creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

A Thesis for the Degree of Doctor of Philosophy

Research on the contamination level of NoV and the combined water treatment system

지하수의 노로바이러스 오염실태 조사 및 노로바이러스 제거를 위한 복합수처리장치 연구

February, 2013

By

JeongWoong Park

Department of Agricultural Biotechnology

College of Agriculture and Life Science

Seoul National University

농학박사학위논문

Research on the contamination level of NoV and the combined water treatment system

지하수의 노로바이러스 오염실태 조사 및 노로바이러스 제거를 위한 복합수처리장치 연구

지도교수 유 상 렬

이 논문을 박사학위논문으로 제출함
2013년 2월

서울대학교 대학원
농생명공학부
박 정 웅

박정웅의 박사학위논문을 인준함
2013년 2월

위 원 장	강 동 현	**(인)**
부 위 원 장	유 상 렬	**(인)**
위 원	장 판 식	**(인)**
위 원	백 순 영	**(인)**
위 원	이 주 훈	**(인)**

Abstract

JeongWoong Park

Department of Agricultural Biotechnology

The Graduate School

Seoul National University

Noroviruses (NoVs) are the most important viruses that cause nonbacterial acute gastroenteritis in humans. One of the virulence factors of NoV, the great genetic diversity attributed to point mutations and recombinations, has brought forth the result of significant changes in the circulating NoV genotype patterns.

In recognition of the necessity for tracking and monitoring of genetic diversity, a novel NoV variant was isolated from a patient in South Korea. To further understand its genetic characteristics and taxonomical classification, its genome was completely sequenced. Complete genome sequence analysis revealed three predicted ORFs; ORF1 (5,100 bp), ORF2 (1,623 bp), and ORF3 (807 bp). Comparative phylogenetic analysis of complete genome sequence of the isolated NoV strain showed that it may be in the Cluster I. Furthermore, subsequent phylogenetic analyses of three ORFs substantiated that this strain belongs to Cluster I/2006b. Therefore, the isolated NoV strain

designated NoV Hu/GII-4/CUK-3/2008/KR.

To investigate the contamination levels of NoVs in groundwater used at food catering facilities in South Korea, 1,090 groundwater samples were collected from Gyeonggi, Chungcheong, Gangwon, Gyeongsang, Jeolla and Jeju provinces and NoVs were detected in seven samples (three GII-4, two GII-17, one GII-3 and one GI-6) using the method based on the guidelines for NoVs Detection in Groundwater (Korea Food and Drug Association, KFDA), suggesting that NoVs are present in the groundwater of South Korea. Therefore, NoV reduction system is urgently required.

To develop the water treatment system, electropositive filters were tested and 3M filter was selected due to its high fidelity of human NoV reduction in water. Using this 3M filter, the combined water treatment system was developed, consisting of electropositive filtration, UV irradiation, and chlorine disinfection. To evaluate this system, spiking samples containing $>10^5$ copies of human NoV particles were used and quantitative real-time RT-PCR using the eluted sample showed that no human NoV particle was detected in the samples, suggesting that this system indeed work for human NoV reduction in water. Therefore, this combined water treatment system should be used for human NoV reduction from the groundwater.

Keywords: Norovirus sequence, Phylogenetic tree, Groundwater, Combined water treatment system, Electropositive filter, Norovirus reduction

Student number : 2005-31092

Table of Contents

Abstract i

Table of Contents iv

List of Tables vii

List of Figures ix

PART I: Overall Introduction 1

PART II: Full sequence analysis and characterization of the South Korean Norovirus GII-4 variant CUK-3

2.1. Materials and Methods 14

2.1.1. Stool specimen 14

2.1.2. RNA extraction 14

2.1.3. Detection of NoV 14

2.1.4. NoV genome sequencing 16

2.1.5. Bioinformatics analysis 16

2.1.6. Nucleotide sequence accession number 17

2.2. Results and Discussion 19

2.2.1. Complete genome sequence analysis of the isolated NoV strain 19

2.2.2. Comparison of ORF2 with reference strains 26

2.2.3. Analysis of comparative evolutionary relationship in NoV 28

2.3. Conclusion 38

PART III: Prevalence of NoVs from Groundwater in South Korea and Their Evolutionary Relationships

3.1. Materials and Methods 40

3.1.1. Overview of detection of NoVs in groundwater 40

3.1.2. Sampling 40

3.1.3. Virus concentration and nucleic acids extraction 45

3.1.4. One-step reverse transcription-PCR (RT-PCR) and nested PCR 46

3.1.5. Determination of genogroup of NoV using PCR with genogroup-specific primer sets 47

3.1.6. Sequencing and phylogenetic analysis 52

3.2. Results and Discussion 53

3.2.1. Detection of NoVs in various provinces and periods 53

3.2.2. Genogrouping of detected NoVs in South Korea 58

3.2.3. Genotyping of detected NoVs in South Korea 60

3.2.4. BLAST functional analysis of predicted ORFs in DNA sequences of the nested PCR products 68

3.3. Conclusion 72

PART IV: Research on the Combined Water Treatment System

4.1. Materials and Methods 74

4.1.1. Evaluation of NoV reduction yield in electropositive filters 74

4.1.1.1. Test sample 74

4.1.1.2. Electropositive filter 74

4.1.1.3. Quantification of human NoV using Real-time RT-PCR 74

4.1.2. Evaluation of NoV reduction yield in combined water treatment system 79

4.1.2.1. Test sample 79

4.1.2.2. Combined water treatment system 79

4.1.2.3. Quantification of human NoV using Real-time RT-PCR 83

4.2. Results and Discussion 84

4.2.1. Evaluation of NoV reduction yield in electropositive filters 84

4.2.2. Evaluation of NoV reduction yield in combined water treatment system 87

4.3. Conclusion 91

PART V: References 92

List of Tables

PART II: Full sequence analysis and characterization of the South Korean Norovirus GII-4 variant CUK-3

Table 2.1. Diagnosis primer sets 15

Table 2.2. Designed primer sets 18

Table 2.3. Nucleotide and amino acid sequence similarities between CUK-3 strain and the full length ORF2 sequences from different Genogroup GII reference strains 27

PART III: Prevalence of NoVs from Groundwater in South Korea and Their Evolutionary Relationships

Table 3.1. Primer sequence for NoV GI detection..... 48

Table 3.2. Primer sequence for NoV GII detection 49

Table 3.3. 1st PCR(one-step RT-PCR) conditions 50

Table 3.4. 2nd PCR (nested PCR) conditions..... 51

Table 3.5. The genogrouping and genotyping of detected NoV positive samples in South Korea 59

Table 3.6. BLAST functional analysis of predicted ORFs in the nested PCR sequences..... 69

PART IV: Research on the Combined Water Treatment System

Table 4.1. Primer and probe of Human NoV used in Real-time RT-PCR 77

Table 4.2. Real-time RT-PCR conditions 78

Table 4.3. Human NoV reduction yield of electropositive filters 86

Table 4.4. Human NoV reduction yield of combined water treatment system..... 90

List of Figures

PART I: Overall Introduction

Figure 1.1. Full genome map of NoV 4

Figure 1.2. Modes of transmission of enteric viruses 6

Figure 1.3. Incidence of water-borne and food-borne outbreaks 9

PART II: Full sequence analysis and characterization of the South Korean Norovirus GII-4 variant CUK-3

Figure 2.1. GenBank deposition of complete genome sequence of the isolated NoV strain 20

Figure 2.2. DNA sequence of ORF1 in the genome sequence of the isolated NoV strain 21

Figure 2.3. DNA sequences of ORF2 and ORF3 in the genome sequence of the isolated NoV strain 22

Figure 2.4. Complete genome sequence of the isolated NoV strain 23

Figure 2.5. Phylogenetic tree constructed on the nucleotide sequence of ORF1 (5.1 kb) 30

Figure 2.6. Phylogenetic tree constructed on the nucleotide sequence of ORF2 (about 1.6 kb) 31

Figure 2.7. Phylogenetic tree constructed on the nucleotide sequence of ORF3 (about 0.8 kb) 32

Figure 2.8. Phylogenetic tree constructed on the nucleotide sequence of complete genome (about 7.5 kb) 33

Figure 2.9. Partial sequence comparison of the RdRp gene 35

PART III: Prevalence of NoVs from Groundwater in South Korea and Their Evolutionary Relationships

Figure 3.1. Process flow chart for detection and identification of NoV in water 41

Figure 3.2. Standard sampling system 42

Figure 3.3. Details of sampling period 43

Figure 3.4. Details of sampling regions 44

Figure 3.5. The number of NoV positive samples by area 54

Figure 3.6. The percentage of NoV positive samples by area 55

Figure 3.7. The number of NoV positive samples by month 56

Figure 3.8. The percentage of NoV positive samples by month 57

Figure 3.9. The nucleic acid sequences of nested PCR products in GII 62

Figure 3.10. The nucleic acid sequence of nested PCR product in GI 64

Figure 3.11. Phylogenetic tree analysis of sample 2010042241, 2010060371 and 2010071392 65

Figure 3.12. Phylogenetic tree analysis of sample 2010082762, 2010100772 and 2010101112 66

Figure 3.13. Phylogenetic tree analysis of sample 2010081812 67

PART IV: Research on the Combined Water Treatment System

Figure 4.1. Procedure for quantification of human NoV 76

Figure 4.2. Combined water treatment system 80

Figure 4.3. Composition and flow diagram of combined water treatment system 81

Figure 4.4. Overall evaluation procedure of reduction yield of NoV after water treatment using combined water treatment system 82

Figure 4.5. Real time RT PCR result of human NoV using electropositive filters 85

Figure 4.6. Amount of NoV in spiking sample using quantitative real-time RT-PCR 89

PART I: Overall Introduction

Noroviruses (NoVs) are the clinically most important viruses that cause nonbacterial acute gastroenteritis in humans. People of all age groups are susceptible to these viruses, especially the elderly people are at increased risk of more severe disease and death, in addition to the very young and the immune compromised individuals (Donaldson *et al.* 2008; Yoon *et al.* 2006). They are small and non-enveloped viruses and belong to the family *Caliciviridae*, genus *Norovirus*. NoVs have a single positive-strand NoV RNA genome of about 7.6 kb in size (Figure 1). Three open reading frames (ORFs) have been identified in the NoV genomes. ORF1 encodes a polyprotein that is cleaved into six non-structural (NS) proteins, which carry amino acid sequence motifs conserved in NTPase, protease and RNA-dependent RNA polymerase (RdRp) (Lee *et al.* 1977; Rueckert *et al.* 1984). ORF2 encodes a major structural protein, Viral Protein (VP1) that consists of two domains; the shell domain (S) and the protruding arm (P) which is again divided into two subdomains, P1 and P2. The S domain is highly conserved while the P domain is variable. P2 of the P domain is hypervariable and carries immune and cellular recognition sites (Tan *et al.* 2003; Nilsson *et al.* 2003; Lochridge *et al.* 2005). ORF3 encodes minor capsid protein, VP2, which is rich in basic amino acids and is proposed to have a role in viral stability (Glass *et al.* 2000; Bertolotti-Ciarlet *et al.*

2003). Based on nucleotide sequence analysis that encodes the capsid proteins, NoVs are classified into five genogroups, GI to GV, each of which can be further divided into several clusters or genotypes (Motomura *et al.* 2008). Among the five genogroups, three genogroups (GI, GII and GIV) are known to cause clinical illness in humans, and genotype GII-4 has been the predominant circulating strain (Hale *et al.* 2000). The error-prone RNA replication and recombination between viruses is primarily responsible for great diversity among NoVs. Furthermore, the accumulated mutations of the hypervariable P2 domain of the VP1 protein produced different GII-4 NoVs (Glass *et al.* 2009). The most representative of the resulting variant of GII-4 strain, GII-4/2006b, with 3 nucleotide insertions in the P2 domain at position 6,265 nt, emerged in the summer of 2002 which led to the major gastroenteritis outbreak as well as an epidemic gastroenteritis worldwide in the winter of 2002/2003 (Lopman *et al.* 2004; Allen *et al.* 2009). In South Korea, gastroenteritis outbreaks by GII-4/2006b variants have been reported during the period from September of 2007 to July of 2008 (Chung *et al.* 2010).

Figure 1.1. Full genome map of NoV (Figure from : Glass *et al*. 2009).

ORF1 - encodes non-structural proteins, including N-terminal protein (N-term), NTPase, 3A-like protein (3A-like), genome-linked viral protein (VPg), 3C-like protease (Pro) and RNA-dependent RNA polymerase (RdRp).

ORF2 - encodes the major capsid protein (VP1).

ORF3 - encodes a small protein (VP2) associated with VP1 stability.

The incubation period of NoV is generally between one and three days. After incubation period, infected persons may develop low-grade fever, vomiting, diarrhea and headache as prominent symptoms. These symptoms usually subside within 2 or 3 days, although the course of illness may be protracted in the elderly. The average attack rate is typically 45% or more. The NoV is shed via stool and vomit, starting during the incubation period and lasting up to 10 days and possibly longer. Foods can be contaminated by five routes (Koopmans *et al.* 2004) (Figure 2). (i) The first route is a direct contact with feces or water contaminated with feces. (ii) The second route is a secondary contact with feces contaminated materials (including hands). (iii) The third route is a contact with vomit or water contaminated with vomit. (iv) The fourth route is a contact with environments (except for aerosols) in which infected people were present. (v) The last route is aerosols generated by infected people. The main route of transmission is suspected to be person-to-person contact, but food and water-based transmissions are also important routes (Donaldson *et al.* 2008; Kearney *et al.* 2007; La Rosa *et al.* 2008; Koopmans *et al.* 2004).

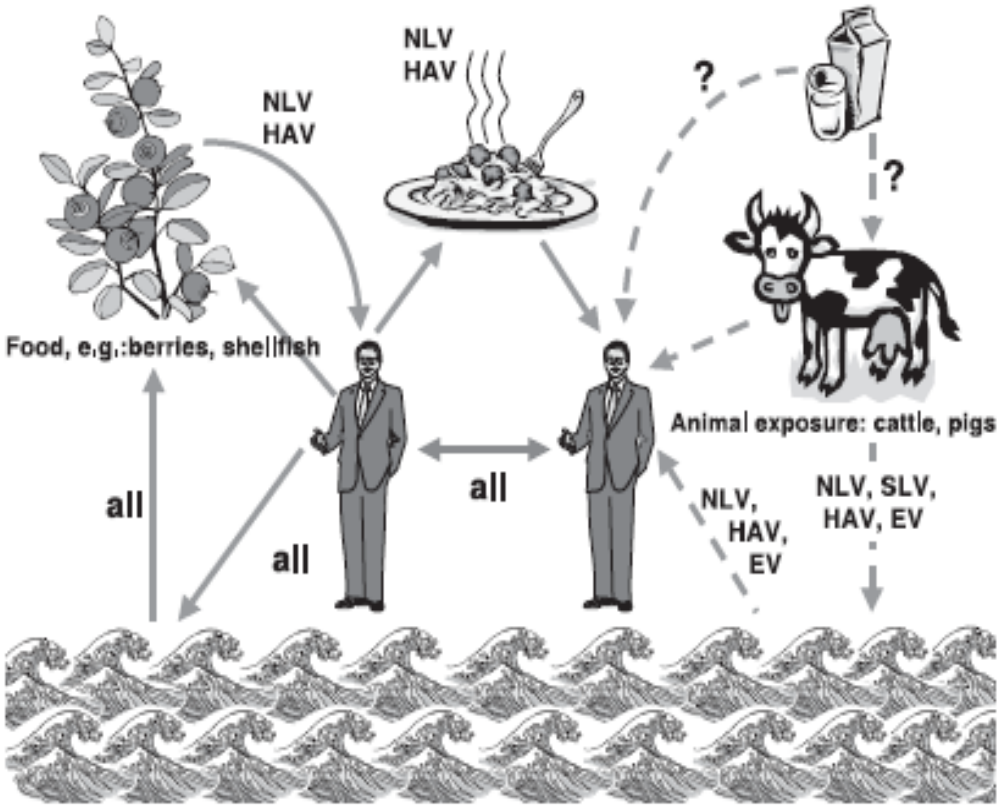
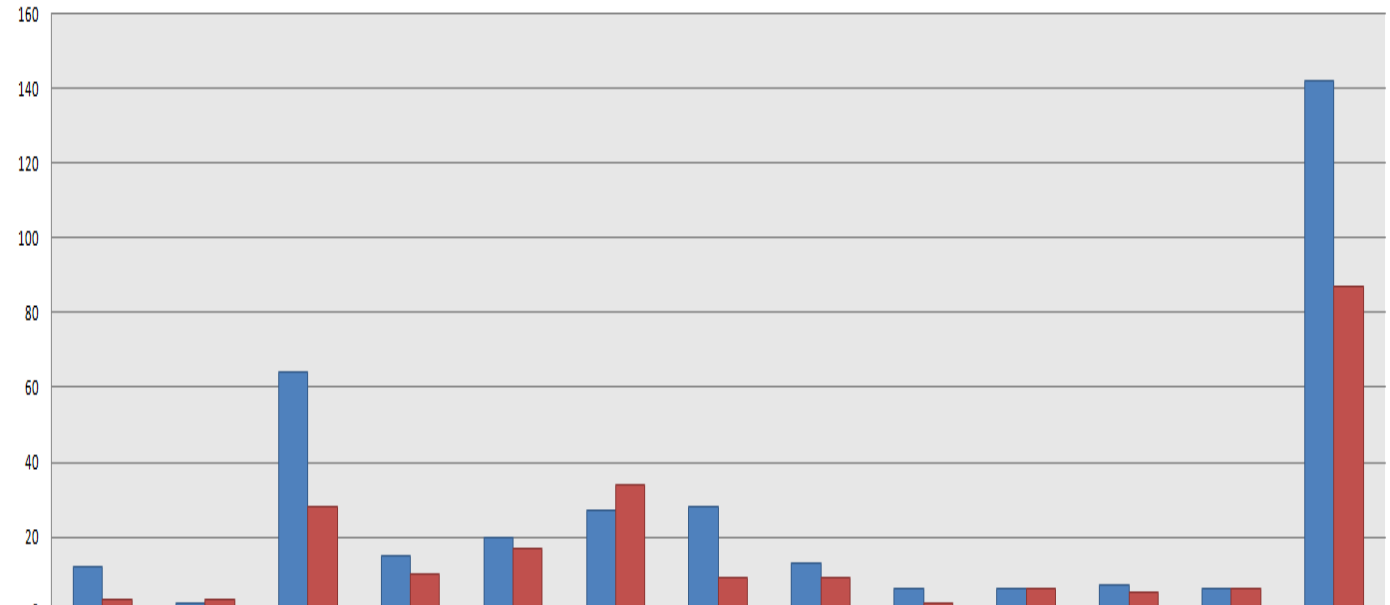


Figure 1.2. Modes of transmission of enteric viruses, showing proven (continuous) and suspected (dashed lines) routes of exposure (Reference, Koopmans *et al.*2004).

Outbreaks caused by NoVs have been reported in a variety of settings (Marshall *et al.* 2003; Kim *et al.* 2005; Medici *et al.* 2009; Widdowson *et al.* 2005; Yoon *et al.* 2008). NoV has the five characteristics that facilitate its rapid, extensive spread (Lee SG *et al.* 2011). The first characteristic is that an extremely low infectious dose of approximately 18 to 1,000 viral particles. This tendency is generally found in various routes such as person-to-person contact, aerosol, fomites, and even environmental contaminations. The low infectious dose can be reflected in the secondary attack rate of 30% or more for those in close contact with an infected individual. The second characteristic is that a prolonged shedding period of NoV. Thus, as the risk of secondary infection increases, spread of NoVs from food handlers already recovered from the symptoms to family members without gastroenteritis could be an epidemic. The third characteristic is that NoVs are stable. This characteristic supports the surviving ability of NoVs in a wide range of environmental conditions including drinking water and various food items such as raw oysters, fruits and vegetables. The great diversity of NoV strains can bring repeated infections due to the lack of complete cross-protection and long-term immunity. Interestingly, frequent Nov genome mutations via antigenic shift and recombination events can endow new NoVs with the capacity to infect NoVs-resistant hosts.

According to the report of Korean Centers for Disease Control and

Prevention (KCDC), NoV is the leading cause of water-borne and food-borne outbreaks of gastroenteritis in South Korea (Figure 3). Several cases of broad-range contamination by NoVs have been reported in South Korea (Lee CH *et al.* 2008). In 2007, water quality test results for NoVs contamination-suspicious food catering facilities showed that 28% of groundwater was contaminated by the NoVs. According to the nationwide survey, known as “Nationwide Groundwater Surveillance of NoVs in South Korea”, funded by National Institute of Environmental Research (NIER) in 2008, NoVs have been detected in 64 sites out of total 176 NoVs contamination-suspicious sites (36.4%) where people use groundwater as the source of drinking water. In 2006, the largest food poisoning outbreak via NoV contamination with low infection dose and rapid transmission was reported in South Korea. From this outbreak, people recognized the importance of groundwater management. Therefore, it was suggested that sophisticated monitoring of human NoV contamination in the environmental resources, especially via water, is highly required in South Korea.



	Shigellosis	Group I etc.	Norovirus	Staphylococcus aureus	Salmonellosis	Escherichia coli	Vibrio parahaemoly ticus	Bacillus cereus	Clostridium perfringens	Campylobacter	Etc.	Multi	Unknown
2008	12	2	64	15	20	27	28	13	6	6	7	6	142
2009	3	3	28	10	17	34	9	9	2	6	5	6	87

Figure 1.3. Incidence of water-borne and food-borne outbreaks (Reference, KCDC, 2010).

To prevent frequent food poisoning from NoV positive groundwater sources, extensive research work has been carried out to eliminate viruses from water. The most significant treatment approach is chlorination. Chlorine is widely used as a disinfectant in groundwater treatment because of its effectiveness in virus inactivation. According to the report of Shin *et al.* (2008), two log reduction after 3 min or three log reduction of NoV after 10 min were observed in the presence of 1 mg free chlorine per liter at 5°C, pH 6, respectively. These results of that study indicate that NoV is not highly resistant to free chlorine disinfection. However, Koopmans *et al.* (2004) reported that NoV is quite resistant to chlorine treatment; NoV remains to be infectious even after 30 min in the presence of 0.5 to 1 mg free chlorine per liter. At higher concentrations (>2 mg free chlorine per liter), NoV is reported to be inactivated, suggesting that NoVs may be sensitive to higher concentration of chlorine than the general treatment concentrations. To evaluate the survival and inactivation of NoVs, Keswick *et al.* (1985) has performed human challenge study. This study showed that 3.75 mg free chlorine per liter could inactivate other viruses but failed to inactivate NoV. Furthermore, it has been reported that chlorine disinfection containing high concentration of free residual chlorine can in turn induce the production of carcinogenic substance like Trihalomathane (Rook *et al*., 1974). Therefore, there is a growing need to find alternative approaches to inactivate NoV. In

addition, UV irradiation method has been studied as a substitute and its inactivation efficiency has been proved by many other scientific researches. Tree *et. al.* (2004) reported that, for a one log reduction of feline calicivirus (a surrogate for NoV), a dose of 4.75 mW s^{-1} cm^{-2} was required. According to the report of Lee *et. al.* (2008), three log reduction of NoV was observed by 254-nm UV irradiation in 200 sec exposure time. Furthermore, Jung *et al.* (2010) found that combined water treatments including chlorine/UV irradiation and chlorine/electropositive filter are more effective than only chlorine water treatment for inactivation of NoV. This combined water treatment method has advantages in preventing offensive odors and THM formation as well as minimizing chlorine concentration for NoV inactivation.

In this study, firstly, a GII-4 strain which is responsible for highest outbreaks in South Korea was selected and its genome sequence was completely sequenced and analyzed in Chapter II. Interestingly, the complete genome sequence revealed that this variant is a novel type of NoV, designating GII-4 variant CUK-3. Comparative analysis of this strain with other known GII-4 variants using phylogenetic analysis revealed that while this GII-4 variant CUK-3 is genetically related with other strains, it is positioned in the different evolutionary group in the tree, supporting its taxonomical novelty. Secondly, total 1,090 groundwater samples were collected from schools and preschools and the collected groundwater

samples were analyzed precisely according to the analysis method in Korean Food Code to determine the prevalence of the NoV contamination. Thirdly, NoV elimination efficiencies of the commercial electropositive filters and its combination with chlorine and UV irradiation were determined to develop the effective NoV elimination system from groundwater.

PART II: Full sequence analysis and characterization of the South Korean Norovirus GII-4 variant CUK-3

2.1. Materials and Methods

2.1.1. Stool specimen

NoV positive-stool sample was originally collected from a patient with acute gastroenteritis in Daejeon, South Korea in November 2008. The isolate was obtained from the Waterborne Virus Bank (Seoul, Korea).

2.1.2. RNA extraction

A 10% fecal suspension was prepared with phosphate buffered saline (PBS, pH 7.4) buffer (Invitrogen, Carlsbad, CA, USA) and the NoV particles were obtained by centrifugation at 10,000 x g for 10 min. The viral genomic RNA was extracted from the NoV particles using a QIAamp Viral RNA Mini kit (Qiagen, Hilden, Germany) in accordance with the manufacturer's instructions.

2.1.3. Detection of NoV

For the detection of NoV, reverse transcription PCR (RT-PCR) was performed using the One-step RT-PCR kit (Qiagen, Hilden, Germany) with primers based on the sequence of the NoV capsid region (Table 2.1).

Table 2.1. Diagnosis primer sets

Primer	Sequence (5'→3')	Polarity	Region
GIF1M	CTG CCC GAA TTY GTA AAT GAT GAT	+	5342-5365
GIR1M	CCA ACC CAR CCA TTR TAC ATY TG	-	5649-5671
GIIF1M	GGG AGG GCG ATC GCA ATC T	+	5049-5067
GIIR1M	CCR CCI GCA TRI CCR TTR TAC AT	-	5367-5389

2.1.4. NoV genome sequencing

To sequence the whole genome of the detected NoV strain, reverse transcription PCR (RT-PCR) was performed using the One-step RT-PCR kit (Qiagen) with 10 pairs of newly designed primer sets (Table 2.2). Each of primer binding regions is indicated with Lordsdale strain (GenBank accession no. X86557). Ten PCR products were amplified and PCR products were then analyzed by electrophoresis on 1.5% agarose gel containing ethidium bromide. The amplified PCR products were purified using HiYield™ Gel/PCR DNA Extraction kit (RBC, Taipei, Taiwan). Then, the products were cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA) and were sequenced (Genotech, Daejeon, South Korea). From sequencing of PCR products, they were assembled by sequence overlapping to complete genome sequencing using DNASTAR (DNASTAR, Madison, WI, USA).

2.1.5. Bioinformatics analysis

The open reading frames (ORFs) were predicted using GeneMarkS program (http://exon.gatech.edu/genemarks.cgi). The annotation and functional analysis of the predicted ORFs were conducted using BLASTX program (http://blast.ncbi.nlm.nih.gov/blast). Comparative phylogenetic

analysis was performed using MEGA5 with neighbor joining method (http://www.megasoftware.net/).

2.1.6. Nucleotide sequence accession number.

Complete genome sequence of the isolated NoV is available under GenBank accession number, FJ514242.

Table 2.2. Designed primer sets

Primer	Sequence (5'→3')	Polarity	Region
ORF1-1F	GTG AAT GAA GAT GGC GTC TAA C	+	1-22
ORF1-1R	AGT TCC ACT GCA AGG TCC TCA G	-	999-1020
ORF1-2F	CTG AGG ACC TTG CAG TGG AAC T	+	999-1020
ORF1-2R	ATG AGG GAA CCA GTG GTG AGA GT	-	2018-2040
ORF1-3F	CGT GCT CGA GGC GCA TCG ATT T	+	1821-1842
ORF1-3R	TTG TAC TCA TCA TAC TCT TCA	-	2701-2721
ORF1-4F	CAT TGC TCG AGC ATC AGG GCT AC	+	2038-2060
ORF1-4R	TTG ACT ATC CTC GAC CAG ATG CT	-	3038-3060
ORF1-5F	AGC ATC TGG TCG AGG ATA GTC AA	+	3038-3060
ORF1-5R	GTG GCA CAT ATG ACA GTG TTT CC	-	3518-3540
ORF1-6F	GGA AAC ACT GTC ATA TGT GCC AC	+	3518-3540
ORF1-6R	CAG TCG TTC TTC CGC ATG TGG TGC GG	-	3935-3960
ORF1-7F	CTC AGC ACC AAG ACT AAA TTC TGG A	+	3635-3659
ORF1-7R	TGG GCG ATG GAA TTC CAT TGA GAG G	-	4485-4509
ORF1-8F	ATG GTT AAA TTC TCC CCA GAA CC	+	4355-4377
ORF1-8R	CCA CCT GCA TAA CCA TTG TAC AT	-	5367-5389
ORF2-F	ATG AAG ATG GCG TCG AAT GGC	+	5085-5105
ORF2-R	TTA TAA TGC ACG CCT GCG CCC CGT	-	6681-6704
ORF3-F	ATG GCT GGA GCT TTC TTT GCT	+	6704-6724
ORF3-R	AAA GAC ACT AAA GAA AGG AAA GAT	-	7532-7555

2.2. Results and Discussion

2.2.1. Complete genome sequence analysis of the isolated NoV strain

After assembly of ten PCR products in the cloned plasmids, the complete genome sequence shows that the single-stranded RNA genome contains 7,559-bp length with a GC content of 49.24% with three ORFs. The sizes of these ORFs are 5,100 bp (ORF1), 1,623 bp (ORF2), and 807 bp (ORF3). This complete genome sequence was deposited into GenBank database with accession number FJ514242 (Figure 2.1). DNA sequences of these three ORFs are shown in Figure 2.2 and Figure 2.3. In addition, complete genome sequence of this strain is also shown in Figure 2.4.

Norovirus Hu/GII-4/CUK-3/2008/KR, complete genome

GenBank: FJ514242.1

FASTA Graphics

Go to:

```
LOCUS       FJ514242               7559 bp    RNA     linear   VRL 18-APR-2011
DEFINITION  Norovirus Hu/GII-4/CUK-3/2008/KR, complete genome.
ACCESSION   FJ514242
VERSION     FJ514242.1  GI:219810312
KEYWORDS    .
SOURCE      Norovirus Hu/GII-4/CUK-3/2008/KR
  ORGANISM  Norovirus Hu/GII-4/CUK-3/2008/KR
            Viruses; ssRNA positive-strand viruses, no DNA stage;
            Caliciviridae; Norovirus.
REFERENCE   1  (bases 1 to 7559)
  AUTHORS   Park,J.W., Lee,S.G., Lee,Y.M., Jheong,W.H., Ryu,S. and Paik,S.Y.
  TITLE     Full Sequence Analysis and Characterization of the South Korean
            Norovirus GII-4 Variant CUK-3
  JOURNAL   Virol. J. 8 (1), 167 (2011)
   PUBMED   21489302
  REMARK    Publication Status: Online-Only
REFERENCE   2  (bases 1 to 7559)
  AUTHORS   Lee,S.-G., Park,S.-W. and Paik,S.-Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-2008) Microbiology, College of Medicine, The
            Catholic University of Korea, Banpo4-dong, Seocho-gu, Seoul
            137-701, South Korea
FEATURES             Location/Qualifiers
     source          1..7559
                     /organism="Norovirus Hu/GII-4/CUK-3/2008/KR"
                     /mol_type="genomic RNA"
                     /strain="Hu/GII-4/CUK-3/2008/KR"
                     /isolation_source="stool"
                     /host="Homo sapiens"
                     /db_xref="taxon:587432"
                     /country="South Korea"
                     /collection_date="2008"
                     /note="genotype: GII-4"
```

Figure 2.1. GenBank deposition of complete genome sequence of the isolated NoV strain.

CDS 5..5104

/note="contains N-terminal, NTPase, 3A, VPg, 3Cpro and 3Dpol; ORF1"
/codon_start=1
/product="polyprotein"
/protein_id="ACL36374.1"
/db_xref="GI:219810313"
/translation="MKMASNDASAAAVANSNNDTAKSSSDKMFSSMAVTFKRALGARP
KQPPPREIPQRPPRPPTPELIKKIPPPPPNGEDEVVVSYSAKDGVSGLPELSTVRQPE
ETNTAFSVPPLNQRENRDAKEPLTGTILEMWDGEIYHYGLYVERGLVLGVHKPPAAIS
LAKVELTPLSLFWRPVYTPQYLISPDTLKKLHGETFPYTAFDNNCYAFCCWVLDLNDS
WLSRRMIQRTTGFFRPYQDWNRKPLPTMDDSKLKKVANIFLCTLSSLFTRPIKDIIGK
LRPLNIINILASCDWTFAGIVESLILLAELFGVFWTPPDVSAMIAPLLGDFELQGPED
LAVELVPVVMGGIGLVLGFTKEKIGKMLSSAASTLRACKDLGAYGLEILKLVMKWFFP
KKEEANELAMVRSIEDAVLDLEAIENNHMTTLLKDKDSLATYMRTLDLEEEKARKLST
KSASPDIVGTINALLARIAAARSLVHRAKEELSSRPRPVVLMISGRPGIGKTHLAREV
AKRIAASLTGDQRVGLIPRNGVDHWDAYKGERVVLWDDYGMSNPIHDALRLQELADTC
PLTLNCDRIENKGKVFDSDVIIITTNLANPAPLDYVNFEACSRRIDFLVYAEAPEVEK
AKRDFPGQPDMWKNAFSSDFSHIKLALAPQGGFDKNGNTPHGKGVMKTLTTGSLIARA
SGLLHERLDEFELQGPALTTFNFDRNKVLAFRQLAAENKYGLMDTMKVGRQLKDVKTM
PELKQALKNISIKKCQIVYSGCTYTLESDGKGNVKVDRVQSTSVQTNNELAGALHHLR
CARIRYYVKCVQEALYSIIQIAGAAFVTTRIIKRVNIQDLWSKPQVENTEEATNKDGC
PKPKDDEEFVISSDDIKTEGKKGKNKTGRGKKHTAFSSKGLSDEEYDEYKRIREERNG
KYSIEEYLQDRDKYYEEVAIARATEEDFCEEEEAKIRQRIFRPTRKQRKEERASLGLV
TGSEIRKRNPEDFKPKGKLWADDDRSVDYNEKLSFEAPPSIWSRIVNFGSGWGFWVSP
SLFITSTHVIPQGAKEFFGIPIKQIQVHKSGEFCRLRFPKPIRTDVTGMILEEGAPEG
TVVTLLIKRSTGELMPLAARMGTHATMKIQGRTVGGQMGMLLTGSNAKSMDLGTTPGD
CGCPYIYKRGNDYVVIGVHTAAARGGNTVICATQGGEGEATLEGGDSKGTYCGAPILG
PGSAPKLSTKTKFWRSSTAPLPPGTYEPAYLGGKDPRVKDGPSLQQVMRDQLKPFTEP
RGKPPKPSVLEAAKKTIINVLEQTIDPPEKWSFAQACASLDKTTSSGHPHHMRKNDCW
NGESFTGKLADQASKANLMFEEGKNMTPVYTGALKDELVKTDKIYGKIKKRLLWGSDL
ATMIRCARAFGGLMDELKAHCVTLPIRVGMNMNEDGPIIFEKHSRYRYHYDADYSRWD
STQQRAVLAAALEIMVKFSPEPHLAQVVAEDLLSPSVVDVGDFTISINEGLPSGVPCT
SQWNSIAHWLLTLCALSEVTNLSPDIIQANSLFSFYGDDEIVSTDIKLDPEKLTAKLK
EYGLKPTRPDKTEGPLVISEDLDGLTFLRRTVTRDPAGWFGKLEQSSILRQMYWTRGP
NHEDPSESMIPHSQRPIQLMSLLGEAALHGPTFYSKISKLVIAELKEGGTDFYVPRQE
PMFRWMRFSDLSTWEGDRNLAPSFVNEDGVE"

Figure2.2. DNA sequence of ORF1 in the genome sequence of the isolated NoV strain.

```
CDS             5085..6707
                /note="VP1; ORF2"
                /codon_start=1
                /product="capsid protein"
                /protein_id="ACL36375.1"
                /db_xref="GI:219810314"
                /translation="MKMASNDANPSDGSAANLVPEVNNEVMALEPVVGAAIAAPVAGQ
                QNVIDPWIRNNFVQAPGGEFTVSPRNAPGEILWSAPLGPDLNPYLSHLARMYNGYAGG
                FEVQVILAGNAFTAGKIIFAAVPPNFPTEGLSPSQVTMFPHIIVDVRQLEPVLIPLPD
                VRNNFYHYNQSNDPTIKLIAMLYTPLRANNAGEDVFTVSCRVLTRPSPDFDFIFLVPP
                TVESRTKPFTVPILTVEEMTNSRFPIPLEKLFTGPSSAFVVQPQNGRCTTDGVLLGTT
                QLSPVNICTFRGDVTHIAGSRNYTMNLASLNWNNYDPTEEVPAPLGTPDFVGKIQGVL
                TQTTKGDGSTRGHKATVYTGSAPFTPKLGSVQFSTDTENDFETHQNTKFTPVGVIQDG
                STTHRNEPQQWVLPSYSGRNVQNVHLAPAVAPTFPGEQLLFFRSTMPGCSGYPNMDLD
                CLLPQEWVQHFYQEAAPAQSDVALLRFVNPDTGRVLFECKLHKSGYVTVAHTGQHDLV
                IPPNGYFRFDSWVNQFYTLAPMGNGTGRRRAL"
CDS             6707..7513
                /note="ORF3"
                /codon_start=1
                /product="VP2"
                /protein_id="ACL36376.1"
                /db_xref="GI:219810315"
                /translation="MAGAFFAGLASDVLSSGLGSLINAGAGAINQKIDFENNRKLQQA
                SFQFSSNLQQASFQHDKEMLQAQIEATKKLQQEMMKVKQAMLLEGGFSETDAARGAIN
                APMTKALDWSGTRYWAPDARTTTYNAGRFSTPQPSGALPGRNNPRTPTPARGSPSTSS
                NASTATSIYSNQTVSTRLGSTAGSGTNVSSPPSTARTRSWVEDQNRNLSPFMRGAHNI
                SFVTPPSSRSSSQGTVSTVPKEVLDSWTGAFNTRRQPLFAHIRRRGESRV"
```

Figure 2.3. DNA sequences of ORF2 and ORF3 in the genome sequence of the isolated NoV strain.

```
ORIGIN
        1 gtgaatgaag atggcgtcta acgacgcttc cgctgccgct gttgctaaca gcaacaacga
       61 caccgcaaaa tcttcaagtg acaaaatgtt ttctagcatg gctgtcactt ttaaacgagc
      121 cctcgggggcg cggcctaaac agccccccccc gagggaaata ccacaaagac ccccacgacc
      181 acctactcca gaactgatca aaaagatccc tcctccccccg cccaacggag aggatgaagt
      241 agtggtttct tatagtgcca aagatggcgt ttccggttta cctgagcttt ccaccgtcag
      301 gcaaccggaa gaaaccaata cggccttcag tgtccctcca ctcaatcaga gggagaatag
      361 ggatgctaag gaaccactga ctggaacaat cctggaaatg tgggatggag aaatctacca
      421 ttatggcctg tatgttgagc gaggtcttgt gctgggtgtg cacaaaccac cagctgccat
      481 tagcctcgcc aaggtcgaac taacaccact ctccttgttc tggagacctg tgtacactcc
      541 tcagtacctc atctctccag acactctcaa gaaattacac ggagaaacat tcccctacac
      601 agccttcgac aacaattgct atgccttttg ttgctgggtc ctggatctaa acgactcgtg
      661 gctgagtagg agaatgatcc agagaacaac tggcttcttc agaccctacc aagattggaa
      721 taggaaaccc ctccccacta tggatgattc caaattgaag aaggtagcta acatatttct
      781 gtgcaccctg tcttcgctat ttaccaggcc cataaaagac ataatagggа agttaaggcc
      841 tctcaacatc atcaacatcc tggcttcatg tgattggact ttcgcaggca tagtggagtc
      901 cttgatactc ttggcagagc tctttggagt cttctggaca cccccagatg tgtctgcgat
      961 gattgccccc ttactcggtg atttcgagtt acaaggacct gaggaccttg cagtggaact
     1021 cgtccctgta gtaatggggg gaattggttt ggtgctagga ttcaccaaag agaagattgg
     1081 aaaaatgttg tcatctgctg catctacctt gagagcttgt aaagatcttg gtgcgtatgg
     1141 gctagagatc ctaaagttag tcatgaagtg gttcttcccg aagaaagagg aagcaaatga
     1201 actggctatg gtgagatcca tcgaggatgc agtactggac cttgaggcaa ttgaaaacaa
     1261 ccatatgacc accttgctca aagacaaaga cagcctggca acctatatga gaacccttga
     1321 cctcgaggaa gagaaagcca gaaaactctc aaccaagtct gcttcacctg acatcgtggg
     1381 cacaatcaac gcccttctgg cgagaatcgc cgctgcacgc tccctggtgc accgagcgaa
     1441 ggaggagctt tccagcagac caagacctgt agtcttgatg atatcaggca gaccaggaat
     1501 agggaagacc caccttgcta gggaagtggc taagagaatc gcagcctccc tcacaggaga
     1561 ccagcgtgta ggcctcatcc cacgcaatgg cgtcgatcac tgggatgcgt acaaggggga
     1621 gagggtcgtc ctatgggacg actatggaat gagcaatccc atccacgacg ccctcaggct
     1681 gcaagaactc gctgacactt gccccctcac tctaaattgt gacaggattg agaataaagg
     1741 aaaggtcttt gacagcgatg tcattattat cactactaac ctggccaacc cagcaccact
     1801 ggactatgtc aactttgaag cgtgctcgag gcgcatcgat ttcctcgtgt atgcagaagc
     1861 ccccgaggtc gaaaaggcaa agcgtgactt cccgggtcaa cctgacatgt ggaaaaacgc
     1921 ttttagttct gatttctcac acataaaatt ggcactggct ccacaaggtg gttttgataa
     1981 gaacgggaac accccacacg ggaagggcgt catgaagact ctcaccactg gttccctcat
     2041 tgctcgagca tcagggctac tccatgagag gttggatgag tttgaactac agggcccagc
     2101 tcttaccacc ttcaactttg accgcaacaa agtgcttgct ttcaggcagc ttgctgctga
     2161 aaacaaatac gggttgatgg acacaatgaa agttgggagg cagctcaagg atgtcaaaac
     2221 catgccagaa cttaaacaag cactcaagaa tatctcaatc aagaagtgcc agattgtgta
     2281 tagtggttgc acctacacac ttgagtctga tggcaagggc aatgtgaaag ttgacagagt
     2341 tcagagcacc tccgttcaga ccaacaatga gctggctggc gccctgcacc atctaaggtg
     2401 cgccagaatc aggtactatg tcaagtgtgt tcaggaggcc ctgtattcta tcatccagat
     2461 cgctggggct gcatttgtca ccacgcgcat catcaagcgc gtgaacattc aagacttatg
     2521 gtccaagcca caagtggaaa acacagagga ggccaccaac aaggacgggt gcccgaaacc
     2581 caaagatgat gaggagttcg tcatttcatc tgacgacatt aaaactgagg gtaagaaagg
     2641 gaagaacaag actggccgtg gcaagaagca cacagccttc tcaagtaaag gtctcagtga
     2701 tgaagagtat gatgagtaca agagaattag agaggaaagg aatggcaagt actccataga
     2761 agagtacctt caggacaggg acaaatacta tgaggaggtg gccattgcca gggcgaccga
     2821 ggaagacttc tgtgaagagg aggaggccaa gatccggcaa aggatcttca gaccaacaag
     2881 gaagcaacgc aaggaagaaa gagcttctct cggtttagtc acaggttctg aaattaggaa
```

Figure 2.4. Complete genome sequence of the isolated NoV strain (continued).

```
2941 gaggaaccca gaagacttca agcctaaggg gaaactatgg gctgacgatg acagaagtgt
3001 ggactacaat gaaaaactca gctttgaggc cccaccaagc atctggtcga ggatagtcaa
3061 ctttggctca ggttgggggct tctgggtctc ccccagcctg ttcataacat caacccacgt
3121 cataccccag ggcgcaaagg agttctttgg aatccccatc aaacaaattc aggtacacaa
3181 gtcaggcgaa ttctgtcgct tgaggttccc aaaaccaatc aggactgatg tgactggcat
3241 gatcttggaa gaaggtgcgc ccgaaggcac cgtggtcaca ctactcatca aacggtctac
3301 tggagaactc atgcccctag cagctagaat ggggacccat gcaaccatga aaatccaagg
3361 gcgcactgtt ggaggtcaga tgggcatgct tctgacagga tccaacgcca aaagcatgga
3421 tctaggcacc acaccaggtg attgcggctg tccctacatc tacaagagag gaaatgacta
3481 tgtggtcatt ggagtccaca cggctgccgc tcgtgggggа aacactgtca tatgtgccac
3541 ccaggggggt gaggggggaag ctacacttga aggtggtgac agtaagggaa catactgtgg
3601 tgcaccaatc ctaggcccag ggagtgcccc aaaacttagc accaaaacca aattctggag
3661 atcatccaca gcaccactcc cacctggcac ctatgaacca gcctaccttg gtggcaagga
3721 ccccagagtc aaggatggcc cctcgttgca gcaagtcatg agggaccagc tgaaaccatt
3781 tacagagcct aggggtaagc caccaaagcc aagtgtatta gaagctgcca agaaaaccat
3841 catcaatgtc cttgaacaaa caattgaccc acctgagaag tggtcgttcg cacaagcttg
3901 cgcgtccctt gacaagacca cttctagcgg ccatccgcac cacatgcgga aaaacgactg
3961 ctggaacggg gaatccttca caggcaagct ggcagaccag gcttccaagg ccaacctgat
4021 gtttgaagaa gggaagaaca tgaccccagt ctacacaggt gcacttaagg atgaattagt
4081 caaaactgac aagatttatg gtaagatcaa gaagaggctt ctctgggggct cggatttggc
4141 aaccatgatc cggtgtgctc gagcattcgg aggcctaatg gatgaactca aagcacactg
4201 tgtcacactt cctatcagag ttggtatgaa tatgaatgag gatggcccca tcatcttcga
4261 gaagcattcc aggtacaggt accactatga tgctgattac tctcggtggg attcaacaca
4321 acagagagcc gtgctggcag ccgctctaga aatcatggtt aaattctccc cagaaccaca
4381 tttggctcag gtagtagcag aagaccttct ttctcctagc gtggtggatg tgggtgactt
4441 cacaatatca atcaatgagg gccttccctc tggggtgccc tgcacctccc aatggaactc
4501 catcgcccac tggcttctca ctctctgtgc actctccgaa gttacaaatt tgtccccaga
4561 catcatacag gctaattctc tcttctcctt ctatggtgat gatgaaattg ttagtacaga
4621 cataaaatta gacccagaga agttgacagc aaagcttaag gaatacgggt tgaaaccaac
4681 ccgccctgat aaaactgaag gacctcttgt tatttctgag gacttagatg gtttgacttt
4741 cctgcggaga actgtgaccc gcgacccagc tggttggttt ggaaaactgg agcagagttc
4801 aatactcagg caaatgtact ggactagggg ccccaaccat gaagacccat ctgaatcaat
4861 gattccacac tctcaaagac ccatacaatt gatgtcctta ctgggagagg ccgcactcca
4921 cggcccaaca ttctacagta aaatcagcaa attagttatt gcagagctaa aagaaggtgg
4981 tacggatttt tacgtgccca ggcaagagcc aatgttcaga tggatgagat tctcagacct
5041 gagcacgtgg gagggcgatc gcaatctggc tcccagtttt gtgaatgaag atggcgtcga
5101 atgacgccaa cccatctgat gggtccgcag ccaacctcgt cccagaggtc aacaatgagg
5161 ttatggcttt ggagcccgtt gtcggtgccg ctattgcggc gcctgtagcg ggccaacaaa
5221 atgtgattga cccctggatt agaaataatt ttgtacaagc ccctggtgga gagttcacag
5281 tatcccctag aaacgctcca ggtgaaatac tatggagcgc gcccttaggc cctgatctga
5341 atccctacct atcccatttg gccagaatgt ataatggtta tgcaggtggt tttgaagtgc
5401 aggtgatcct cgcggggaac gcgttcaccg ccggaaaaat tatatttgca gcagtcccac
5461 caaattttcc aactgaaggc ttgagtccca gccaggtcac tatgttcccc cacataatag
5521 tggatgttag gcaattggaa cctgtgttga tccccttacc tgatgttagg aataacttct
5581 accactataa tcagtcaaat gatcctacca ttaaattgat agcaatgctg tatacaccac
5641 ttagggctaa taatgctggg gaggatgtct tcacagtctc ttgtcgagtt ctcacgaggc
5701 catcccctga ttttgatttt atatttttgg tgccacctac agttgagtca agaactaaac
5761 catttactgt cccaatccta actgttgaag aaatgaccaa ttcaagattc cccattcctt
5821 tggaaaaatt gtttacgggt cccagcagtg cctttgttgt tcaaccacaa aatggcagat
5881 gcacgactga tggcgtgctc ttaggcacca cccaactgtc tcctgtcaac atctgcacct
```

Figure 2.4. Complete genome sequence of the isolated NoV strain (continued).

```
5941 tcagagggga tgtcacccac attgcaggtt ctcgtaatta cacaatgaat ttggcttctc
6001 taaattggaa caattatgac ccaacagaag aagttccagc ccctctagga actccagatt
6061 tcgtgggaaa gatccaaggt gtgctcactc aaaccacaaa gggagatggt tcgacccgtg
6121 gccataaagc tacagtttac actgggagtg ccccctttac tccaaagctg ggcagtgttc
6181 aattcagtac tgatacagaa aatgattttg aaactcacca aaacacaaaa ttcaccccag
6241 tcggtgtcat ccaggatggt agcaccaccc accgaaatga accccaacaa tgggtgctcc
6301 caagttattc aggtagaaat gtccaaaatg tacacctagc ccctgctgta gcccccactt
6361 ttccgggtga acaacttctt ttctttaggt ccactatgcc cggatgcagc gggtatccca
6421 acatggattt ggattgccta ctcccccagg agtgggtgca gcacttctac caagaggcag
6481 ctccagcaca atctgatgtg gctctattga gatttgtgaa tccagacacg ggtagggtcc
6541 tgtttgagtg caagcttcat aaatcaggct atgtcacagt ggctcacacc ggccagcatg
6601 atttggtcat ccccccccaat ggttatttta ggtttgattc ctgggttaat caattctaca
6661 cacttgcccc catgggaaat ggaacggggc gcaggcgtgc attataatgg ctggagcttt
6721 ctttgctgga ttggcatctg atgtccttag ctctggactt ggttccctaa tcaatgctgg
6781 ggctggggct atcaaccaaa agattgattt tgaaaataac agaaaattgc agcaagcttc
6841 cttccagttt agcagtaatc tacaacaggc ttcctttcaa catgataaag agatgctcca
6901 agcacaaatt gaggccacta aaaagttgca acaggaaatg atgaaagtca aacaggcaat
6961 gcttttagaa ggtggattct ctgaaacaga tgcagcccgt ggggcaatca acgcccctat
7021 gacaaaggct ttggactgga gcgggacaag gtactgggcc cctgatgcta ggactacaac
7081 atacaatgca ggccgttttt ccacccctca accttcgggg gcactgccag gaagaaacaa
7141 tcccaggact cctacccccg ctcggggctc ccccagcaca tcttctaatg cttctactgc
7201 tacttctata tattcaaatc aaactgtttc aacgagactt ggttctacag ctggttctgg
7261 caccaatgtc tcgagtcccc cgtcaactgc aaggactagg agttgggttg aggatcaaaa
7321 cagaaatttg tcacctttca tgaggggggc tcacaacata tcgtttgtca ccccaccatc
7381 tagcagatcc tctagccaag gcacagtctc aaccgtgcct aaagaagttt tggactcctg
7441 gactggcgct ttcaacacgc gcaggcagcc tctcttcgct cacattcgta ggcgagggga
7501 gtcacgggtg taatgtgaaa aggcaaaatt gattatcttt tcctttcttt agtgtctttt
```

Figure 2.4. Complete genome sequence of the isolated NoV strain.

2.2.2. Comparison of ORF2 with reference strains

The predicted ORF2 may encode the viral capsid protein, which is generally used for comparative analysis of viruses. Therefore, the nucleotide and amino acid sequences of ORF2 were compared with the DNA and protein sequences of viral capsid proteins from other reference viruses. This comparative analysis showed that this isolated NoV strain had the highest sequence similarity to the NoV GII-4 reference strain X76717 (GenBank accession no.). The nucleotide similarity of the isolated NoV strain was from 60.1 to 88.0% with the selected reference strains. A protein database search (BLASTX) with the capsid protein encoded by ORF2 yielded 59.3 to 92.0% similarity. However, because the highest similarity (92%) between ORF2 of the isolated NoV strain and Bristol-GBR-93 of GII-4 genotype is lower than the cut-off value of amino acid similarity, the isolated NoV strain is not identical to Bristol-GBR-93, but it belongs to GII-4 genotype (Table 2.3).

Table 2.3. Nucleotide and amino acid sequence similarities between CUK-3 strain and the full length ORF2 sequences from different Genogroup GII reference strains

Genotype	Strain	Accession No.	Similarity (%)	
GII-1	H awaii-USA94	U07611	64.9[a]	64.6[b]
GII-2	Msham-GBR95	X81879	62.8	63.9
GII-3	Toronto-CAN93	U02030	64.8	67.4
GII-4	Bristol-GBR-93	X76716	88.0	92.0
GII-5	Hilingd-GBR00	AJ277607	64.6	63.9
GII-6	Seacrof-GBR00	AJ277620	63.0	63.9
GII-7	Leeds-GBR00	AJ277608	64.3	64.8
GII-8	Amstdam-NLD99	AF195848	64.3	64.8
GII-9	VABeach-USA01	AY038599	63.7	63.5
GII-10	Eufurt-DEU01	AF421118	63.5	64.8
GII-11	SW918-JPN01	AB074893	62.3	64.6
GII-12	Wortley-GBR00	AJ277618	63.3	63.9
GII-13	Faytvil-USA02	AY113106	64.5	64.3
GII-14	M7-USA03	AY130761	64.3	64.8
GII-15	J23-USA02	AY130762	60.1	59.3
GII-16	Tiffin-USA03	AY502010	64.1	65.9
GII-17	CSE1-USA03	AY502009	63.7	65.2

[a], The nucleotide sequences of the ORF2 regions starting at nucleotide 1620 were analyzed.
[b], The amino acid sequences of the ORF2 regions starting at amino acid 540 were analyzed.

2.2.3. Analysis of comparative evolutionary relationship in NoV

To analyze the genetic relationship between the isolated NoV strain and the other NoV variants reported worldwide, the sequences of three ORFs and the whole genome were processed with multiple sequence alignment using phylogenetic tree. Each DNA sequence of ORF1 (5,100 bp), ORF2 (1,623 bp), ORF3 (807 bp) as well as the whole genome sequence (about 7.5 kb) was compared with other DNA sequences of NoV strains in the Cluster I, which was reported in 2006. From the comparative analysis, the NoV strains in Cluster I, which have ORFs homologous to ORF1 of the isolated NoV strain, contain Hiroshoma1, Kumamoto1, Saga1, Hokkaido1, Aomori4, Miayagi5, Akita1, Aichi3 and showed the highest similarity range (98.0 to 98.9%) (Figure 2.5). In the case of ORF2, the GII-4 variants from Japan in Cluster I (2006b) contain very similar ORFs with the highest similarity range (97.7 to 99.0%). However, other GII-4 variants in Cluster II (2006a) contain relatively low similarity range (91.7 to 92.3%) (Figure 2.6). Furthermore, like ORF2, comparative sequence analysis of ORF3 showed the highest similarity to ORFs of GII-4 variants in the Cluster I with 97.9% to 99.3% similarity. Interestingly, 02/03 group contains the GII-4 variants in Europe showed higher similarity (93.1 - 93.6%) than other GII-4 variants from Japan in the Cluster II (Figure 2.7). Comparative analysis of the complete genome sequences of the isolated NoV strain and other NoV strains revealed that the

genome sequence of the isolated NoV strain has highest similarity (98.3%) to the 2006 epidemic strain, Aichi3/2006/JP strain. However, Ehime/5-30/2005/JP strain and Lordsdale/1993/UK strain showed the lowest similarity (89.2%) (Figure 2.8).

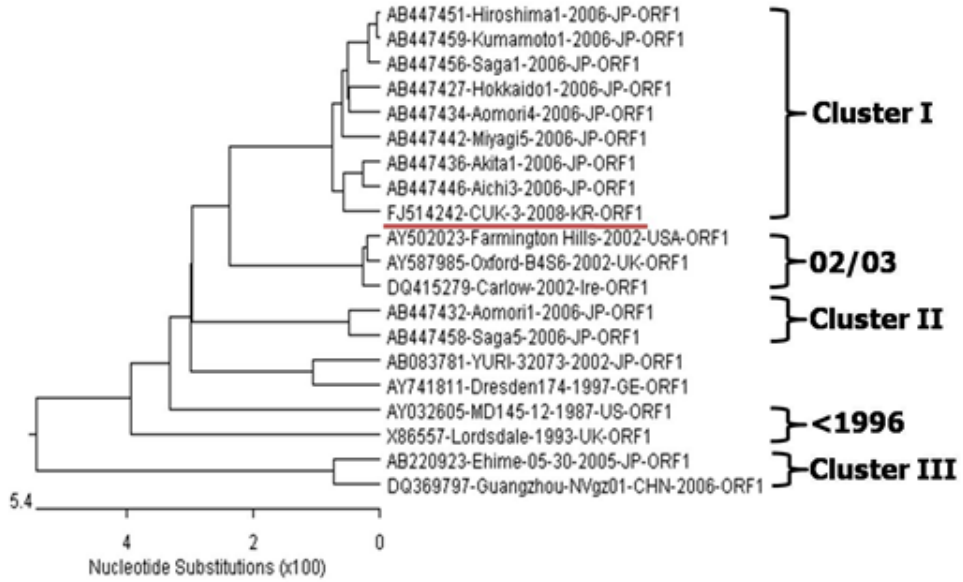


Figure 2.5. Phylogenetic tree constructed on the nucleotide sequence of ORF1 (5.1 kb).

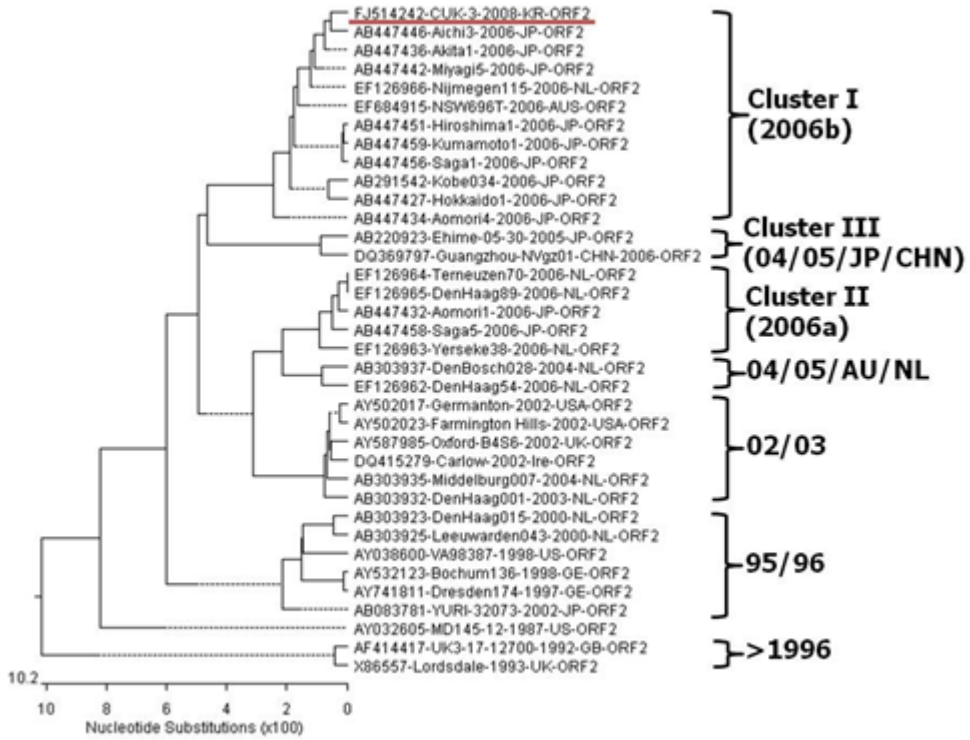


Figure 2.6. Phylogenetic tree constructed on the nucleotide sequence of ORF2 (about 1.6 kb).

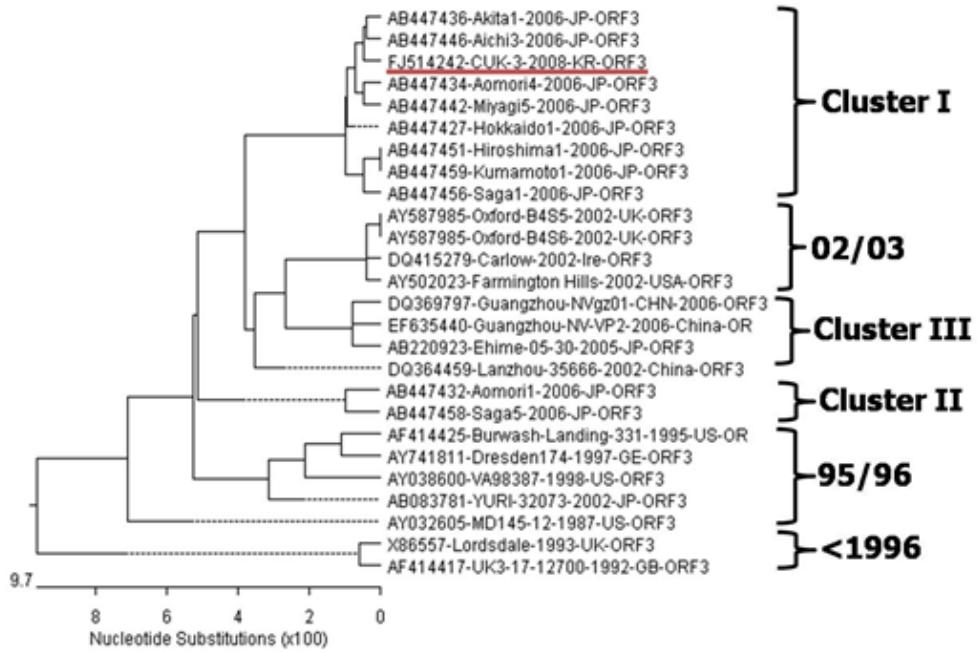


Figure 2.7. Phylogenetic tree constructed on the nucleotide sequence of ORF3 (about 0.8 kb).

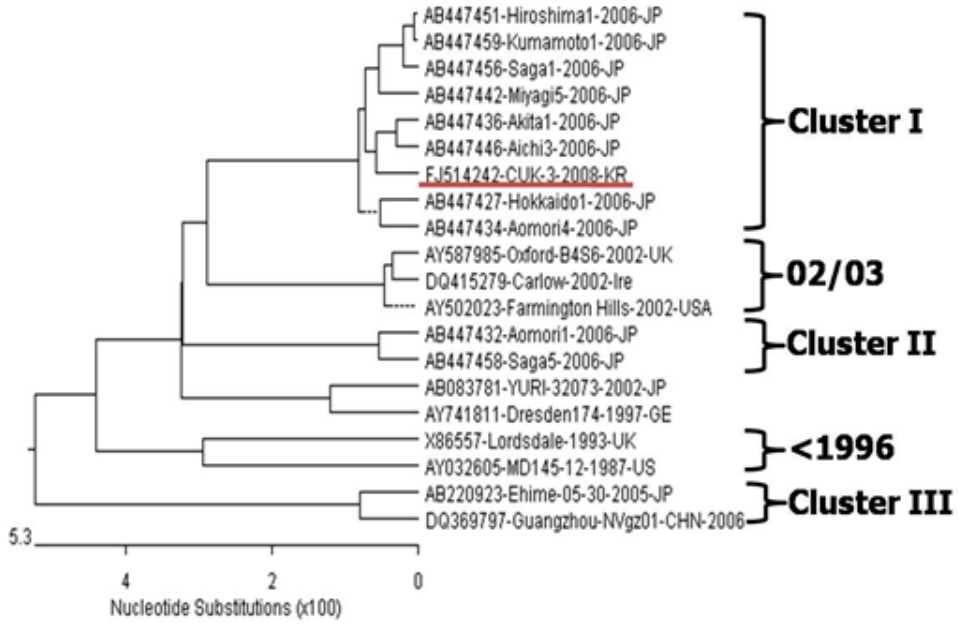


Figure 2.8. Phylogenetic tree constructed on the nucleotide sequence of complete genome (about 7.5 kb).

These phylogenetic tree analyses of three ORFs and complete genome sequence of the isolated NoV strain suggests that it could belong to the GII-4/2006b cluster. Furthermore, subsequent multiple sequence alignment of RdRp in various NoV strains also suggests that the isolated NoV strain is highly similar to NoV strains in the GII-4/2006b cluster, substantiating this (Figure 2.9). Additional comparative analysis of the GDD motifs (ANNNTG) regarding catalytic domain of RdRp protein in various NoV strains also supports this close relationship (Figure 2.9). In this Cluster, these NoV strains were common in Hong Kong, Japan, and other regions of North East Asia during the same period, 2006 (Motomura *et al.* 2008; Ho *et al.* 2007). However, the NoV strain in this study was isolated in 2008, indicating the reemergence of GII-4 variant in South Korea.

In Europe and Asia, most of NoV outbreaks occurred since 2002 have been caused by GII-4 variant strains (Figure 2.9). Interestingly, emergences of pandemic NoV GII-4 variants have been reported every two years such as 2002, 2004, 2006 and 2008 (Motomura *et al.* 2008; Ho *et al.* 2007; Tu *et al.* 2008). In 2008, the GII-4 NoV variant has also been firstly identified in South Korea and this type of NoV variants is now rapidly spreading worldwide.

Figure 2.9. Partial sequence comparison of the RdRp gene. The boxed sequence is the conserved GDD motif, ANNNTG. Nucleotide positions are referenced to the Lordsdale-1993-UK sequence.

Amino acid sequence of the capsid region encoded by ORF2 was compared with a total of 39 GII-4 variants reported from various countries including US, UK, Japan, China, Germany etc. Comparative analysis of 540 amino acids encoded in ORF2 with those of GII-4 variants showed the pattern of changes of two amino acid in first P1 domain, aa238 (S $_{\text{GII-4->2002}}\rightarrow$ T $_{\text{GII-4-<2002}}$), aa250 (Y $_{\text{GII-4->2002}}\rightarrow$ F $_{\text{GII-4-<2002}}$). The P2 domain showed five substitutions of amino acids, aa280 (A $_{\text{GII-4->2002}}\rightarrow$ P $_{\text{GII-4-<2002}}$), aa346 (A $_{\text{GII-4->2002}}\rightarrow$ P $_{\text{GII-4-<2002}}$), aa367(F $_{\text{GII-4->1995}}\rightarrow$ Y $_{\text{GII-4->2002}}\rightarrow$ F $_{\text{GII-4-<2002}}$), aa376(Q $_{\text{GII-4->2002}}\rightarrow$ E $_{\text{GII-4-<2002}}$), and aa394 (G or S $_{\text{GII-4->2006}}\rightarrow$ T $_{\text{GII-4-<2006}}$). In addition, the second P1 domain also showed two amino acid substitutions, aa459 (L $_{\text{GII-4->1995}}\rightarrow$ Q $_{\text{GII-4->2002}}$) and aa504 (P $_{\text{GII-4->2002}}\rightarrow$ Q $_{\text{GII-4-<2002}}$). These analytic data showed that there had already been numerous mutations, which led to the production of a wide range of variants, already before 2002 and produced various variants. And as shown on the case of aa394 of P2 domain which deleted one of amino acids, G or S, and then substituted to T after 2006.

Although previous studies have indicated that the sequence for the NoV capsid region is useful for the diagnosis of NoV variants, no studies have analyzed the RdRp region of the South Korean NoV GII-4 variant. This is the first report that describes the full-length sequence of a NoV GII-4 variant that was isolated from clinical samples in South Korea. We suggest

that this sequence can be used as a standard for the comparison of full-length NoV GII-4 variant sequence with other NoV strains.

The information acquired from whole genome sequencing in this study can be useful not only for a more accurate diagnoses of NoVs but also for the basic research for the elucidation of NoV genetic functions. Furthermore, it will be helpful for the prediction of newly appearing pandemic variants through comparison with GII-4 variants in neighboring countries, for fundamental research for vaccine development, and eventually, for the field of public health through with the provision of new emerging strains of NoV.

2.3. Conclusion

A novel NoV strain was isolated in 2008 from a patient in South Korea and to further understand this strain, complete genome sequencing was conducted. The comparative genomic analysis of this isolated NoV strain with many reference NoV strains and their phylogenetic analysis revealed that this strain belongs to GII-4 genotype. Furthermore, ORF prediction of the genome of this isolated NoV strain showed that this genome contains three ORFs. Comparative phylogenetic analyses of these ORFs with genes encoded in other reference strains showed that these ORFs are very similar to genes in Cluster I (2006b), suggesting that this isolated NoV strain may be in the cluster of GII-4/2006b. This strain is the first NoV strain identified in South Korea. Interestingly, while other reference NoV strains were found worldwide in 2006, this isolated NoV strain was isolated in 2008, South Korea, suggesting that this GII-4 type NoV strains are rapidly spreading worldwide.

PART III: Prevalence of NoVs from Groundwater in South Korea and Their Evolutionary Relationships

3.1. Materials and Methods

3.1.1. Overview of detection of NoVs in groundwater

The process of identification of NoV from groundwater is mainly involved in several steps of sampling, virus concentration, nucleic acids extraction, reverse transcription-PCR (RT-PCR), sequencing and phylogenetic analysis. And the whole steps were based on the guidelines for NoVs Detection in Groundwater (Korea Food and Drug Association, KFDA) (Figure 3.1).

3.1.2. Sampling

To investigate NoV occurrence in South Korea, 1,090 sampling sites were chosen, located in schools and preschools. The standard sampling system was previously developed by collaboration of KFDA, 3M and Sanigen (Figure 3.2). Using this standard sampling system, most of groundwater samples were collected during 8 months period from April, 2010 to January, 2011 (Figure 3.3). Sampling regions include Gyeonggi, Gangwon, Chungcheong, Jeolla, Gyeongsang and Jeju provinces (Figure 3.4).

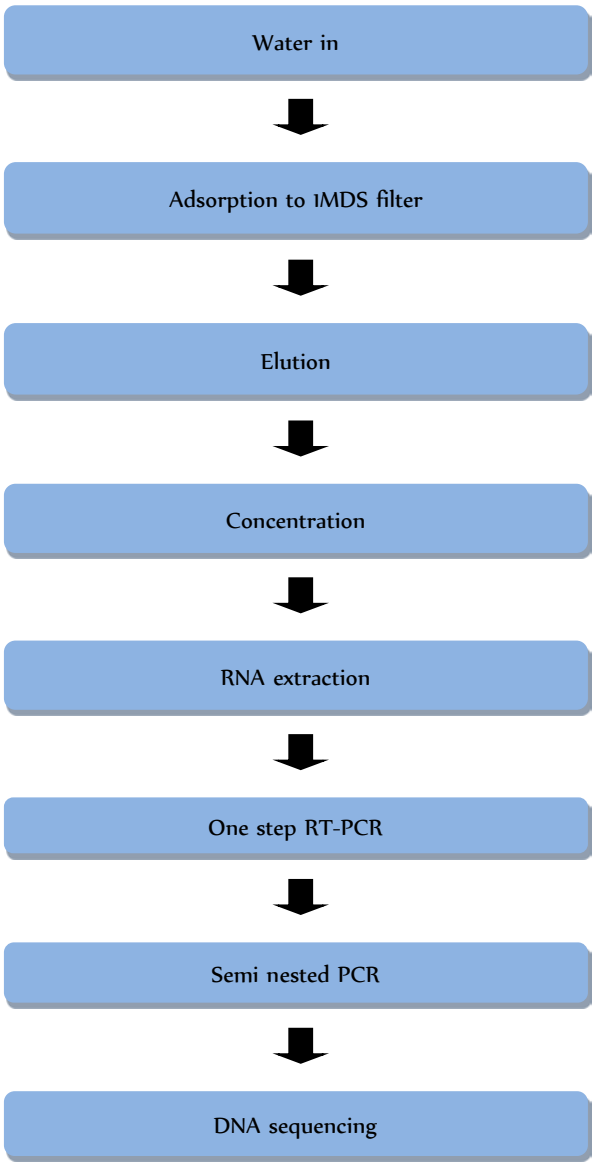


Figure 3.1. Process flow chart for detection and identification of NoV in water.

Figure 3.2. Standard sampling system.

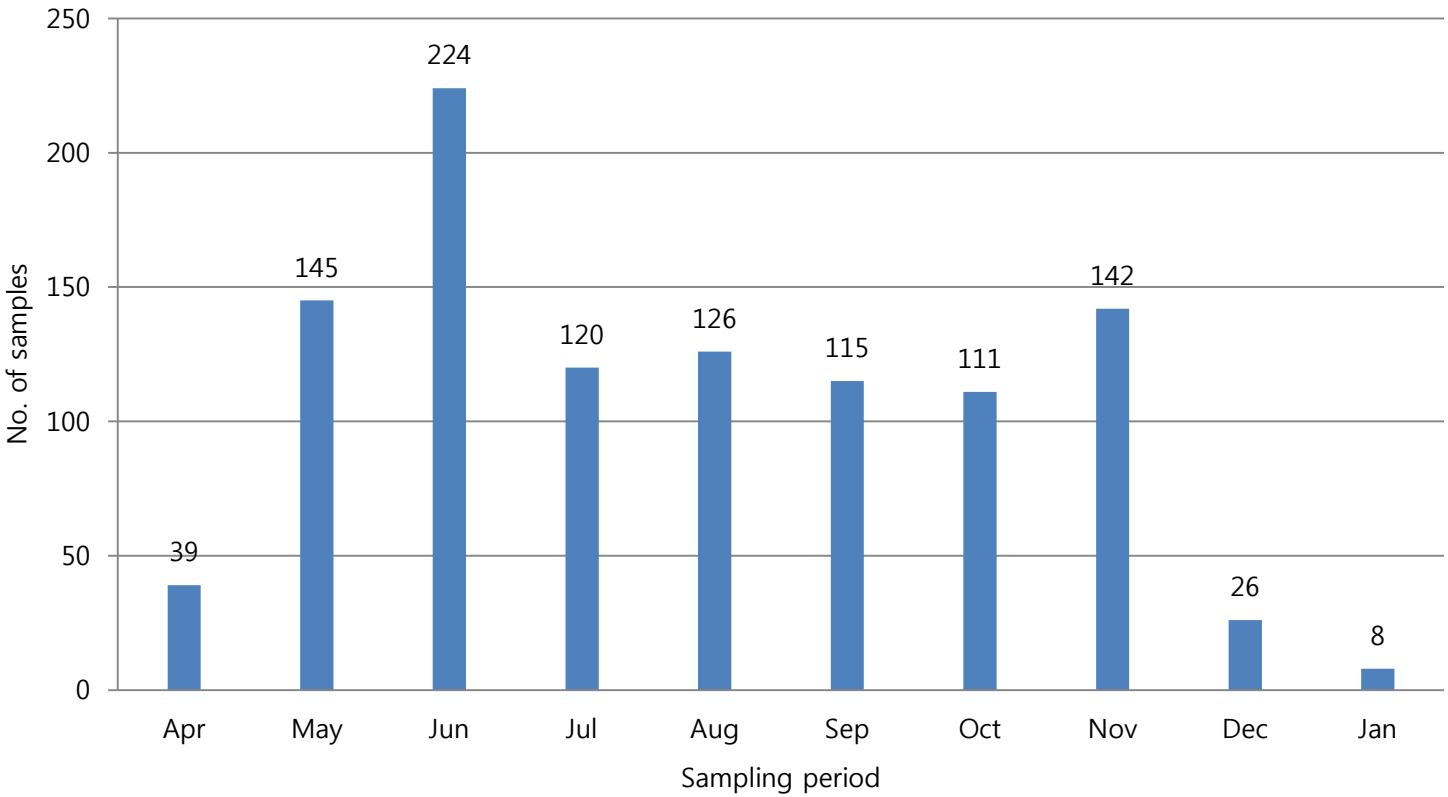


Figure 3.3. Details of sampling period.

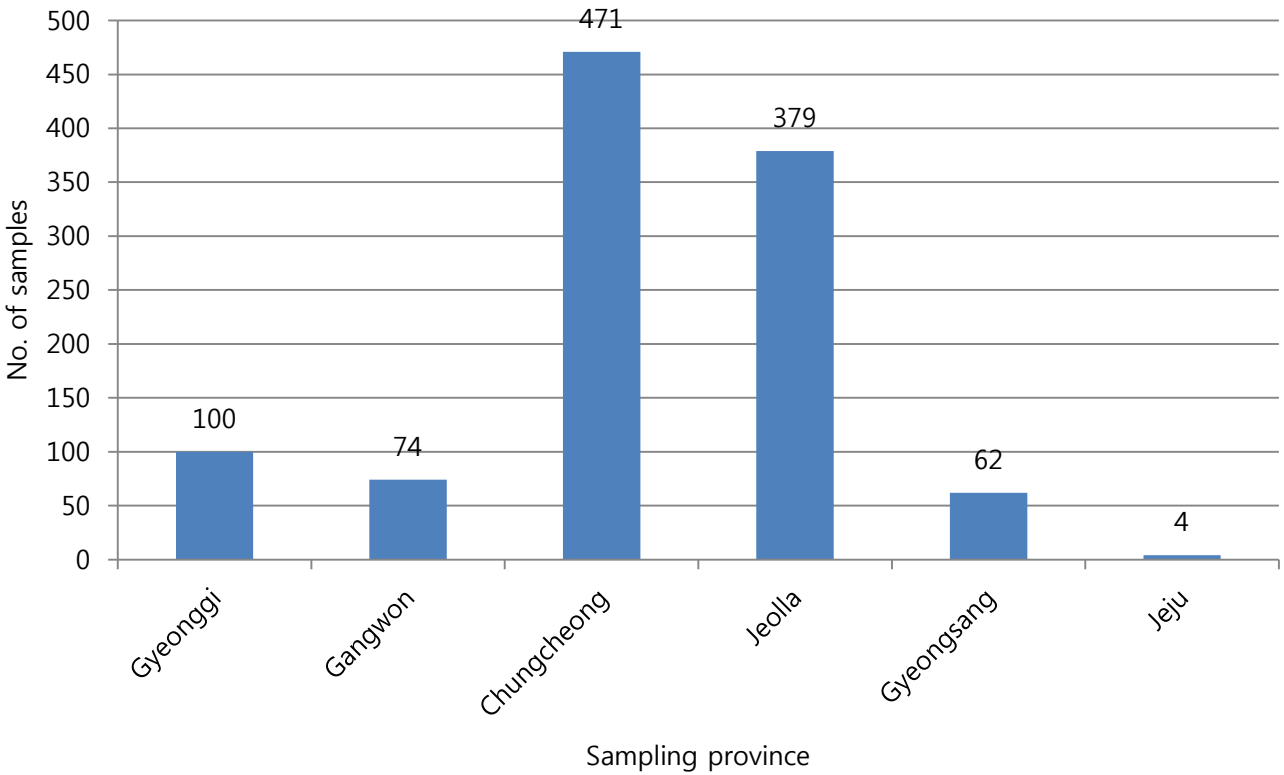


Figure 3.4. Details of sampling regions.

3.1.3. Virus concentration and nucleic acids extraction

Adsorption–elution method was employed to concentrate virus particles present in the water samples based on the Guidelines for NoVs Detection in Groundwater (KFDA). Briefly, potential virus in the water samples was collected bypassing from 1,500 L to 1,800 L of groundwater through a 1MDS filter in the standard sampling system. After bypass of the groundwater, NoVs were eluted from the filter by submerging in 1.5% beef extract (pH 9.5) for 1 to 5 min in the filter housing, and the elute containing NoVs was transferred to a sterile beaker. After the elution process with the beef extract was repeated, the pH of elute sample was adjusted to 7.0 to 7.5 by using 1 M HCl. The pH of elute sample was gradually adjusted to 3.5± 0.1 and stirred slowly for 30 min at room temperature. Then, the elute sample was centrifuged at 2,500 × g for 15 min at 4°C, and the pellet was resuspended in 30 mL of 0.15 M sodium phosphate buffer, pH 9.0 to 9.5. The resuspended solution was centrifuged again at 7,000 × g for 10 min at 4°C and the supernatant was filtered by using sterile, 0.22-μm pore size cellulose acetate filter (Millipore, Billerica, MA, USA). Finally, the concentrated samples were stored at -70°C until nucleic acid extraction. Viral nucleic acid extraction was carried out using the QIAamp® Viral RNA Mini kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol.

3.1.4. One-step reverse transcription-PCR (RT-PCR) and nested PCR

One-step RT–PCR reaction was performed as described in the manual for the Guidelines for NoVs Detection in Groundwater (KFDA). In brief, 5μL of RNA extract was mixed with 20 μL of RT–PCR pre-mix, containing 20 pmol of each GI-F1M/GI-R1M primer set for detection of GI type (Table 3.1) or GIIF1M/GII-R1M primer set for detection of GII type(Table 3.2). Expected PCR product sizes of G1 type and GII type are 313 bp and 310 bp, respectively. For conversion of total RNA to cDNA, reverse transcription was conducted: the mixture was incubated for 30 min at 45°C followed by reverse transcriptase inactivation and subsequent cDNA denaturation for 5 min at 94°C. And then, cDNA was routinely amplified by general PCR with the following conditions: 35 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 90 s. A final extension step was done at 72°C for 7 min (Table 3.3).

To further increase PCR-amplified NoV DNA, the nested PCR was conducted using the PCR-amplified NoV DNA as a template. For this second-round nested PCR, 2μL of the first-round PCR product was added to 48 μL of the second-round PCR mixture containing 20 pmol of each GIF2/GIR1M or GIIF3M/FIIR1M primer sets (Table 3.1 and 3.2). Denaturation of DNA templates was carried out for 5 min at 94°C, and the subsequent PCR amplification was carried out for 25 cycles of 94°C for 30 s,

55°C for 30 s and 72°C for 90 s. A final extension step was done at 72°C for 7 min (Table 3.4).

3.1.5. Determination of genogroup of NoV using PCR with genogroup-specific primer sets

To determine the genogroup of NoV between GI and GII, genogroup-specific primer sets (Table 3.1 and 3.2), 1.5% standard agarose gel electrophoresis of amplified PCR products using two primer sets was conducted. According to the presence of specific DNA band (313-bp for GI and 310-bp for GII) from PCR with GI- or GII-genogroup specific primer in the agarose gel, the genogroup was determined.

Table 3.1. Primer sequence for NoV GI detection

Primer	Sequence (5'→3')*	Polarity
GIF1M	CTG CCC GAA TTY GTA AAT GAT GAT	+
GIF2	ATG ATGATG GCG TCT AAG GAC GC	+
GIR1M	CCA ACC CAR CCA TTR TAC ATY TG	-

*(Y=C/T, R=A/G, I=C/G/A/T)

Table 3.2. Primer sequence for NoV GII detection

Primer	Sequence (5'→3')*	Polarity
GIIF3M	TTG TGA ATG AAG ATG GCG TCG ART	+
GIIF1M	GGG AGG GCG ATC GCA ATC T	+
GIIR1M	CCR CCI GCA TRI CCR TTR TAC AT	-

* (Y=C/T, R=A/G, I=C/G/A/T)

Table 3.3. 1st PCR(one-step RT-PCR) conditions

Reaction	Conditions	
Reverse Transcription	45°C, 30min	1 cycle
One step RT-PCR	94°C, 5min	1 cycle
	94°C, 30sec 55°C, 30sec 72°C, 1min 30sec	35 cycle
	72°C, 7min	1 cycle

Table 3.4. 2nd PCR (nested PCR) conditions

Reaction	Conditions	
Nested PCR	94°C, 5min	1 cycle
	94°C, 30sec 55°C, 30sec 72°C, 1min 30sec	25 cycle
	72°C, 7min	1 cycle

3.1.6. Sequencing and phylogenetic analysis

After nested PCR using PCR products from one-step RT-PCR as PCR templates, the nested PCR products were analyzed by 1.5% standard agarose gel electrophoresis. To sequence the nested PCR products, they were purified using QIAquick PCR Purification Kit (Qiagen, Valencia, CA, USA). The purified nested PCR products were sequenced using Sanger sequencing method by Macrogen, Seoul, Korea. DNA sequences of nested PCR products were edited and handled with DNASTAR (DNASTAR, Madison, WI, USA), and BLASTN program was used for sequence analysis (http://www.ncbi.nlm.nih.gov/blast/). Multiple sequence alignments were carried out with the MegAlign program in DNASTAR (DNASTAR) and ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Phylogenetic trees were generated using MEGA5 program with neighbor-joining method (http://www.megasoftware.net/). The open reading frames (ORFs) were predicted using GeneMarkS program (http://exon.gatech.edu/genemarks.cgi).

3.2. Results and Discussion

3.2.1. Detection of NoVs in various provinces and periods

In 2010, 1,090 final concentrated groundwater samples were collected and analyzed using RT-PCR/nested-PCR assay for detection of NoV. Consequently, seven of the 1,090 samples were shown to be positive for NoV. As a percentage of detected NoV in each province, 1.4% of samples from Gangwon province, 0.4% of samples from Chungcheong province and 1.4% of samples from Jeolla province were positive for NoV, respectively (Figure 3.5 and 3.6). However, NoV positive sample was not detected in Gyeonggi, Gyeongsang and Jeju provinces. The monthly prevalence of NoV positive samples was as follows; 2.6% of samples were positive in March, 0.4% in June, 0.8% in July, 1.6% in August and 1.8% in October, respectively (Figure 3.7 and 3.8).

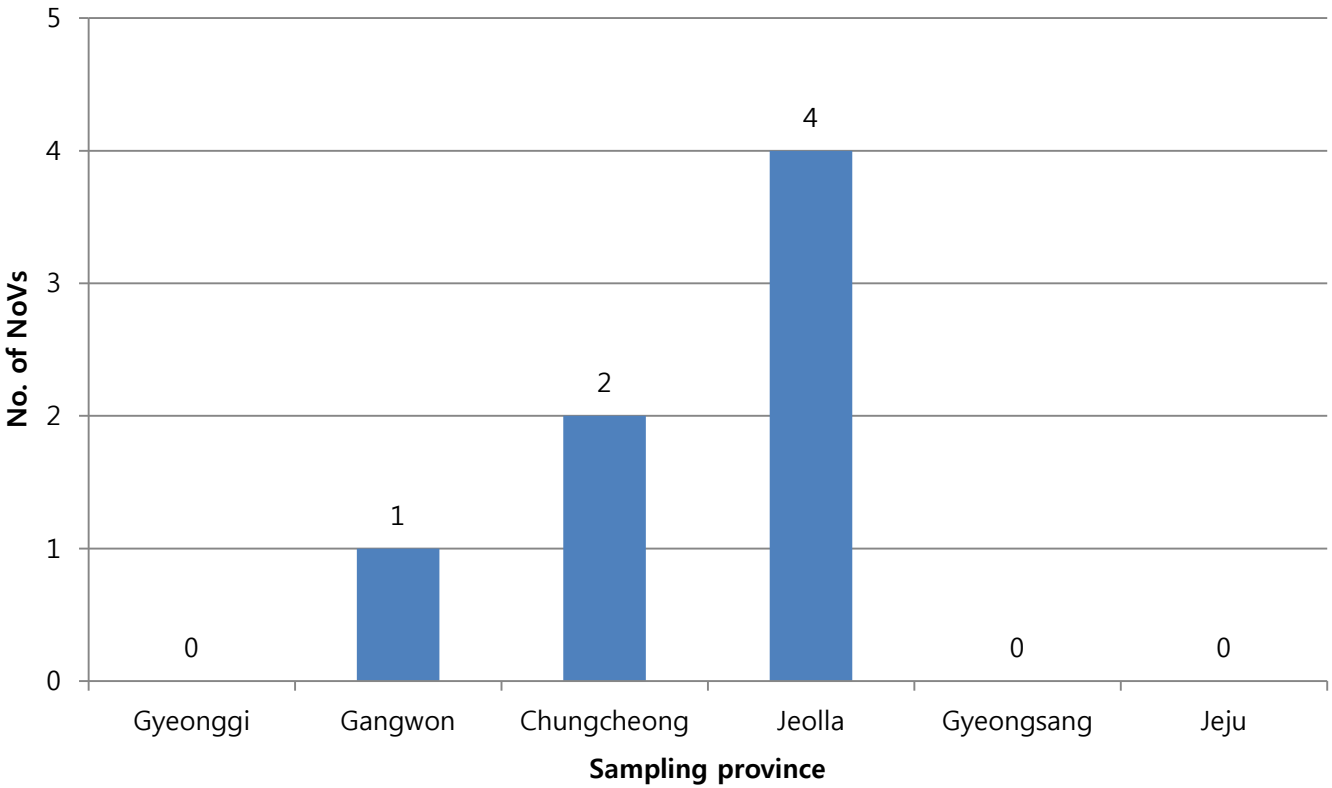


Figure 3.5. The number of NoV positive samples by area.

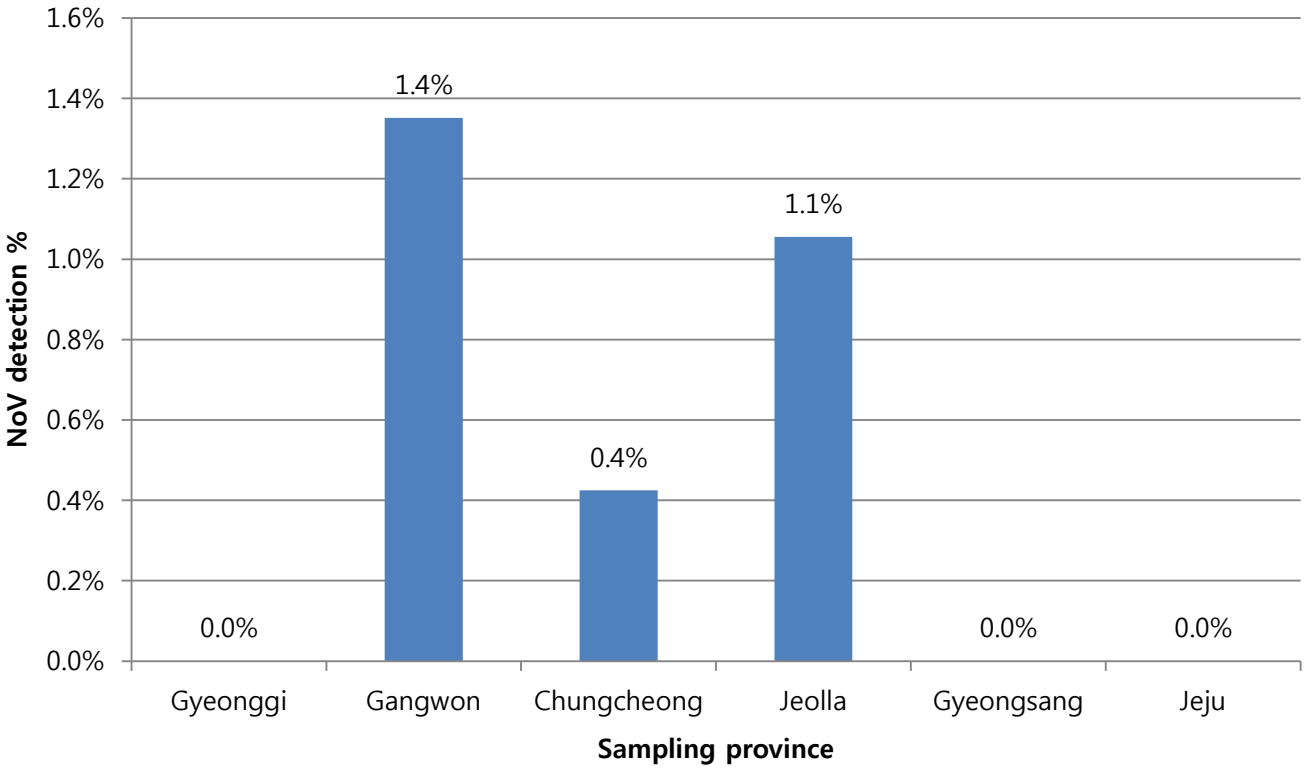


Figure 3.6. The percentage of NoV positive samples by area.

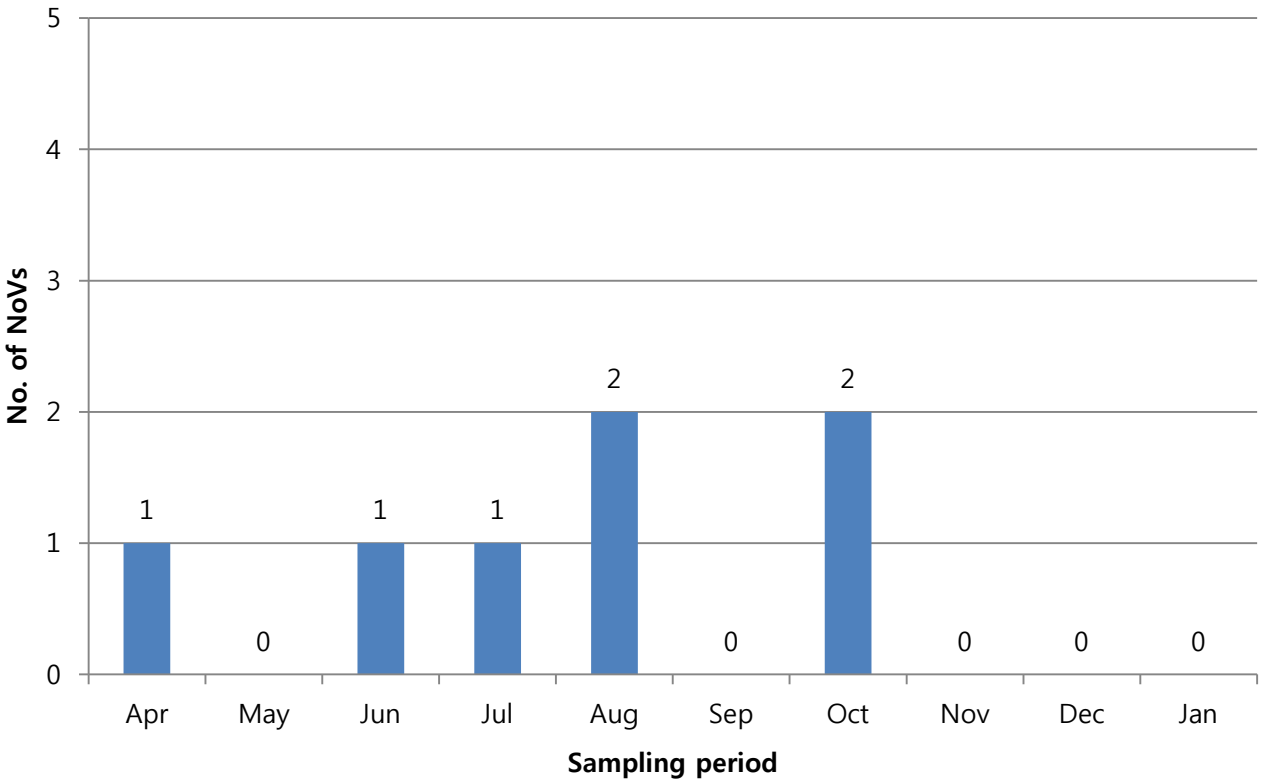


Figure 3.7. The number of NoV positive samples by month.

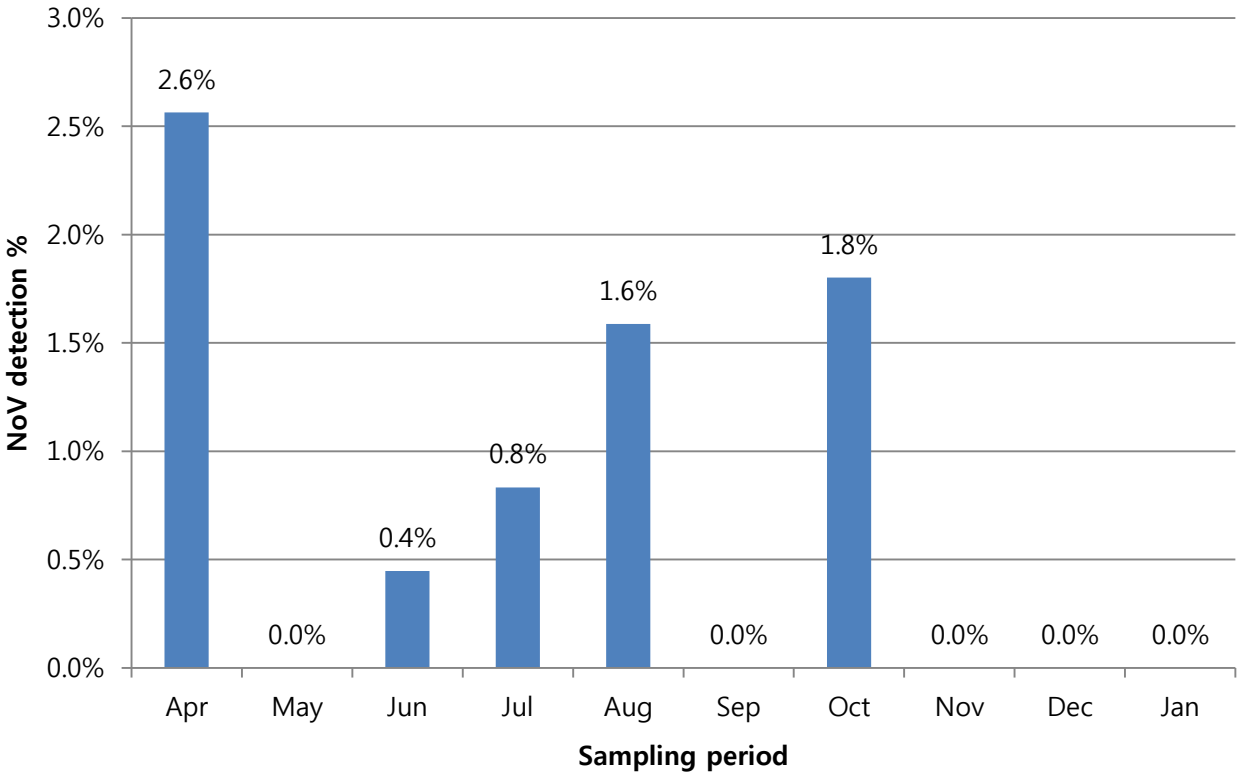


Figure 3.8. The percentage of NoV positive samples by month.

3.2.2. Genogrouping of detected NoVs in South Korea

NoVs have been classified into five NoV genogroups (GI to GV) and three human NoV genogroups (GI, GII, and GIV). Each of the human NoV genogroups is further classified into genotypes; there are 8, 17, and 1 genotypes in GI, GII, and GIV, respectively (Zheng et al, 2006). NoVs exhibit broad genomic diversity, and human NoV genogroups (GI, GII, and GIV) are comprised of at least 25 genotypes and many subgroups. The diversity of NoVs is maintained by the accumulation of point mutations from the error-prone nature of RNA replication and by genetic recombination due to sequence exchange between related RNA viruses. Therefore, phylogenetic analysis of genetically diverse NoVs is very important for their epidemiology studies.

In this study, the isolated NoVs were genetically analyzed and compared with other reference strains to reveal the genetic relationship of isolated NoV in South Korea. For genogrouping of detected seven NoVs, one-step RT-PCR and standard agarose gel electrophoresis were performed. In the agarose gel electrophoresis analysis of amplified PCR products by one-step RT-PCR, PCR products of six samples were detected in agarose gel lanes of GII PCR products (310 bp), determined as GII and the rest PCR product of one sample was detected in agaorse gel lanes of GI PCR products (313 bp), determined as GI (Table 3.5).

Table 3.5. The genogrouping and genotyping of detected NoV positive samples in South Korea

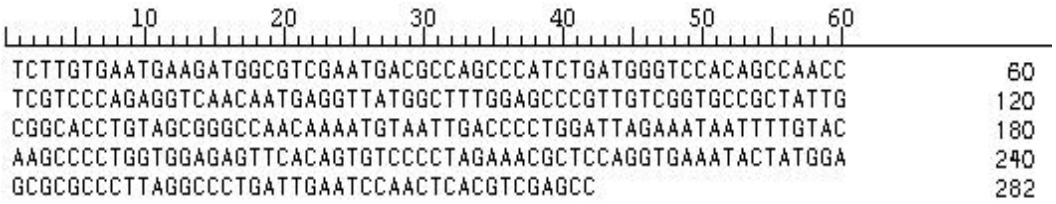
Samples	PCR		Nucleic acid sequence
	GI	GII	
2010042241	-	+	GII-4
2010060371	-	+	GII-4
2010071392	-	+	GII-3
2010081812	+	-	GI-6
2010082762	-	+	GII-17
2010100772	-	+	GII-17
2010101112	-	+	GII-4

3.2.3. Genotyping of detected NoVs in South Korea

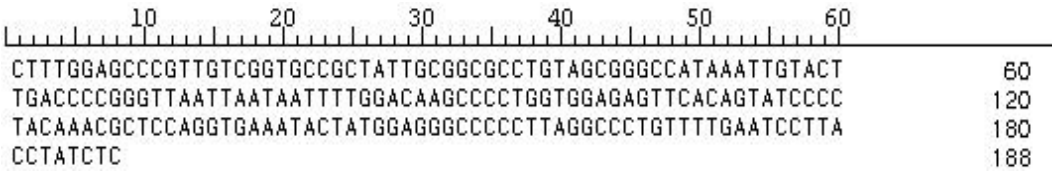
To sequence PCR products encoding NoV capsid protein, additional nested PCR was performed and the PCR primers for nested PCR were used as sequencing primers. DNA sequences of six GII PCR products and one GI PCR product were shown in Figure 3.9 and 3.10, respectively. Based on the sequences of nested PCR products, their comparative phylogenetic analyses were conducted to determine their genotypes. All DNA sequences of the nested PCR products and reference DNA sequences were multiple-aligned using ClustalW2 and multiple alignment results were used for generation of comparative phylogenetic analyses. These analyses showed that GII NoVs were classified into one GII-3, three GII-4, and two GII-17 and GI NoV was classified into one GI-6 (Table 3.5). The phylogenetic analysis of sample 2010071392 showed that this sample was highly homologous to GII-3 Tronto-CAN1993, suggesting that this sample may belong to GII-3 genotype (Figure 3.11). In addition, the phylogenetic analysis of samples 2010042241, 2010060371, and 2010101112 showed that they were very similar to GII-4 Bristol-GBR1993, suggesting that they may belong to GII-4 genotype (Figure 3.11 and 3.12). However, the phylogenetic analysis of samples 2010082862 and 2010100772 showed that they were very similar to GII-17 CSE1-USA2002, suggesting that they may belong to GII-17 genotype (Figure 3.12). Finally, the phylogenetic

analysis of sample 2010081812 showed that it was highly homologous to GI-6 Hesse-DEU1998, suggesting belonging to GI-6 genotype (Figure 3.13). Therefore, Major genotypes of NoVs in South Korea may belong to GII-4 and it is similar to findings in other counries (Buesa *et al.* 2002; Foley *et al.* 2001; Gallimore *et al.* 2004; Greening *et al.* 2001; Kearney *et al.* 2007; Tu *et al.* 2006).

(A)



(B)



(C)

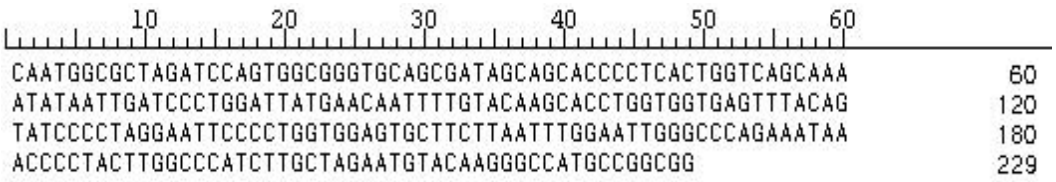
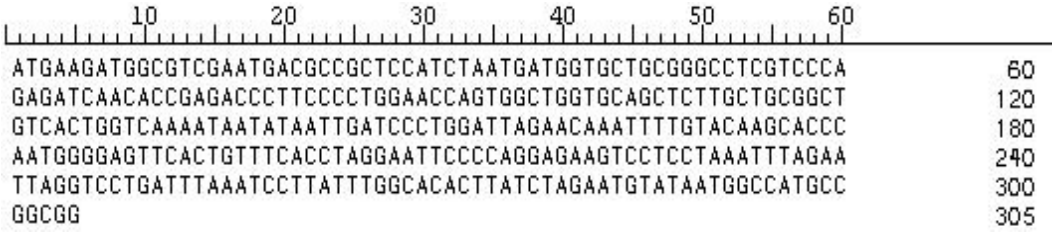


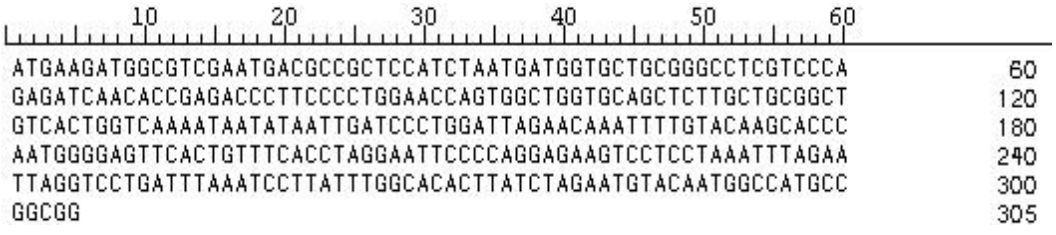
Figure 3.9. The nucleic acid sequences of nested PCR products in GII (Continued).

(A) sample No.2010042241, (B) sample No.2010060371, (C) sample No.2010071392,

(D)



(E)



(F)

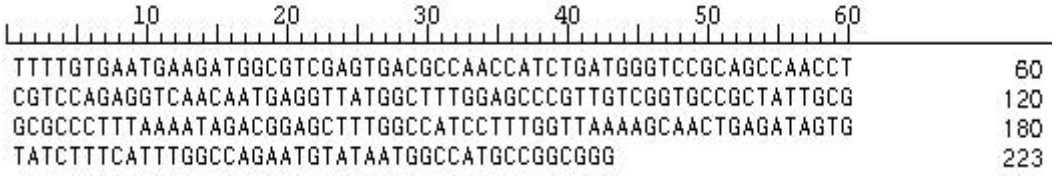


Figure 3.9. The nucleic acid sequences of nested PCR products inGII.

(D) sample No.2010082762, (E) sample No.2010100772, (F) sample No.2010101112.

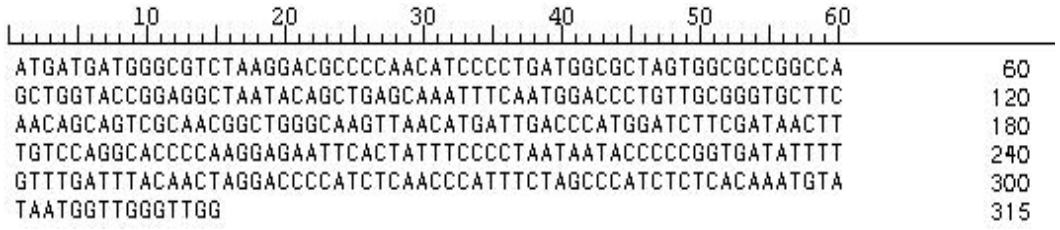


Figure 3.10. The nucleic acid sequence of nested PCR product in GI.

sample No.2010081812.

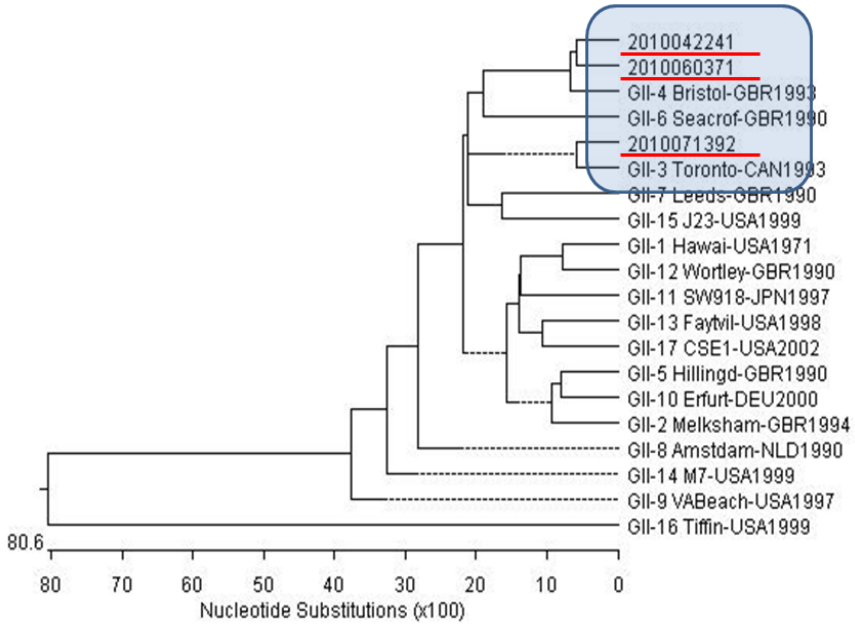


Figure 3.11. Phylogenetic tree analysis of sample 2010042241, 2010060371 and 2010071392.

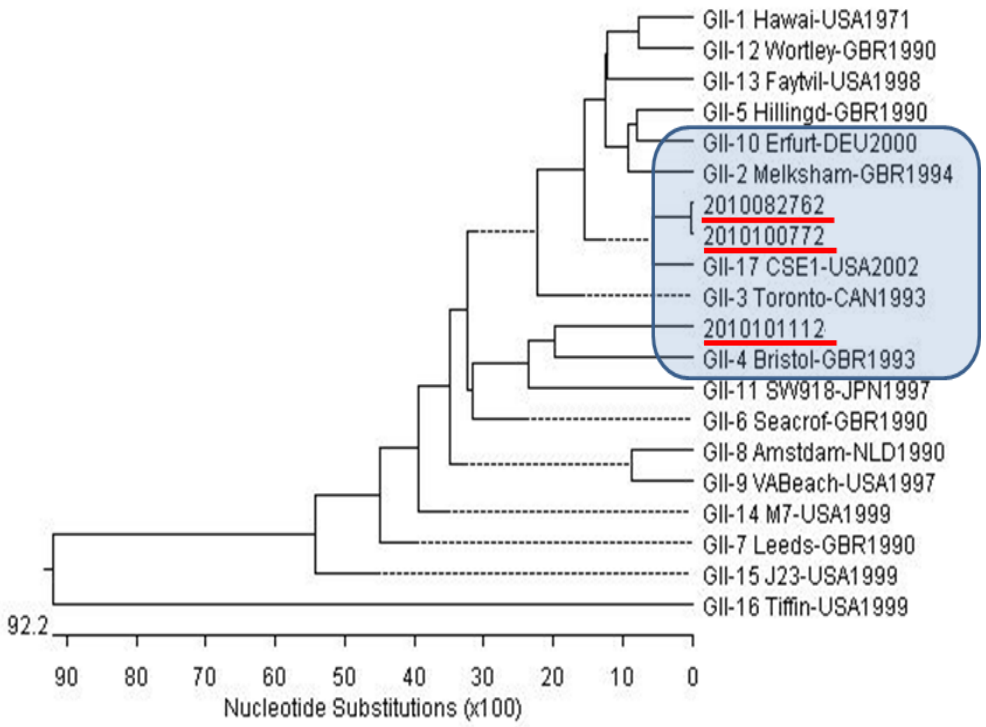


Figure 3.12. Phylogenetic tree analysis of sample 2010082762, 2010100772 and 2010101112.

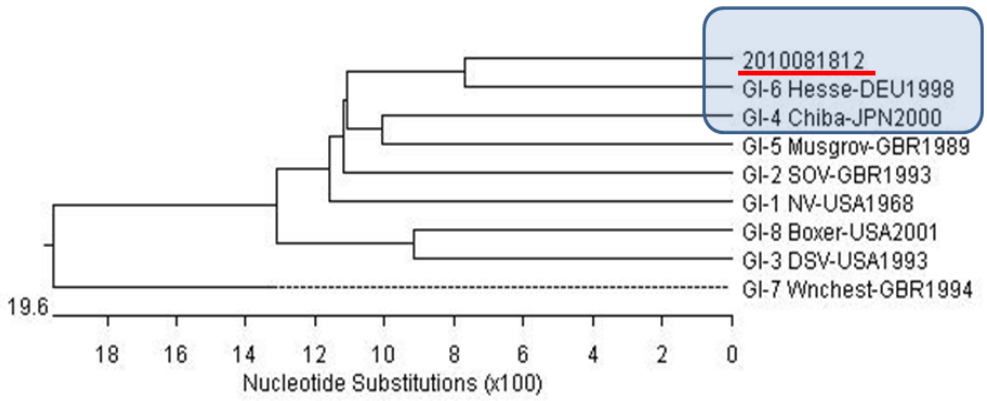


Figure 3.13. Phylogenetic tree analysis of sample 2010081812.

3.2.4. BLAST functional analysis of predicted ORFs in DNA sequences of the nested PCR products

The prediction of open reading frames (ORFs) in all DNA sequences of seven nested PCR products revealed that they all have three predicted ORFs (Table 3.6). The annotation of predicted ORFs showed that they may encode capsid proteins, RNA-dependent RNA polymerases (RdRp), and sometimes hypothetical proteins in the nested PCR products. The functions of capsid protein and RdRp have been known to be probably involved in the NoV structure and its RNA replication. The sequence identities of all ORFs are extremely high to other known NoV genes with the range of 92 to 100%, suggesting that those ORFs are highly conserved in various NoVs (Table 3.6).

Table 3.6. BLAST functional analysis of predicted ORFs in the nested PCR sequences (continued)

Samples	Accession number	Blast	Max score	E value	Max identity
2010042241	AB246162.1	NorovirusCor/Gunma/Dec-3/GII/2004/JP gene for capsid protein, partial cds	260	1e-132	99%
	AB240187.1	Norovirus Hu/NV/Hokkaido/194/2004/JP genes for RNA dependent RNA polymerase, capsid protein, partial cds	260	1e-132	99%
	AB240180.1	Norovirus Hu/NV/Hokkaido/186/2004/JP genes for RNA dependent RNA polymerase, capsid protein, partial cds	260	1e-132	99%
2010060371	EU003964.1	Norovirus Hu/GII-4/C5-150/South Korea nonfunctional capsid protein mRNA, partial sequence.	270	2e-69	92%
	HQ003288.1	Norovirus Hu/GII.4/Nizhny Novgorod/1921/2008/RUS capsid protein gene, partial cds	265	8e-68	92%
	AB541340.1	Norovirus Hu/GII-4/Sakai1/2008/JP genomic RNA, complete genome	265	8e-68	92%
2010071392	EU249135.1	Norovirus Hu/GII/C7-212/2005/KOR capsid protein mRNA, partial cds	412	3e-112	99%
	EU249140.1	Norovirus Hu/GII/C1-307/2005/KOR capsid protein mRNA, partial cds	392	4e-106	97%
	EU249132.1	Norovirus Hu/GII/C7-307/2006/KOR capsid protein mRNA, partial cds	392	4e-106	97%

Table 3.6.BLAST functional analysis of predicted ORFs in the nested PCR sequences (continued)

Samples	Accession number	Blast	Max score	E value	Max identity
2010081812	EU249146.1	Norovirus Hu/GI/C7-167/2005/KOR capsid protein mRNA, partial cds	564	1e-157	99%
	EU003971.1	Norovirus Hu/GI-8/C6-52/South Korea nonfunctional capsid protein mRNA, partial sequence	555	8e-155	98%
	EU003970.1	Norovirus Hu/GI-8/C6-49/South Korea nonfunctional capsid protein mRNA, partial sequence	555	8e-155	98%
2010082762	FJ383845.1	Norovirus Hu/GII/2005/8093/Chelyabinsk/RUS RNA-dependent RNA polymerase and capsid genes, partial cds	531	1e-147	100%
	DQ438972.1	Norovirusgenogroup 2 strain Hu/NoV/Katrina-17/2005/US VP1 and VP2 genes, complete cds	531	1e-147	99%
	FJ264887.1	Norovirus Hu/GII/Chelyabinsk/9418/2006/RUS capsid gene, partial cds	520	3e-144	100%
2010100772	FJ383845.1	Norovirus Hu/GII/2005/8093/Chelyabinsk/RUS RNA-dependent RNA polymerase and capsid genes, partial cds	534	1e-148	100%
	EF529741.1	Norovirus Hu/Briancon870/2004/France RNA-dependent RNA polymerase and capsid protein genes, partial cds	525	6e-146	98%
	DQ438972.1	Norovirusgenogroup 2 strain Hu/NoV/Katrina-17/2005/US VP1 and VP2 genes, complete cds	525	6e-146	98%

Table 3.6.BLAST functional analysis of predicted ORFs in the nested PCR sequences

Samples	Accession number	Blast	Max score	E value	Max identity
2010101112	AB541356.1	Norovirus Hu/GII-4/Toyama2/2008/JP genomic RNA, complete genome	222	6e-55	98%
	AB541261.1	Norovirus Hu/GII-4/Hokkaido1/2008/JP genomic RNA, complete genome	222	6e-55	98%
	AB541233.1	Norovirus Hu/GII-4/Chiba4/2008/JP genomic RNA, complete genome	222	6e-55	98%

3.3. Conclusion

Six out of seven NoV positive samples (85.7%) were identified as GII while the other (14.3%) was GI (Table 3.5). This is consistent with a general recognition of GII dominance in Asian countries. According to the phylogenetic analysis based on the NoV DNA sequences of the nested PCR products in the capsid protein region, three of six sequences of GII were clustered with GII-4, two were GII-17, and one was GII-3 genotypes. GI-related sequence was clustered with genotype GI-6. This result suggests that GII genogroup is highly dominant in NoVs of South Korea. The subsequent genotyping via comparative phylogenetic analysis also suggests that major genotype in GII genogroup is GII-4 in South Korea.

This extensive-scale investigating NoVs in ground water is very rare worldwide and when considering the population density in Korea, the results collected through this investigation is very valuable and could be practically used in the near future.

PART IV: Research on the Combined Water Treatment System

4.1. Materials and Methods

4.1.1. Evaluation of NoV reduction yield in electropositive filters

As described in Fig. 4.1, NoV reduction yield was evaluated by measuring human NoV absorption degree of each electropositive filter, respectively.

4.1.1.1. Test sample

After spiking 2.5 X 10^5 copies of human NoV which was isolated from patient's stool sample, it was put into 50 liters of sterilized distilled water to produce test sample.

4.1.1.2. Electropositive filter

1MDS filter (3M, St. Paul, MN, USA), 3M filter (not for sale, 3M) and Nanoceram filter (Argonide, Sanford, FL, USA) were selected as the electropositive filters for this study.

4.1.1.3. Quantification of human NoV using Real-time RT-PCR

Water processed with the 1-MDS filter produced a final elute of 20 ml. Viral RNA was extracted from 140 μl of that elute using the QIAamp® Viral RNA Mini kit (Qiagen, Hilden, Germany) according to the manufacturer's

protocol to obtain the final volume of 60 µl. Real-time PCR was conducted using a Bio-Rad iScript One-step RT-PCR Kit (Hercules, CA, USA). To quantify the NoV, real-time RT-PCR amplifications were performed with primer sets (Table 4.1) and real-time RT-PCR conditions (Table 4.2). Twenty microliter of viral RNA as the template and 30µl of the premixed kit solution were mixed for the reaction. The real-time RT-PCR was carried out using MyiQ real-time PCR machine (Bio-Rad), according to the following protocol.

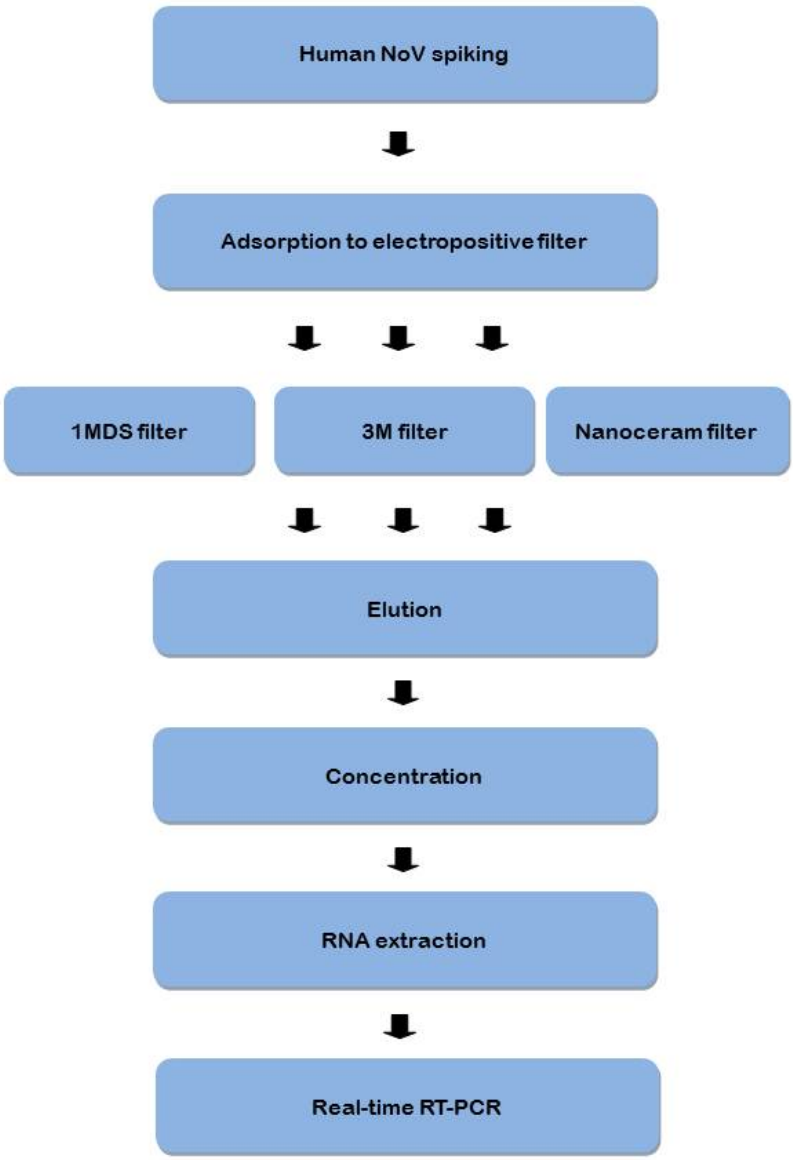


Figure 4.1. Procedure for quantification of human NoV.

Table 4.1. Primer and probe of Human NoV used in Real-time RT-PCR

Primer & probe	Sequence (5'→3')
Cog2F (GII)	CAR GAR BCN ATG TTY AGR TGG ATG AG
Cog2R (GII)	TCG ACG CCA TCT TCA TTC ACA
Ring2 (GII)	FAM-TGG GAG GGC GAT CGC AAT CT-BHQ

Table 4.2. Real-time RT-PCR conditions

Reaction	Conditions	
Reverse transcription	50℃ 10 min	1 cycle
Initial denaturation	95℃ 5 min	1 cycle
	95℃ 10 sec	
Amplification	55℃ 30 sec	45 cycle

4.1.2. Evaluation of NoV reduction yield in combined water treatment system

As described in Figure 4.2, NoV reduction yield was evaluated by measuring human NoV absorption and elimination degree in combined water treatment system.

4.1.2.1. Test sample

After spiking, 8.0 X 10^6 copies (1st trial) and 9.1 X 10^6 copies (2nd trial) of human NoV which was isolated from patient's stool sample and itwas dispersed into 50 liters of sterilized distilled water to produce test sample.

4.1.2.2. Combined water treatment system

Combined water treatment system was composed of chlorine feeder, UV irradiation device and electropositive filter cartridge (Figure 4.2) and overall water treatment system consists of filtration using 3M filter, UV irradiation (39 W/m^2)and chlorine disinfection (0.5mg/L) (Figure 4.3). Virus spiking sample was passed through the combined water treatment system with dechlorinated tap water at flow velocity of 10 liter per min. Overall procedure of evaluation of combined water treatment system in groundwater using the combined water treatment system and real-time RT-PCR is shown at Figure 4.4.

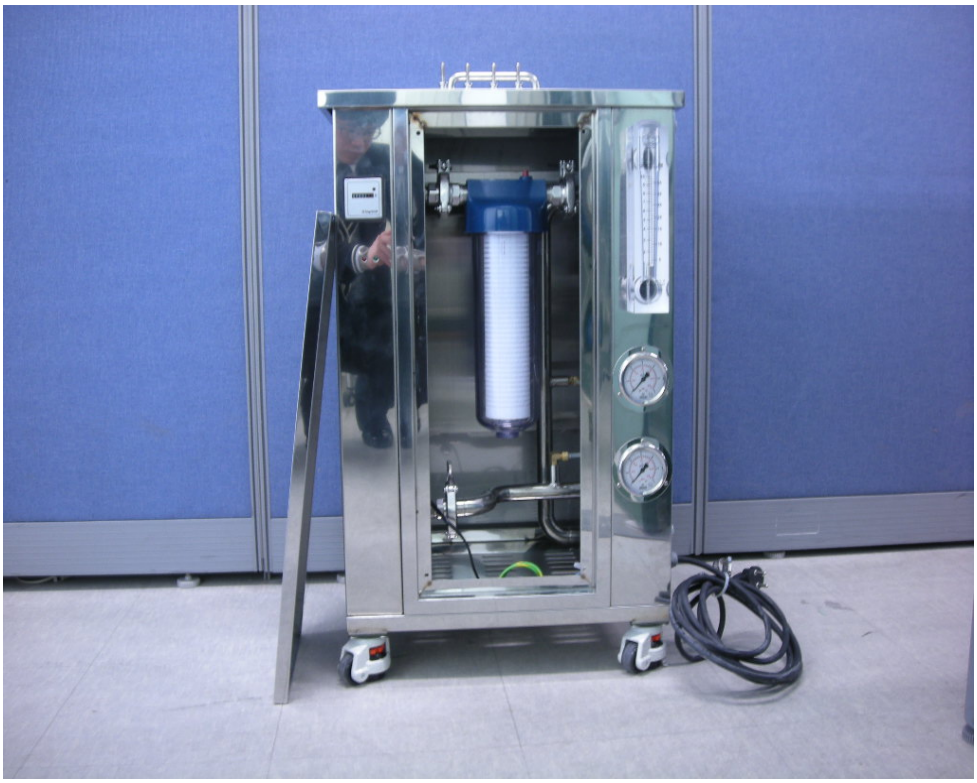


Figure 4.2. Combined water treatment system.

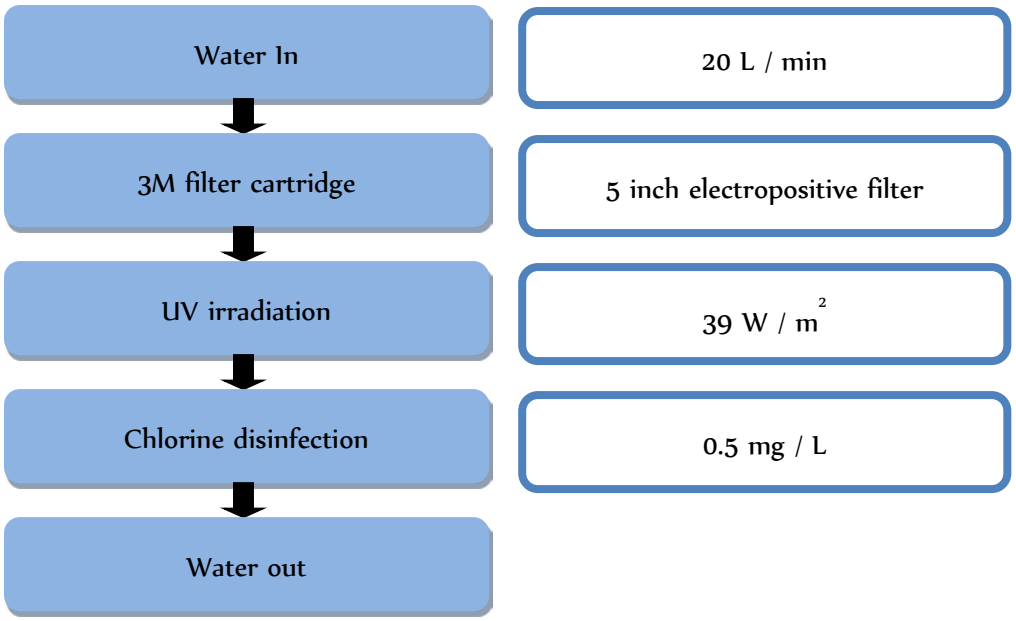


Figure 4.3. Composition and flow diagram of combined water treatment system.

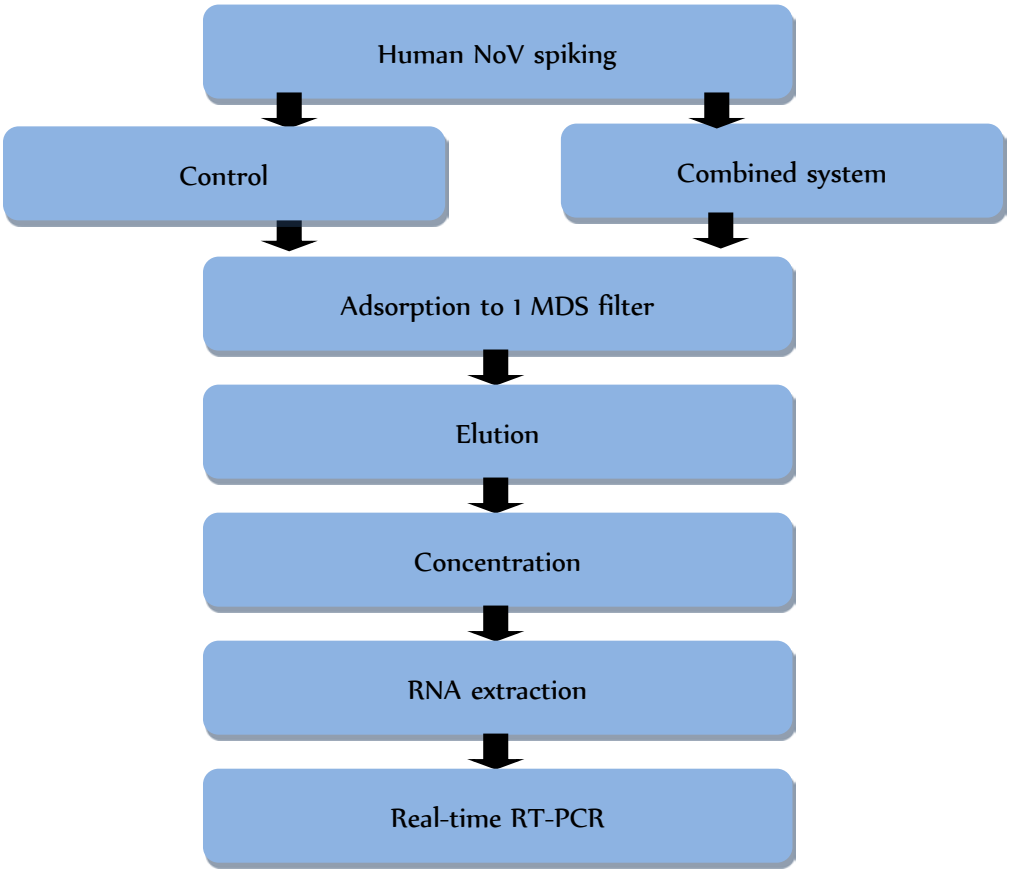


Figure 4.4. Overall evaluation procedure of reduction yield of NoV after water treatment using combined water treatment system.

4.1.2.3. Quantification of human NoV using Real-time RT-PCR

Water processed with the 3M filter produced a final eluate of 20 ml. Viral RNA was extracted from 140 µl of that elute using the QIAamp® Viral RNA Mini kit (Qiagen) according to the manufacturer's protocol to obtain the final volume of 60 µl. Real-time PCR was conducted using a Bio-Rad iScript One-step RT-PCR Kit.

To quantify the NoV, Real-time RT-PCR amplification was performed with primer sets (Table 4.1) and Real-time RT-PCR conditions (Table 4.2). We used 20µl of viral RNA as the template and 30µlof the premixed kit solution. The real-time RT-PCR was carried out using MyiQreal-time PCR machine (Bio-Rad), according to the following protocol.

4.2. Results and Discussion

4.2.1. Evaluation of NoV reduction yield in electropositive filters

Electropositive filters are effective to capture the majority of virus particles from the water passing through the filter and several electropositive filters of various compositions are now commercially available. Real-time RT-PCR result of human NoV, described in Figure 4.5, shows that these electropositive filters work efficiently for removal of NoVs. From the filter test twice for statistical analysis, the average human NoV reduction yields of 3M filter, 1MDS filter and Nanoceram filter were 42.05%, 33.85 % and 35.30%, respectively (Table 4.3). And these two tests of each electropositive filter showed consistent result for reduction of NoV in water and very low standard deviations (especially for 3M filter), suggesting their high fidelity for NoV removal. Capture capacity of human NoV on the 3M filter was equivalent or better than those of others. Therefore, 3M filter was used for further water treatment and evaluation of this combined water treatment system.

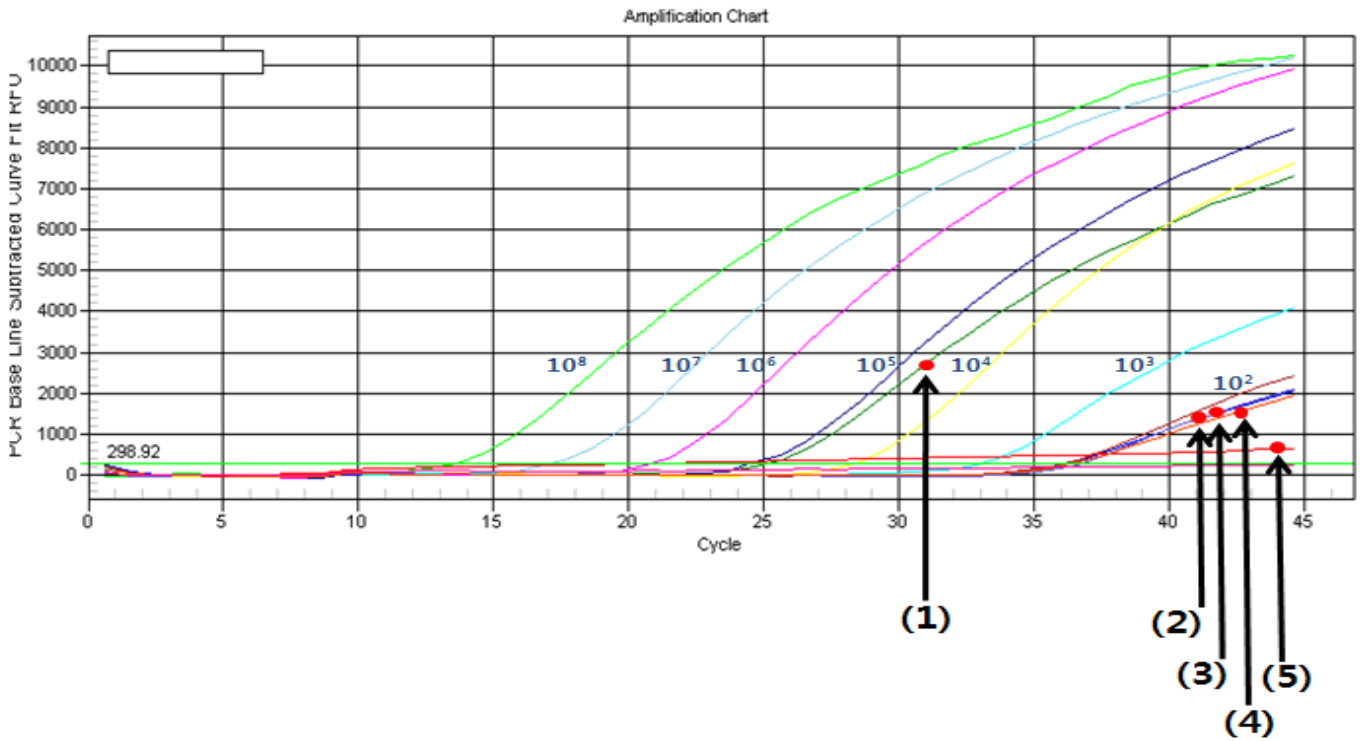


Figure 4.5. Real-time RT-PCR result of human NoV using electropositive filters.

(1) Seed titer (2) 3M fiter (3) 1MDS filter (4) Nanoceram filter (5) Negative control without NoV

Table 4.3. Human NoV reduction yield of electropositive filters

Filter type	1st trial (%)	2nd trial (%)	Average	Standard deviation
1-MDS	32.0	35.7	33.85	2.616
3M	42.2	41.9	42.05	0.212
Nanoceram	34.0	36.6	35.30	1.838

4.2.2. Evaluation of NoV reduction yield in combined water treatment system.

Due to high fidelity of 3M filter for reduction of human NoVs in water, 3M filter was selected and used for development of combined water treatment system. This system consists of 3M filter cartridge, UV irradiator and chlorine feeder. To evaluate this system, the human NoV reduction yield was tested. For quantification of human NoV particles, quantitative real-time RT-PCR was conducted. To test this, 6.9 X 10^5 copies of human NoV particles in the spiking sample were quantified using this quantitative real-time RT-PCR system. The RT-PCR result showed that quantified amount of the spiking sample is very close to the original amount, suggesting that this method is very useful for quantification of human NoVs in water (Figure 4.6). To evaluate this combined water treatment system, two spiking samples, 6.9 X 10^5 and 7.9 X 10^5 copies of human NoV particles, were prepared (Table 4.4). These two spiking samples were passed by the treatment system and amount of human NoV particles in eluted samples after water treatment were quantified using quantitative real-time RT-PCR. However, no human NoV particle was detected in both eluted samples, suggesting that the number of human NoV particles in the eluted samples may be lower than detection limit or possibly human NoV particles were completely removed

by this system. Therefore, this combined water treatment system can be very useful for water treatment to significantly reduce more than 5 logs of human NoV particles from the groundwater.

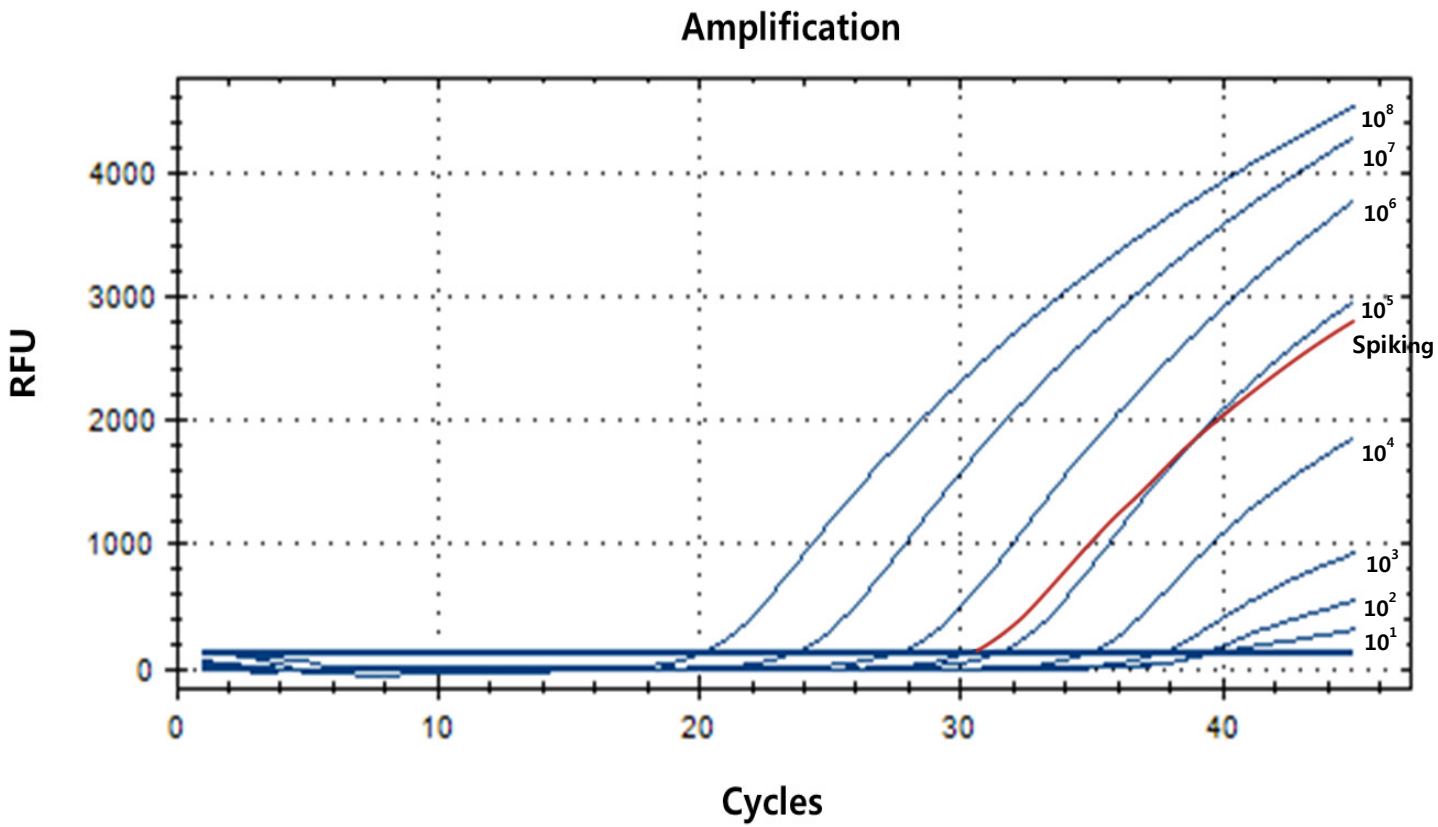


Figure 4.6. Amount of NoV in spiking sample using quantitative real-time RT-PCR.

Table 4.4. Human NoV reduction yield of combined water treatment system

	Control	Combined system*
1st trial	6.9 X 10^5 copies	ND
2nd trial	7.9 X 10^5 copies	ND

* , ND, not detected.

4.3. Conclusion

To develop the novel combined water treatment system using electropositive filtration, UV irradiation, and chlorine disinfection, three commercially available electropositive filters were tested and 3M electropositive filter was selected, due to high fidelity of human NoV reduction in water. This filter can reduce more than 40% of human NoVs in water, confirmed by quantitative real-time RT-PCR. After selection of electropositive filter, the combined water system was developed and evaluated using spiking samples containing more than 10^5 copies of human NoV particles. After treatment of the spiking samples, no human NoV was detected in both eluted samples from the system, suggesting that this combined water treatment system can reduce more than 10^5 copies of human NoV particles or can remove human NoV particles completely. Therefore, this system should be applied for groundwater treatment in schools and industries to make sure the safety of groundwater for drinking.

PART V: References

Donaldson EF, Lindesmith LC, Lobue AD, Baric RS. 2008. Norovirus pathogenesis: mechanisms of persistence and immune evasion in human populations. Immunol Rev 225: 190–211.

Yoon JS, Lee SG, Hong SK, Lee SA, Jheong WH, Oh SS, Oh MH, Ko GP, Lee CH, Paik SY. 2008. Molecular epidemiology of norovirus infections in children with acute gastroenteritis in South Korea in November 2005 through November 2006. J ClinMicrobiol 46: 1474–1477.

Lee YF, Nomoto A, Detjen BM, Wimmer E. 1977. A protein covalently linked to poliovirus genome RNA. ProcNatlAcadSci USA 74: 59–63.

Rueckert RR, Wimmer E. 1984. Systematic nomenclature of picornavirus proteins. J Virol 50: 957–959.

Tan M, Huang P, Meller J, Zhong W, Farkas T, Jiang X. 2003. Mutations within the P2 domain of norovirus capsid affect binding to human histoblood group antigens: evidence for a binding pocket. J Virol 77: 12562–12571.

Nilsson MK, Hedlund O, Thorhagen M, Larson G, Johansen K, Ekspong A, Svensson L. 2003. Evolution of human calicivirus RNA in vivo:

accumulation of mutations in the protruding P2 domain of the capsid leads to structural changes and possibly a new phenotype. J Virol 77: 13117–13124.

Lochridge VP, Jutila KI, Graff JW, Hardy ME. 2005. Epitopes in the P2 domain of norovirus VP1 recognized by monoclonal antibodies that block cell interactions. J Gen Virol 86: 2799–2806.

Glass PJ, White LJ, Bull JM, Leparc-Goffart I, Hardy ME, Estes MK. 2000. Norwalk virus open reading frame 3 encodes aminor structural protein. J Virol 74: 6581–6591.

Bertolotti-Ciariet A, Crawford SE, Hutson AM, Estes MK. 2003. The 3' end of Norwalk virus mRNA contains determinants that regulate the expression and stability of the viral capsid protein VP1: a novel function for the VP2 protein. J Virol 77: 11603–11615.

Motomura K, Oka T, Yokoyama M, Nakamura H, Mori H, Ode H, Hansman GS, Katayama K, Kanda T, Tanaka T, Takeda N, Sato H. 2008. Identification of monomorphic and divergent haplotypes in the 2006-2007 norovirus GII/4 epidemic population by genome wide tracing of evolutionary history. J Virol 82: 11247–11262.

Hale A, Mattick K, Lewis D, Estes M, Jiang X, Green J, Eglin R, Brown D. 2000. Distinct epidemiological patterns of Norwalk-like virus infection. J Med Virol 62: 99–103.

Glass RI, Parashar UD, Estes MK. 2009. Norovirus gastroenteritis. N Engl J Med 361: 1776–1785.

Lopman B, Vennema H, Kohli E, Pothier P, Sanchez A, Negredo A, Buesa J, Schreier E, Reacher M, Brown D, Gray J, Iturriza M, Gallimore C, Bottiger B, Hedlund KO, Torvén M, Bonsdorff CH, Maunula L, Poljsak-Prijatelj M, Zimsek J, Reuter G, Szücs G, Melegh B, Svennson L, Duijnhoven Y, Koopmans M. 2004. Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. Lancet 363: 682–688.

Allen DJ, Noad R, Samuel D, Gray JJ, Roy P, Iturriza-Gómara M. 2009. Characterisation of a GII-4 norovirus variant-specific surface-exposed site involved in antibody binding.Virol J 6: 150.

Chung JY, Han TH, Park SH, Kim SW, Hwang ES. 2010. Detection of GII-4/ 2006b Variant and Recombinant Noroviruses in Children With Acute Gastroenteritis, South Korea. J Med Virol 82: 146–152.

Kearney K, Menton J, Morgan JG. 2007. Carlow virus, a 2002 GII-4 variant norovirus strain from Ireland. Virol J 4: 61.

La Rosa G, Pourshaban M, Iaconelli M, Muscillo M. 2008. Detection of genogroup IV noroviruses in environmental and clinical samples and partial sequencing through rapid amplification of cDNA ends. Arch Virol 153: 2077–2083.

Koopmans M, Duizer E. 2004. Foodborne viruses: an emerging problem. International Journal of Food Microbiology 90: 23– 41.

Marshall JA, Hellard ME, Sinclair MI, Fairley CK, Cox BJ, Catton MG, Kelly H, Wright PJ. 2003. Incidence and characteristics of endemic Norwalk-like virus-associated gastroenteritis. J Med Virol 69: 568-578.

Kim SH, Cheon DS, Kim JH, Lee DH, Jheong WH, Heo YJ, Chung HM, Jee YM, Lee JS. 2005. Outbreaks of gastroenteritis that occurred during school

excursions in Korea were associated with several waterborne strains of norovirus. J ClinMicrobiol 43: 4836-4839.

Medici MC, Morelli A, Arcangeletti MC, Calderaro A, DeConto F, Martinelli M, Abelli LA, Dettori G, Chezzi C. 2009. An outbreak of norovirus infection in an Italian residential-care facility for the elderly.ClinMicrobiol Infect 15: 97-100.

Widdowson MA, Sulka A, Bulens SN, Beard RS, Chaves SS, Hammond R, Salehi ED, Swanson E, Totaro J, Woron R, Mead PS, Bresee JS, Monroe SS, Glass RI. 2005. Norovirus and foodborne disease, United States, 1991-2000. Emerg Infect Dis 11:95-102

Lee SG, Jheong WH, Suh CI, Kim SH, Lee JB, Jeong YS, Ko GP, Jang KL, Lee GC, Paik SY. 2011. Nationwide groundwater surveillance of Noroviruses in South Korea, 2008. ApplEnvMicrobiol 77: 1466-1474.

Lee CH, Kim SJ. 2008. The genetic diversity of human noroviruses detected in river water in Korea. Water Res 42:4477-4484

Shin GA, Mark S. 2008. Inactivation of norovirus by chlorine disinfection of water.Water Research. 1-7

Keswick BH, Satterwhite TK, Johnson PC, DuPont HL, Secor SL, Bitsura JA, Gary GW, Hoff JC. 1985. Inactivation of Norwalk virus in drinking water by chlorine. Appl Environ Microbiol 50:261-264

Rook JJ. 1974.Formation of haloforms during chlorination of natural waters. Water Treatment and Examination 23: 234-243.

Tree JA, Adams MR, Lees DN. 2004. Disinfection of feline calicivirus (a surrogate for Norovirus) in wastewaters. J Appl Microbiol 98: 155-162.

Lee JE, Zoh KD, Ko GP. 2008. Inactivation and UV disinfection of murine Norovirus with TiO_2 under various environmental conditions. Appl Environ Microbiol 74: 2111-2117.

Jeong WH, Park SJ, Park JY. 2010. A study on water quality management about small scale water supply system. NEIR 2010- 48-1223.

Ho EC, Cheng PK, Lau AW, Wong AH, Lim WW. 2007. A typical norovirus epidemic in Hong Kong during summer of 2006 caused by a new genogroup II/4 variant. J ClinMicrobiol 45: 2205-2211.

Tu ET, Bull RA, Greening GE, Hewitt J, Lyon MJ, Marshall JA, McIver CJ, Rawlinson WD, White PA. 2008. Epidemics of gastroenteritis during 2006 were associated with the spread of norovirus GII.4 variants 2006a and 2006b. Clin Infect Dis 46: 413-420.

Zheng DP, Ando T, Fankhauser RL, Beard RS, Glass RI, Monroe SS. 2006. Norovirus classification and proposed strain nomenclature. Virol 346: 312 - 323.

Buesa J, Collado B, Lopez-Andujar P. 2002. Molecular epidemiology of caliciviruses causing outbreaks and sporadic cases of acute gastroenteritis in Spain. J Clin Microbiol 40: 2854 - 2859.

Foley B, O'Mahony J, Hill C, Morgan JG. 2001. Molecular detection and sequencing of "Norwalk-like viruses" in outbreaks and sporadic cases of gastroenteritis in Ireland. J Med Virol 65: 388-394

Gallimore CI, Green J, Richards AF. 2004. Methods for the detection and characterisation of noroviruses associated with outbreaks of gastroenteritis: outbreaks occurring in the north-west of England during two norovirus seasons. J Med Virol 73: 280-288.

Greening GE, Mirams M, Berke T. 2001. Molecular epidemiology of ‘Norwalk-like viruses’ associated with gastroenteritis outbreaks in New Zealand. J Med Virol 64: 58-66.

초 록

노로바이러스는 사람에게 급성 장염을 일으키는 식중독 바이러스로서 최근에 전세계적으로 관심이 높아지고 있는 상황이다. 노로바이러스의 유전학적 변종을 확인하기 위해 한국에서 발생한 환자로부터 분리된 노로바이러스에 대해 complete genome sequence 분석을 실시하였고 ORF1 (5,100 bp), ORF2 (1,623 bp), ORF3 (807 bp) 로 구성되어 있음을 확인하였다. 분리된 균종에 대해 phylogenetic analysis 를 실시하여 아시아지역에서 가장 빈번하게 나타나는 노로바이러스 GII-4 의 변종임을 확인하였고 그에 대한 complete genome 을 Genbank 등록과 함께 노로바이러스 GII-4 variant CUK-3 로 명명하였다.

한국에서 식수로 사용되는 지하수 중의 노로바이러스 오염도를 조사하기 위해 1,090 개의 지하수 시료를 경기도, 충청도, 강원도, 경상도, 전라도, 제주도로부터 채취하여 분석한 결과 7 개의 시료로 부터 노로바이러스를 검출하여(3 개의 GII-4, 2 개의 GII-17, 1 개의 GII-3, 1 개의 GI-6) 한국에서도 노로바이러스 오염으로부터 지하수를 관리할 수 있는 대책이 필요함을 확인하였다.

세가지 종류의 양전하 필터에 대해 노로바이러스 제거능을 비교 평가하여 가장 성능이 우수한 3M 필터를 선정하였고 선정된

양전하 필터와 자외선 살균기, 염소소독기를 복합적으로 이용하는 복합수처리장치를 제작하였다. 10^5 copies 이상의 노로바이러스를 식수에 spiking 하여 시료를 만들고 복합수처리장치를 이용한 처리구에 대해 real-time RT-PCR 을 이용하여 정량분석하였다. 그 결과 물에 존재하는 노로바이러스에 대하여 10^5 이상의 제거효과가 있음을 확인하였고 지하수의 노로바이러스 오염에 대한 해결책으로써 복합수처리장치가 효과적으로 적용될 수 있음을 확인하였다.

주요어 : 노로바이러스 Sequence, Phylogenetic tree, 지하수, 복합수처리장치, 양전하필터, 노로바이러스 제거

학번 : 2005-31092