University of Cincinnati

Date: 8/21/2014

I, RANJANI RAVI, hereby submit this original work as part of the requirements for the degree of Master of Science in Environmental Engineering.

It is entitled:
A NOVEL RNA VIRUS DETECTION SYSTEM BASED ON DUPLEX SPECIFIC NUCLEASE

Student’s name: RANJANI RAVI

This work and its defense approved by:

Committee chair: David Wendell, Ph.D.

Committee member: Drew McAvoy, Ph.D.

Committee member: George Sorial, Ph.D.
A Novel RNA Virus Detection System

Based on Duplex Specific Nuclease

A thesis submitted to the

Graduate School

of the University of Cincinnati

in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

In the department of Environmental Engineering

of the College of Engineering and Applied Sciences

2014

By

Ranjani Ravi

B.Tech Electronics and Communication

Amrita School of Engineering, Coimbatore, Tamil Nadu, India, 2011

Committee Chair: Dr. David Wendell
Abstract

The Norwalk viral agent or Norovirus is the most prevalent cause for non-bacterial gastroenteritis worldwide [1-4] and is named for the northern Ohio town where it was first identified [5]. Every year 21 million cases of acute gastroenteritis are caused by norovirus infection, leading to 70,000 hospitalizations and 800 deaths, mainly young children and the elderly [6]. According to the Centers for Disease Control and Prevention (CDC), 11% of deaths in the United States from food-borne illnesses were caused by norovirus in 2011 and outbreaks frequently occur in semi-closed communities such as schools, hospitals, retirement homes and cruise ships [7, 8, 9], producing the nickname “cruise ship virus”. A particularly virulent Australian strain has emerged as a global threat in 2012, readily passing between individuals, making Norovirus detection of paramount interest [10-11].

Efforts to study the impact of norovirus infection have been hindered by the difficulty of controlled laboratory cultivation. In addition, norovirus culture techniques are fraught with the risk of exposure to the real virus. However, non-culture based techniques such as Reverse-Transcription Polymerase Chain Reaction (RT-PCR) have confirmed norovirus involvement in epidemics and sporadic outbreaks of gastroenteritis. While RT-PCR has proven useful for detection it requires the essential enzymatic step of reverse transcription which increases the experimental cost, time and complexity. The detection of norovirus is also hindered by factors like inappropriate sample storage (light exposure), low virus concentrations and ineffective viral RNA extraction methods [12], particularly for surface waters. As a result there remains a need for an improved Norovirus assay, one which is faster, more cost effective and more reliable in environmental waters. The main goal of this study was to develop a method which is not
limited by the current drawbacks in detection of norovirus. An ideal system would enable quick and accurate detection without increase in cost. We have used the inherent specificity of the Duplex-Specific Nuclease (DSN) for cleaving single stranded DNA in a DNA-RNA duplex since norovirus contains a single stranded RNA genome. DSN, recently isolated from the Kamchatka (King) Crab [13], was combined with FRET (Forsters Resonance Energy Transfer) based DNA probes (Taqman™) that have been designed to target the most common norovirus RNA genogroups (I and II). The fundamental idea is to have the FRET probes bind to the viral RNA template to form a DNA: RNA duplex. DSN would then preferentially digest the DNA probe within the duplex, resulting in a fluorescent signal when the quencher and reporter dye are separated. As the RNA template usually endures this procedure, an amplified fluorescent signal will result proportional to the target, as multiple probe molecules are cleaved for individual RNA template.

The total reaction time for this strategy is less than 30 minutes, compared to the 4-6 hours required for quantitative RT-PCR [14], and can be achieved on a simple heat block, rather than a more costly thermocycler. The real-world efficacy of our assay was tested using environmental water (Little Miami River) and the results demonstrate a detection limit of 10 norovirus copies per assay, below the minimum concentration required for infection (18 viral copies) [15-17]. Norovirus is included in the EPA Contaminant Candidate List 3 (CCL3) [18], which means that presently it is not subject to any proposed national drinking water regulations but may require regulation in future under the Safe Drinking Water Act (SDWA).
Acknowledgement

I would like to express my heartfelt and sincere gratitude to my advisor, Dr. David Wendell who has provided invaluable support and guidance throughout my research. He has been a great source of inspiration and his timely advice and suggestions have been key motivating factors throughout the course of my Master’s program.

I would also like to thank my lab mates Vikram Kapoor and Beth Wurtlzer for helping me with the experiments and for all the technical discussions as a part of my research. I would also like to profusely thank Dr. Drew McAvoy and Dr. George Sorial for serving on my thesis committee.

I would like to thank Dr. Shay Fout (USEPA, Cincinnati, Ohio) who kindly donated us with norovirus RNA, which was very crucial for my research and also for agreeing to validate the assay. He has been a great support in the progress of the project and has constantly guided through his knowledge and wisdom in the field of virology.

I appreciate and thank the support I have received from my friends and colleagues at the University of Cincinnati.

I would lastly like to thank my parents and family back home. Their love and constant motivation has always helped me stay focused on my goals.
Table of Contents

Chapter 1 Introduction ............................................................................................................. 1
  1.1 Introduction to Norovirus ................................................................................................. 2
  1.2 Impact of Norovirus Outbreaks ....................................................................................... 3
  1.3 Viral Detection Methods ................................................................................................. 5
Chapter 2 Duplex Specific Nuclease ...................................................................................... 12
  2.1 DSN Origin .................................................................................................................... 12
  2.1 DSN Application and Uses ............................................................................................ 15
Chapter 3 Methods .................................................................................................................. 17
  3.1 Sampling Technique ..................................................................................................... 17
  3.2 RNA Oligonucleotides and DNA Probes ...................................................................... 17
  3.3 Environmental RNA Extraction ..................................................................................... 18
  3.4 DSN Based detection for RNA oligonucleotides .......................................................... 19
  3.5 Viral RNA Extraction .................................................................................................. 19
  3.6 Large Scale RNA Production ....................................................................................... 20
  3.7 Virus Concentration by Skimmed Milk Flocculation .................................................. 21
  3.8 Optimized DSN Reaction ............................................................................................. 21
  3.9 RT-PCR based detection .............................................................................................. 21
Chapter 4 Experimental Setup and Results .......................................................................... 23
  4.1 Primary Optimization .................................................................................................... 27
  4.2 Secondary Optimization ............................................................................................... 31
Chapter 5 Discussion and Conclusion .................................................................................. 42
  5.1 Future Work .................................................................................................................. 46
Bibliography .......................................................................................................................... 47
List of Tables

Table 3.1 Sequence of RNA Oligos and Taqman™ Probes .................................................. 18

Table 3.2 Primers for Large Scale RNA Production ................................................................ 20

Table 3.3 Primers and Probe Sequence for RT-PCR ................................................................. 22

Table 4.1 Effect of Select Additives on the DSN based Detection ............................................. 29

Table 5.1 Final Trials of Optimization ....................................................................................... 43

Table 5.2 Future RNA Virus Targets for DSN ......................................................................... 45
List of Figures

Figure 1. 1 Electron micrograph of Norovirus. Scale bar = 50 nanometers ....................... 2

Figure 1. 2 Taqman™ probe ........................................................................................................... 8

Figure 2.1 The native distribution of the Kamchatka Crab ...................... .................. ......12

Figure 2. 2 Paralithodes camtschaticus (Kamchatka Crab) .............................. ............................. 13

Figure 4.1 Schematic representation of norovirus RNA detection based on Duplex-Specific Nuclease .......................................................................................................................... 23

Figure 4.2 Plot of norovirus DSN assay with RNA oligos for (A) Norovirus Genogroup I and (B) Genogroup II. 10-fold serial dilutions of RNA ($10^4$ to $10^0$ genomic equivalents per assay) were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=3). ............................................................................................................................. 25

Figure 4.3 Plot of norovirus DSN assay with RiboMax RNA for Norovirus Genogroup I and ($10^4$ to $10^0$ genomic equivalents per assay) were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=3). ......................... 26

Figure 4.4 Scatter plot of relative fluorescence units (RFU) as a function of norovirus RNA copies for different concentrations of ATA. Error bars depict to standard deviation (n=3). .................................................................................................................................................. 30

Figure 4.5 3D bar chart of ATA variance. 10-fold serial dilutions of RNA ($10^4$ to $10^0$ genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying ATA concentration. .......................................................................................................................... 30

Figure 4.6 3D surface plot of ATA variance. 10-fold serial dilutions of RNA ($10^4$ to $10^0$ genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying ATA concentration. .......................................................................................................................... 31
Figure 4.7 Plot of norovirus DSN assay with 950 bp norovirus RNA. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) of Norovirus GI were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=3).

Figure 4.8 Scatter plot of relative fluorescence units (RFU) as a function of norovirus RNA copies for different probe concentrations. Error bars depict to standard deviation (n=3).

Figure 4.9 Scatter plot of relative fluorescence units (RFU) as a function of norovirus RNA copies for varying sheared DNA copies. Error bars depict to standard deviation (n=3).

Figure 4.10 3D bar chart on Impact of Sheared DNA. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying sheared DNA concentration.

Figure 4.11 3D surface chart on Impact of Sheared DNA. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying sheared DNA concentration.

Figure 4.12 Scatter plot of relative fluorescence units (RFU) as a function of norovirus RNA copies for varying Mg concentration. Error bars depict to standard deviation (n=3).

Figure 4.13 3D bar chart of Mg Variance. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying Mg concentration.

Figure 4.14 3D Surface plot on Mg Variance. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying Mg concentration.

Figure 4.15 Plot of norovirus DSN assay with 950 bp norovirus RNA. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) of Norovirus GI were plotted...
versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=3).

Figure 4.16 (A) Plot of norovirus DSN assay with RNA extracted from water spiked with norovirus particles. 10-fold serial dilutions of RNA ($10^4$ to $10^0$ genomic equivalents per assay) of Norovirus GI were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=6). (B) Plot of norovirus RT-PCR with RNA extracted from water spiked with norovirus particles. 10-fold serial dilutions of RNA ($10^4$ to $10^0$ genomic equivalents per assay) of Norovirus GI were plotted versus $1/C_t$ value. Error bars show standard deviations (n=3).

Figure 4.17 Plot of norovirus DSN assay with RNA extracted from water spiked with norovirus particles. 10-fold serial dilutions of RNA ($10^4$ to $10^0$ genomic equivalents per assay) of Norovirus GI were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=6).
Acronyms and Abbreviations

μg microgram
μl microliter
ml milliliter
ng nanogram
ATA Aurintricarboxylic Acid
CCL3 Contaminant Candidate List 3
CDC Centers for Disease Control and Prevention
CHPV Chandipura Virus
DMSO Dimethyl sulfoxide
DNA deoxyribonucleic acid
DSN Duplex Specific Nuclease
EPA Environmental Protection Agency
FRET Forsters Resonance Energy Transfer
NFQ Non-fluorescent quencher
NUC Nonspecific endonuclease
PCR polymerase chain reaction
PFU plaque-forming units
qPCR Quantitative Polymerase Chain Reaction
RNA ribonucleic acid
RT-PCR Reverse-Transcription Polymerase Chain Reaction
RT-qPCR Reverse Transcriptase Quantitative Polymerase Chain Reaction
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARS</td>
<td>Severe Acute Respiratory Syndrome</td>
</tr>
<tr>
<td>SDWA</td>
<td>Safe Drinking Water Act</td>
</tr>
<tr>
<td>SM</td>
<td>skimmed milk</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction

The majority of waterborne diseases are caused by pathogenic viruses and are the primary reason for mortality worldwide [19-22]. The most prevalent waterborne viruses include the RNA-based calicivirus, astrovirus, human immunodeficiency virus (HIV), influenza, chikungunya virus and rotavirus. These viruses are the main source of infection in water used for drinking and recreation [12, 24-26], causing diseases such as gastroenteritis, hepatitis, flu and respiratory illnesses. Norovirus is an RNA virus belonging to the Caliciviridae family and is the major cause for gastroenteritis world over [1-4]. Caliciviridae are a family of single stranded, positive sense RNA viruses that are 35-40nm in diameter and have an icosahedral symmetry. The calciviruses infect humans and are classified into four different genera namely Norovirus, Sapovirus, Vesivirus and Lagovirus on the basis of sequence diversity [27].

In the Norovirus subgroup five genogroups (GI-V) have been recognized based on geographical area where they were first identified. Genogroup I consists of the Norwalk, Southampton, Desert Shield, and Chiba virus. Genogroup II includes Lordsdale, Grimsby virus, Mexico virus and the Snow Mountain agent [28]. Genogroups I, II and IV infect humans, while GIII and GV infect bovine and murine species respectively [15]. Norovirus has a relatively simple, single stranded RNA genome which is approximately 7.7 kb with three open reading frames [29], encoding for structural proteins which self-assemble around the genomic RNA, to form an icosahedral capsid [30-31]. Sapoviruses and Noroviruses are known to cause gastroenteritis in young children and adults and spread via fecal-oral-route. It has been reported that children develop a long lasting immunity
following sapovirus infection whereas both younger and older adults are at risk of norovirus infection throughout life. The vesiviruses and lagoviruses primarily infect animal hosts and pose a lesser risk to human beings.

1.1 Introduction to Norovirus

Norovirus is the most prevalent cause for non-bacterial gastroenteritis worldwide [1-4]. It was first identified in 1968 when an outbreak occurred among elementary school children in Norwalk, Ohio [5].

Norovirus is extremely contagious and is transmitted directly via fecal-oral-route, as well as via air borne virons resulting from vomiting and sneezing [33, 17]. Norovirus also spreads through food [34] and survives over a wide range of temperatures (0 °C to 60 °C), thriving in recreational and drinking water leading to large community outbreaks [24-26].

Norovirus spreads easily as it requires low number of viral copies (18-1000) to cause infection [15-17] and a short incubation period (12-48 hrs). Once ingested, it enters the human

Figure 1.1 Electron micrograph of Norovirus. Scale bar = 50 nanometers [32]
body and attaches itself to the mucosal epithelium of the small intestine, where it starts to replicate and causes gastroenteritis [17]. Noroviruses infect humans by binding to specific histo-blood group antigens which are complex carbohydrate structures expressed on gastrointestinal epithelial cells [31]. Upon infection by norovirus, these cells get damaged which reduces the enzymatic activity of the small intestine, leading to malabsorption of carbohydrates and excessive discharge of fats in feces [37, 15]. There is also a delay in the movement of food through the digestive tract, causing reduced gastric motility which may cause symptoms such as diarrhea and vomiting. Inactivation of the virus begins once the immune system of the host releases antibodies. However, the virus mutates frequently during replication allowing it to evade the defense system of the host.

The common symptoms of norovirus infection are vomiting, watery diarrhea, nausea and fever. The age groups which are mainly affected by norovirus infection are the pediatric and geriatric populations. Symptoms become perceptible within a few days of norovirus ingestion, with dehydration the most common symptom leading to hospitalization and even death [38-39]. According to the CDC from 1999-2007, norovirus infection was responsible for 797 deaths annually in the U.S.[40]. The cost incurred in direct health care due to hospitalizations caused by norovirus infection in U.S. is around $493 million.

1.2 Impact of Norovirus Outbreaks

Norovirus outbreak in cruise ships has become prevalent recently, producing the nickname “cruise ship virus”. The cruise ship Royal Caribbean's Explorer of the Seas Liner was forced to return earlier than the schedule date of arrival after nearly 700 passengers on board developed gastrointestinal illnesses. The restricted spaces of cruise ships are easy targets for infections like norovirus, which easily spread through unhygienic food or water [40, 41].
Around the same time frame, another severe outbreak of norovirus infection affected nearly 200 passengers and crew members of the Princess Cruises' vessel that returned to Houston after cutting short its journey [41].

The above events have led to stricter enforcement of Vessel Sanitation Program (VSP) initiated by CDC, in an effort to prevent and curb the transmission and spread of acute gastrointestinal infection on cruise ships. The VSP provides comprehensive sanitation program and required course of action to mitigate the risk of a widespread illness on cruise ships. According to their guidelines any ship traveling to a foreign land and carrying more than 13 passengers will have to undergo bi-annual inspection and re-inspection if required.

The cruise ship companies that are a part of the VSP have to submit a detailed report indicating the number of illness cases caused by the Norovirus (GI). The report will be examined by the medical personnel before the ships embark a U.S. point of entry, especially when the vessel has sailed from a foreign destination. In case, the outbreak is caused by GII strain of norovirus and exceeds 2% of the total number of travelers, then a separate document has to be presented detailing the extent of infection.

The objectives of the VSP plan are:

a) To inspect cruise ships on a periodic basis and also perform simultaneous surprise checks to ensure that the vessels meet all the sanitary requirements.

b) Tracking gastrointestinal infections and responding to epidemics. Also, training the crew ship employees and staff on public health practices.

c) Offering up to date public health information and preventive measures to the cruise ship industry, the passengers and to the state and local health authorities in case of an outbreak [42].

Presently, there are no vaccines for preventing norovirus, thus treatment must be taken in the form of replacing body fluids and other palliative care [43–44]. One of the best courses of action
remains to prevent infection by early detection, and disinfection or avoidance of contaminated water. Thus there is a need for a sensitive and rapid virus detection method which can be used as an early warning system for infection prevention.

An ideal virus detection system for water borne disease would enable quick and accurate warning of the most common human pathogens. The technique should be inexpensive and fast so that it can be used in developing nations where large populations and underdeveloped treatment infrastructure enable the virus to spread quickly. The virus detection system should work as an early warning system for water quality administration organizations and assist them in taking rapid measures to prevent pandemics.

1.3 Viral Detection Methods

There are numerous techniques used in waste water treatment plants for virus detection. At present the most widely used methods for concentration and discovery involves adsorption followed by elution [45], using electropositive filters [45-46], glass wool columns [47], direct flocculation [48] and ultra filtration [46]. Membrane filtration is known to be effective for obtaining viruses from substantially large volumes of waste water [49-50]; however, efficacy in turbid water is limited by filter clogging.

Adsorption followed by elution is a frequently used method for virus recovery. The most widely used filters include either electropositive or electronegative filters. The positively charged filters are mainly used for enteric viruses as they are negatively charged at surrounding pH [51]. Also, electropositive filters have low virus recovery rate in sea water as they get easily clogged due to the high salt content and alkalinity of marine water [51]. The use of negatively charged filter for waters with high turbidity has shown higher rates of virus recovery than positively charged filters [52]. Conversely, negatively charged viruses will adsorb to electronegative membrane only in the
presence of cations which will make their net charge positive. Katyama et al (2002), has developed a method based on adsorption followed by elution using electronegative filter for concentrating poliovirus from sea water. Their results showed a recovery of 73% poliovirus from 1 liter of sea water [49].

The use of glass wool for concentrating viruses is a promising method as it is cost effective and can be used for large volumes of water. Lambertini et al [47] has described the method for concentrating viruses from drinking water using glass wool. Glass wool coated with mineral oil was used for the experiments which exposed electropositive sites on the surface [47]. The virus particles get entrapped in the wool fibers when the sample is passed through the glass wool columns. The drawback of this method is that the efficiency of virus recovery is highly variable, 21% for adenovirus, 29% for norovirus and up to 70% for poliovirus illustrating that recovery is affected by the strain of the virus [47]. However, the virus recovery efficiency using glass wool is comparable to other standard methods and is available at a minimal cost thus making this method highly desirable [47]. Direct flocculation method is used for virus detection when the samples have high organic load. The organic matter is allowed to settle via gravity and concentrated using flocculation methods. Calgua et al (2008), has described a cost effective and efficient procedure for concentrating adenovirus from sea water using skimmed milk. In this technique, viruses get adsorbed to pre-flocculated acidified skimmed milk proteins and form flocs [48]. The flocs are allowed to settle and are separated, followed by detection for viruses using PCR techniques [48]. The method does not require any special apparatus and can be used for treating large volumes of water sample. The mean virus recovery was 42-49% in sea water as compared to the glass wool columns which yielded a recovery of 0.77% [48].
Ultra filtration (UF) is a proven method, more effective than the conventional methods for water treatment which include clarification followed by disinfection [53]. UF is employed mainly for removal of colloidal particles in the size range of 0.001 micron-1.0 micron. The pore size in a UF membrane plays a major role in determining the size and category of contaminant to be removed. The primary advantages of UF are the absence of any chemicals such as coagulants, disinfectants or flocculates for water treatment. The type of contaminant to be removed can be manipulated depending on the membrane pore size. UF is an attractive procedure for removal of virus from water because of its high throughput and the technique can be carried out at ambient temperature with minimum damage to viral proteins [53]. These concentration methods discussed above yield the environmental filtrate which can be used for nucleic acid based viral detection.

The environmental filtrate is analyzed for the presence of specific nucleic acid sequences via an amplification technique called Polymerase Chain Reaction (PCR). PCR is used to identify pathogenic organisms by detecting particular genes in them. The major steps involved in PCR detection requires extraction of DNA from water samples to be tested. The extracted DNA is amplified using short gene specific primers and the amplified product is visualized using gel electrophoresis. PCR has been used widely for detection of viruses in groundwater, drinking water and surface waters. Abbaszadegan et al [54] has reported use of PCR for detection of enteric viruses in groundwater. In his study, 400 gallons of water from different sources was utilized and used for concentrating viruses by a filter adsorption and elution method. The concentrated virus obtained after filtration was used in PCR to screen for pathogens. PCR has various advantages as compared to the traditional cell culture techniques. The time required for PCR is a few hours compared to the several days required for obtaining results from culture
based techniques. Also the initial and subsequent cost involved in PCR is much lesser than culture techniques. While traditional PCR is generally binary for the presence/absence of the virus of interest, quantitative PCR (qPCR) techniques can quantify contamination level [4, 55-56]. When the intended target is RNA based, reverse transcription is required to translate the viral genome into cDNA for the DNA-dependent polymerase used in the PCR reaction. This technique, Reverse Transcriptase qPCR (RT-qPCR), uses a fluorescent signal to report how much of the viral target RNA is present.

In the Taqman™ qPCR method an oligonucleotide probe is used which consists of a fluorescent dye attached to a non-fluorescent quencher (NFQ) [57-58].

The quencher prevents the reporter dye from fluorescing until it is used in the PCR reaction. The selection of reporter and quencher pair is based on their emission and absorption spectral overlap, with the ideal quencher dissipating the reporter’s excitation photons as heat. FRET probes may also include a DNA minor groove binder (MGB) for additional specificity [4].

During conventional qPCR, the polymerase digests the Taqman™ probe releasing the reporter dye fluorescence in proportion to the amount of viral target being synthesized. The PCR cycle in which the fluorescent signal is above the background level is known as the threshold
crossover or Ct value [59]. If a higher number of DNA copies are available at the beginning of
the reaction, then less amplification cycles are needed to reach Ct.

RT-qPCR is not used regularly for virus detection in waste water treatment plants
because it involves expensive equipment, significant time/labor cost, and skilled technicians to
optimize and analyze the results [60-61]. The factors which limit the efficiency of RT-qPCR are
quality of the extracted RNA, effectiveness of the reverse transcriptase conversion step, and
inconsistencies in environmental PCR inhibitors within the viral concentrate. Environmental
inhibitors can include humic acids, divalent cations, and polysaccharides that copurify during the
concentration and purifications steps [62]. Although RT-qPCR is accurate for testing relatively
high viral concentration it can be unreliable for samples with very low viral counts, particularly
when inhibitors are present.

Serological testing and Virus Assays are additional methods used for norovirus
detection. In serological testing, a patient’s blood sample is subject to an antibody screening, a
method generally reserved for outbreaks [63-64]. The primary drawbacks to serological testing
are the length of time required for results (at least 24 hours) and the material costs.

In the plaque assay, norovirus attaches itself to a monolayer of macrophage cells and is
then covered by a culture media and agarose [65]. The spread of newly formed viral copies to
neighboring cells is facilitated by diffusion, and can be observed by holes formed in the
monolayer from lysed cells [65]. Based on the number of cell clearings, or plaques, the number
of initially infected cells can be determined, a metric commonly referred to as plaque-forming
units (PFU) [66]. This method poses serious safety concerns of exposure to real virus to qualify
as a primary method of detection.
EPA Method 1615 [67] was conceptualized by USEPA for detecting enterovirus, human norovirus and mammalian orthoreovirus occurrence in water by culture and RT-qPCR techniques. The method combines both conventional cell culture with latest molecular procedures that facilitates rapid testing of large samples of groundwater for the aforementioned viruses. Orthoreoviruses belong to the *Reoviridae* family and are known to infect plants, vertebrates and invertebrates. They possess double stranded RNA genomes and have a non enveloped viral capsid 80-82nm in diameter. The cell culture procedure is used only for enterovirus and orthoreovirus which can infect a special cell line, the Buffalo Green monkey kidney (BGM) and produce cytopathic effects. There is no ascertained cell line for human norovirus detection. But the molecular procedure integrated into 1615 method is effective in the detection of norovirus and enterovirus.

As a first step of EPA Method 1615, viruses are concentrated by passing water samples through electropositive filters and then are eluted using a beef extract reagent and organic flocculation. The concentrated eluate obtained is used to inoculate BGM cells, which are observed for the development of cytopathic effects for couple of weeks and verified by transferring to fresh cultures. Virus concentration is determined using most probable number (MPN) method. RNA is extracted from the concentrated eluate and analyzed for the presence of norovirus or enterovirus using RT-qPCR. Virus concentration is determined based on the standard curve obtained in terms of genomic equivalents of viral RNA per liter.

A recent study by Straub et al [69] shows the possibility of having a suitable model to learn and understand the cytopathic effects of norovirus infection on human beings. Cell culture models for norovirus infecting murine species have been reported in the past [68]. However, pathogenic pathways for norovirus infecting humans, has not been well documented due to the
difficulty of controlled laboratory cultivation. Straub et al [69] developed a 3D model of the human small intestine, which is the chief target for viral attachment to the host epithelial cells. Research has also indicated the presence of a specific histo-blood group antigen [31] that is crucial in facilitating the attachment. The model was achieved by cultivating cells on microbeads coated porous collagen. To replicate the human gastrointestinal epithelium the cells were grown in a biological fluid environment which was contained in a rotating bioreactor. The results have shown vacuole formation and shortening of the microvilli in the infected tissues 24 hours post infection. After 48 hours the infected cells were observed to be elongated and disfigured. Till date, there has not been a clear estimate to show the effectiveness of our prevention methods against norovirus. The studies that have been carried out in the past to inactivate or kill the virus or to mitigate contamination have always involved a proxy. Hence, this study has helped in knowing the norovirus pathogenesis cycle in humans.

Virus detection methods, especially for norovirus, have been hindered by the difficulty of controlled laboratory cultivation [68, 70]. In addition, norovirus culture techniques are fraught with the risk of exposure to the real virus. However, non-culture based techniques such as Reverse-Transcription Polymerase Chain Reaction (RT-PCR) have confirmed norovirus involvement in epidemics and sporadic outbreaks of gastroenteritis [9, 12]. While RT-PCR has proven useful for detection it requires the essential enzymatic step of reverse transcription which increases the experimental cost, time and complexity. The detection of norovirus is also hindered by factors like inappropriate sample storage (light exposure), low virus concentrations and ineffective viral RNA extraction methods [5], particularly for surface waters. Clearly, there is a need for an improved detection method, one which is faster, more cost effective and more reliable in environmental waters.
Chapter 2 Duplex Specific Nuclease

Nucleases, a primary requirement for cellular development, play a crucial role in various biological processes such as DNA replication, mismatch repair, defense and digestion. They also facilitate strict conservation of genetic material and protect DNA from environmental stresses, carcinogens, reactive species and other factors that modify DNA. The occurrence of any defect in these nucleases can lead to severe genetic mutation and instability in the organism. Being small motifs possessing unique properties, they are of vital importance to the evolution of novel microbiological assays.

2.1 DSN Origin

In our assay, we have efficiently utilized one such nuclease called DSN, first identified from hepatopancreas of the Red King (*Kamchatka*) Crab (*Paralithodes camtschaticus*) [13]. The Red King Crab although native to the Bering, Okhotsk and Japan Sea can also be found along the northern regions of the Pacific Ocean, as indicated in yellow marking in Figure: 2.1.

![The native distribution of the Kamchatka Crab](image)

**Figure 2.1 The native distribution of the Kamchatka Crab [72]**
The Kamchatka crabs possess a dark red/brownish body and are one of the largest arthropods, measuring up to 22cm in length and weighing over 22 pounds [71-73].

![Image of Kamchatka Crab](image)

**Figure 2. 2 Paralithodes camtschaticus (Kamchatka Crab) (Source: internet)**

The unique DSN nuclease, found in the Kamchatka crab has found great potential for use in many novel assays. The DSN nuclease is very closely associated with the *Serratia* family of non specific nucleases. The discovery of DSN was incited by the examination of the novel nuclease secreted by the kuruma shrimp. A detailed study further revealed the existence of a similarity between the nuclease secreted from the shrimp heptopancreas and *Serratia*. [74].

DSN from Kamchatka crab was found to be homologous to kuruma shrimp nuclease bearing 82 % similarity in structure and demonstrating around 67% comparable amino acid homology [13]. Also, similar to the shrimp nuclease the DSN sequence was found to have the DNA-RNA nonspecific endonuclease (NUC) domain. The NUC domain is a conserved
characteristic of *Serratia* DNA/RNA endonucleases. These endonucleases hydrolyze RNA as well as single and double stranded DNA and require the presence of a cation for their effective functioning [75]. Despite possessing the NUC domain, the specificity of DSN is distinctive as it cleaves only dsDNA or DNA in DNA/ RNA hybrid and is relatively inactive in the presence of a single stranded DNA or double stranded RNA unlike other members of the *Serratia* family [13]. This specificity is believed to stem from the structural dissimilarity observed between DSN and the shrimp nuclease.

Sequence analysis of DSN reported by Anisimova et al (2008), confirmed a nucleotide sequence, 1221 bases long that encodes for a 407 amino acid protein. The protein had a molecular mass of 41.5kD and was heat stable over a wide range of temperatures with an optimum between 58°C-60°C [13]. DSN was also observed to be stable over a wide range of pH values (4-12) with an ideal working pH of 6.6. The nuclease requires a minimum of ten pairs of complementary nucleotides and a maximum of twenty five pairs to generate significant activity [13,76], as well as the presence of divalent cations like Mn^{2+}, Co^{2+}, or Mg^{2+}, with a minimum of 5mM Mg necessary for cleavage. This activity is further enhanced by the presence of cations in addition to CaCl$_2$. The ion concentration of up to 20 mM is found to improve the DSN activity drastically, whereas higher concentrations only serve to restrain it [13]. Similarly the optimum temperature found for maximum DSN activity was 60° C and further increase caused a sharp decline. The presence of denaturing agents like mercaptoethanol, NaCl and SDS reagents leads to decrease in specific DSN activity while the presence of EDTA completely inhibits it [13].
2.1 DSN Application and Uses

This unique nuclease activity of DSN has been employed as a molecular tool for micro RNA detection [77], full length cDNA library normalization [78] and single nucleotide polymorphism (SNP) detection [76], and in this work we extend detection to RNA viruses.

Yin et al [77] has described the use of DSN as a means of a signal amplifying mechanism to augment the use of traditional Taqman probes. DSN nuclease is utilized efficiently to aid the process of target directed cleaving of Taqman probes to produce an amplified fluorescent signal proportional to the target, as multiple probe molecules are cleaved for each target template. He has further illustrated the efficacy of this method by employing it for quantification of several miRNAs in biological samples. Shagin et al [76] has described a useful method for SNP detection using the specificity of DSN. SNP is a single genetic variation in nucleotide sequences constituting DNA. SNP variance may be associated with fatal ailments such as diabetes, cancer and coronary arterial diseases [76]. SNP detection plays a key role in studies related to genetic disorders, population studies and microbial source tracking. SNPs can also be helpful in determining a person’s reaction to specific drugs, immunity towards toxins and susceptibility to contract certain diseases. The ability of DSN to distinguish between complementary sequences (8-10bp) and to effectively report the presence of a single mismatch was used to design the novel assay. Duplexes were created in the form of targets (containing one mismatch) and their complementary FRET based probes. The results showed a drastic decrease in fluorescence for all mismatched hybrids. The decrease in fluorescence intensity was at least 10 fold times compared to the perfectly matched duplexes [76].

Bogdanova et al [79] has devised a method for normalization of full length enriched cDNA using DSN specificity. cDNA normalization has always been hindered by the presence of excessive transcripts and their considerable variation within cells. This method is based on
degrading double stranded DNA strands that form a great fraction of the transcripts to be queried. The normalization using DSN does not involve cumbersome physical separation step. Standardization of cDNA library with the help of DSN helps in bringing the rate of occurrence of each transcript within an acceptable range that can be analyzed easily. Yi et al [80] has further extended the application of DSN for full length library normalization of bacterial transcripts. Prokaryotes, such as bacteria do not have a distinguished mechanism for rRNA selection over mRNA in the total RNA present. Hence for mRNA enrichment, various methods like synthetic microbial hybridization and exonuclease digestion have been conventionally used. However, since mRNA constitutes only 1-5% of total RNA, Yi et al has implemented removal of non-mRNA constituents for mRNA enrichment. DSN based removal technique achieves efficient removal of rRNA component while ensuring adequate enrichment of mRNA. Since the binding preference of DNA varies as a square of its concentration [78, 80], the abundant rRNA after conversion to cDNA is readily cleaved by the DSN enzyme during renaturation [80].
Chapter 3  Methods

3.1 Sampling Technique

Water samples were obtained from the Little Miami River (+39° 7’ 37.49", -84° 24’ 36.39"), using sterile 1000ml Nalgene sample bottles in accordance with Environmental Protection Agency (EPA) sample collection protocol (EPA Method 1669). The bottles were labeled with date and time of sampling. The samples were immediately placed in a lightproof insulated box containing ice packs to ensure rapid cooling and delivered to the lab within 6 hours of sampling. Cold and dark conditions were maintained throughout transportation to the lab.

3.2 RNA Oligonucleotides and DNA Probes

Short RNA oligonucleotides and probes corresponding to the norovirus genogroups GI and GII were synthesized (Operon, USA) for the initial stage of this study. The probe sequences were obtained from Hoehne et al [12] with required modifications to ensure the absence of secondary structure formation. This was achieved using RNAfold, an online tool which predicts secondary structures of single stranded RNA and DNA [81]. The software provides a map of all possible secondary structures for a given RNA or DNA sequence input. Using an iterative design process, linear sequences without secondary structures were identified (Table 3.1).

![RNA Probe Diagram]

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norovirus GI Taqman™ probe</td>
<td>Cy3-CCTTAGACGCCATATCA-BHQ</td>
</tr>
</tbody>
</table>

Figure 3.1 Optimization technique using RNAfold software
The above figure shows the optimization technique used for obtaining modified probe sequences. The probe sequence has been engineered to avoid the possibility of self-hybridization and eliminating any possible secondary structure that could be formed.

The lengths of the probes were designed to meet a melting temperature of 60°C, the optimum working temperature for the DSN enzyme. The GI probe was labeled at 5'-end with the Cyanine3 (Cy3 fluorophore) and at the 3'-end with BHQ (Black Hole Quencher). The GII probe contained 6-FAM (6-Fluorescein amidite) and BHQ at the same respective locations.

**Table 3.1 Sequence of RNA Oligos and Taqman™ Probes used in this work**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’-3’)</th>
<th>Location*</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVGI RNA</td>
<td>AAAUGAUGAUGGCGUCAAGGACG</td>
<td>5355-5379</td>
</tr>
<tr>
<td>NVGI probe</td>
<td>Cy3-CCTTAGACGCCATCATCA-BHQ</td>
<td>5376-5359</td>
</tr>
<tr>
<td>NVGII RNA</td>
<td>UGUGAAUGAAGAUGGCGUCGAAUG</td>
<td>5080-5103</td>
</tr>
<tr>
<td>NVGII probe</td>
<td>6-FAM-TCGACGCCATCTTCATT-C-BHQ</td>
<td>5100-5083</td>
</tr>
</tbody>
</table>

Cy3: Cyanine3   BHQ: Black Hole Quencher   6-FAM: 6-Fluorescein amidite

* Genome location of sequences are based for GI on the sequence of Norwalk/68/US [GenBank: M87661] and for GII on the sequence of Lordsdale/93/UK [GenBank: X86557]

### 3.3 Environmental RNA Extraction

Environmental RNA was extracted from water samples obtained from Little Miami River, Ohio (+39° 7' 37.49", -84° 24' 36.39”). All samples were collected using sterilized 1L bottles (Nalgene, Rochester, NY), and transported on ice to the laboratory at the University of Cincinnati within 2 hours. The collected water samples were filtered through 0.45-μm-pore-size
filters and stored at -80°C until RNA extraction. RNA was extracted from filter samples using the PowerSoil™ RNA Isolation Kit (MO BIO Laboratories, CA) according to the manufacturer's protocol. RNA purity and concentration was obtained using NanoDrop Spectrophotometer. RNA extracts were stored at -140°C for subsequent steps.

3.4 DSN Based detection for RNA oligonucleotides

For the DSN assay, a 30µl reaction mixture containing 1×DSN buffer A (50mM Tris-HCl, pH 8.0; 7mM MgCl₂, 1mM DTT), 0.1 U DSN enzyme (dissolved in 25 mM Tris-HCl, pH 8.0; 50% glycerol), RNA mixture and complimentary probe (50 nM) was prepared. The reaction mixture was heated at 60°C for 25 min. Subsequently, 30µl 10 mM Ethylenediaminetetraacetic acid (EDTA) (stop solution) was added to the reaction mixture and incubated at 60°C for 5 min to inactivate the DSN enzyme and end point fluorescence was recorded. 10–10,000 copies per 30 µl reaction with additional no template controls (NTCs; zero copies) were examined in triplicates. Dilution series for each of the synthetic RNAs were made in either RNase-free and DNase-free H₂O or environmental RNA solution to provide a constant background of extraneous RNA.

3.5 Viral RNA Extraction

Norovirus GI.1 RNA was kindly donated by Dr. Shay Fout (USEPA, Cincinnati, Ohio). The strain has been isolated from a human stool sample and contained 10⁶ virons/ml. Purification of viral RNA was carried out using QIAamp™ Viral RNA kit (Qiagen, USA) according to manufacturer's protocol. The RNA purity and concentration was obtained using NanoDrop Spectrophotometer (Thermo Scientific, DE). RNA extracts were stored at -140°C in the presence of 1U/µL RNase inhibitor (Promega, USA).
3.6 Large Scale RNA Production

Additional RNA viral target was generated from a DNA PCR template using the RiboMAX Large Scale RNA Production System (Promega, USA). To obtain the DNA template, extracted viral RNA was amplified with primers flanking a 950 bp genomic target using One step RT-PCR kit (Qiagen, USA). The total reaction mix was 25 µl containing 50 ng RNA template and 0.6 µM of each forward and reverse primers. Thermal cycling was performed as follows: reverse transcription for 30 min at 50°C, PCR inactivation of the reverse transcriptase and the activation of the polymerase for 15 min at 95°C, 40 cycles of denaturation at 94°C for 30s, annealing at 55°C for 30 seconds and extension at 72°C for 60s. The 950 bp DNA product obtained was used in a second PCR reaction to attach a T7 promoter on the 5’ end of the amplicon to create a linear template suitable for in-vitro transcription. The T7 PCR DNA template was used to create the 950 bp norovirus genome RNA replica by incubating the following reaction at 37°C for 2 hours.

Table 3.2 Primers for Large Scale RNA Production

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
<th>Location*</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVGI950-Fwd</td>
<td>ATGGTGATGATGAGATTGTG</td>
<td>4866-4885</td>
</tr>
<tr>
<td>NVGI950-Rev</td>
<td>CTAACATCAGCAATCACATG</td>
<td>5815-5796</td>
</tr>
<tr>
<td>NVGI950-Fwd_T7 (with T7 promoter)</td>
<td>CAATTCCCCTCTAGTAATACGACTCACTATAGGGAGAATGGTG ATGATGAGATTGTG</td>
<td>4866-4885</td>
</tr>
</tbody>
</table>

*Genome location of primers are based on the sequence of Norwalk/68/US [GenBank: M87661]
3.7 Virus Concentration by Skimmed Milk Flocculation

Water samples (Little Miami River) were spiked with norovirus particles ranging from $10^4$ to $10^1$ copies per liter of water. Skim milk solution (1% (w/v)) was prepared by dissolving 1 g of milk powder (Difco) in 100 ml of water and the pH was set to 3.5 with 1N HCl. The spiked water samples were acidified (pH 3.5) and 10 ml of the prepared skimmed milk solution was added to each 1 L sample. The environmental samples were then stirred for 8 hours at room temperature and flocs were allowed to settle by gravity overnight. The supernatant was decanted to avoid unsettling the sediment. The final sediment volume (approx. 100 ml) was centrifuged at 7000×g for 30 min at 12°C. The supernatant was poured out and the pellet resuspended in 10 ml of 0.2 M phosphate buffer at pH 7.5 (1:2, v/v of Na$_2$HPO$_4$ 0.2 M and NaH$_2$PO$_4$ 0.2 M). The concentrate was used for viral RNA extraction using QIAamp™ Viral RNA kit (Qiagen, USA) according to manufacturer's protocol.

3.8 Optimized DSN Reaction

A volume of 30 μL reaction mixture containing 1×DSN buffer B (50 mM Tris-HCl, pH 8.0; 10 mM MgCl$_2$, 1 mM DTT), 0.1U DSN (dissolved in 25 mM Tris-HCl, pH 8.0; 50% glycerol), 0.45 mg/ml ATA, sheared dsDNA (10 copies), 10 nM probe and RNA template, was incubated in a thermal cycler at 60°C for 25 min. Subsequently, 30μL10 mM EDTA was added to the reaction mixture and incubated at 60°C for 5 min to inactive DSN enzyme. No template RNA was used as controls to test for cross contamination.

3.9 RT-PCR based detection

For RT-PCR assay, water RNA extracts (5 μL) were used as templates in a final reaction volume of 25 μL using the Onestep RT-PCR kit (Qiagen, USA) with 200 nM each of the forward and
reverse primer and 100 nM probe. The samples were arranged in MicroAmp Optical 96-well plates and amplified in a 7500 Real-Time PCR System (Applied Biosystems, NY). The conditions used for amplification are as follows: reverse transcription for 30 min at 50°C, PCR inactivation of the reverse transcriptase and the activation of the polymerase for 15 min at 95°C, 40 cycles of denaturation at 94°C for 30s, annealing at 55°C for 30 seconds and extension at 72°C for 60s. (signal collection step). No-template controls were used to check for cross contamination, and 10-fold dilutions of each RNA extract were used to test for PCR inhibition.

Table 3.3 Primers and Probe Sequence for RT-PCR

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
<th>Location*</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVG1-Fwd</td>
<td>GCCATGTTCCGCTGGATGC</td>
<td>5282-5300</td>
</tr>
<tr>
<td>NVG1-Rev</td>
<td>CGTCCTTAGACGCCATCATCA</td>
<td>5379-5359</td>
</tr>
<tr>
<td>G1-MGB Probe</td>
<td>6-FAM-TGGACAGGAGATCGC-MGB-NFQ</td>
<td>5345-5359</td>
</tr>
</tbody>
</table>

*Genome location of primers and probes are based on the sequence of Norwalk/68/US [GenBank: M87661]

6-FAM= 6-Fluorescein amidite; MGB= minor groove binder; NFQ= non fluorescent quencher
Chapter 4 Experimental Setup and Results

In our isothermal assay, DSN is used to preferentially cleave single stranded DNA in a DNA-RNA duplex. The basic strategy is to have the Taqman-based DNA probes hybridize with the viral RNA template to form a DNA:RNA hybrid and allow DSN to specifically digest the DNA probe within the duplex, thereby generating a fluorescent signal upon the separation of quencher and reporter dye. Since the RNA template survives this process, an amplified fluorescent signal will result proportional to the target, as multiple probe molecules are cleaved for individual RNA templates. The strategy of DSN assay has been summarized in the Figure 4.1

Figure 4.1 Schematic representation of norovirus RNA detection based on Duplex-Specific Nuclease
This unique nuclease activity of DSN has been employed for other applications such as micro RNA detection [77], full length cDNA library normalization [78] and single nucleotide polymorphism detection [76], and in this work we extend this to detection of RNA viruses.

For the initial experiments, we used serial dilutions of synthetic oligoribonucleotides representing norovirus genogroup I (GI) and II (GII). Furthermore, we prepared each dilution series using, in parallel, environmental RNA extracted from river water (Little Miami River, Ohio: +39° 7' 37.49", -84° 24' 36.39"), to determine performance in the setting of pure template as well as in a complex background matrix. Through the addition of environmental RNA to the synthesized RNA, non-target RNA and natural enzymatic inhibitors were introduced which are expected to co-elute from natural samples. The detection of norovirus RNA was carried out by modifying a method described earlier for microRNA [77]. The RNA samples and the complementary Taqman probes were combined with DSN enzyme (Evrogen, Moscow, Russia) in a 30 µL reaction volume containing 7mM Mg\(^{2+}\) and incubated at 60°C for 25 min. Based on our probe melting temperatures, the incubation temperature was set at the optimum (60°C), which enabled hybridization while the probe is intact, but dissociation when it is cut, ensuing separation of RNA template from the probe.

A plot of fluorescent intensity against 10-fold serial dilutions of oligo RNA (10\(^4\) to 10\(^0\) genomic equivalents per assay) is shown in Figure 4.2 for Genogroup I (A) and Genogroup II (B).
Figure 4.2 Plot of norovirus DSN assay with RNA oligos for (A) Norovirus Genogroup I and (B) Genogroup II. 10-fold serial dilutions of RNA ($10^4$ to $10^0$ genomic equivalents per assay) were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=3).
Although, the use of short RNA template helped with the initial optimization of the assay, it has been shown that DSN has been ineffective in distinguishing targets which are longer than 18-25 bp [13]. To test the efficacy of our method on longer targets and to minimize the risk of exposure to the real virus, we obtained a 950 bp RNA template synthesized from a de novo constructed DNA fragment, through an in-vitro transcription system (RiboMAX, Promega, USA). To construct the DNA template for the in-vitro transcription, pure noroviral RNA was extracted from norovirus particles (GI.1) obtained through USEPA, Cincinnati, Ohio and subjected to RT-PCR. The obtained norovirus RNA product represents approximately 14% of the total viral genome, and includes homologous regions for DNA probes which were used for our DSN assay, as well as the primer and probe binding sites used by Hoehne et al. [4] for real time RT-PCR which we used for secondary detection and comparison.

![Graph](image)

**Figure 4.3** Plot of norovirus DSN assay with RiboMax RNA for Norovirus Genogroup I and (10⁴ to 10⁰ genomic equivalents per assay) were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=3).

The longer RNA target also contains some of the secondary structures that can be expected from the native Norovirus genome, making it a more real-like target. The Figure 4.3 illustrates the results obtained using the longer RNA target.
4.1 Primary Optimization

Given the length of our viral target RNA, we found it necessary to increase the specificity of the nuclease cleavage by challenging the DSN-nucleic acid duplex with a number of compounds. We found sensitivity could be improved by ensuring specific hybridization between the probe and the target RNA, thus reducing non-specific DSN activity.

Strong secondary structure of some RNAs can inhibit target-probe binding in many ways [81]. DSN enzyme can get separated from the RNA if it has complex secondary structures. In other cases, the enzyme may skip looped structures, resulting in improper binding and cleavage [82, 83]. To reduce the chances of non-specific binding with our relatively short Taqman probe, we tested several common nuclease inhibitors to challenge DSN: including Dimethyl sulfoxide (DMSO), Betaine and Aurintricarboxylic Acid (ATA).

DMSO is a colorless chemical compound miscible in a wide range of organic solvents. It is extensively used in PCR reactions to inhibit secondary structures in DNA or primers by disrupting hydrogen bonding. DMSO obstructs DNA strands from curling back onto itself and thereby reduces interfering interactions [84]. This enables the polymerase to read through the region without getting entangled in the secondary structure.

Betaine is a chemical found in living organisms as an organic osmolyte, and is widely used in PCR reactions to improve DNA amplification by reducing the secondary structures of GC rich DNA segments. Betaine is also an isostabilizing agent, as it dislocates secondary structures by balancing the strength of GC and AT bonds [85] rather than interfering with the bonds like DMSO.

It is well known that divalent cations are critical to the enzymatic activity of nucleases. DSN as well as many other enzymes that act on phosphate containing substrates require magnesium for their activity [13]. In literature it has been shown that presence of Magnesium Chloride (MgCl₂)
and Cobalt Chloride (CoCl$_2$) can increase DSN activity [13]. In our experiments, increasing the amount of Mg from 5 mM to 7 mM improved specific nuclease activity of probe cleavage.

ATA is a triphenylmethane dye which is normally used for inhibition of protein synthesis through disruption of protein-nucleic acid complexing [86]. ATA binds to positively charged amino acid residues since it is negatively charged at neutral pH [87]. The presence of divalent cations (Mg$^{2+}$, Ca$^{2+}$, and Mn$^{2+}$) triggers maximum binding of ATA [88]. EDTA is also known to improve the binding capacity of ATA by either altering the charge or by bringing about conformational changes in the protein structure so as to expose new binding sites [89]. Yap et al have reported that ATA is very strongly correlated to polymerase activity and binds favorably to structurally conserved twin aspartic acid residues (DD), a motif which forms the catalytic center for RNA polymerases and integrases [90,91]. The DD amino acid motif occurs at residues 302, 374 and 387 in the structure of DSN [92].

Thus, we hypothesized that ATA would interact with DSN in these regions and make it more specific to the nucleic acid duplex interaction as it competes with the target duplex for binding. Furthermore, DSN as well as many other enzymes that act on phosphate containing substrates require divalent cations (Mg$^{2+}$, Ca$^{2+}$) for their activity [13]. The impact of varying Mg$^{2+}$ concentration on DSN activity has been well documented [13]. DSN digestion efficiency reaches 100% at ~20mM Mg$^{2+}$ and begins to reduce beyond this concentration [13]. Notably, increasing the amount of Mg$^{2+}$ in the reaction mix aided in improving the efficiency of detection in the presence of ATA. We investigated the effect of aforementioned additives to make our assay more specific, and found that ATA and Mg$^{2+}$ improved the sensitivity of the assay with long RNA target. We tried various combinations and concentrations of the above compounds to make our assay more specific, and found that ATA contributed the most significant benefit. The
addition of Mg to reaction mixture showed significant improvement in the goodness of fit. We could achieve $R^2$ value of $> 0.8$ using the above additives. These results are significant since this method depends on developing a linear fit with known copy numbers to facilitate easy correlation for quantifying the unknown virus. The experimental results are tabulated below in Table 4.1

<table>
<thead>
<tr>
<th>Additive</th>
<th>Concentration range$^a$</th>
<th>R-squared value$^b$</th>
<th>Coefficient of variation$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMSO</td>
<td>10-20% (w/v)</td>
<td>$&lt; 0.7$</td>
<td>10 - 30%</td>
</tr>
<tr>
<td>Betaine</td>
<td>0.5-1.5 M</td>
<td>$&lt; 0.7$</td>
<td>5 - 25%</td>
</tr>
<tr>
<td>ATA</td>
<td>0.3-0.6 mg/ml</td>
<td>$&gt; 0.8$</td>
<td>0.2 - 15%</td>
</tr>
<tr>
<td>Mg$^{2+}$</td>
<td>5-10 mM</td>
<td>$&gt; 0.8$</td>
<td>0.2 - 15%</td>
</tr>
<tr>
<td>Ca$^{2+}$</td>
<td>5-10 mM</td>
<td>$&lt; 0.7$</td>
<td>15 - 30%</td>
</tr>
</tbody>
</table>

$^a$ - Concentration ranges are for a 30 µL reaction volume. All reactions already contained 5 mM Mg$^{2+}$; additional Mg$^{2+}$ was added to further optimize the reaction conditions.

$^b,c$ - Values based on norovirus DSN assay with 950 bp norovirus GI RNA. Triplicate samples of 10-fold serial dilutions of RNA ($10^4$ to $10^9$ genomic equivalents per assay) were plotted versus fluorescence intensity.

Further experiments were carried out to determine specific values of ATA concentration to obtain easily distinguishable fluorescence values with respect to various norovirus copy numbers and with $R^2$ values greater than 0.95 ensuring a linear fit. The following graphical analysis suggests the optimal working window of ATA to be in the range of 0.45-0.47 mg/ml with less scattered error bars, good line fitment and appropriate slope (as shown in Figure 4.4)
The correlation of experimental data for norovirus copy numbers plotted against fluorescence for varying ATA concentrations in the following graphs (Figure 4.6 and Figure 4.7) highlights the best results produced by 0.47mg/ml of ATA.

Figure 4.4 Scatter plot of relative fluorescence units (RFU) as a function of norovirus RNA copies for different concentrations of ATA. Error bars depict standard deviation (n=3).

The correlation of experimental data for norovirus copy numbers plotted against fluorescence for varying ATA concentrations in the following graphs (Figure 4.6 and Figure 4.7) highlights the best results produced by 0.47mg/ml of ATA.

Figure 4.5 3D bar chart of ATA variance. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying ATA concentration.
Figure 4.6 3D surface plot of ATA variance. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying ATA concentration.

It is seen from the bars in Figure 4.5 and the surface plot in Figure 4.6 that the sensitivity of the experiment with 0.4mg/ml and 0.44mg/ml is lower and hence the differentiation between closer values of norovirus copy numbers would be challenging. In the case of 0.47 mg/ml ATA concentration, the slope of the bar is steeper allowing higher levels of differentiation for norovirus copies. The surface area between ATA concentrations of 0.44mg/ml and 0.47mg/ml shows definitive and gradual increase in fluorescence values indicating favorable conditions for developing a standard reference for determining the norovirus copy numbers in an unknown sample.

4.2 Secondary Optimization

Through primary optimization effective working window for ATA concentration was established. However, the results showed larger than desirable variance for lower norovirus copy numbers. This phenomenon is illustrated in the following graph (Figure 4.7).
In order to reduce the error bars and improve virus copy enumeration we optimized our assay for specificity. The probe concentration was decreased in order to reduce the number of non probe-target binding events. This was done to ensure that there are just enough probe molecules available in the reaction mixture to bind to the RNA target. All the three probe concentrations, viz 10nM, 30 nM and 50 nM, produced consistent results. However the fitment of 10 nM probe was superior with R² value of 0.9442 and adequate fluorescence sensitivity. The results are depicted in Figure 4.8.
Figure 4.8 Scatter plot of relative fluorescence units (RFU) as a function of norovirus RNA copies for different probe concentrations. Error bars depict standard deviation (n=3).

Zhao et al [73] have reported that DSN can demonstrate minor activity against single stranded DNA (which is the probe in our case) when both DSN enzyme and substrate (DNA-RNA duplex) are present in high concentrations. This activity, however, is not evident in the presence of competitive dsDNA [73]. Thus we optimized the reaction with addition of varying amounts of sheared dsDNA to improve accuracy. The DNA viral template used for the RiboMAX RNA production system (Promega) was used for obtaining the sheared dsDNA, since it effectively competed with the RNA target for the probe. Thus we optimized the reaction with addition of varying amounts of sheared dsDNA to improve accuracy.

Numerous lab trials were carried out to standardize the sheared DNA copies in the assay so that fluorescence from single stranded DNA cleavage does not influence the norovirus detection procedure and results. We used sheared DNA values of 1, 10 and 100 copies for this experiment. The results are depicted in the following graphs: It can be seen from Figure 4.9, the presence of 10 copies of sheared DNA produced the highest $R^2$ value. The slope of this line also allows moderate sensitivity for detection.
Figure 4.9 Scatter plot of relative fluorescence units (RFU) as a function of norovirus RNA copies for varying sheared DNA copies. Error bars depict standard deviation (n=3).

The 3D representation of the experimental data in the form of bars and surface chart at Figure 4.10 and Figure 4.11 shows that the results are consistent for all the three trials. In the case of 1 copy, the fluorescence sensitivity decreases rapidly for higher virus copy numbers. However, this phenomenon is exactly opposite in the case of 100 copies of sheared DNA. The surface graph clearly shows the smooth transition of increasing fluorescence intensity with increase in viral load for 10 copies of sheared DNA. Hence this is the desired concentration of sheared DNA for the assay.
Figure 4.10 3D bar chart on Impact of Sheared DNA. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying sheared DNA concentration.

Figure 4.11 3D surface chart on Impact of Sheared DNA. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying sheared DNA concentration.
DSN requires the presence of divalent cation such as Mg for its enzymatic activity [13]. It has been shown that addition of Mg facilitates binding of the enzyme to the substrate, at the site of cleavage, with proper orientation [74]. The impact of varying Mg concentration on DSN activity has been well documented. DSN digestion efficiency reaches 100% at ~20mM Mg and begins to reduce beyond this concentration [13]. Thus, we increased the amount of Mg in the final reaction to ascertain the optimum Mg concentration for balancing increased activity and specificity [13]. Optimization of addition of Mg in the assay proved to be complex posing difficulty in ascertaining correct dosage. Initial trials were carried out on both oligo RNA group I and II. As the environmental waters in and around Cincinnati contained only GI further trials were restricted to this group alone. It is seen that concentration above 5 mM produced consistent results for better fitment of graphs. 10mM of Mg addition proved to be with best fit with good slope for detection sensitivity. The results are depicted in Figure 4.12-4.14.

Figure 4.12 Scatter plot of relative fluorescence units (RFU) as a function of norovirus RNA copies for varying Mg concentration. Error bars depict standard deviation (n=3).
Figure 4.13 3D bar chart of Mg Variance. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying Mg concentration.

Figure 4.14 3D Surface plot on Mg Variance. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) were plotted versus Relative Fluorescence Units (RFU) and varying Mg concentration.
It can be seen from the 3D bars at Figure 4.13 that 5mM of Mg behaves inversely in comparison to higher Mg concentrations. The surface plot reveals the low sensitivity of 7mM concentration with respect to viral copy numbers. In contrast the gradually increasing tip of the surface for 10mM Mg proves its suitability for optimization.

Through this approach, we could establish clear ranges for ATA, probe concentration, sheared DNA and Mg additions to achieve clear separation between norovirus copy numbers, and most importantly strong separation between lowest detectable viral counts and no-template control (NTC) as shown in Figure 4.15.

![Figure 4.15](image_url)

Figure 4.15 Plot of norovirus DSN assay with 950 bp norovirus RNA. 10-fold serial dilutions of RNA (10⁴ to 10⁰ genomic equivalents per assay) of Norovirus GI were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=3).
4.3 Validation and Verification

In order to verify the efficacy of our optimized assay we performed an experiment spiking environmental water with norovirus particles ($10^4$-10$^1$ copies per liter of water) to mimic a real world application. The total RNA was extracted using the previously described skim milk flocculation method, to concentrate the virus from environmental waters containing known viron numbers. In this way we obtained RNA-water extracts replete with non-target RNA and natural inhibitors which co-elute with the extraction. A plot of change in fluorescence intensity in response to the different concentrations of RNA extracted from river water spiked with norovirus particles is shown in Figure 4.16 A.
Figure 4.16 (A) Plot of norovirus DSN assay with RNA extracted from water spiked with norovirus particles. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) of Norovirus GI were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=6). (B) Plot of norovirus RT-PCR with RNA extracted from water spiked with norovirus particles. 10-fold serial dilutions of RNA (10^4 to 10^0 genomic equivalents per assay) of Norovirus GI were plotted versus 1/C_t value. Error bars show standard deviations (n=3).

Performed as a blind test, the operator was able to correctly identify the concentration of all but two of the RNA samples (16/18 or 89% correct) with no false positives. We also compared the quantification of the same RNA samples with conventional RT-PCR method (Figure 4.16B) using the primers and probe developed previously [4]. This enabled us to test the accuracy of the DSN based quantification against the present standard. As illustrated in Figure 4.16B, we observed larger coefficient of variation for C_t values of RT-PCR indicating the higher sensitivity of our assay for the same RNA samples.
The present assay was independently verified in the lab of Dr. Shay Fout, USEPA, Cincinnati, Ohio. We could achieve a detection limit of 10 norovirus copies per assay as shown in Figure 4.17, which is below the minimum concentration required for infection (18 viral copies) [15-17]. However, some variation with $10^3$ copies was observed.

Figure 4.17 Plot of norovirus DSN assay with RNA extracted from water spiked with norovirus particles. 10-fold serial dilutions of RNA ($10^4$ to $10^9$ genomic equivalents per assay) of Norovirus GI were plotted versus normalized Relative Fluorescence Units (RFU). Error bars show standard deviations (n=6).
Chapter 5  Discussion and Conclusion

The currently established methods to detect norovirus employ reverse transcription as an essential step to determine the viral load. Converting RNA to cDNA may involve some loss or biased amplification leading to lower accuracy and reliability. Recently, advanced PCR techniques such as digital PCR have been described for absolute RNA quantification [93], but the high equipment cost and time limits their applicability. Since the viral dose of norovirus infection is as low as 18 copies and the current methods are unable to penetrate the threshold, the DSN system proves efficient and cost effective even at 10 copies of virus.

The main challenge in the development of the assay was to determine the optimization parameters which augment the principle of achieving increased fluorescence sensitivity for viral copies along with acceptable variance for low virus concentration. Exhaustive trials with various combinations of these experiment parameters (ATA, probe, Mg and sheared DNA) were conducted to establish their working range. Optimization results have proved the desired effect can be achieved with appropriate combination of probe (10nM-50nM), ATA (0.40mg/ml-0.47mg/ml), Mg (5mM-10mM) and sheared DNA (1, 10 and 100 copies) concentration. A set of final trials as given in Table 5.2 were done to validate specific values for the above parameters in order to standardize the final assay.
Table 5.1 Final Trials of Optimization

<table>
<thead>
<tr>
<th>TRIAL</th>
<th>ATA (mg/ml)</th>
<th>Mg (Mm)</th>
<th>PROBE (nM)</th>
<th>SHEARED DNA (ng/µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.4mg/ml</td>
<td>5mM</td>
<td>50nM</td>
<td>-------</td>
</tr>
<tr>
<td>2.</td>
<td>0.4mg/ml</td>
<td>7mM</td>
<td>50nM</td>
<td>-------</td>
</tr>
<tr>
<td>3.</td>
<td>0.45mg/ml</td>
<td>7mM</td>
<td>50nM</td>
<td>-------</td>
</tr>
<tr>
<td>4.</td>
<td>0.45mg/ml</td>
<td>10mM</td>
<td>50nM</td>
<td>-------</td>
</tr>
<tr>
<td>5.</td>
<td>0.47mg/ml</td>
<td>10mM</td>
<td>50nM</td>
<td>-------</td>
</tr>
<tr>
<td>6.</td>
<td>0.47mg/ml</td>
<td>10mM</td>
<td>30nM</td>
<td>-------</td>
</tr>
<tr>
<td>7.</td>
<td>0.47mg/ml</td>
<td>10mM</td>
<td>10nM</td>
<td>-------</td>
</tr>
<tr>
<td>8.</td>
<td>0.47mg/ml</td>
<td>10mM</td>
<td>10nM</td>
<td>1.82 X 10^{-5}</td>
</tr>
<tr>
<td>9.</td>
<td>0.47mg/ml</td>
<td>10mM</td>
<td>10nM</td>
<td>1.82 X 10^{-6}</td>
</tr>
<tr>
<td>10.</td>
<td>0.47mg/ml</td>
<td>10mM</td>
<td>10nM</td>
<td>1.82 X 10^{-7}</td>
</tr>
</tbody>
</table>

The results of ATA concentration between 0.4mg/ml -0.47mg/ml are consistent for consideration. However, within this range an ATA concentration of 0.45mg/ml and 0.47mg/ml gave the best results as they exhibited superior linear curve fitment and low error bar dispersion. This ATA concentration window is suitable for establishing the reference curve for determining copy numbers of the unknown sample. Earlier, effective functioning of DSN has been largely limited to short RNA targets, thus the optimization of our reaction with the use of ATA was crucial, as it significantly improved the DSN binding to the target RNA by competing with DSN for binding sites. After fixing the ATA dose, the Mg concentration was increased in steps from 5mM through 10mM. This increase not only yielded a higher coefficient of determination ($R^2$)
but also improved the distinction between lower copy numbers and NTC. The trials done with varying probe concentration revealed that the results are consistent in the range of 10nM-30nM. At 10nM probe concentration, with ATA (0.47mg/ml) and Mg (10mM) concentration resulted in higher sensitivity. The final optimization step was the addition of sheared DNA copies to reduce false positive signals. The entire range of 1, 10 and 100 copies produced similar trends but the \( R^2 \) values and error dispersion were best at 10 copies. It is concluded that the values of these four parameters enumerated in trial 9 in Table 5.1 are best suited for using DSN assay.

The application of the assay involves establishing a standard curve, comprising of known norovirus copy numbers using the above validated parameters. The virus copy numbers in unknown sample can be readily determined by directly correlating the resultant fluorescence to the standard curve. The total reaction time for our strategy is less than 30 minutes, compared to the several hours required for quantitative RT-PCR [29], and can be achieved on a simple heat block, rather than a more costly thermocycler.

The advantage of targeting the viral RNA directly with DSN assay gives the opportunity to extend this detection system to other single stranded RNA viruses such as Astrovirus [95,96], HIV[97,98], Chikungunya [99], Influenza [100-101] and, Chandipura [102-103] and Rotavirus [104]. The genetic diversity of the RNA targets makes it crucial to design the probes such that they are complementary to the most conserved region in the viral genome. The other possible RNA virus targets for DSN are listed in Table 5.2.
Table 5.2 Future RNA Virus Targets for DSN

<table>
<thead>
<tr>
<th>Name</th>
<th>Genome</th>
<th>Size (nm)</th>
<th>Disease Associated</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astrovirus</td>
<td>Positive sense single stranded RNA</td>
<td>25-28</td>
<td>Gastroenteritis in children and adults</td>
<td>95,96</td>
</tr>
<tr>
<td>Human immunodeficiency virus</td>
<td>Positive sense single stranded RNA</td>
<td>120</td>
<td>Acquired Immunodeficiency Syndrome (AIDS)</td>
<td>97,98</td>
</tr>
<tr>
<td>Chikungunya</td>
<td>Positive sense single stranded RNA</td>
<td>60-70</td>
<td>Similar to dengue fever followed by severe joint pain</td>
<td>99</td>
</tr>
<tr>
<td>Influenza A virus</td>
<td>Negative sense single stranded RNA</td>
<td>6-9</td>
<td>Flu in birds and mammals</td>
<td>100,101</td>
</tr>
<tr>
<td>Chandipura virus</td>
<td>Negative sense single stranded RNA</td>
<td>50-60</td>
<td>Neurologic dysfunctions especially in children</td>
<td>102,103</td>
</tr>
<tr>
<td>Rotavirus</td>
<td>Double stranded RNA virus</td>
<td>76.5</td>
<td>Acute diarrhea among infants and young children</td>
<td>104</td>
</tr>
</tbody>
</table>

In conclusion, we have described a novel isothermal assay for the detection of RNA targets using the specificity of Duplex Specific Nuclease. Our assay has the capacity to go beyond RT-PCR as a more cost effective and faster detection method. The technique is a quick, sensitive one-step method for detection of long viral RNA targets. Since norovirus is directly transmitted by fecal-oral-route, the extent of infection is higher in emerging nations, which tend to have larger populations and inadequate water treatment infrastructure. Thus, inexpensive virus detection with our method would be of great benefit to developing countries where sporadic outbreaks of norovirus infections have caused catastrophic effects.
5.1 Future Work

The assay developed in this work has been optimized with the use of ATA, for specific binding of the enzyme to the substrate. ATA has been reported to have strong preference for binding to ‘DD’ amino acid residues [91] hence; we plan to mutate the corresponding binding sites on the DSN structure. The DD amino acid motif occurs at residues 302, 374 and 387 in the DSN sequence [13]. The mutation of DSN binding sites would enable proper hybridization without the external addition of ATA and also make the experimental setup less cumbersome. To enable widespread use of the DSN detection method, a commercial production system for the enzyme is required. Since, DSN enzyme is the major constituent of the assay and significantly increases the experimental cost, we plan to devise a recombinant protein production system. To accomplish this we intend to clone and express codon optimized DSN in *Ecoli*, and perform site directed mutagenesis at the aforementioned positions. This will serve the dual purpose of reducing the experimental cost as well as eliminating the need for ATA addition in the reaction mixture.
Bibliography


18 http://water.epa.gov/scitech/drinkingwater/dws/ccl/ccl3.cfm


Williams, F.P. *Negative-stain transmission electron micrograph of Norwalk virus isolated from an individual with gastroenteritis. Bar = 50 nm, U.S. EPA*


40 www.cdc.gov/norovirus/php/illness-outbreaks.html

41 http://www.foxnews.com/health/2014/01/31/norovirus-should-be-afraid-to-go-on-cruise/


49 Katayama H, Shimasaki A, Ohgaki S: Development of a Virus Concentration Method and Its Application to Detection of Enterovirus and Norwalk Virus from Coastal Seawater. Appl


60 Smith CJ, Osborn AM: Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. FEMS microbiology ecology 2009, 67(1):6-20


62 Wilson IG: Inhibition and facilitation of nucleic acid amplification. Applied and


67 http://www.epa.gov/nerlcwww/documents/Method1615v1_1.pdf


70 Updated Norovirus Outbreak Management and Disease Prevention Guidelines, Morbidity and Mortality Weekly Report, Centers for Disease Control and Prevention, March 4, 2011


http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi


94 Hindson CM, Chevillet JR, Briggs HA, Galichotte EN, Ruf IK, Hindson BJ, Vessella RL, Tewari M: Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat Meth 2013, advance online publication


Bhatt Pn Fau - Rodrigues FM, Rodrigues FM: Chandipura: a new Arbovirus isolated in India from patients with febrile illness. (0971-5916 (Print)).