to survive for up to 140 days on soft surfaces (Yeargin *et al.*, 2016), murine norovirus for at least 7 days on spinach leafs (Hirneisen and Kniel, 2013), HAdV for over 70 days in water (Prevost *et al.*, 2016), and one study detected survival of HAdV for over 300 days in virus spiked sterile water (Rigotto *et al.*, 2011). In addition, the non-enveloped viruses typically show a high resistance to disinfection measurements (Prevost *et al.*, 2016) and common treatment procedures at WWTPs (Rodriguez-Manzano *et al.*, 2012). The inability to replicate outside a host cell makes viruses rare in the environment, and their small genome size compared to bacterial or protozoal species (Hodgkin, 2001) means only a small fraction of the total nucleic acid within a sample is of viral origin, making them hard to detect in a water sample. However, the low infectious dose of particular enteric viruses (Koopmans and Duizer, 2004) makes them a health risk even when they are present at low concentrations.

Viruses can infect a wide range of hosts including humans, animals, plants, bacteria, or even other viruses (Acheson, 2007). This thesis will focus on the enteric viruses those which infect via the gastrointestinal tract and therefore can infect through food and drinking water. The environment within the human gastrointestinal tract is harsh and therefore most of the enteric viruses lack the lipid-containing envelope which otherwise would make them susceptible to environmental degradation. The genomic structure and specific symptoms from some of the most common foodborne viruses are listed in Table 2.

Table 2: List of important enteric viruses including: genome size (in kilo bases (kb)), sense of reading frame(positive/negative (+/-)) and nucleic acid format (ss or ds, RNA or DNA), as well as main symptoms of infection and mostly applied detection method used in environmental samples.

Family	Virus	Genome structure	Envelope	Genome Size (kb)	Symptoms	Detection method used in environmental samples
Caliciviridae	Norovirus	+ ssRNA	no	7-8	Gastroenteritis	Molecular
	Sapovirus	+ ssRNA	no	7-8	Gastroenteritis	Molecular
Adenoviridae	Human Adenovirus 40 and 41	dsDNA	no	30-36	Gastroenteritis	Culture Molecular
Picornaviridae	Hepatitis A virus	+ ssRNA	no	7-8	Hepatitis	Culture Molecular
	Enterovirus	+ ssRNA	no	7-8	Gastroenteritis Meningitis Encephalitis Common cold Conjunctivitis	Culture Molecular
Reoviridae	Rotavirus	dsRNA	no	7-8	Gastroenteritis	Molecular Culture
Hepeviridae	Hepatitis E virus	+ ssRNA		7.2	Hepatitis	Molecular Culture

Information obtained from (Rodríguez-Lázaro *et al.*, 2012)

Among the most common foodborne viruses are norovirus (NoV), sapovirus, enterovirus, aichi virus, hepatitis A and E viruses (HAV and HEV, respectively), astrovirus, rotavirus, and HAdV. All these viruses with the exception of HAdV and rotavirus are ssRNA viruses. HAdV has a dsDNA genome and rotavirus a dsRNA genome. The most common symptom upon infection by enteric viruses is gastroenteritis (GE) but hepatitis and neurological symptoms are caused by some of the viruses (Rodríguez-Lázaro *et al.*, 2012). Below is more detailed presentation of the most common foodborne viruses organized by their disease symptoms.

2.1.1.1 Enteric viruses causing gastroenteritis

GE is the term for an infection of the gastrointestinal tract, with symptoms including diarrhea, vomiting, and abdominal pain. Viral induced GE is often self-limiting in healthy individuals but could cause more severe symptoms in risk groups such as elderly, children, or hospitalized persons (Lopman *et al.*, 2004). NoV is the leading cause of GE infections in adults and is the largest source of foodborne disease worldwide (WHO, 2015). During 2016, NoV was the cause of 37 % of the foodborne outbreaks in Denmark (Anonymous, 2016), and was thus the most common cause of foodborne outbreaks within the country. NoVs belongs to the *Caliciviridae* family and are classified into five genogroups (GI, GII, GIII, GIV, and GV) where GI, GII, and GIV infect humans (Patel *et al.*, 2009). The *Caliciviridae* family also includes human sapoviruses that are increasingly being recognized as an etiological agent in foodborne outbreaks (Kobayashi *et al.*, 2012).

Another virus causing a large number of GE infections and the most common cause of diarrheal disease associated death in children (Lanata *et al.*, 2013) is rotavirus, from the *Reoviridae* family. In industrialized countries rotavirus is responsible for 40 % of hospitalizations associated with child diarrhea. Rotaviruses mainly infect children and by the age of five most children have had the infection. Rotavirus consists of eight different species (A – H) where humans can be infected by rotavirus A (the most common), B, and C. Rotavirus are rarely associated with foodborne outbreaks and when they are it is usually a coinfection with another enteric virus such as NoV (Räsanen *et al.*, 2010). Since 2005, a vaccine against rotavirus A has been available and as of January 2017 the rotavirus vaccine have been introduced in 92 countries (O'Ryan, 2017). HAdV is another common source of childhood GE worldwide (Grimwood *et al.*, 1995; Lee *et al.*, 2012; Banerjee *et al.*, 2017). There are several species of HAdV within the family *adenoviridae*. HAdVs are the cause of several diseases, of which GE is mainly caused by HAdV-F serotype 40 and 41, the latter is the most common serotype detected in wastewater (Ogorzaly *et al.*, 2015).

2.1.1.2 Enteric viruses causing hepatitis

HAV and HEV spreads through the fecal oral route but when they reach the intestine the virus enters the bloodstream through the epithelium. The blood will transfer HAV and HEV to the liver where it infects the hepatocytes and Kupffer cells. HAV infection causes acute viral hepatitis in most adult cases, and these infections are often self-limiting (Stanaway et al., 2016), although in some cases it can cause acute liver failure. Shedding of viruses in feces starts 12-14 days before onset of symptoms both for HEV (WHO, 2017) and HAV (Koopmans and Duizer, 2004), and children below five years of age often get asymptomatic HAV infections but are still shedding the viruses in high number which could contribute to the possible spread of an outbreak. A HAV infection will leave the patient immune for the rest of their life and outbreaks are therefore rare in endemic regions where most persons are infected at an early age (Koopmans and Duizer, 2004). HAV was the second most common cause of viral foodborne outbreaks within the EU in 2014 (EFSA, 2015). However there is an effective vaccine for HAV which has reduced the incidence rate of HAV infection by 90 % in regions that have implemented routine childhood vaccination (Jacobsen and Wiersma, 2010). HAV belongs to the family *picornaviridae* and only infects humans in contrast to HEV which belongs to the *hepeviridae* family and is a zoonotic virus. HEV infects a number of mammals including humans, pigs, wild boars and deer. Four HEV genotypes can infect humans where genotypes 1 and 2 are exclusive for humans and genotypes 3 and 4 are zoonotic. Genotypes 1 and 2 are most common in areas with low sanitation, spreading through contaminated drinking water (Guerra *et al.*, 2017). In areas with access to clean drinking water, HEV infection is only seen in sporadic cases mostly due to zoonotic spread of HEV genotype 3 (WHO 2017). In contrast to HAV, a HEV infection more often leads to acute liver failure if infecting pregnant women during their second or third trimester, with case fatality rates as high as 20-25 % (WHO, 2017). In addition, fetal loss is another symptom of HEV infection in pregnant women. Yearly, HEV causes

20 million infections where 3.3 million of these are symptomatic, and in 2015 HEV was the cause of 44000 deaths worldwide (WHO, 2017).

2.1.1.3 Enterovirus

Enteroviruses belong as HAV to the *picornavirus* family and are a special group within the enteric viruses causing a wide range of symptoms: from common cold to meningitis. Historically, poliovirus, belonging to the species human enterovirus C, has been the most important enterovirus. However, since the introduction of the eradication campaign poliovirus is now only endemic in a few countries. Among the enteroviruses, coxsackievirus has been associated with GE from consumption of oysters (Iritani *et al.*, 2014). Even though enterovirus does not have a big health burden they have been proposed as indicator virus for fecal contamination since effective cell culture models exist to assess concentrations.

2.1.2 Bacteria

Traditionally more focus has been on the presence of bacteria in reused water. To better put the results from this PhD thesis into perspective the concentration of the bacteria *Salmonella* and *Campylobacter* were investigated throughout the wastewater treatment in **Hellmér II.** *Salmonella* and *Campylobacter* yearly cause many foodborne infections in Denmark (Anonymous, 2016), which therefore makes their presence in wastewater likely. Bacteria that transmit through water can in general be classified in two types, the first being enteric bacteria adapted for replication in the intestine and the second being aquatic bacteria. Enteric bacteria can survive for long periods in water (Moore *et al.*, 2003) but it is often under starving conditions and their replication is slow. Aquatic bacteria, like *Legionella* described elsewhere (Leclerc, Schwartzbrod and Dei-Cas, 2002), are outside the scope of this thesis.

2.1.2.1 Salmonella

Salmonella is a Gram-negative anaerobe and is a member of the *Enterobacteriaceae* family. It is further a zoonosis with more than 2500 identified serotypes and a wide host range. It can survive for months in water (Moore *et al.*, 2003) and is therefore a risk when reusing wastewater. Depending on the clinical presentations, *Salmonella* infections are often classified as typhoidal or non-typhoidal. A *Salmonella* infection often causes mild GE symptoms, but depending on host factors the infections can be more severe (WHO, 2018b). Some human-specific serotypes may result in enteric fever, often referred to as typhoid fever caused by *Salmonella* Typhi. *Salmonella* Typhi can spread through water and was the cause of 70 % of all waterborne outbreaks between 1920 and 1940 in the United States (Craun *et al.*, 2006). With increased sanitation, the number of infections with *Salmonella* Typhi has decreased in the developed world but it is still a problem in areas with poor sanitation and is the cause of 21 million infections and 200 000 deaths yearly (Crump, Luby and Mintz, 2004). Today, non- typhoidal *Salmonella* is one of the world's most common causes of GE (WHO, 2018b). The transmission is often through contaminated food. In 2016 non-

typhoidal *Salmonella* caused 24 % of the foodborne outbreaks and a total of 1074 human cases of salmonellosis in Denmark (Anonymous, 2016).

2.1.2.2 Campylobacter

Campylobacter is considered the most common cause of bacterial GE in the world (WHO, 2018a). They are Gram-negative and include 4 species that are significant human pathogens of which C*ampylobacter jejuni* is the most common. Like most enteric pathogens, C*ampylobacter* spreads through contaminated food or water most often resulting in a self-limiting mild GE. Like *Salmonella*, *Campylobacter* species are zoonotic and the main route of infection is through contaminated poultry (Anonymous, 2016) but transmission through water also occurs (Kuhn *et al.*, 2017). During 2016 *Campylobacter* was the most common cause of foodborne disease with 4677 confirmed cases distributed over 6.1 % of the foodborne outbreaks in Denmark (Anonymous, 2016).

2.1.2.3 Escherichia coli

E. coli is a commensal bacterium that is always present in the human gut. It is often non-pathogenic but some strains can cause mild GE and it is a common cause of urinary tract infections (Nielsen *et al.*, 2017). There are some strains associated with a higher virulence due to toxin production. *E. coli* producing vero toxin (VTEC) or shiga toxin (STEC) cause haemorrhagic colitis that might progress to hemolytic uremic syndrome (Leclerc, Schwartzbrod and Dei-Cas, 2002).

2.2 Indicator organisms

Knowledge of the total content of enteric pathogens in a sample would be desirable. However, to monitor all possible pathogens in water is costly and time consuming. Instead the microbiological quality of water is assessed by monitoring fecal indicator organisms. An ideal fecal indicator should be representative of the presence of enteric pathogens in the sample and by monitoring its concentration, the level of fecal contamination could be estimated. Traditionally, *E. coli* and coliform bacteria are used as fecal indicators to determine the microbiological quality of water. Coliform bacteria are defined as gram-negative, oxidase-negative, non-spore forming rods, that ferment lactose (Cabral, 2010). However, a number of studies show a lack of correlation between *E. coli* levels and the presence of human pathogenic viruses and bacteria (Ferguson *et al.*, 2012; Zhou *et al.*, 2015; McMinn, Ashbolt and Korajkic, 2017). In a review by McMinn et al. (2017) the removal during wastewater treatment of the traditional fecal indicator bacterial (FIB) were in average one log higher than the removal of enteric viral pathogens. Treated wastewater tested negative for FIB could therefore still contain pathogenic viruses and cause disease if used in food production. Instead some abundant viruses and bacteriophages have been proposed as indicators for the presence of enteric viruses (Bofill-Mas *et al.*, 2006).

The two double stranded DNA viruses, HAdV and JCPyV,have due to their stability in the water environment been suggested as indicators for human fecal contamination (Bofill-Mas *et al.*, 2006). Both can be analyzed with molecular methods, and the infectivity of HAdV can be assessed with culture methods (Bofill-Mas *et al.*, 2006). Besides the environmental stability of HAdV in water, Rames et al. (2016) list a number of traits which makes HAdV a good candidate to assess the human fecal contamination in water: stable and persistent in a wide distribution of water matrixes, stable detection in wastewater all year round, high resistance to ultraviolet light, higher abundance than other enteric viruses, propagation in cell cultures, and human specific. In **Hellmér II** the concentrations of HAdV and JCPyV were measured with quantitative polymerase chain reaction (qPCR) throughout the wastewater treatment at both the urban and hospital WWTP. In the urban WWPTs, HAdV followed the same seasonality pattern as NoV GI and GII with a decrease in viral concentration during the summer. The same study showed no significant difference in the concentrations of HAdV, JCPyV, and NoV GI and GII at the urban WWTP.

Bacteriophages are viruses that infect and replicates in bacteria. They exist in high numbers in wastewater and their ability to function as fecal indicators to monitor viral removal in WWTP are reviewed by Amarasiri et al. (2017) and McMinn et al. (2017). Several authorities recommend the use of MS2, an F-pilli specific RNA coliphage infecting *E. coli*, to validate and monitor water reclamation processes due to its morphological and structural resemblance to enteric viruses (Sarkis, 2013; California State Water Resources Control Board, 2015). Amarasiri et al. (2017) concludes that MS2 is the best suited indicator bacteriophage for enteric viruses based on its correlation to NoVs. However, lots of water cleaning technologies depend on charged filters and their effect on viruses relay on the surface charge of the virus. It is therefore almost impossible to find an indicator that can represent all viruses. Chaudhry et al. (2015) propose therefore to test the effect of wastewater treatment methods on multiple enteric viruses and not only the indicator organisms.

2.3 Wastewater treatment plants and microbiology

Within the urban wastewater directive (EEC Council, 1991) there are no regulations regarding the reduction of microorganisms during the treatment of wastewater. Thus most conventional treatment processes are designed for elimination of solids, organic substances and nutrients, referred to as BOD (biochemical oxygen demand). Some more advanced WWTPs have a final disinfection step of the wastewater with the aim to eliminate microbes, including viruses. Wastewater treatment methods are divided into different levels depending on the removal of BOD from the water. Primary treatment of wastewater refers to the physical or chemical processes that reduce BOD5 (the 5 day BOD) of incoming wastewater by at least 20 % and suspended solids by at least 50 % before discharge (EEC Council 1991). The secondary treatment is the processes that reduces the BOD5 to 25 mg/L O_2, chemical oxygen demand (COD) to 125 mg/L O_2 and the total suspended solids to 35 mg/L (EEC Council 1991). Tertiary treatment refers to additional treatment steps after secondary treatment to further improve quality of the effluent prior to its release to the catchment area.

WWTPs can have one or several tertiary treatments, such as filtration, relaying in lagoons or ponds, removal of biological nutrient, nitrogen or phosphorus, and disinfection. Tertiary treatments are costly, especially in large urban WWTPs that treat large volumes of wastewater daily (Heinonen-Tanski *et al.*, 2003) and are therefore mostly used if the cleaned water is released to a sensitive catchment area.

In the first chapter of the book Wastewater Treatment: advanced processes and technologies (Rao *et al.*, 1997) the wastewater treatment methods are divided into four main classes. The first consists of physical methods that clean the water without changing the chemical structure of the substances within the wastewater, often used during primary treatment. This can include sedimentation, flotation, adsorption or physical barriers as filters, bar racks, and membranes. The second includes the chemical methods where added chemicals are used to clean the wastewater. The most common ones being chemical precipitation, coagulation, and oxidation. The third class combines the physical and chemical methods. For example use of a physical filter to remove solid particles formed during chemical precipitation. The fourth class includes the biological methods where microorganisms are used to degrade contaminants in the wastewater and thereby decrease BOD.

Even in WWTPs where no specific disinfection process with the aim to reduce microorganisms is used, bacteria and viruses are removed from the water throughout the process mainly by removal of solid particles which they adhere to. For example during the most common wastewater treatment strategy, activated sludge, which is a biological treatment process where added oxygen will produce biological flocs that then are left to settle, bacteriophages are reduced by an average 1.9 log units (Amarasiri *et al.*, 2017).

If the water is meant to be reused, sand or membrane filters are often used. Sand filters can remove 10-98 % of viruses by separating solids and activated sludge flocs from the liquid (Zhang *et al.*, 2016). Membrane filtration works by size exclusion and if the filter size is below 10^{-2} µM the filter is thought to be small enough to remove viruses. Depending on pore size these are referred to as microfiltration membranes with a pore size ≤ 0.1 µm or ultrafiltration membrane with a pore size between 0.01-0.1 µm. The reduction of MS2 coliphage was found to be 1.38 and 3.69 log plaque-forming units (PFU) on average for microfiltration and ultrafiltration membranes, respectively (Amarasiri *et al.*, 2017).

The use of membrane bioreactor (MBR) is increasing in WWTPs as an advance treatment to improve the water quality. The MBR is a version of the conventional activated sludge treatment but instead of using settling to remove the solids the water passes through a micro or ultrafilter membrane. In addition the MBR membrane is coated with a cake layer that harbors enzymes that break down the microorganisms. There have been reports of up to 5-log removal of viruses by using MBR systems (Simmons, Kuo and Xagoraraki, 2011). Even though high removal rates have been reported, viruses have still been detected in MBR effluents (Zhou et al. 2015, **Hellmér II**). The effectiveness of an MBR system, and other membrane-based treatments,

is dependent on the membrane pore size, membrane integrity, solution environment, membrane charge, and water pH (Antony, Blackbeard and Leslie, 2012). The different surface properties of the viral capsid influence the viral attachment to the added biomass and thereby the level of viral reduction during MBR (Chaudhry *et al.*, 2015). Due to the importance of operational conditions, the monitoring protocol developed by the WaterVal program assigned a virus log removal of 1.5 for MBR after investigating full scale WWTPs (WaterSecure, 2017).

The removal of microorganisms via disinfection is done just prior to discharge of the effluent. The most commonly used methods are either chemical disinfection (chlorination or ozonation) or disinfection by UV irradiation. The chemical disinfection will damage the viral protein capsid whereas UV irradiation will degrade the nucleic acids. In the EU directive for drinking water (NSF, 2009) a UV dose of 40 mJ /cm^2 is required to ensure a 4-log inactivation of viruses. However, HAdV has been shown to need a UV dose of 93 mJ / cm^2 to be inactivated (Prevost *et al.*, 2016).

2.4 Current regulations for reused wastewater

The European Union is currently working towards a legislation on water reuse at EU level (Alcalde-Sanz and Gawlik, 2017). The proposed EU minimal quality requirements (EU MQR) for water reuse in agricultural irrigation and aquifer recharge is based on reclaimed water quality class A – D. All water classes should have been treated with secondary treatment, filtration, and disinfection, and measured *E. coli* should be below 10 cfu / 100 mL or below the detection limit for class A, below 100 cfu / 100 mL for class B, below 1000 cfu / 100 mL for class C, and below 10000 cfu / mL for class D (Alcalde-Sanz and Gawlik, 2017). The proposal also lists the minimum reclaimed water quality class required for a number of different crop categories listed in Table 3.

Since the publication of the EU MQR, scientific advice from the European Food Safety Authority (EFSA) and the Scientific Committee on Health, Environmental and Emerging Risks (SCHEER) were requested. Both EFSA (Allende *et al.*, 2017) and SCHEER (Rizzo *et al.*, 2018) answered the request and both reports concluded that a more in-depth revision of the subject is needed to provide sufficient protection both to environmental and human health. SCHEER proposed that a case-by-case approach should be used to assess the effectiveness of each WWTP. Both SCHEER and EFSA also mention the need to use indicator organisms other than *E. coli,* e.g. a viral or protozoa indicator if the risk of these pathogens should be addressed.

Table 3. Minimum reclaimed water quality class and irrigation method for three food crop categories. Table adapted from (Alcalde-Sanz and Gawlik, 2017).

Crop category	Minimum reclaimed water quality class	Irrigation method
All food crops, including root crops consumed raw and food crops where the edible portion is in direct contact with reclaimed water	Class A	All irrigation methods
Food crops consumed raw where the edible portion is produced above ground and is not in direct contact with reclaimed water	Class B	All irrigation methods
	Class C	Drip irrigation
Processed food crops	Class B	All irrigation methods
	Class C	Drip irrigation

There are only few regulations or guidelines regarding levels of viruses in reused water. The water reuse regulation in California, United States, requires a disinfection combination in tertiary treated water that inactivates or remove 4-log plaque forming units of MS2 or poliovirus to use the water for crop irrigation (California State Water Resources Control Board, 2015). To use treated wastewater for groundwater replenishment the regulation requires a combined water treatment that reduce viruses with a total of 12 log (California State Water Resources Control Board, 2015). The hospital WWTP investigated in **Hellmér II** has theoretical reduction of enteric viruses by 5-12 log (NSF, 2009; US-EPA, 2012) and would therefore be allowed to use its effluent water for irrigation according to the water reuse regulation in California.

3. Analysis of the microbiome in complex water samples

Today there is no unified method to analyze a water sample for all pathogenic microorganisms of interest (Straub and Chandler, 2003). Instead several techniques have been developed, most of them with the aim to identify one or a group of specific bacteria (Wesolowska-Andersen *et al.*, 2014a) or viruses (Calgua *et al.*, 2008; Albinana-Gimenez *et al.*, 2009; John *et al.*, 2011a; Pei *et al.*, 2012; Bibby and Peccia, 2013b). Even though a large number of methods and combination exist, most analyses of viruses and bacteria from water follows these main steps: sampling of water, concentration of the target microbe or microbes, extraction of the nucleic acid (if a molecular detection method is used), detection of the microbe or microbes and finally analysis of data. The general workflow and some of the most common methods are included in the flowchart in Figure 4.

The next generation sequencing (NGS) tools existing today makes it theoretically possible to investigate all nucleic acids (NAs) within a sample. In short all the NAs within a sample are fragmented into millions of small NA fragments. These millions of NA fragments are then sequenced in parallel on the NGS platform. Each sequenced fragment produce a read that can be further process with bioinformatics (Behjati and Tarpey, 2013). For a perfect analysis, this demands an unbiased preparation of the given sample. However, until the existence of a one-step sequencing system that without errors can sequence and align the reads to the right organisms, the biases of each used method needs to be taken into account. It is therefore important to know the advantages and disadvantages with each method before the start of a NGS project or when comparing the results with previous studies. To increase the knowledge of method dependent biases **Hellmér I** investigated in total 16 combinations of viral concentration and extraction procedures and their impact on the detected viral metagenome in raw sewage. This chapter briefly presents each step throughout the workflow and discusses its effect on the detected metagenome. Even though this thesis mainly focuses on viruses, some methods to analyze bacteria will be presented to illustrate the differences and challenges with virus detection from water samples.
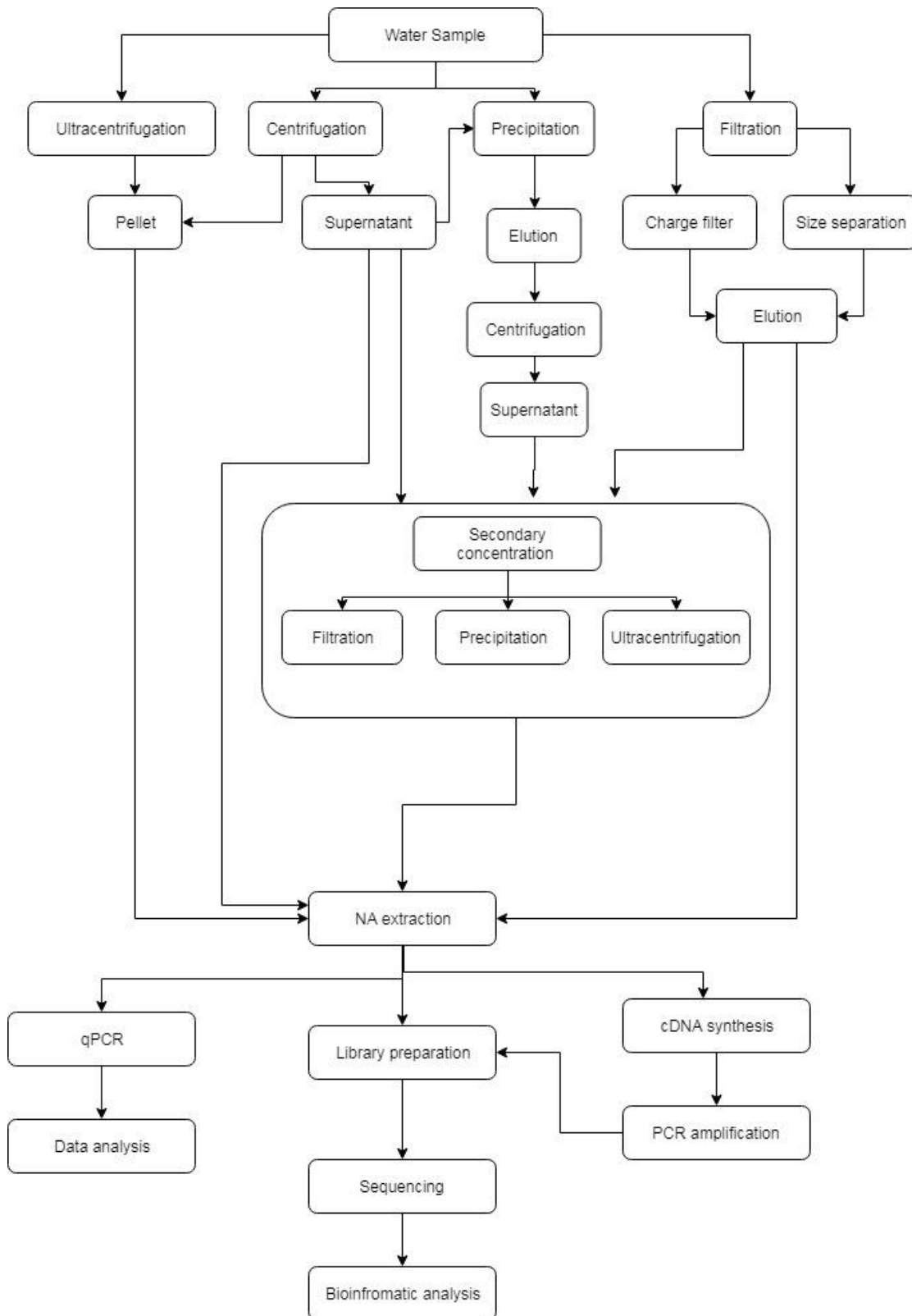
Figure 4: Flowchart of some of the most commonly applied molecular techniques to analyze microorganisms from water samples.

3.1 Sampling

Sampling is the first and one of the most important steps in any project. The sampling will directly affect the quality of the obtained data and is therefore crucial to the success of the project. It is important that the sample represent the whole microbial community in that water source (Council, 2007; Wooley, Godzik and Friedberg, 2010; Thomas, Gilbert and Meyer, 2012; Felczykowska *et al.*, 2015). To ensure that the sample is representative there are three main factors to consider when sampling water: sampling point, timing, and volume. The sampling point should be chosen in a way that answers the hypothesis of the specific project. In **Hellmér I** the aim was to compare the effect of concentration and extraction methods on highly contaminated samples and therefore a point where we could sample large volumes of raw sewage was chosen. In **Hellmér II** the aim was to analyze the effect of different sewage treatment procedures and the sampling points thus needed to be spread throughout the treatment process. The timing of sample collection can affect the detected microbial community composition due to different concentrations of pathogens dependent on seasons (Rusinol *et al.*, 2015).

A sample can be collected via grab or composite sampling. Grab sampling is when the entire sample is collected at once and is often the most convenient sampling method since it does not require any specific instruments, a bucket and a string is often enough. A composite sample consists, on the other hand, of a number of small samples taken over a period of time. The composite sample could either be time dependent e.g. a fixed volume every five minutes, or flow dependent e.g. a fixed volume for every passing cubic meter of water (Hellmér *et al.*, 2014). A composite sample will give a more representative view of the sampling period compared to a grab sample that only reflects the time point when the sample is taken. Frequently collected samples e.g. monthly (**Hellmér II**), weekly, or daily would account for changes in the community over time even when using grab samples.

To analyze the microbial community, the sampling volume has to be large enough to include even the rare species present at low concentrations in the water. Hence the type of water has to be taken into account when decide upon a sampling volume, for example a higher volume needs to be sampled from ground or drinking water that has an expected low level of contamination (Shi *et al.*, 2013) compared to sewage water with high level of contamination. In practice the sampling volume is often limited due to the concentration method in hand, the practical lab environment, or the available time. Even though a larger sample volume in theory could enable detection of rare and uncommon species, it also has a larger proportion of inhibitors (Schrader *et al.*, 2012) that could prevent detection. In **Hellmér I** a larger sample volume was associated with PCR inhibition.
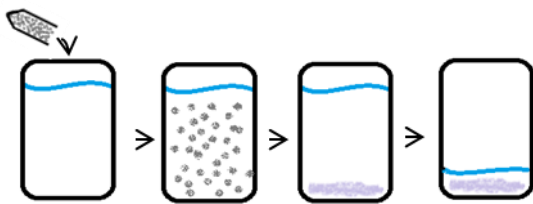
3.2 Concentration

The next step after sampling is to concentrate the viruses and bacteria. Most protocols are developed for the concentration of either virus or bacteria but work is being conducted to evaluate the possibility for simultaneous concentration of these two classes of microbes (Abd-Elmaksoud *et al.*, 2014; Gonzales-Gustavson *et al.*, 2017). As an example, from raw wastewater sample with high levels of contamination, pelleting down bacteria by centrifugation could be sufficient (Nordahl Petersen *et al.*, 2015). The viral particles can only be pelleted by high speed (90,000×g) ultracentrifugation (**Hellmér I**) and would stay in the supernatant after regular centrifugation. Alternatively, bacteria can be concentrated from water by filtration through a 20-45 µM membrane (Kisand *et al.*, 2012; Chao *et al.*, 2015; Krustok *et al.*, 2015; Newton *et al.*, 2015; Satinsky *et al.*, 2015). The bacteria will thus stick to the membrane and subsequently be eluted. The small size of viruses and the fact that viral particles often are present in low abundance within a sample requires the use of viral concentration methods prior to nucleic acid extraction (Daly *et al.*, 2011). Even in samples with a high concentration of viruses their genomes are small in comparison to bacteria and therefore only represent a small fraction of the total nucleic acids within the sample, hence additional viral purification is often needed to detect them (Ng *et al.*, 2012). A wide array of methods and steps to concentrate viruses from water samples has been developed. Figure 4 illustrates some of the most common. Most concentration methods rely on the same principles that are used to eliminate viruses throughout the wastewater treatment, described in chapter 2, the two main ones being filtration or precipitation.

Filtration can be further divided into size-dependent filtration, as hollow fiber filtration (ultrafiltration), or charge-dependent filtration that relies on the viral isoelectric point, as glass wool filtration (Albinana-Gimenez *et al.*, 2009) or monolithic adsorption filtration (MAF) (Pei *et al.*, 2012). In charge-dependent filtration, the charged filter will bind viruses of the opposite charge, e.g. a positively charged filter will bind negatively charged viruses. The charge of the virus dependents on its isoelectric point depending on its surface structure which can be changed by changing the pH of the water. Viruses in the sample attach to the filter and are then eluted from the filter using an elution buffer. The elution buffer can either change the charge of the virus which makes it detach from the filter, or by changing the direction of the flow push the viruses out of the filter. Filtering allows for processing of large volumes of clean water, e.g. ground and drinking water (Kunze *et al.*, 2015). A filtering system can, if combined with composite sampling, perform real time concentration of water. Due to the small pore sizes, more turbid water samples like raw wastewater and some environmental waters could clog the filters thereby limiting the analyzed volume.

Concentration of viruses using precipitation relies on the addition of large charged particles or proteins to the water samples, such as polyethylene glycol (Bibby and Peccia, 2013a), $FeCl_3$ (John *et al.*, 2011b), or skimmed milk proteins (Calgua *et al.*, 2008). Viruses binds to the particles and after a period of mixing the

particles are left to precipitate, Figure 5. The supernatant is removed and the particles are pelleted through centrifugation. The pellet is suspended and saved for nucleic acid extraction (Calgua *et al.*, 2008). Secondary concentration of viruses from the precipitate could be done by adding a buffer to make the viruses release from the particles, remove particles by centrifugation and then concentrate viruses from the supernatant by high speed ultracentrifugation (Hellmér et al. 2014, **Hellmér I**). Precipitation-based concentration methods are cheap and do not require any expensive equipment. For practical reasons, these methods have a maximum volume of approximately 10 L and are thus not the best choice for metagenomics analysis of water samples with a low level of contamination, such as ground and drinking water.

Figure 5: Schematic picture of the precipitation principle.
Charge or organic material is added to the water container. The added material is mixed and then left to precipitate. Excess water is removed.

Previously, the influence of concentration method on viral recovery has been evaluated on sea water (Hurwitz *et al.*, 2013), spiked tap water (Albinana-Gimenez *et al.*, 2009; Kunze *et al.*, 2015) and raw wastewater (Calgua *et al.*, 2013). All studies warn of method-associated biases. To further investigate the effect of concentration method on the detected viral metagenome, four viral concentration methods were compared in **Hellmér I**: two filtration-based methods, glass wool filtration and MAF, and two precipitation methods, PEG and skimmed milk flocculation. Sample volume and applied secondary viral concentration for each method were performed based on the published procedures. Glass wool filtrate was secondary concentrated with PEG, MAF concentrate was secondary concentrated by ultrafiltration, and skimmed milk flocculation was secondary concentrated with ultracentrifugation. PEG precipitation did not have a secondary concentration step. Instead, the sample was pre-processed by passing it through a 0.45 μm filter. The choice of concentration method affected the composition of the detected viral community and the methods had different advantages and disadvantages summarized in Table 4.

Table 4: List of advantages and disadvantages for the different concentration methods tested in Hellmér I

Method	Advantages	Disadvantage
PEG	High proportion of viral reads	Low sampling volume Labor intensive
Skimmed Milk	Good detection of RNA viruses	Inhibition in qPCR analysis Maximum volume of 10L Labor intensive
MAF	Good detection of DNA viruses Possibility to filter large volumes of clean water	Clogging of filter in dirty samples
Glass wool	High viral richness Possibility to filter large volumes of clean water	Clogging of filter in dirty samples

MAF was chosen for viral concentration in **Hellmér II** based on the results of **Hellmér I** where MAF showed low inhibition during qPCR detection and fair detection of ssRNA viruses in combination with the ability to filter large volumes of water in a short time period enabling easier sampling. MAF was also the only method able to concentrate NoV from a natural contaminated sample in **Hellmér III.**

3.3 Extraction

After concentrating the microbes from the water sample their NAs have to be extracted to allow further processing. There are a number of commercial kits developed for specific extraction of viruses or bacteria. All processes start with lysis of the bacterial cell or the viral capsid. The next step is to purify the NAs and wash away proteins, cell debris, and polysaccharides. The NAs is purified by either binding to a spin column or magnetic beads (Berensmeier, 2006), or by phase separation followed by precipitation. NA is then eluted with either RNase free water or kit-specific elution buffer. The structural and genomic differences between bacteria and viruses have led to a development of specific kits for either bacteria or viral NA extraction. In addition, the diverse viral genome structure consisting of either DNA or RNA genomes puts extra pressure on the extraction kit. Most kits can extract both DNA and RNA. But the ratio between the extracted DNA and RNA differ as shown in **Hellmér I.** There are some kits that favor the extraction of RNA. Steps to increase the extracted RNA includes treating the extracted NAs with DNase which degrades loose DNA, adding extra denaturation agents to the lysis buffer to destroy RNases, or adding poly-A carrier RNA to facilitate extraction of RNA and reduce degradation by RNase. The chosen nucleic acid extraction method will have a great impact on the detected microbial community. There are a number of studies investigating

biases caused by nucleic acid extraction kits in both bacteria (Kennedy *et al.*, 2014; Wesolowska-Andersen *et al.*, 2014b) and viruses (Petrich *et al.*, 2006; Iker *et al.*, 2013, **Hellmér I**). In addition, contaminants have been found in some extraction kits (Salter *et al.*, 2014) and laboratory reagents (Newsome *et al.*, 2004). This could give rise to false positive results if a negative extraction control is not sequenced together with the tested samples (Naccache *et al.*, 2013; Rosseel *et al.*, 2014).

These biases are important to take into account when comparing observed metagenomes and to decide which methods to be used in a study. If the aim is to compare the metagenome with previous studies, then the use of the same extraction kit is of essence. If it, on the other hand, is to look at a specific community within the sample, then an extraction kit that favors the extraction of e.g. RNA for enteric viruses could be the right one.

3.4 Detection

The ideal detection method should be reliable, accurate, rapid, simple, sensitive, selective and cost effective (Zhao *et al.*, 2014). Detection of microorganisms in water samples has traditionally been through culturing bacteria in liquid or solid growth media, and viruses have been detected through cell culture and titration. With the development of nucleic acid-based methods, previously non-cultivable viruses such as NoV, which are not possible to detect in cell culture, could be detected. Sequencing further improved the possibility to type the detected virus. The development of NGS made it possible to detect all nucleic acids, the metagenome, and thereby microorganisms within a sample. Below I will go through the three main detection methods: culturing, PCR, and sequencing and discuss their advantages and disadvantages.

3.4.1 Culture

Routine analysis and detection of bacteria in water samples is conducted using traditional microbiological culturing methods, the most common of which is multiple tube fermentation in which a series of tubes containing broth are inoculated with a dilution series of the water sample (Rompré *et al.*, 2002; Deshmukh *et al.*, 2016). A positive sample shows production of acid and gas within 48 hours at 35°C. The results are interpreted in terms of most probable number (MPN) which provides an estimate of the mean coliforms present in the sample (Rompré *et al.*, 2002; Deshmukh *et al.*, 2016). Since viruses do not grow outside a host cell, cell culture systems are a requirement to grow viruses in laboratory settings. With the discovery of human cell cultures in the early 1900s, viruses could for the first time be grown outside embryonated eggs and laboratory animals (Leland and Ginocchio, 2007). But it was not until the early 1970s, and the availability of commercial cell lines, that viral diagnostic became more wide-spread (Hsiung, 1984). Since then, cell culture systems have been the gold standard for viral diagnosis. Cell cultures are also used for viral discovery (Pan *et al.*, 2017) and infectivity analysis (Li *et al.*, 2011). The ability to grow and propagate viruses is of great importance to further understand the viral mechanisms and to produce material to use for spiking during method development and as process controls during investigations. Cell cultures are also an

important tool in vaccine development. To maintain a cell culture system is time consuming and requires technical expertise, in addition, the diagnosis procedure takes at least 24 hours and all viruses cannot be detected in cell culture systems (Leland and Ginocchio, 2007). With the introduction of molecular methods the use of cell cultures for viral diagnostics has decreased.

3.4.2 PCR and quantitative PCR (qPCR)

The invention of PCR amplification of DNA fragments (Saiki *et al.*, 1985) opened new possibilities to detect microorganisms. With specific primers, unique stretches can be amplified and a successful amplification is identified by agarose gel electrophoresis of the PCR product. To enable amplification and thereby detection of RNA, e.g. RNA viruses, the genome fragment needs to be transcribed to complementary DNA (cDNA) by the use of reverse transcription (RT) prior to PCR amplification (Gibson, Heid and Williams, 1996). PCR detection is a fast and specific method but it is unable to quantify the detected microbe. By adding fluorophores that either bind all dsDNA molecules, such as SYBR green, or to a specific sequence, such as TaqMan probes, the amount of PCR-generated DNA can be measured. With this, the initial concentration of genomic material can be estimated and qPCR is now one of the most used methods to detect and analyze enteric viruses. By using TaqMan probes with different fluorophores it is possible to detect multiple organisms in one qPCR well (Hellmér *et al.*, 2014; Irshad *et al.*, 2016). By combining RT and qPCR it is possible to detect and quantify the genomes extracted from the present RNA viruses in a sample (Gibson, Heid and Williams, 1996).

PCR reactions are sensitive to inhibitory substances that decrease the sensitivity of the PCR assay and underestimate the initial concentration or give false negative results. Inhibitory substances could originate from the sample or the sample processing and could be of both organic or inorganic origin (Schrader *et al.*, 2012). The concentration and extraction methods used can affect the level of inhibition within a sample (Borgmästars et al. 2017, **Hellmér I**) as can the initial sample volume (Schrader et al. 2012, **Hellmér I**). Pre-PCR processing of the sample has been shown to increase the viral recovery in highly inhibited samples (Borgmästars *et al.*, 2017).

Today it is not possible to differentiate between infectious and non-infectious virus particles or bacterial cells (Cangelosi and Meschke, 2014). This problem could lead to overestimating the microbial threat within a sample and give false positive results. There is ongoing work to make a molecular method that only detects nucleic acids from whole and unbroken viral capsids. One method uses pre-PCR treatment with intercalating dyes, e.g. propidium monoazide, that binds to free nucleic acid and prevents it from amplifying (Prevost *et al.*, 2016; Quijada *et al.*, 2016). These methods are based on two assumptions: the first that a virus with a broken capsid is noninfectious and the second that the intercalating dyes will bind to and block amplification of free nucleic acids.

3.4.3 Sequencing

The development of sequencing technologies has made it possible to read the genome of individual organisms. The first generation of sequencing technologies, called Sanger sequencing, is based on chain-terminating inhibition in which every new PCR cycle adds a single base to the read sequence (Sanger, Nicklen and Coulson, 1977). Sequencing has made it possible to in-depth characterize a detected organism and to link patients from foodborne outbreaks with the source (Müller *et al.*, 2015). However, sequencing large genomes is time-consuming when using Sanger-based sequencing methods. To increase the sequencing data output the second generation of sequencing technologies was developed. These new technologies are referred to as NGS. NGS allows massive parallel sequencing of PCR amplicons or environmental DNA. The NGS technology has three major improvements compared to the Sanger sequencing (van Dijk *et al.*, 2014). First, they rely on a cell free library preparation of the DNA fragments. Second, all sequencing actions are run in parallel. Third, the output from the NGS platforms are detected directly as reads. The first NGS platform was the 454 pyrosequencing (Margulies *et al.*, 2005) from Roche (Life Sciences), followed by the NGS platform developed by Illumina (Solexa) (Turcatti *et al.*, 2008). The Illumina platform generates a larger number of reads compared to 454 but the produced reads are shorter (van Dijk *et al.*, 2014). The new sequencing technologies made it possible to sequence all nucleic acids, the metagenome, within a sample. In contrast to other molecular detection methods no prior knowledge of the target organisms is needed since in theory all nucleic acid within the sample will be sequenced. But just as for PCR-based detection mechanisms, inhibition can be an issue since most library kits include a PCR step. Today NGS is mostly used for research purposes (Bibby and Peccia, 2013a; Fernandez-Cassi *et al.*, 2017). However, NGS approaches have been proposed as routine analysis of foodborne bacterial pathogens (Deng, den Bakker and Hendriksen, 2016) and as a tool for biopreparedness in outbreak scenarios (Karlsson *et al.*, 2013)

In **Hellmér I, II** and **III** the Illumina MiSeq sequencing platform was used to sequence the metagenomes from water samples. This platform was chosen due to a combination of its large data output and relative low cost per base pair. Another reason for choosing the Illumina MiSeq was that it is one of the most commonly used platforms for environmental metagenomics and therefore has a lot of already made bioinformatics pipelines.

3.5 Bioinformatics

To transform the sequencing data to a list of microbes a number of bioinformatic steps are required. The data output from a sequencing platform is in the form of raw reads 70-1000 base pair (bp) long, depending on sequencing platform used (van Dijk *et al.*, 2014). The first step is to ensure the quality of the data by so-called preprocessing. First each read is assigned to the respective sample based on the read ID-tag or barcode that was added during the library preparation. The next step is to analyze the quality of the data to get an overview of the dataset, this can be done with e.g. FastQC (Andrews S, 2018) that will summarize the per base sequence quality, GC content, sequence length distribution, overrepresented sequences and kmer

content. Overrepresented sequences are often the sequence adaptors and are removed by trimming so as not to course false alignment. There are a number of programs that can be used to trim sequences, in this thesis the program Cutadapt was used (Martin, 2011). Paired end reads can be merged to extend the read length prior to de novo assembly and thereby decrease the error rate (Magoč and Salzberg, 2011).

When looking for a specific virus within the dataset the reads can be mapped to a reference genome, this can be of use in an outbreak setting were a specific agent is of interest (**Hellmér III**). Data from metagenomics sequencing consists of reads from a number of unknown organisms so it is not possible to align the reads to one reference genome. Instead the individual reads can be directly mapped to a reference database or the reads can be de novo assembled into contigs. Contigs are produced by finding overlaps between reads and connect them to longer fragments. The produced contigs can then be mapped against the chosen database. Direct mapping of the trimmed reads is the fastest method to detect the present viruses especially when using large datasets (Petersen *et al.*, 2017). The National Center for Biotechnology Information (NCBI) hosts the most applied tool for direct mapping of sequencing reads, Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990). The BLAST algorithm is time consuming and as the output from sequencing platforms has increased, the need for faster algorithms has resulted in a number of new methods. In **Hellmér I** and **II** the open metagenomics pipeline MGmapper (Petersen *et al.*, 2017) which uses direct mapping of reads to chosen databases was used to map the sequencing reads. A drawback of direct mapping of the short reads is an increased possibility of false assignment to a genome if it contains conserved regions, which is common in large DNA viruses from families as *Herpesviridae* and *Poxviridae* (Rosseel *et al.*, 2014). The alternative to direct mapping is to first assemble the reads contigs. This step is time consuming and requires high computer power. In **Hellmér III** the assembler SPAdes (SPAdes v. 3.11.1) was used to assemble the reads into contigs with the aim to see if the de novo assembly increased the likelihood of detecting human enteric viruses.

It is important to remember that the read count abundance is biased since more reads are sequenced from larger genomes compared to small. This can be adjusted for during the bioinformatic analysis. In MGmapper this is adjusted for by a normalization of reference sequence size and read count abundance (Petersen *et al.*, 2017).

4. Discussion

The elimination of enteric viruses from wastewater prior to reuse in food production is of great public health interest. Enteric viruses have in several studies shown higher resistance to wastewater treatment procedures compared to FIB (Rodriguez-Manzano *et al.*, 2012; Rusinol *et al.*, 2015). To rely on the detection of FIB when evaluating the microbial water quality is therefore insufficient. In this PhD thesis I have looked at the possibility to use viral metagenomics to assess the microbial water quality of reused wastewater. This has been done by analyzing the removal of viruses along the wastewater treatment process and their final presence in the wastewater effluents.

The first objective of this PhD thesis was to develop a methodological pipeline covering all steps from sampling of water to detection of viral communities by use of metagenomics sequencing. Sixteen different combinations of concentration and extraction methods were therefore evaluated in **Hellmér I**. The choice of concentration and extraction method was shown to affect the detected viral richness, specificity, pathogen detection, and community composition. However, none of the tested method combinations appeared to be superior in all tested parameters. Each method and method combination showed strengths and weaknesses as summarized in table 4. To choose which methodological pipeline to use for **Hellmér II** two main criteria was evaluated: the ability of processing of large water volumes and the efficiency to detect RNA viruses. The initial main focus was to investigate the enteric viruses which are the cause of most water-related outbreaks (Koopmans and Duizer, 2004). Samples extracted with the RNA-specific extraction kit Nucleospin RNA XS were shown to give a higher proportion of reads mapped to the enteric RNA viruses compared to the other tested extraction kits (**Hellmér I**). This kit has been successfully used for detection of RNA viruses by metagenomics (Leeuwen *et al.*, 2010). The investigation of wastewater effluent, which was hypothesized to contain low levels of contamination, needed a viral concentration method suitable for large volumes of water. The two tested precipitation methods PEG and skimmed milk flocculation were therefore excluded. At this time we received drinking water suspected to be the cause of a NoV outbreak (**Hellmér III**). The received volume of water was sufficiently large to enable a small comparison between the filtering methods. MAF was the only of the tested methods that could concentrate NoV from the naturally contaminated water. Based on the results from **Hellmér I** and **III,** MAF concentration followed by Nucleospin RNA XS extraction was used in **Hellmér II**.

The second objective was to analyze the viral population before and after different treatment steps in one conventional urban and one highly advanced hospital WWTP. This study (**Hellmér II**) highlighted the challenges when using metagenomics sequencing to evaluate the presence of specific viral species within a water sample. Even though HAdV and NoV GI and GII were found by qPCR in concentrations up to 10^5 gc / L they were not always detected by using the open metagenomics pipeline MGmapper. Detecting NoV by metagenomics sequencing has proven hard (Bibby and Peccia, 2013; Mee *et al.*, 2016; Yang *et al.*, 2017;

Hellmér II), which could be a result of NoV's capsid or genome structure (Ding *et al.*, 2014). The diversity within viral species also affects the efficacy of pre-metagenomics processing (Li *et al.*, 2015). Previous studies have demonstrated that increased sequencing depths could increase the viral detection (Wylie et al. 2012). In the studies presented in this thesis the sequencing depth has ranged from one to six million raw reads per sample. The sequencing depth can be increased by sequencing fewer samples per Illumina MiSeq run or by using another sequencing platform, e.g. the Illumina NextSeq 500 that has a higher data output. Increased sequencing depth may increase the chances of detecting specific pathogens but it will also increase the analysis cost of each sample. Targeted pre-processing by the use of poly(A)-capture (Fonager *et al.*, 2017) or amplification with NoV genotyping primers (Kundu *et al.*, 2013; Cotten *et al.*, 2014; Park *et al.*, 2015) have shown to enhance the number of NoV reads. This approach requires previous knowledge of the suspected contaminant. By favoring the sequencing of a particular virus one of the biggest advantages of using metagenomics sequencing, the fact that no prior knowledge of the viral content is needed, is lost. However, in an outbreak setting as described in **Hellmér III** this approach could increase the ability to detect the causative agent by NGS. Another approach to enhance the detection of NoV is to use targeted mapping tools (Yang *et al.*, 2017). As with pre-processing, this requires prior knowledge of the contaminant, and larger bioinformatic knowledge is needed to test and troubleshoot extensive analysis of the data. However, targeted mapping will not affect the output sequencing data, thus additional information can still be obtained.

In addition, results from **Hellmér II** showed that analysis of the viral community structure could differentiate clean water from contaminated water without looking at specific pathogens. However, the results from **Hellmér I** demonstrated that the detected community structure were dependent on the concentration method used. To use the community structure to assess the microbial quality of water could therefore be influenced by the laboratory methods used. This should not replace existing control measures but could be an aid in the evaluation of a water source.

In the proposed European Union minimal quality requirements for water reuse in agricultural irrigation and aquifer recharge (Alcalde-Sanz and Gawlik, 2017) a *E. coli* concentration below 10 cfu / 100 mL in wastewater effluents is accepted for irrigation water for all purposes. By this regulation the MBR effluent from the hospital WWTP would be allowed to be use for irrigation of ready to eat crops. These proposed water safety classes were questioned for being too weak and only rely on FIB for the assessment of microbial water quality (Allende *et al.*, 2017; Rizzo *et al.*, 2018). Instead an individual approach was suggested were the effect of a multiple barrier approach is taken into account. At the hospital WWTP the MBR effluent is subjected to ozone, granular activated carbon, and UV irradiation before discharge of the effluent. As these treatments account for a combined of 5-12 log reduction of viruses (NSF, 2009; US-EPA, 2012). The hospital effluent would thereby be regarded as microbial class A.

The third objective was to investigate the usefulness of the broadly proposed viral indicators HAdV and JCPyV. In **Hellmér II** HAdV showed a seasonal pattern similar to NoV GI and GII such that the HAdV concentration decreased during the summer months by about 2-log. One of the reasons why HAdV has been proposed as a fecal indicator is its stable detection in wastewater all year long (Rusinol *et al.*, 2015; Rames *et al.*, 2016). The results from **Hellmér II** are the first study that shows seasonality of HAdV concentrations in wastewater. Further studies are needed to evaluate if this observed seasonality is specific for Denmark or if the investigated season August 2016 to August 2017 was atypical.

The last objective was to validate the developed pipeline during the investigation of contaminated water incriminated in a disease outbreak (**Hellmér III**). NoV could be detected in the tested water both with RT-qPCR and Sanger sequencing but not with NGS. However, the applied metagenomics methods allowed two positive viral process controls to be detected with NGS. One of them was detected by the open metagenomics pipeline MGmapper, while the other was only detected when using a viral targeting approach. An enhanced approach to detect NoV was also tested by developing three specific NoV databases and linking them to the MGmapper tool. The significance of NoV as a waterborne pathogen stresses the importance of more stable tools to detect it.

Despite that the main focus and time have been invested on laboratory sample processing the three studies within this PhD thesis also highlight the challenges with detecting specific enteric viruses in the metagenomics dataset. Another major disadvantage is the problem estimate the initial viral concentration of the water sample. Until these obstacles are solved analysis of enteric viruses by metagenomics will not be implemented in routine assessments of water quality. However, viral metagenomics have a great potential and the presented work showed that viral metagenomics could differentiate microbial clean water from contaminated water without looking at specific pathogens. In combination with existing control measures this could be a powerful tool to analyze the effectiveness of the process in a wastewater treatment plant.

5. Conclusion and further perspectives

In conclusion this PhD thesis has extended the knowledge of viral metagenomics for microbial water quality assessment and outbreak analysis. The extensive evaluation of the effect of sample processing of water for viral metagenomics has increased our understanding of method dependent biases. The influence of viral concentration and extraction methods on the observed viral richness, specificity, pathogen detection, and community structure stress the importance of the chosen laboratory methods. This knowledge is essential when interpreting published results and conducting meta-studies. A metagenomics dataset without attached metadata describing the used laboratory methods is therefore not of great use for future comparison and analysis.

With the results of this work I propose to further investigate the viral community structure in different water sources. With the goal to create a viral community profiles for different stages of contaminated and microbial clean water. Today there is no general agreement on when wastewater is safe to reuse for production of ready to eat foods. New methods to assess the microbial quality of water are needed and the usefulness of these viral community profiles should therefore be further investigated.

The investigation of the urban and hospital WWTP demonstrated that the effluent from the hospital WWTP were cleaned from microbes and could be reused. However, most WWTPs, including the urban WWTP, do not have advanced disinfection treatments to reduce microorganisms sufficiently to avoid pathogens. To enable reuse of a larger quantity of wastewater extensive investments are therefore needed to update the current WWTPs.

Finally the rapid development within the sequencing field, with portable sequencing devices and user friendly open bioinformatics pipelines, will make sequencing more accessible for a wider group of users. The application of such methods to identify microbe reductions in WWTPs may enable development of better treatment procedures and thus, turning wastewater into a stable freshwater resource.

References

Abd-Elmaksoud, S., Spencer, S. K., Gerba, C. P., Tamimi, A. H., Jokela, W. E. and Borchardt, M. a. (2014) 'Simultaneous Concentration of Bovine Viruses and Agricultural Zoonotic Bacteria from Water Using Sodocalcic Glass Wool Filters', *Food and Environmental Virology*, pp. 253–259. doi: 10.1007/s12560-014-9159-z.

Acheson, N. H. (2007) *Fundamentals of molecular virology*. John Wiley and sons.

Albinana-Gimenez, N., Clemente-Casares, P., Calgua, B., Huguet, J. M., Courtois, S. and Girones, R. (2009) 'Comparison of methods for concentrating human adenoviruses, polyomavirus JC and noroviruses in source waters and drinking water using quantitative PCR', *Journal of Virological Methods*, 158, pp. 104–109. doi: 10.1016/j.jviromet.2009.02.004.

Alcalde-Sanz, L. and Gawlik, B. . (2017) 'Minimum quality requirements for water reuse in agricultural irrigation and aquifer recharge - Towards a legal instrument on water reuse at EU level', *EUR 28962 EN, Publications Office of the European Union*. Luxembourg, (December), pp. 1–63. doi: 10.2760/804116.

Allende, A., Barceló Culleres, D., Gironés Llop, R., Laval, A., Robertson, L., da Silva Felício, M. T., Gervelmeyer, A., Ramos Bordajandi, L. and Liebana, E. (2017) 'Request for scientific and technical assistance on proposed EU minimum quality requirements for water reuse in agricultural irrigation and aquifer recharge', *EFSA Supporting Publications*, 14(7), pp. 1–19. doi: 10.2903/sp.efsa.2017.EN-1247.

Altschul, S., Gish, W., Miller, W., Myers, E. and Lipman, D. (1990) 'Basic local alignment search tool.', *J Mol Biol*, pp. 403–410. doi: 10.1016/S0022-2836(05)80360-2.

Amarasiri, M., Kitajima, M., Nguyen, T. H., Okabe, S. and Sano, D. (2017) 'Bacteriophage removal efficiency as a validation and operational monitoring tool for virus reduction in wastewater reclamation: Review', *Water Research*. Elsevier Ltd, 121, pp. 258–269. doi: 10.1016/j.watres.2017.05.035.

Andrews S (2018) *Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data*. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (Accessed: 18 February 2018).

Anonymous (2016) *Annual report on zoonoses in denmark 2016, National Food Instutute; Technical University of Denmark*. doi: 10.1128/AEM.02991-09.

Antony, A., Blackbeard, J. and Leslie, G. (2012) 'Removal efficiency and integrity monitoring techniques for virus removal by membrane processes', *Critical Reviews in Environmental Science and Technology*, 42(9), pp. 891–933. doi: 10.1080/10643389.2011.556539.

Araud, E., DiCaprio, E., Ma, Y., Lou, F., Gao, Y., Kingsley, D., Hughes, J. H. and Li, J. (2016) 'Thermal inactivation of enteric viruses and bioaccumulation of enteric foodborne viruses in live oysters (Crassostrea virginica)', *Applied and Environmental Microbiology*, 82(7), pp. 2086–2099. doi: 10.1128/AEM.03573-15.

Banerjee, A., De, P., Manna, B. and Chawla-Sarkar, M. (2017) 'A clinically relevant, syngeneic model of spontaneous, highly metastatic B16 mouse melanoma', *Anticancer Research*, 89(4), pp. 606–614. doi: 10.1002/jmv.24672.

Behjati, S. and Tarpey, P. S. (2013) 'What is next generation sequencing?', *Archives of disease in childhood. Education and practice edition*. BMJ Publishing Group, 98(6), pp. 236–8. doi: 10.1136/archdischild-2013-

304340.

Berensmeier, S. (2006) 'Magnetic particles for the separation and purification of nucleic acids', *Applied Microbiology and Biotechnology*, 73(3), pp. 495–504. doi: 10.1007/s00253-006-0675-0.

Bernard, H., Faber, M., Wilking, H., Haller, S., Höhle, M., Schielke, A., Ducomble, T., Siffczyk, C., Merbecks, S., Fricke, G., Hamouda, O., Stark, K. and Werber, D. (2014) 'Large multistate outbreak of norovirus gastroenteritis associated with frozen strawberries, Germany, 2012', *Eurosurveillance*, 19(8). doi: 10.2807/1560-7917.ES2014.19.8.20719.

Bibby, K. and Peccia, J. (2013a) 'Identification of viral pathogen diversity in sewage sludge by metagenome analysis', *Environmental Science and Technology*, 47(4), pp. 1945–1951. doi: 10.1021/es305181x.

Bibby, K. and Peccia, J. (2013b) 'Identification of viral pathogen diversity in sewage sludge by metagenome analysis.', *Environmental science & technology*, 47(4), pp. 1945–51. doi: 10.1021/es305181x.

Bofill-Mas, S., Albinana-Gimenez, N., Clemente-Casares, P., Hundesa, A., Rodriguez-Manzano, J., Allard, A., Calvo, M. and Girones, R. (2006) 'Quantification and stability of human adenoviruses and polyomavirus JCPyV in wastewater matrices', *Applied and Environmental Microbiology*, 72(12), pp. 7894–7896. doi: 10.1128/AEM.00965-06.

Borgmästars, E., Jazi, M. M., Persson, S., Jansson, L., Rådström, P., Simonsson, M., Hedman, J. and Eriksson, R. (2017) 'Improved Detection of Norovirus and Hepatitis A Virus in Surface Water by Applying Pre-PCR Processing', *Food and Environmental Virology*, 9(4), pp. 395–405. doi: 10.1007/s12560-017-9295-3.

Cabral, J. P. S. (2010) 'Water microbiology. Bacterial pathogens and water', *International Journal of Environmental Research and Public Health*, 7(10), pp. 3657–3703. doi: 10.3390/ijerph7103657.

Calgua, B., Fumian, T., Rusiñol, M., Rodriguez-Manzano, J., Mbayed, V. a., Bofill-Mas, S., Miagostovich, M. and Girones, R. (2013) 'Detection and quantification of classic and emerging viruses by skimmed-milk flocculation and PCR in river water from two geographical areas', *Water Research*, 47, pp. 2797–2810. doi: 10.1016/j.watres.2013.02.043.

Calgua, B., Mengewein, A., Grunert, A., Bofill-Mas, S., Clemente-Casares, P., Hundesa, A., Wyn-Jones, A. P., López-Pila, J. M. and Girones, R. (2008) 'Development and application of a one-step low cost procedure to concentrate viruses from seawater samples', *Journal of Virological Methods*, 153(2), pp. 79–83. doi: 10.1016/j.jviromet.2008.08.003.

California State Water Resources Control Board (2015) *Regulations Related to Recycled Water*. Available at: http://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/Lawbook.shtml.

Callejón, R. M., Rodríguez-Naranjo, M. I., Ubeda, C., Hornedo-Ortega, R., Garcia-Parrilla, M. C. and Troncoso, A. M. (2015) 'Reported Foodborne Outbreaks Due to Fresh Produce in the United States and European Union: Trends and Causes', *Foodborne Pathogens and Disease*, 12(1), pp. 32–38. doi: 10.1089/fpd.2014.1821.

Cangelosi, G. A. and Meschke, J. S. (2014) 'Dead or alive: Molecular assessment of microbial viability', *Applied and Environmental Microbiology*, 80(19), pp. 5884–5891. doi: 10.1128/AEM.01763-14.

Cantalupo, P. G., Calgua, B., Zhao, G., Hundesa, A., Wier, A. D., Katz, J. P., Grabe, M., Hendrix, R. W.,

Girones, R., Wang, D. and Pipas, J. M. (2011) 'Raw sewage harbors diverse viral populations', *mBio*, 2(5), pp. 1–11. doi: 10.1128/mBio.00180-11.

Ceylan, E., McMahon, W. and Garren, D. M. (2017) 'Thermal Inactivation of *Listeria monocytogenes* and *Salmonella* during Water and Steam Blanching of Vegetables', *Journal of Food Protection*, 80(9), pp. 1550–1556. doi: 10.4315/0362-028X.JFP-16-517.

Chao, Y., Mao, Y., Wang, Z. and Zhang, T. (2015) 'Diversity and functions of bacterial community in drinking water biofilms revealed by high-throughput sequencing', *Scientific Reports*. Nature Publishing Group, 5(May 2014), p. 10044. doi: 10.1038/srep10044.

Chaudhry, R. M., Holloway, R. W., Cath, T. Y. and Nelson, K. L. (2015) 'Impact of virus surface characteristics on removal mechanisms within membrane bioreactors', *Water Research*. Elsevier Ltd, 84, pp. 144–152. doi: 10.1016/j.watres.2015.07.020.

Chaudhry, R. M., Nelson, K. L. and Drewes, J. E. (2015) 'Mechanisms of pathogenic virus removal in a full-scale membrane bioreactor', *Environmental Science and Technology*, 49(5), pp. 2815–2822. doi: 10.1021/es505332n.

Cotten, M., Petrova, V., Phan, M. V. T., Rabaa, M. A., Watson, S. J., Ong, S. H., Kellam, P. and Baker, S. (2014) 'Deep Sequencing of Norovirus Genomes Defines Evolutionary Patterns in an Urban Tropical Setting', *Journal of Virology*, 88(19), pp. 11056–11069. doi: 10.1128/JVI.01333-14.

Council, N. R. (2007) *The New Science of Metagenomics*. Washington DC: The National Academies Press. doi: 10.17226/11902.

Craun, G. F., Craun, M. F., Calderon, R. L. and Beach, M. J. (2006) 'Waterborne outbreaks reported in the United States', *Journal of Water and Health*, 4(Suppl 2), p. 19. doi: 10.2166/wh.2006.016.

Crump, J. A., Luby, S. P. and Mintz, E. D. (2004) 'The global burden of typhoid fever', *Bulletin of the World Health Organization*, 82(5), pp. 346–353. doi: 10.1590/S0042-96862004000500008.

Daly, G. M., Bexfield, N., Heaney, J., Stubbs, S., Mayer, A. P., Palser, A., Kellam, P., Drou, N., Caccamo, M., Tiley, L., Alexander, G. J. M., Bernal, W. and Heeney, J. L. (2011) 'A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing', *PLoS ONE*, 6(12). doi: 10.1371/journal.pone.0028879.

Deng, X., den Bakker, H. C. and Hendriksen, R. S. (2016) 'Genomic Epidemiology: Whole-Genome-Sequencing–Powered Surveillance and Outbreak Investigation of Foodborne Bacterial Pathogens', *Annual Review of Food Science and Technology*, 7(1), pp. 353–374. doi: 10.1146/annurev-food-041715-033259.

Deshmukh, R. A., Joshi, K., Bhand, S. and Roy, U. (2016) 'Recent developments in detection and enumeration of waterborne bacteria: a retrospective minireview', *MicrobiologyOpen*, 5(6), pp. 901–922. doi: 10.1002/mbo3.383.

van Dijk, E. L., Auger, H., Jaszczyszyn, Y. and Thermes, C. (2014) 'Ten years of next-generation sequencing technology', *Trends in Genetics*, 30(9), pp. 418–426. doi: 10.1016/j.tig.2014.07.001.

Ding, Y., Tang, Y., Kwok, C. K., Zhang, Y., Bevilacqua, P. C. and Assmann, S. M. (2014) 'In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features', *Nature*. Nature Publishing Group, 505(7485), pp. 696–700. doi: 10.1038/nature12756.

EEC Council (1991) '91/271/EEC of 21 May 1991 concerning urban waste-water treatment', *EEC Council Directive*, (L), p. 10. doi: http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:31991L0271.

EFSA (2015) *The European Union summary report on trends and sources of zoonoses, zoonotic agents and food‑borne outbreaks in 2014*, *EFSA Journal*. doi: 10.2903/j.efsa.2015.4329.

Ethelberg, S., Lisby, M., Böttiger, B., Schultz, A. C., Villif, A., Jensen, T., Olsen, K. E., Scheutz, F., Kjelsø, C. and Müller, L. (2010) 'Outbreaks of gastroenteritis linked to lettuce, Denmark, January 2010', *Eurosurveillance*, 15(6), p. 1. doi: 19484 [pii].

European Union (2017) *Commission notice on guidance document on addressing microbial risks in fresh fruits and vegetables at primary production through good hygiene 2017/C 163/01.*

FAO/WHO (2004) *Fruit and Vegetables for Health*, *Handbook of Plant Food Phytochemicals*. doi: 10.1002/9781118464717.ch5.

Felczykowska, A., Krajewska, A., Zielińska, S. and Łos, J. M. (2015) 'Sampling, metadata and DNA extraction - Important steps in metagenomic studies', *Acta Biochimica Polonica*, 62(1), pp. 151–160. doi: 10.18388/abp.2014_916.

Ferguson, A. S., Layton, A. C., Mailloux, B. J., Culligan, P. J., Williams, D. E., Smartt, A. E., Sayler, G. S., Feighery, J., McKay, L. D., Knappett, P. S. K., Alexandrova, E., Arbit, T., Emch, M., Escamilla, V., Ahmed, K. M., Alam, M. J., Streatfield, P. K., Yunus, M. and van Geen, A. (2012) 'Comparison of fecal indicators with pathogenic bacteria and rotavirus in groundwater', *Science of the Total Environment*, 431, pp. 314–322. doi: 10.1016/j.scitotenv.2012.05.060.

Fernandez-Cassi, X., Timoneda, N., Gonzales-Gustavson, E., Abril, J. F., Bofill-Mas, S. and Girones, R. (2017) 'A metagenomic assessment of viral contamination on fresh parsley plants irrigated with fecally tainted river water', *International Journal of Food Microbiology*, 257(June), pp. 80–90. doi: 10.1016/j.ijfoodmicro.2017.06.001.

Fonager, J., Stegger, M., Rasmussen, L. D., Poulsen, M. W., Rønn, J., Andersen, P. S. and Fischer, T. K. (2017) 'A universal primer-independent next-generation sequencing approach for investigations of norovirus outbreaks and novel variants', *Scientific Reports*, 7(1), pp. 1–11. doi: 10.1038/s41598-017-00926-x.

Gerba, C. P., Goyal, S. M., LaBelle, R. L. and Bodgan, G. F. (1979) 'Failure of indicator bacteria to reflect the occurrence of enteroviruses in marine waters', *American Journal of Public Health*, 69(11), pp. 1116–1119. doi: 10.2105/AJPH.69.11.1116.

Gibson, U. E. M., Heid, C. A. and Williams, P. M. (1996) 'A novel method for real time quantitative RT-PCR.', *Genome research*, 6(10), pp. 995–1001. doi: 10.1101/gr.6.10.995.

Gonzales-Gustavson, E., Cárdenas-Youngs, Y., Calvo, M., da Silva, M. F. M., Hundesa, A., Amorós, I., Moreno, Y., Moreno-Mesonero, L., Rosell, R., Ganges, L., Araujo, R. and Girones, R. (2017) 'Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa', *Journal of Microbiological Methods*, 134, pp. 46–53. doi: 10.1016/j.mimet.2017.01.006.

Grimwood, K., Carzino, R., Barnes, G. L. and Bishop, R. F. (1995) 'Patients with enteric adenovirus gastroenteritis admitted to an Australian pediatric teaching hospital from 1981 to 1992', *Journal of Clinical*

Microbiology, 33(1), pp. 131–136.

Guerra, J. A. de A. A., Kampa, K. C., Morsoletto, D. G. B., Junior, A. P. and Ivantes, C. A. P. (2017) 'Hepatitis E: A Literature Review', *Journal of Clinical and Translational Hepatology*, X(X), pp. 1–8. doi: 10.14218/JCTH.2017.00012.

Heinonen-Tanski, H., Juntunen, P., Rajala, R., Haume, E. and Niemelä, A. (2003) 'Costs of tertiary treatment of municipal wastewater by rapid sand filter with coagulants and UV', *Water Science and Technology: Water Supply*, 3(4), pp. 145–152.

Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., Bergström, T. and Norder, H. (2014) 'Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks.', *Applied and environmental microbiology*, 80(21), pp. 6771–81. doi: 10.1128/AEM.01981-14.

Hirneisen, K. A. and Kniel, K. E. (2013) 'Norovirus Surrogate Survival on Spinach During Preharvest Growth', *Phytopathology*, 103(4), pp. 389–394. doi: 10.1094/PHYTO-09-12-0231-FI.

Hodgkin, J. (2001) 'Genome Size', *Encyclopedia of Genetics*, p. 865. doi: 10.1006/rwgn.2001.0557.

Hsiung, G. D. (1984) 'Diagnostic virology: From animals to automation', *Yale Journal of Biology and Medicine*, 57(5), pp. 727–733.

Hurwitz, B. L., Deng, L., Poulos, B. T. and Sullivan, M. B. (2013) 'Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics.', *Environmental microbiology*, 15(5), pp. 1428–40. Available at: http://doi.wiley.com/10.1111/j.1462-2920.2012.02836.x (Accessed: 9 March 2015).

Iker, B. C., Bright, K. R., Pepper, I. L., Gerba, C. P. and Kitajima, M. (2013) 'Evaluation of commercial kits for the extraction and purification of viral nucleic acids from environmental and fecal samples', *Journal of Virological Methods*. Elsevier B.V., 191(1), pp. 24–30. doi: 10.1016/j.jviromet.2013.03.011.

Iritani, N., Kaida, A., Abe, N., Kubo, H., Sekiguchi, J.-I., Yamamoto, S. P., Goto, K., Tanaka, T. and Noda, M. (2014) 'Detection and genetic characterization of human enteric viruses in oyster-associated gastroenteritis outbreaks between 2001 and 2012 in Osaka City, Japan', *Journal of Medical Virology*, 86(12), pp. 2019–2025. doi: 10.1002/jmv.23883.

Irshad, M., Gupta, P., Mankotia, D. S. and Ansari, M. A. (2016) 'Multiplex qPCR for serodetection and serotyping of hepatitis viruses: A brief review', *World Journal of Gastroenterology*, 22(20), pp. 4824–4834. doi: 10.3748/wjg.v22.i20.4824.

Jacobsen, K. H. and Wiersma, S. T. (2010) 'Hepatitis A virus seroprevalence by age and world region, 1990 and 2005', *Vaccine*. Elsevier Ltd, 28(41), pp. 6653–6657. doi: 10.1016/j.vaccine.2010.08.037.

John, S. G., Mendez, C. B., Deng, L., Poulos, B., Kauffman, A. K. M., Kern, S., Brum, J., Polz, M. F., Boyle, E. A. and Sullivan, M. B. (2011a) 'A simple and efficient method for concentration of ocean viruses by chemical flocculation', *Environmental Microbiology Reports*, 3(2), pp. 195–202. doi: 10.1111/j.1758-2229.2010.00208.x.

John, S. G., Mendez, C. B., Deng, L., Poulos, B., Kauffman, A. K. M., Kern, S., Brum, J., Polz, M. F., Boyle, E. A. and Sullivan, M. B. (2011b) 'A simple and efficient method for concentration of ocean viruses

by chemical flocculation', *Environmental Microbiology Reports*, 3(2), pp. 195–202. doi: 10.1111/j.1758-2229.2010.00208.x.

Karlsson, O. E., Hansen, T., Knutsson, R., Löfström, C., Granberg, F. and Berg, M. (2013) 'Metagenomic detection methods in biopreparedness outbreak scenarios.', *Biosecurity and bioterrorism : biodefense strategy, practice, and science*, 11 Suppl 1, pp. S146-57. doi: 10.1089/bsp.2012.0077.

Kennedy, N. A., Walker, A. W., Berry, S. H., Duncan, S. H., Farquarson, F. M., Louis, P., Thomson, J. M., Satsangi, J., Flint, H. J., Parkhill, J., Lees, C. W. and Hold, G. L. (2014) 'The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing', *PLoS ONE*, 9(2), pp. 1–9. doi: 10.1371/journal.pone.0088982.

Kirk, M. D., Pires, S. M., Black, R. E., Caipo, M., Crump, J. A., Devleesschauwer, B., Döpfer, D., Fazil, A., Fischer-Walker, C. L., Hald, T., Hall, A. J., Keddy, K. H., Lake, R. J., Lanata, C. F., Torgerson, P. R., Havelaar, A. H. and Angulo, F. J. (2015) 'World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis', *PLoS Medicine*, 12(12), pp. 1–21. doi: 10.1371/journal.pmed.1001921.

Kisand, V., Valente, A., Lahm, A., Tanet, G. and Lettieri, T. (2012) 'Phylogenetic and functional metagenomic profiling for assessing microbial biodiversity in environmental monitoring', *PLoS ONE*, 7(8). doi: 10.1371/journal.pone.0043630.

Kobayashi, S., Fujiwara, N., Yasui, Y., Yamashita, T., Hiramatsu, R. and Minagawa, H. (2012) 'A foodborne outbreak of sapovirus linked to catered box lunches in Japan', *Archives of Virology*, 157(10), pp. 1995–1997. doi: 10.1007/s00705-012-1394-8.

Koopmans, M. and Duizer, E. (2004) 'Foodborne viruses: An emerging problem', *International Journal of Food Microbiology*, 90(1), pp. 23–41. doi: 10.1016/S0168-1605(03)00169-7.

Krustok, I., Truu, J., Odlare, M., Truu, M., Ligi, T., Tiirik, K. and Nehrenheim, E. (2015) 'Effect of lake water on algal biomass and microbial community structure in municipal wastewater-based lab-scale photobioreactors', *Applied Microbiology and Biotechnology*, pp. 6537–6549. doi: 10.1007/s00253-015-6580-7.

Kuhn, K. G., Falkenhorst, G., Emborg, H. D., Ceper, T., Torpdahl, M., Krogfelt, K. A., Ethelberg, S. and Mølbak, K. (2017) 'Epidemiological and serological investigation of a waterborne Campylobacter jejuni outbreak in a Danish town', *Epidemiology and Infection*, 145(4), pp. 701–709. doi: 10.1017/S0950268816002788.

Kundu, S., Lockwood, J., Depledge, D. P., Chaudhry, Y., Aston, A., Rao, K., Hartley, J. C., Goodfellow, I. and Breuer, J. (2013) 'Next-generation whole genome sequencing identifies the direction of norovirus transmission in linked patients', *Clinical Infectious Diseases*, 57(3), pp. 407–414. doi: 10.1093/cid/cit287.

Kunze, A., Pei, L., Elsässer, D., Niessner, R. and Seidel, M. (2015) 'High performance concentration method for viruses in drinking water', *Journal of Virological Methods*. Elsevier B.V., 222, pp. 132–137. doi: 10.1016/j.jviromet.2015.06.007.

Lanata, C. F., Fischer-Walker, C. L., Olascoaga, A. C., Torres, C. X., Aryee, M. J. and Black, R. E. (2013) 'Global Causes of Diarrheal Disease Mortality in Children <5 Years of Age: A Systematic Review', *PLoS ONE*, 8(9). doi: 10.1371/journal.pone.0072788.

Leclerc, H., Schwartzbrod, L. and Dei-Cas, E. (2002) 'Microbial agents associated with waterborne diseases', *Critical Reviews in Microbiology*, 28(4), pp. 371–409. doi: 10.1080/1040-840291046768.

Lee, J. I., Lee, G.-C., Chung, J. Y., Han, T. H., Lee, Y. K., Kim, M. S. and Lee, C. H. (2012) 'Detection and molecular characterization of adenoviruses in Korean children hospitalized with acute gastroenteritis', *Microbiology and Immunology*, 56(8), pp. 523–528. doi: 10.1111/j.1348-0421.2012.00469.x.

Leeuwen, M. Van, Williams, M. M. W., Simon, J. H., Smits, S. L., Albert, D. M., Osterhaus, E., Koraka, P. and Osterhaus, A. D. M. E. (2010) 'Human picobirnaviruses identified by molecular screening of diarrhea samples.', *Journal of clinical microbiology*, 48(5), pp. 1787–94. doi: 10.1128/JCM.02452-09.

Leland, D. S. and Ginocchio, C. C. (2007) 'Role of cell culture for virus detection in the age of technology', *Clinical Microbiology Reviews*, 20(1), pp. 49–78. doi: 10.1128/CMR.00002-06.

Li, D., Gu, A. Z., Zeng, S., Yang, W., He, M. and Shi, H. (2011) 'Evaluation of the infectivity, gene and antigenicity persistence of rotaviruses by free chlorine disinfection', *Journal of Environmental Sciences*. The Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences, 23(10), pp. 1691–1698. doi: 10.1016/S1001-0742(10)60623-7.

Li, L., Deng, X., Mee, E. T., Collot-teixeira, S., Anderson, R., Schepelmann, S., Minor, P. D. and Delwart, E. (2015) 'Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent', *Journal of Virological Methods*. Elsevier B.V., 213, pp. 139–146. doi: 10.1016/j.jviromet.2014.12.002.

Lopman, B. A., Reacher, M. H., Vipond, I. B., Sarangi, J. and Brown, D. W. G. (2004) 'Clinical Manifestation of Norovirus Gastroenteritis in Health Care Settings', *Clinical Infectious Diseases*, 39(3), pp. 318–324. doi: 10.1086/421948.

Magoč, T. and Salzberg, S. L. (2011) 'FLASH: Fast length adjustment of short reads to improve genome assemblies', *Bioinformatics*, 27(21), pp. 2957–2963. doi: 10.1093/bioinformatics/btr507.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J. B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. and Rothberg, J. M. (2005) 'Genome sequencing in microfabricated high-density picolitre reactors', *Nature*, 437(7057), pp. 376–380. doi: 10.1038/nature03959.

Martin, M. (2011) 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.journal*, 17(1), p. 10. doi: 10.14806/ej.17.1.200.

McMinn, B. R., Ashbolt, N. J. and Korajkic, A. (2017) 'Bacteriophages as indicators of faecal pollution and enteric virus removal', *Letters in Applied Microbiology*, 65(1), pp. 11–26. doi: 10.1111/lam.12736.

Mee, E. T., Preston, M. D., Participants, S., Minor, P. D., Schepelmann, S., Huang, X., Nguyen, J., Wall, D., Hargrove, S., Fu, T., Xu, G., Li, L., Cote, C., Delwart, E., Li, L., Hewlett, I., Simonyan, V., Ragupathy, V., Voskanian-Kordi, A., Mermod, N., Hill, C., Ottenwälder, B., Richter, D. C., Tehrani, A., Weber-Lehmann,

J., Cassart, J.-P., Letellier, C., Vandeputte, O., Ruelle, J.-L., Deyati, A., Neve, F. La, Modena, C., Mee, E., Preston, M., Minor, P., Eloit, M., Muth, E., Lamamy, A., Jagorel, F., Cheval, J., Anscombe, C., Misra, R., Wooldridge, D., Gharbia, S., Rose, G., Ng, S. H. S., Charlebois, R. L., Gisonni-Lex, L., Mallet, L., Dorange, F., Chiu, C., Naccache, S., Kellam, P., Van Der Hoek, L., Cotten, M., Mitchell, C., Baier, B. S., Sun, W. and Malicki, H. D. (2016) 'Development of a candidate reference material for adventitious virus detection in vaccine and biologicals manufacturing by deep sequencing', *Vaccine*, 34(34), pp. 2035–2043. doi: 10.1016/j.vaccine.2015.12.020.

Moore, B. C., Martinez, E., Gay, J. M. and Rice, D. H. (2003) 'Survival of Salmonella enterica in Freshwater and Sediments and Transmission by the Aquatic Midge Chironomus tentans (Chironomidae: Diptera)', *Applied and Environmental Microbiology*, 69(8), pp. 4556–4560. doi: 10.1128/AEM.69.8.4556-4560.2003.

Müller, L., Schultz, A. C., Fonager, J., Jensen, T., Lisby, M., Hindsdal, K., Krusell, L., Eshøj, A., Møller, L. T., Porsbo, L. J., Böttiger, B. E., Kuhn, K., Engberg, J. and Ethelberg, S. (2015) 'Separate norovirus outbreaks linked to one source of imported frozen raspberries by molecular analysis, Denmark, 2010-2011', *Epidemiology and Infection*, 143(11), pp. 2299–2307. doi: 10.1017/S0950268814003409.

Naccache, S. N., Greninger, A. L., Lee, D., Coffey, L. L., Phan, T., Rein-Weston, A., Aronsohn, A., Hackett, J., Delwart, E. L. and Chiu, C. Y. (2013) 'The Perils of Pathogen Discovery: Origin of a Novel Parvovirus-Like Hybrid Genome Traced to Nucleic Acid Extraction Spin Columns', *Journal of Virology*, 87(22), pp. 11966–11977. doi: 10.1128/JVI.02323-13.

Newsome, T., Li, B.-J., Zou, N. and Lo, S.-C. (2004) 'Presence of Bacterial Phage-Like DNA Sequences in Commercial Taq DNA Polymerase Reagents', *Journal of Clinical Microbiology*, 42(5), pp. 2264–2267. doi: 10.1128/JCM.42.5.2264-2267.2004.

Newton, R. J., Mclellan, S. L., Dila, D. K., Vineis, J. H., Morrison, H. G., Eren, a M. and Sogin, M. L. (2015) 'Sewage Reflects the Microbiomes of Human Populations', *mBio*, 6(2), pp. 1–9. doi: 10.1128/mBio.02574-14.Editor.

Ng, T. F. F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, B., Bodhidatta, L., Oderinde, B. S., Wommack, K. E. and Delwart, E. (2012) 'High Variety of Known and New RNA and DNA Viruses of Diverse Origins in Untreated Sewage', *Journal of Virology*, 86(22), pp. 12161–12175. doi: 10.1128/JVI.00869-12.

Nielsen, K. L., Stegger, M., Kiil, K., Godfrey, P. A., Feldgarden, M., Lilje, B., Andersen, P. S. and Frimodt-Møller, N. (2017) 'Whole-genome comparison of urinary pathogenic Escherichia coli and faecal isolates of UTI patients and healthy controls', *International Journal of Medical Microbiology*, 307(8), pp. 497–507. doi: 10.1016/j.ijmm.2017.09.007.

Nieuwenhuijse, D. F. and Koopmans, M. P. G. (2017) 'Metagenomic sequencing for surveillance of food- and waterborne viral diseases', *Frontiers in Microbiology*, 8(FEB), pp. 1–11. doi: 10.3389/fmicb.2017.00230.

Nordahl Petersen, T., Rasmussen, S., Hasman, H., Carøe, C., Bælum, J., Charlotte Schultz, A., Bergmark, L., Svendsen, C. A., Lund, O., Sicheritz-Pontén, T. and Aarestrup, F. M. (2015) 'Meta-genomic analysis of toilet waste from long distance flights; a step towards global surveillance of infectious diseases and antimicrobial resistance', *Scientific Reports*. Nature Publishing Group, 5(April), p. 11444. doi: 10.1038/srep11444.

NSF (2009) 'NSF Guide to the European Standards for Drinking Water CONTENTS NSF Guide to the European Standards for Drinking Water Treatment Units'. Available at: http://www.nsf.org/newsroom_pdf/European_Stds_Guide_LT_EN_LDW10050309.pdf.

O'Ryan, M. (2017) 'Rotavirus Vaccines: a story of success with challenges ahead', *F1000Research*, 6(0), p. 1517. doi: 10.12688/f1000research.11912.1.

Ogorzaly, L., Walczak, C., Galloux, M., Etienne, S., Gassilloud, B. and Cauchie, H. M. (2015) 'Human Adenovirus Diversity in Water Samples Using a Next-Generation Amplicon Sequencing Approach', *Food and Environmental Virology*, 7(2), pp. 112–121. doi: 10.1007/s12560-015-9194-4.

Pan, Y., Tian, X., Qin, P., Wang, B., Zhao, P., Yang, Y. Le, Wang, L., Wang, D., Song, Y., Zhang, X. and Huang, Y. W. (2017) 'Discovery of a novel swine enteric alphacoronavirus (SeACoV) in southern China', *Veterinary Microbiology*. Elsevier, 211(September), pp. 15–21. doi: 10.1016/j.vetmic.2017.09.020.

Park, J., Lee, S., Jin, J., Cho, H., Jheong, W. and Paik, S. (2015) 'Complete Nucleotide Sequence Analysis of the Norovirus GII . 4 Sydney Variant in South Korea', *BioMed Research International*, 2015.

Patel, M. M., Hall, A. J., Vinjé, J. and Parashar, U. D. (2009) 'Noroviruses: A comprehensive review', *Journal of Clinical Virology*, 44(1), pp. 1–8. doi: 10.1016/j.jcv.2008.10.009.

Pei, L., Rieger, M., Lengger, S., Ott, S., Zawadsky, C., Hartmann, N. M., Selinka, H., Tiehm, A., Niessner, R. and Seidel, M. (2012) 'Combination of Cross fl ow Ultra fi ltration , Monolithic A ffi nity Filtration , and Quantitative Reverse Transcriptase PCR for Rapid Concentration and Quanti fi cation of Model Viruses in Water'.

Petersen, T. N., Lukjancenko, O., Thomsen, M. C. F., Maddalena Sperotto, M., Lund, O., Møller Aarestrup, F. and Sicheritz-Pontén, T. (2017) 'MGmapper: Reference based mapping and taxonomy annotation of metagenomics sequence reads', *Plos One*, 12(5), p. e0176469. doi: 10.1371/journal.pone.0176469.

Petrich, A., Mahony, J., Chong, S., Broukhanski, G., Gharabaghi, F., Johnson, G., Louie, L., Luinstra, K., Willey, B., Akhaven, P., Chui, L., Jamieson, F., Louie, M., Mazzulli, T., Tellier, R., Smieja, M., Cai, W., Chernesky, M. and Richardson, S. E. (2006) 'Multicenter comparison of nucleic acid extraction methods for detection of severe acute respiratory syndrome coronavirus RNA in stool specimens.', *Journal of clinical microbiology*, 44(8), pp. 2681–8. doi: 10.1128/JCM.02460-05.

Prevost, B., Goulet, M., Lucas, F. S., Joyeux, M., Moulin, L. and Wurtzer, S. (2016) 'Viral persistence in surface and drinking water: Suitability of PCR pre-treatment with intercalating dyes', *Water Research*. Elsevier Ltd, 91, pp. 68–76. doi: 10.1016/j.watres.2015.12.049.

Quijada, N. M., Fongaro, G., Barardi, C. R. M., Hernández, M. and Rodríguez-Lázaro, D. (2016) 'Propidium monoazide integrated with qPCR enables the detection and enumeration of infectious enteric RNA and DNA viruses in clam and fermented sausages', *Frontiers in Microbiology*, 7(DEC), pp. 1–6. doi: 10.3389/fmicb.2016.02008.

Rames, E., Roiko, A., Stratton, H. and Macdonald, J. (2016) 'Technical aspects of using human adenovirus as a viral water quality indicator', *Water Research*. Elsevier Ltd, 96, pp. 308–326. doi: 10.1016/j.watres.2016.03.042.

Rao, D. G., Senthilkumar, R., Bryne, J. A. and Feroz, S. (1997) *Wastewater Treatment Advanced processes*

and technologies. Edited by D. G. Rao, R. Senthilkumar, J. A. Bryne, and S. Feroz. IWA Publishing, CRC Press.

Rasmussen, L. D., Schultz, A. C., Uhrbrand, K., Jensen, T. and Fischer, T. K. (2016) 'Molecular Evidence of Oysters as Vehicle of Norovirus GII.P17-GII.17', *Emerging Infectious Diseases*, 22(11), pp. 2024–2025. doi: 10.3201/eid2211.161171.

Rigotto, C., Hanley, K., Rochelle, P. A., De Leon, R., Barardi, C. R. M. and Yates, M. V. (2011) 'Survival of adenovirus types 2 and 41 in surface and ground waters measured by a plaque assay', *Environmental Science and Technology*, 45(9), pp. 4145–4150. doi: 10.1021/es103922r.

Rizzo, L., Krätke, R., Linders, J., Scott, M., Vighi, M. and de Voogt, P. (2018) 'Proposed EU minimum quality requirements for water reuse in agricultural irrigation and aquifer recharge: SCHEER scientific advice', *Current Opinion in Environmental Science & Health*. Elsevier Ltd, 2, pp. 7–11. doi: 10.1016/j.coesh.2017.12.004.

Rodríguez-Lázaro, D., Cook, N., Ruggeri, F. M., Sellwood, J., Nasser, A., Nascimento, M. S. J., D'Agostino, M., Santos, R., Saiz, J. C., Rzezutka, A., Bosch, A., Gironés, R., Carducci, A., Muscillo, M., Kovač, K., Diez-Valcarce, M., Vantarakis, A., von Bonsdorff, C. H., de Roda Husman, A. M., Hernández, M. and van der Poel, W. H. M. (2012) 'Virus hazards from food, water and other contaminated environments', *FEMS Microbiology Reviews*, 36(4), pp. 786–814. doi: 10.1111/j.1574-6976.2011.00306.x.

Rodriguez-Manzano, J., Alonso, J. L., Ferrús, M. A., Moreno, Y., Amorós, I., Calgua, B., Hundesa, A., Guerrero-Latorre, L., Carratala, A., Rusiñol, M. and Girones, R. (2012) 'Standard and new faecal indicators and pathogens in sewage treatment plants, microbiological parameters for improving the control of reclaimed water', *Water Science and Technology*, 66(12), pp. 2517–2523. doi: 10.2166/wst.2012.233.

Rompré, A., Servais, P., Baudart, J., De-Roubin, M. R. and Laurent, P. (2002) 'Detection and enumeration of coliforms in drinking water: Current methods and emerging approaches', *Journal of Microbiological Methods*, 49(1), pp. 31–54. doi: 10.1016/S0167-7012(01)00351-7.

Rosseel, T., Pardon, B., De Clercq, K., Ozhelvaci, O. and Van Borm, S. (2014) 'False-positive results in metagenomic virus discovery: A strong case for follow-up diagnosis', *Transboundary and Emerging Diseases*, 61(4), pp. 293–299. doi: 10.1111/tbed.12251.

Rusinol, M., Fernandez-Cassi, X., Timoneda, N., Carratal??, A., Abril, J. F., Silvera, C., Figueras, M. J., Gelati, E., Rod??, X., Kay, D., Wyn-Jones, P., Bofill-Mas, S. and Girones, R. (2015) 'Evidence of viral dissemination and seasonality in a Mediterranean river catchment: Implications for water pollution management', *Journal of Environmental Management*, 159, pp. 58–67. doi: 10.1016/j.jenvman.2015.05.019.

Räsanen, S., Lappalainen, S., Kaikkonen, S., Hämäläinen, M., Salminen, M. and Vesikari, T. (2010) 'Mixed viral infections causing acute gastroenteritis in children in a waterborne outbreak', *Epidemiology and Infection*, 138(9), pp. 1227–1234. doi: 10.1017/S0950268809991671.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and Arnheim, N. (1985) 'Enzymatic Amplification of β-Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia', *Science*, 230(4732), pp. 1350–1354. doi: 10.1126/science.39.1002.398.

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J. and Walker, A. W. (2014) 'Reagent and laboratory contamination can critically impact

sequence-based microbiome analyses', *BMC Biology*, 12(1), p. 87. doi: 10.1186/s12915-014-0087-z.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) 'DNA sequencing with chain-terminating inhibitors', *Proceedings of the National Academy of Sciences*, 74(12), pp. 5463–5467. doi: 10.1073/pnas.74.12.5463.

Sarkis, S. (2013) *Guidelines for validating treatment processes for pathogen reduction - Supporting Class A recycled water schemes in Victoria*.

Satinsky, B. M., Fortunato, C. S., Doherty, M., Smith, C. B., Sharma, S., Ward, N. D., Krusche, A. V., Yager, P. L., Richey, J. E., Moran, M. A. and Crump, B. C. (2015) 'Metagenomic and metatranscriptomic inventories of the lower Amazon River, May 2011', *Microbiome*. Microbiome, 3(1), p. 39. doi: 10.1186/s40168-015-0099-0.

Schrader, C., Schielke, A., Ellerbroek, L. and Johne, R. (2012) 'PCR inhibitors - occurrence, properties and removal', *Journal of Applied Microbiology*, 113(5), pp. 1014–1026. doi: 10.1111/j.1365-2672.2012.05384.x.

Shi, P., Jia, S., Zhang, X. X., Zhang, T., Cheng, S. and Li, A. (2013) 'Metagenomic insights into chlorination effects on microbial antibiotic resistance in drinking water', *Water Research*. Elsevier Ltd, 47(1), pp. 111–120. doi: 10.1016/j.watres.2012.09.046.

Simmons, F. J., Kuo, D. H. W. and Xagoraraki, I. (2011) 'Removal of human enteric viruses by a full-scale membrane bioreactor during municipal wastewater processing', *Water Research*. Elsevier Ltd, 45(9), pp. 2739–2750. doi: 10.1016/j.watres.2011.02.001.

Straub, T. M. and Chandler, D. P. (2003) 'Towards a unified system for detecting waterborne pathogens', *Journal of Microbiological Methods*, 53(2), pp. 185–197. doi: 10.1016/S0167-7012(03)00023-X.

Thomas, T., Gilbert, J. and Meyer, F. (2012) 'Metagenomics - a guide from sampling to data analysis', *Microbial Informatics and Experimentation*. BioMed Central Ltd, 2(1), p. 3. doi: 10.1186/2042-5783-2-3.

Turcatti, G., Romieu, A., Fedurco, M. and Tairi, A. P. (2008) 'A new class of cleavable fluorescent nucleotides: Synthesis and optimization as reversible terminators for DNA sequencing by synthesis', *Nucleic Acids Research*, 36(4). doi: 10.1093/nar/gkn021.

UN-WWAP (2015) *The United Nations World Water Development Report 2015: Water for a Sustainable World*. doi: 978-92-3-100071-3.

United Nations Department of Economic and Social Affairs, P. D. (2017) *World Population Prospects The 2017 Revision Key Findings and Advance Tables*, *Working Paper No. ESA/P/WP/248*. doi: 10.1017/CBO9781107415324.004.

US-EPA (2012) *Guidelines for Water Reuse 2004*.

WaterSecure (2017) *Membrane bioreactor, WaterVal validation protocol*.

Wéry, N., Lhoutellier, C., Ducray, F., Delgenès, J. P. and Godon, J. J. (2008) 'Behaviour of pathogenic and indicator bacteria during urban wastewater treatment and sludge composting, as revealed by quantitative PCR', *Water Research*, 42(1–2), pp. 53–62. doi: 10.1016/j.watres.2007.06.048.

Wesolowska-Andersen, A., Bahl, M. I., Carvalho, V., Kristiansen, K., Sicheritz-Pontén, T., Gupta, R. and Licht, T. R. (2014a) 'Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis.', *Microbiome*, 2, p. 19. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4063427&tool=pmcentrez&rendertype=abstract.

Wesolowska-Andersen, A., Bahl, M. I., Carvalho, V., Kristiansen, K., Sicheritz-Pontén, T., Gupta, R. and Licht, T. R. (2014b) 'Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis.', *Microbiome*, 2, p. 19. doi: 10.1186/2049-2618-2-19.

WHO (2006) *Guidelines for the safe use of wastewater, excreta and greywater - Wastewater used in agriculture*.

WHO (2015) *WHO estimates of the global burden of foodbonre diseases*.

WHO (2017) *Media centre Hepatitis E*. Available at: http://www.who.int/mediacentre/factsheets/fs280/en/ (Accessed: 8 February 2018).

WHO (2018a) *Campylobacter*, *WHO Fact sheet*. Available at: http://www.who.int/mediacentre/factsheets/fs255/en/ (Accessed: 10 February 2018).

WHO (2018b) *Salmonella (non-typhoidal)*, *WHO Fact sheet*. Available at: http://www.who.int/mediacentre/factsheets/fs139/en/ (Accessed: 10 February 2018).

Wooley, J. C., Godzik, A. and Friedberg, I. (2010) 'A primer on metagenomics', *PLoS Computational Biology*, 6(2). doi: 10.1371/journal.pcbi.1000667.

Yang, Z., Mammel, M., Papafragkou, E., Hida, K., Elkins, C. A. and Kulka, M. (2017) 'Application of next generation sequencing toward sensitive detection of enteric viruses isolated from celery samples as an example of produce', *International Journal of Food Microbiology*. Elsevier, 261(July), pp. 73–81. doi: 10.1016/j.ijfoodmicro.2017.07.021.

Yeargin, T., Buckley, D., Fraser, A. and Jiang, X. (2016) 'The survival and inactivation of enteric viruses on soft surfaces: A systematic review of the literature', *American Journal of Infection Control*. Elsevier Inc., 44(11), pp. 1365–1373. doi: 10.1016/j.ajic.2016.03.018.

Zhang, C.-M., Xu, L.-M., Xu, P.-C. and Wang, X. C. (2016) 'Elimination of viruses from domestic wastewater: requirements and technologies', *World Journal of Microbiology and Biotechnology*, 32(4), p. 69. doi: 10.1007/s11274-016-2018-3.

Zhao, X., Lin, C., Wang, J. and Oh, D. (2014) 'Advances in rapid detection methods for foodborne pathogens', *J. Microbiol. Biotechnol*, 24(3), pp. 297–312. doi: 10.4014/jmb.1310.10013.

Zhou, J., Wang, X. C., Ji, Z., Xu, L. and Yu, Z. (2015) 'Source identification of bacterial and viral pathogens and their survival/fading in the process of wastewater treatment, reclamation, and environmental reuse', *World journal of microbiology & biotechnology*, 31(1), pp. 109–120. doi: 10.1007/s11274-014-1770-5.

Manuscripts