Characterization of Norovirus strains and a seroprevalence study to understand the fluctuation of outbreak activity in Alberta, Canada

by

Maria Eloisa Hasing Rodriguez

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Medical Sciences - Laboratory Medicine and Pathology

University of Alberta

Abstract

Norovirus (NoV) is recognized as the leading agent causing gastroenteritis among individuals of all ages. Historical data of NoV outbreak surveillance in Alberta, Canada, from the period July 2000 to June 2008 demonstrated a biennial pattern of NoV outbreak activity, in which peaks of high outbreak activity appeared to be driven by the emergence of new phylogenetic clusters of NoV within genotype GII.4 (GII.4 variants). The work in this thesis aimed to analyze the outbreak activity of NoV in Alberta between July 2008 and June 2015 and examine novel NoV GII.4 variants as well as immune responses at the population level.

Norovirus outbreak surveillance data since July 2008, showed that GII.4 continued to be the most predominant genotype in Alberta, and was responsible for 47.6% to 80.2% of all NoV outbreaks per year. The overall number of NoV outbreaks observed per year demonstrated a loss of the biennial pattern since July 2011 as well as a gradual decline in the number of NoV outbreaks. Two novel GII.4 variants emerged in the 7-year period, New Orleans 2009 and Sydney 2012. Genotype GII.17 was also detected in November 2014. Unlike previous years, peaks of high outbreak activity were not always associated with the emergence of novel GII.4 variants, suggesting that, in addition to antigenic change of NoV GII.4, other factors also contributed to the high levels of NoV outbreak activity in Alberta.

Since NoV strain genotyping for outbreak surveillance in the province has been performed based on a small 320bp region within the major capsid gene (ORF2) that provides limited information about the genetic diversity of NoV, a near-full length genome characterization was performed on a group of NoV strains representing 7 different GII.4 variants that have circulated in the province since 2002. A phylogenetic comparison demonstrated that most of the strains found in Alberta were similar to strains that had been

described in global pandemics. Furthermore, recombination detection analysis as well as an assessment of the intra-host genetic diversity of NoV during chronic infection suggested that genetic changes in other genes besides the major capsid gene (ORF2), such as p22 (ORF1) and VP2 (ORF3) might improve the fitness of NoV GII.4.

In order to examine if waning of herd immune response over time could be a factor in the annual fluctuations of outbreak activity levels, archived sera collected in 2008 and 2012-2014 from two different populations of Alberta, pregnant women and individuals receiving virology-serology testing, were tested for anti-NoV GII.4 IgG in a cross-sectional study. A high seroprevalence of 84.2% and 76.7%, respectively, were observed in each group. No significant differences in prevalence or geometric mean titers were identified across time points in these populations.

In summary, the findings from this study suggest that, in addition to antigenic change of NoV at ORF2, other factors also affect the outbreak activity in Alberta and that ORF1 and ORF3 phylogenetic analysis should be included while monitoring the evolution of pandemic NoV strains. The results from this study provide useful data regarding the seroprevalence of NoV GII.4 in Alberta and the variability of anti-NoV GII.4 IgG responses within a large group of individuals, including pregnant women, providing a framework for further studies of immune responses at the population level.

Preface

This thesis represents collaborative work, led by Dr. Xiaoli Pang at the University of Alberta.

The literature review described in chapter 1 is my original work.

The study described in chapter 2 contains information that has been previously published as: M.E. Hasing, B.E. Lee, J.K. Preiksaitis, R. Tellier, L. Honish, A. Senthilselvan, X.L. Pang. "Emergence of a new norovirus GII.4 variant and changes in the historical biennial pattern of norovirus outbreak activity in Alberta, Canada, from 2008 to 2013". *J Clin Microbiol.* 2013 Jul;51(7):2204-11. I share first authorship of this publication with B.E. Lee. I wrote the first draft of the manuscript and I performed all NoV sequencing analysis. B.E. Lee designed and coordinated the study and analyzed the data from outbreak investigations. J.K. Preiksaitis, R. Tellier and L. Honish reviewed and edited the manuscript. A. Senthilselvan performed the statistical analysis on NoV seasonality and revised the manuscript. X.L. Pang designed, coordinated and supervised the study and revised the manuscript. This study received research ethics approval from the University of Alberta Research Ethics Board, under the Project Name "Molecular surveillance of norovirus causing outbreaks and seroprevalence study in Alberta, Canada", No. Pro00037093, on April 10, 2013.

The study described in chapter 3 has been published as: M.E. Hasing, B. Hazes, B.E. Lee, J.K. Preiksaitis, X.L. Pang. "Detection and analysis of recombination in GII.4 norovirus strains causing gastroenteritis outbreaks in Alberta". *Infect Genet Evol.* 2014 Oct;27:181-92. I am the first author of this publication; I wrote the manuscript and performed all laboratory experiments and bioinformatics analysis. B. Hazes supervised the bioinformatics analysis,

contributed in data interpretation and revised the manuscript. B.E. Lee contributed with sample data and revised the manuscript. J.K. Preiksaitis reviewed and edited the manuscript. X.L. Pang designed, coordinated and supervised the study and revised the manuscript.

The study described in chapter 4 has been published as: M.E. Hasing, B. Hazes, B.E. Lee, J.K. Preiksaitis, X.L. Pang. "A next generation sequencing-based method to study the intra-host genetic diversity of norovirus in patients with acute and chronic infection". *BMC Genomics*. 2016 Jul 1;17(1):480. I am the first author of this publication. I wrote the manuscript, performed the laboratory and bioinformatics analyses. B. Hazes supervised the bioinformatics analysis, contributed with data interpretation and revised the manuscript. B.E. Lee contributed with sample data and revised the manuscript. J.K. Preiksaitis reviewed and edited the manuscript. X.L. Pang designed, coordinated and supervised the study and revised the manuscript. This study received research ethics approval from the University of Alberta Research Ethics Board, under the Project Name "Gastroenteritis Virus Infection in Immunocompromised patients", No. Pro00050079, on November 27, 2015

The study presented in chapter 5 was led by Dr. Pang and had the collaboration of Dr. Jason Jiang at the Cincinnati Children's Hospital. No part of this chapter has been previously published. I prepared the *E. coli* clone used for p-particle expression at Dr. Jiang's lab, performed all ELISA tests, data analysis and data interpretation. Dr. Jiang supplied the NoV P-particles. Drs. B.E. Lee and X. Pang contributed in the study design. This study received research ethics approval from the University of Alberta Research Ethics Board, under the Project Name "Molecular surveillance of norovirus causing outbreaks and a seroprevalence study in Alberta, Canada", No. Pro00037093, on April 10, 2013.

Dedication

To my husband Ricardo and my son Rafael, for all their support and love.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Xiaoli Pang, for her support, guidance and mentorship. I am honoured to have been part of her research team and be her first PhD student. I appreciate her efforts to make my PhD experience productive and encourage my training beyond the lab bench.

I would like to extend my deepest gratitude to each of my committee members. I thank Dr. Bonita Lee for her support and for making herself available whenever I needed help. I am truly grateful to Dr. Bart Hazes for his helpful advice and guidance with bioinformatics analyses. I also thank Dr. Jutta Preiksaitis and Dr. David Evans for their insightful ideas; their contributions were key for the development of this thesis work.

I want to express my sincere thanks to Dr. Jason Jiang and his research team at the Cincinnati Children's Hospital for kindly providing the NoV p-particles used in this study and for letting me visit their lab to learn the process of P-particle production. I especially thank Weiming Zhong, Ming Xia and Jeff Wei for their patience, and for sharing with me some of their expert knowledge on norovirus and protein expression.

I would like to thank all my fellow labmates: Min, Yupin, Judy and Ran who helped me many times during my experiments.

I have to thank my mom, for all her love and support and, my father, whose words and example are still with me and keep inspiring me to achieve dreams. Lastly but not least, I need to thank my husband Ricardo and my son Rafael: My journey through graduate studies has been yours as well; thank you for filling it with unconditional support, patience and love.

Table of Contents

ABSTRACT ..... **II**

PREFACE ..... **IV**

DEDICATION ..... **VI**

ACKNOWLEDGEMENTS ..... **VII**

TABLE OF CONTENTS ..... **VIII**

LIST OF TABLES ..... **XIV**

LIST OF FIGURES ..... **XV**

LIST OF ABBREVIATIONS ..... **XVII**

CHAPTER I: GENERAL INTRODUCTION **1**

1.1. A HISTORIC PERSPECTIVE OF NOROVIRUS 1

1.2 NOROVIRUS TAXONOMY AND VIRION DESCRIPTION 1

1.3. NOROVIRUS GENOME ORGANIZATION 2

1.3.1. Non coding regions and secondary structures *4*

1.3.2. Non-structural proteins *4*

1.3.3. Structural proteins *8*

1.4. NOROVIRUS CLASSIFICATION AND NOMENCLATURE 9

1.5. BURDEN OF NOROVIRUS ILLNESS 11

1.5.1. Disease burden *11*

1.5.2. Economic burden *11*

1.6. METHODS OF DETECTION 12

1.6.1. Immune electron microscopy *12*

1.6.3. RT-PCR *13*

1.6.4. Multi-gastrointestinal-pathogen diagnostic platforms *14*

1.7. MOLECULAR EPIDEMIOLOGY 14

1.7.1. Outbreaks *14*

1.7.2. Sporadic disease *16*

1.7.3. Seasonality *17*

1.8. TRANSMISSION 18

1.9. CLINICAL SYMPTOMS AND PATHOGENESIS 19

1.10. POPULATIONS AT RISK 20

1.10.1. Children *20*

1.10.2. Elderly *21*

1.10.3. Immunocompromised individuals *22*

1.11. GENETIC SUSCEPTIBILITY TO NOROVIRUS 22

1.12. NOROVIRUS EVOLUTION 27

1.12.1. Antigenic drift *27*

1.12.2. Recombination *29*

1.13. IMMUNITY TO NOROVIRUS 30

1.13.1. Humoral responses *31*

1.13.2. Cellular mediated immunity *32*

1.14. NOROVIRUS IN ALBERTA 33

1.15. CURRENT CAVEATS 35

1.16. RATIONALE OF THE PROJECT 36

1.17. REFERENCES (CHAPTER 1) 37

CHAPTER II: MOLECULAR EPIDEMIOLOGY OF NOROVIRUS OUTBREAKS IN ALBERTA 55

2.1. INTRODUCTION 56
2.2. MATERIALS AND METHODS 57
2.2.1. Outbreak investigations *57*
2.2.2. Testing and characterization of Norovirus in stool specimens *57*
2.2.3. Amplification and phylogenetic analysis of GII.4-2012 capsid sequences *60*
2.2.4. Data inclusion and statistical analysis *61*
2.3. RESULTS 62
2.3.1. Norovirus outbreaks by GII and GI *62*
2.3.2. Norovirus GII diversity *64*
2.3.3. Norovirus GI diversity *69*
2.3.4 Demographic data of individuals affected in Norovirus outbreaks *69*
2.3.4. Temporal variations in norovirus outbreak levels *70*
2.4. DISCUSSION 71
2.5. ACKNOWLEDGEMENTS 77
2.6. REFERENCES (CHAPTER 2) 77

CHAPTER III: DETECTION AND ANALYSIS OF RECOMBINATION IN GII.4 NOROVIRUS STRAINS CAUSING GASTROENTERITIS OUTBREAKS IN ALBERTA 81

3.1. INTRODUCTION 82
3.2. MATERIALS AND METHODS 84
3.2.1. Stool specimens *84*
3.2.2. Genomic amplification and DNA sequencing *85*
3.2.3. Preparation of reference dataset *86*
3.2.4. Phylogenetic and recombination detection analyses *86*
3.3. RESULTS 87

3.3.1. Norovirus epidemic activity, outbreak settings and patient demographics associated with samples selected for near full-length genome sequencing *87*

3.3.2. Characterization of recombinant NoV reference sequences from GenBank *88*

3.3.3. Norovirus GII.4 recombinants from outbreaks in Alberta *94*

3.3.4. Comparison of previous ORF2-genotyping results vs. whole genome analysis *95*

3.4. DISCUSSION 97

3.5. ACKNOWLEDGEMENTS 100

3.6. REFERENCES (CHAPTER 3) 101

CHAPTER IV: A NGS BASED METHOD TO STUDY THE INTRA-HOST POPULATION OF NOROVIRUS 104

4.1. INTRODUCTION 105

4.2. MATERIALS AND METHODS 107

4.2.1. Patient samples *107*

4.2.2. RNA extraction *107*

4.2.3. Illumina library preparation and sequencing *108*

4.2.4. Identification, characterization and removal of rRNA reads *109*

4.2.5. Assembly of NoV genomes *109*

4.2.6. NoV sequencing using Sanger's method *110*

4.2.7. Characterization of non-rRNA, non-NoV sequences *111*

4.2.8. Analysis of single nucleotide variants *112*

4.3. RESULTS 113

4.3.1. Characterization of the major components of stool RNA after bacterial rRNA depletion *113*

4.3.2. Non-rRNA-non-NoV RNA *116*

4.3.1. Mapping of NoV sequencing reads *118*

4.3.2. Detection of Single Nucleotide Polymorphisms *120*

4.4. DISCUSSION 124

4.5. ACKNOWLEDGEMENTS 128

4.6. REFERENCES (CHAPTER 4) 128

CHAPTER V: ANALYSIS OF THE ANTIBODY RESPONSES TO NOV GII.4 IN ALBERTA, CANADA 133

5.1. INTRODUCTION 134

5.2. MATERIALS AND METHODS 135

5.2.1. Sample population *135*

5.2.2. Production of NoV GII.4 capsid antigens *136*

5.2.3. ELISA test for anti-NoV IgG *137*

5.2.4. Statistical analysis *137*

5.3. RESULTS 138

5.3.1. Validation of the ELISA anti-NoV IgG test *138*

5.3.2. Seroprevalence of Norovirus *140*

5.3.3. Analysis of anti-NoV IgG titers in time *141*

5.4. DISCUSSION 144

5.5. REFERENCES (CHAPTER 5) 148

CHAPTER VI: DISCUSSION AND FUTURE DIRECTIONS 151

6.1. DISCUSSION 151

6.2. FUTURE DIRECTIONS 161

6.2.1. Genetic and antigenic diversity of NoV at ORF1 and ORF3 ..... *161*

6.2.2. The elderly as potential reservoirs of novel NoV strains *161*

6.2.3. Immune responses to NoV and antigenic seniority *162*

6.3. REFERENCES (CHAPTER 6) 163

REFERENCES 166

APPENDIX A 192

APPENDIX B 193

APPENDIX C 195

APPENDIX D 196

List of Tables

Table 1.1. Proteins encoded by Norovirus 5

Table 1.2. Patterns of NoV binding to synthetic oligosaccharide conjugates and human saliva 26

Table 2.1. Reference sequences used for NoV genotyping. 58

Table 2.2.Number of NoV outbreaks by genogroups and genotypes, from July 1 2008 to June 30 2015 65

Table 2.3. Age and sex of subjects involved in NoV outbreak investigations by outbreak setting 70

Table 3.1. List of primers used for near full-length norovirus genome sequencing 85

Table 3.2. Near full-length NoV GII.4 sequences obtained in this study 88

Table 3.3. Recombinant NoV GII.4 sequences identified in this study 89

Table 4.1. Samples analyzed by NGS 107

Table 4.2. Primers used for sequencing NoV strain OU3 using Sanger's method 111

Table 4.3. Percentage of rRNA reads per source 113

Table 4.4. Norovirus reads and coverage per sample 114

Table 4.5. Distribution of single-nucleotide variants in NoV by gene 123

Table 5.1. Samples included in this study 140

Table 5.2. Anti-NoV GII.4 IgG GMT titers in PW and VS populations 142

Table 5.3. Seroprevalence of NoV GII.4 reported by previous studies. 145

List of Figures

Figure 1.1. NoV genome organization, proteins and capsid structure 3

Figure 1.2. Phylogenetic analysis of complete ORF2 sequences of NoV 10

Figure 1.3. NoV regions frequently used for genotyping purposes 14

Figure 1.4. Biosynthetic pathways of HBGAs type 1 24

Figure 1.5. Mechanism of recombination in NoV 29

Figure 1.6. NoV outbreaks in Alberta between July 2000 and June 2008 34

Figure 2.1. Monthly distribution of Norovirus outbreaks by genogroup in Alberta, Canada. 63

Figure 2.2. GII.4 variants identified in Norovirus outbreaks in Alberta, Canada, between July 1st, 2002 and June 30, 2015. 66

Figure 2.3. Phylogenetic analysis of complete VP1 amino acid sequences of GII.4 variants identified in Alberta 67

Figure 2.4. Neighbor-joining tree of NoV GI sequences (region C) from samples identified in Alberta 68

Figure 2.5. Numbers of Norovirus positive and negative outbreaks in Alberta, Canada by observation period. 71

Figure 2.6. Neighbor-joining tree of NoV GII.4 sequences (region C) including GII.4 Sydney 2012 strains identified in Alberta. 76

Figure 3.1. Maximum likelihood phylogenies of NoV GII.4 sequences in ORF1 90

Figure 3.2. Maximum likelihood phylogenies of NoV GII.4 sequences in ORF2 91

Figure 3.3. Maximum likelihood phylogenies of NoV GII.4 sequences in ORF3 93

Figure 3.4. Simplot analysis of recombinant sequences identified in this study 94

Figure 3.5. Comparison of AlbertaEI337/2011 with recombinant sequences from Europe 96

Figure 3.6. Proposed origin of NoV GII.4 Alberta 2011 and Sydney 2012 strains 98

Figure 4.1. Taxonomic fingerprint of bacterial rRNA 23S sequences present in stool RNA of individuals with NoV infection *115*

Figure 4.2. Characterization of non-rRNA-non-NoV RNA in stool RNA. *117*

Figure 4.3. Coverage per NoV genome position..... *119*

Figure 4.4. Distribution of NoV SNV frequencies across the viral genome *121*

Figure 4.5. Distribution per gene of the amino acid residues affected by non-synonymous single nucleotide variants in NoV..... *123*

Figure 5.1. Detection of seroconversion in individuals suffering natural infection by Norovirus. ... *139*

Figure 5.2: Seroprevalence of NoV in two populations in Alberta at different time points..... *141*

Figure 5.3. Reverse cumulative distribution plots of anti-NoV IgG titers in two adult populations in Alberta *143*

Figure 5.4. Time of sera collection in relationship to NoV GII.4 outbreak levels in the province..... *147*

Figure 6.1. Trend in the evolution of VP1 (ORF2b) in Norovirus GII.4 ..... *156*

List of abbreviations

aa	Amino acid
ELISA	Enzyme-linked immunosorbent assay
FCV	Feline Calicivirus
GI	Genogroup I
GII	Genogroup II
HBGA	Histo-blood group antigen
HuNoV	Human Norovirus
LSU	Large subunit (ribosomal RNA)
LTC	Long term care
MNV	Murine Norovirus
NGS	Next generation sequencing
NoV	Norovirus
nt	nucleotide
NV	Norwalk Virus
NRNNR	non-rRNA-non-NoV reads
nsSNV	non-synonymous single nucleotide variants
ORF	Open reading frame
OTU	Operational taxonomic unit
PV	Picornavirus
RCD	Reverse Cumulative Distribution
RdRp	RNA dependent RNA polymerase
RHDV	Rabbit Hemorrhagic Disease Virus
rRNA	Ribosomal RNA
RT-PCR	Reverse transcription polymerase chain reaction

SMV	Snow Mountain Virus
SNV	Single nucleotide variants
ssRNA	Single-stranded RNA
SSU	Small subunit (ribosomal RNA)
UTR	Untranslated region
VLP	Virus-like particle

Chapter I: General Introduction

1.1. A historic perspective of Norovirus

Early descriptions of Norovirus (NoV) disease date back to 1929 when Dr. Zahorsky referred to it as 'winter vomiting disease' [1]. In subsequent years, the disease was identified based on its symptoms, as the etiologic agent remained elusive. The pathogen was known to be nonbacterial and was presumed to be a virus because bacteria-free stool filtrates were able to replicate the disease in volunteer studies. NoV was finally discovered in 1972 by Kapikian *et al.* [2] who identified the virus in samples from an outbreak occurring at an elementary school in Norwalk, Ohio. NoV is a fastidious agent, -still- challenging to culture [3] and difficult to visualize by electron microscopy due to low number of viral particles in samples and its relatively small size [4]. NoV biology was greatly benefited in the 1990s by advances in molecular techniques, specifically DNA sequencing and PCR.

1.2 Norovirus taxonomy and virion description

The genus Norovirus, formerly known as "Norwalk' or 'Norwalk-like viruses' [5], belongs to the *Caliciviridae* family along with other four genera: Sapovirus, Nebovirus, Lagovirus and Vesivirus [6]. Nebovirus has been found in cattle and includes Bovine enteric calicivirus. Lagovirus, identified in hares and rabbits, includes Rabbit Hemorrhagic Disease Virus (RHDV), a pathogen known to cause a systemic deadly infection in its hosts. Vesivirus has been identified in cats, sea lions, and dogs and its most well characterized strain is Feline Calicivirus (FCV). Norovirus and Sapovirus are the only caliciviruses known to infect humans.

NoV is a nonenveloped virus with a 27-38nm diameter [2, 7] and a T=3 icosahedral symmetry [8] resulting in 32 cup-like structures featured by all caliciviruses (*calici* is derived from the latin word '*calyx*' meaning 'cup') [6].

1.3. Norovirus genome organization

Noroviruses have a linear, single-stranded, positive-sense RNA genome of ~7,600 bp [9]. The majority of NoVs have their genomes divided into 3 consecutive ORFs (Figure 1.1) except murine noroviruses (MNV) which represent an exceptional case carrying an additional ORF within ORF2 (ORF4) [10]. ORF1 encompasses about two-thirds of the viral genome and encodes a large polyprotein that is processed into 6 non-structural proteins. ORF2 shares a small overlapping region of 17-20nt with ORF1 and encodes the major capsid protein (VP1), and ORF3 encodes the minor capsid protein (VP2) (Figure 1.1).

During its life cycle, NoV produces a subgenomic RNA that excludes ORF1 and encompasses ORF2 and ORF3 [11]. The 5' ends of both, NoV genomic and subgenomic RNAs present a covalently linked VPg molecule which allows a cap-independent translation process [11].

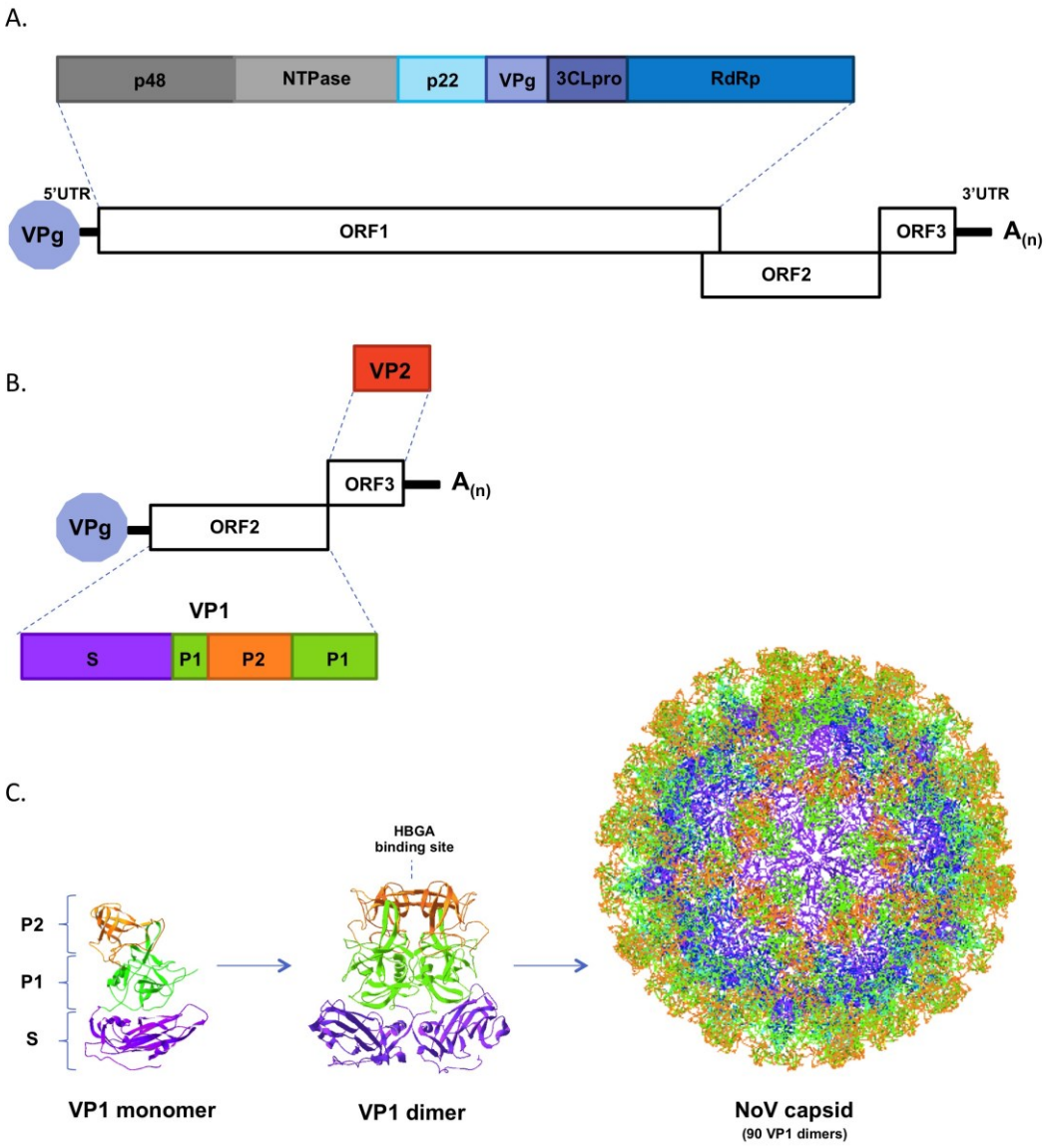


Figure 1.1. NoV genome organization, proteins and capsid structure

A) NoV genomic RNA is used for expression of non-structural proteins and as template during genome replication. B) Expression of capsid proteins VP1 and VP2 occurs from the subgenomic RNA. C) VP1 has two domains (S and P) and spontaneously forms dimers. The P2 subdomain (orange) forms the outermost part of the capsid and contains the interface that binds the receptor molecules (HBGAs). The capsid structure of NoV was prepared based on the crystal protein structure of Norwalk Virus (PBD ID: 1IHM) using the software DeepView – Swiss-PdbViewer.

1.3.1. Non coding regions and secondary structures

The non-coding regions of HuNoV include a 4nt long 5' untranslated region (UTR), a 45-54nt long 3' UTR and a 3' poly-A tail [9, 12, 13]. The 5' and 3' end of the genome as well as the start of ORF2 present stem-loop secondary structures that participate in NoV replication [14-17]. Multiple cellular proteins interact with the 5' and 3' secondary structures although the effect and relevance of such interactions remains to be studied in more detail for HuNoV [16, 18, 19]. Some of these interactions promote genome circularization of HuNoV, a conformation that allows *cis*-regulation of translation and replication in several RNA viruses [20]. Moreover, the 3'end of MNV presents a polypyrimidine tract that has been associated with viral fitness and virulence but studies confirming similar associations with HuNoV are still pending [15].

1.3.2. Non-structural proteins

NoV has six non-structural proteins (Table 1.1) that are expressed from ORF1 as a single 1,789-aminoacid polyprotein and are cleaved co- and post-translationally by the virus protease, 3CLpro [21]. Norovirus shares large sequence similarity and genome organization with picornavirus (PV) [9, 22], hence the names of several non-structural proteins of NoV have been adopted from their PV counterparts.

Table 1.1. Proteins encoded by Norovirus

ORF	Protein	Other names	Approximate Size (kDa)
1	NS1/2	p48, p37*, N-terminal	37-48
1	NS3	NTPase, p41, p40*, 2C-like protein	40-41
1	NS4	p22, p20*, 3A-like	20-22
1	NS5	VPg, p16	15-16
1	NS6	3CLpro, p19	19
1	NS7	RdRp, p57, 3D-like RNA polymerase	57
1	VF1**	Virulence Factor 1	24
2	VP1	Major capsid protein	57-60
3	VP2	Minor capsid protein	23

* Names used for NoV genogroup II only due to slight differences in size

**VF1 is present in MNV only.

Protein sizes are given based on Genbank accession number DQ223042.

Data presented in this table were obtained from references [8, 12, 21, 23-25].

NS1/2 (p48)

The function of p48 remains largely unknown. The p48 protein shares the same genomic position as 2A and 2B proteins in PV, however, p48 shows no sequence similarity with these proteins or with any other protein sequence in public databases [24]. Among human caliciviruses, p48 is the least conserved protein in ORF1 [24] and has been identified as an 'intrinsically disordered protein' with potential structural flexibility to interact with multiple molecules [26]. The most conserved region of p48 constitutes a predicted transmembrane domain at the C-terminus which appears to confer localization in the Golgi Apparatus [24]. Congruently, p48 has been associated with cellular proteins that mediate vesicle trafficking [24]. Since NoV replication occurs on intracellular membranes, it has been suggested that p48 might act as a scaffold protein to anchor viral replication complexes to specific membranes [24]. Hbox/NC motifs have also been identified in p48 [21]; these motifs are also present in the H-rev107 and TIG3 family of cellular proteins that are involved in cell growth control [27] but there is no experimental confirmation of this function for p48. Recently, a single amino acid change in MNV NS1/2 has been associated with increased enteric viral persistence and tropism to colon [28].

NS3 (NTPase)

The activity of this protein has been confirmed experimentally using in vitro assays [29]. The NoV NTPase belongs to superfamily 3 of RNA helicases and contains 4 motifs: motifs A and B are present in NTP-binding proteins of various functions, motif C is characteristic of helicases encoded by small DNA and RNA viruses whereas motif D, is conserved in caliciviruses and picornaviruses [29]. The NTPase of NoV can bind and hydrolyze all NTPs (ATP, CTP, GTP and UTP) in vitro and is more resistant to guanidine hydrochloride inhibition compared to the 2C helicase of PV [29]. However, in contrast with protein 2C, NoV NTPase is unable to unwind DNA-RNA heteroduplex in vitro, and therefore, it seems to lack helicase activity [29].

NS4 (p22)

The function of p22 is still unknown; its homolog in PV is protein 3A which recruits replication complexes to the cellular membrane [21]. NoV p22 contains two motifs: MAD (Membrane Association Domain) and MERES (Mimic of an Endoplasmic Reticulum Export Signal). The MAD motif is well conserved among caliciviruses, however, the MERES motif is less conserved and presents variations among NoV genotypes [30]. The MERES motif appears to induce trafficking from the endoplasmic reticulum towards the Golgi, disrupt of the Golgi apparatus, and inhibit cellular protein secretion [30, 31]. The p22 protein displays the highest evolutionary rates among all non-structural proteins [32] and the biological bases behind this observation are yet unknown.

NS5 (VPg)

VPg is a minor component of NoV, present at 1-2 molecules per virion [33], but it is considered a non-structural protein because of its role in viral replication. VPg is covalently linked to the 5' end of genomic and subgenomic RNA where it appears to allow a protein-

primed RNA replication [34]. VPg promotes the recruitment of the host translation machinery [35-38] through a HEAT-1 domain that interacts with the translation initiation factor eIF4G. Vpg and eIF4G in turn interact with eIF3 and other initiation factors to mediate the binding of NoV RNA to the ribosomal complex [37]. VPg can inhibit cap-dependent and IRES-dependent translation processes in experiments using cell-free systems [36], but no studies have confirmed whether NoV uses this function of VPg strategically to decrease the expression of host mRNA during infection.

NS6 (3CLpro)

This protein is the homolog of the 3C protease of PV (hence the name 3C-like protease) and is a cysteine protease with a chymotrypsin-like fold [39]. 3CLpro cleaves the NoV polyprotein following a specific order that is regulated by the aminoacid sequences surrounding the cleavage sites [40]. The junctions between proteins NS1/2-3 and NS3-4 are cleaved early in the process whereas NS4-5, NS5-6 and NS6-7 are cleaved later [40]. During its life cycle, NoV produces a precursor form composed of 3CLpro and RdRp (ProPol) which presents protease and polymerase functions [41]. The ProPol form also participates in NoV replication by catalyzing the nucleotidylylation of VPg, a modification that is considered the first step in VPg-mediated priming [42]. Being indispensable for NoV replication, 3CLpro has been studied extensively as potential target of antiviral drug development [43].

NS7 (RdRp)

Due to its essential role in virus replication, the RNA-dependent-RNA polymerase is another potential target for antiviral therapy [43, 44]. The crystallographic structures of NoV genogroups I and II polymerases have been determined and have provided insights about the function of these enzymes [45-47]. The RdRp of NoV has three major domains (thumb, fingers and palm) and its active form appears to have a homodimer arrangement although

monomers might enhance RNA synthesis as well [45]. Unlike other non-structural proteins, the RdRp of MNV does not show any association with any cell compartment, and localizes in the cytoplasm and the nucleus, following a pattern that is common of other viral RdRps [48].

1.3.3. Structural proteins

NoV has two structural proteins, VP1 and VP2, also known as the major and minor capsid proteins, respectively. Both proteins are translated from subgenomic RNA (Figure 1) [49]. Each virion is constructed by 180 copies of VP1 organized into 90 dimers [8]; the number of VP2 copies per virion seems to vary but it has been estimated to be around 1.4 to 8 in average [33, 50, 51]

VP1

The major capsid protein, VP1, has two domains: the shell (S) domain, which forms the innermost part of the capsid, and the protruding (P) domain, which contains two subdomains, P1 and P2 (Figure 1.1). P2 forms the outermost part of the capsid and contains a hypervariable region that seems to evolve in response to immune pressure [52, 53]. In addition, P2 also carries binding sites for receptors that mediate the entry of NoV in host cells [54]. The histoblood group antigens (HBGAs) are widely recognized as putative receptor molecules although experimental confirmation is still pending. Moreover, there is evidence that other molecule(s) might also act as receptor(s) or co-receptor(s) (a more detailed description regarding NoV receptors is provided under section 1.11).

VP2

The function of the minor capsid protein, VP2, remains elusive. VP2, has a molecular weight of 23kDa but higher molecular weight forms have also been identified in virions isolated from infected patients [50]. VP2 associates with VLPs but is not essential for VLP

formation and appears to have a short half-life (<1h) [50, 55]. In contrast with VP1, which is acidic (isoelectric point = 5.64), VP2 is a basic protein (isoelectric point ≥9.5) able to interact with the negatively charged viral RNA, thus, it has been hypothesized that VP2 might assist in packaging NoV genomic RNA during the encapsidation process [50, 55]. Consistent with this hypothesis, VP2 also seems to bind VP1 at the S-domain and localize at the interior surface of the capsid [55]. Other functions that have been attributed to VP2 include: up-regulation of VP1 expression and protection of VP1 from degradation by proteases [55, 56]. VP2 expression is also essential for the production of infectious FCV [57] but the possible mechanism behind this association has not been elucidated.

1.4. Norovirus classification and nomenclature

Genotype-based classifications are widely used for NoV and are preferred over serotyping due to their reproducibility. Genotyping is commonly performed on ORF2 (VP1) sequences; however, due the existence of numerous recombinant strains of NoV with breakpoints at the ORF1-ORF2 junction (further described in section 1.12.2), concomitant ORF1 (RdRp)-based genotyping is also used although in less extend.

Based on complete amino acid VP1 sequences NoV are classified into genogroups that are further classified into genotypes (Figure 1.2A). Seven genogroups (GI-GVII) and 40 genotypes have been reported to date [58, 59] of which at least 29 have been identified in humans [6]. NoV GII.4 is the most prevalent genotype among humans and is also highly diverse as new clusters or 'variants' emerge periodically (Figure 1.2B). For the purpose of ORF2-genotype allocation, Zheng *et al.* proposed the use of amino acid pairwise distances on complete VP1 sequences: strains with uncorrected amino acid pairwise distances below 45%, 15% and 5% would be considered to be within the same genogroup, genotype and GII.4 variant, respectively [59, 60]. However, a more recent in-depth analysis identified overlaps

in distance values, implicating that pre-defined cut-off values are not reliable references [61]. Instead, a '2xSD criterion' has been proposed for identification of new genotypes, where 'the average distance between all sequences within a newly identified cluster and its nearest established cluster should not overlap within 2 standard deviations' [61]. Since the 2xSD criterion is applicable for genotypes only, other criteria has been proposed to identify new GII.4 variants which include phylogenetic analysis and evidence of epidemic circulation in at least two geographically different locations [61].

Similar considerations have not yet been proposed for ORF1-genotypes. ORF1-genotypes have a slightly different nomenclature compared to ORF2-genotypes. Firstly, ORF1 genotypes are denoted with a 'P' in order to indicate that the designation is based on polymerase sequences. ORF1-genotypes are designated with the same number of their accompanying ORF2-genotypes in prototype strains, and cases of orphan sequences that have been associated with multiple ORF2 genotypes are designated with letters (e.g. GII.Pe).

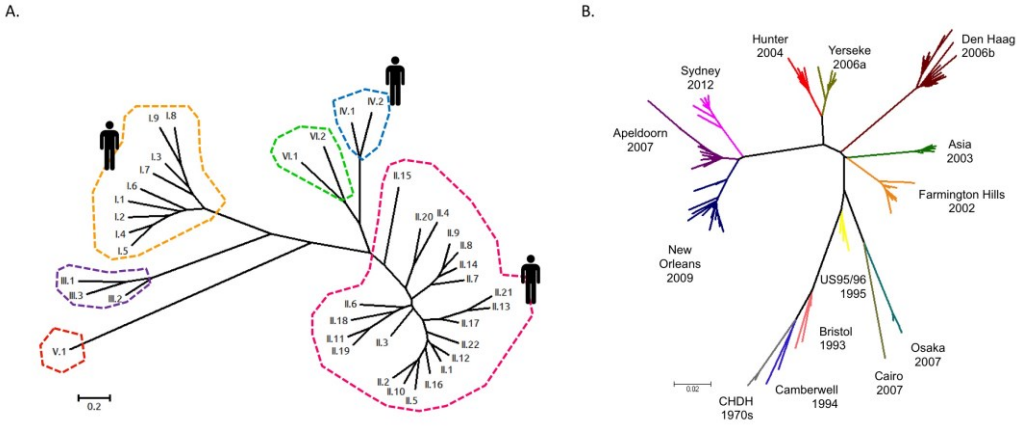


Figure 1.2. Phylogenetic analysis of complete ORF2 sequences of NoV

A) Maximum likelihood tree of VP1 amino acid sequences of NoV strains from genogroups GI-GVI. Genogroups are shown with different colors. The reference sequences are the same used in [58] except GVII because they were not publicly available. The phylogenetic tree was constructed in Mega

6.0 using the model Le Gascuel (LG+F), gamma distribution of rates among sites with invariant sites (G+I), and 1000 bootstrap replicates. B) Maximum likelihood tree of VP1 nucleotide sequences of NoV GII.4 strains. A total of 104 sequences listed in [62] were included in the dataset. The tree was constructed using the Kimura 2-parameter model, gamma distribution of rates among sites with invariant sites (G+I).

1.5. Burden of Norovirus illness

1.5.1. Disease burden

Norovirus disease is spread worldwide and has high prevalence. Norovirus causes about 18% (95%CI 17-20) of all gastroenteritis cases worldwide [63]. Estimates from the World Health Organization identified NoV as the leading cause of morbidity and mortality from diarrhoeal disease, being responsible for 677 million (95%CI 468-1,153 million) cases of diarrhoea and 213,515 (95%CI 171,783-266,561) deaths worldwide during 2010 [64]. A large proportion of NoV-associated deaths (43%) are estimated to occur in children less than five years of age [64]. Moreover, NoV causes the fourth greatest burden in terms of Disability-Adjusted Life Years among foodborne diseases (2,496,078 DALYs, 95%CI 1,175,658-5,511,092) [65]. While the incidence rates of NoV disease do not seem to vary greatly across countries [63, 64, 66], there is a large difference in mortality rates, being higher in low-income countries [65].

1.5.2. Economic burden

Due to its ubiquitous prevalence, NoV disease affects a large number of individuals resulting in large economic impact. The cost of NoV disease in the United States between 2001 and 2009 has been estimated as 284 million dollars per year [67] considering healthcare expenses only. When including expenses such as medications, transport to health clinics, cleaning products, special foods, productivity losses due to missed work days

(including time parents stay at home or at hospitals to care of children) and mortality, the estimated cost was 2,896 million dollars for NoV foodborne disease during 2010 [68]. In the United Kingdom, the total annual cost to patients and health services combined based on data from the years 2008 and 2009, were 81 million pounds (95%CI 63-106 millions) [69]. Norovirus also has a large impact on the cruise ship industry, being responsible for about 92% of its outbreaks in the United States [70]. There are no published estimates of the costs of NoV in cruise ship industry which produced a gross revenue of 21 billion dollars and carry about 11 million passengers in the United States alone, during 2014 [71].

1.6. Methods of detection

1.6.1. Immune electron microscopy

The first method used for NoV detection was immune electron microscopy which successfully lead to the discovery of the virus in 1972 by Kapikian *et al.* [2]. In this method, stool samples are mixed with anti-NoV serum to allow NoV precipitate in aggregates (during initial experiments leading to NoV discovery, Kapikian *et al.* used inactivated convalescent sera from experimentally infected volunteers). The mixture of serum and stool is centrifuged and the pellet is examined under electron microscopy for the presence of NoV virions. A disadvantage of this method is its cost, which includes expensive reagents and equipment as well as highly trained personnel. Also, its sensitivity is low; the visualization and identification of virions is difficult due to their small size and similar morphology with that of other enteric viruses.

1.6.2. ELISA

The main advantages of enzyme immunoassays are its low cost and good specificity. Moreover, ELISA can be adapted into immunochromatographic tests to provide short

turnaround times and testing in field settings. A major disadvantage of ELISA is that its sensitivity depends on the ability of the chosen antibodies to recognize the virus, which is usually modest given the large genetic/antigenic diversity of NoV. In order to increase sensitivity, ELISA tests are usually designed with cocktails of antibodies capable, together, to bind multiple NoV genotypes. Yet, commercial ELISA assays have reported sensitivities of 17-90% compared to RT-PCR and have limited or completely absent reactivity to several NoV genotypes [72-79].

1.6.3. RT-PCR

This method is largely preferred over all others due to its high sensitivity, high specificity and relatively low cost. The RdRp gene located at the 3' end of ORF1 as well as the ORF1/ORF2 junction are two well-conserved regions commonly targeted for NoV detection (shown in Figure 1.3); however, different sets of primers are required for each genogroup because of the large genetic diversity of the virus. For diagnostic purposes, real time RT-PCR is frequently preferred over conventional RT-PCR for its higher sensitivity, faster turnaround times, semi-quantitative estimation of viral load, lower risk of contamination (due to its closed system approach) and ability to include multiple targets within one assay [80]. On the other hand, conventional RT-PCR is used mostly for genotyping purposes. The regions targeted for NoV genotyping (shown in Figure 1.3) are more informative (i.e. less conserved) than those for used for diagnostic purposes.

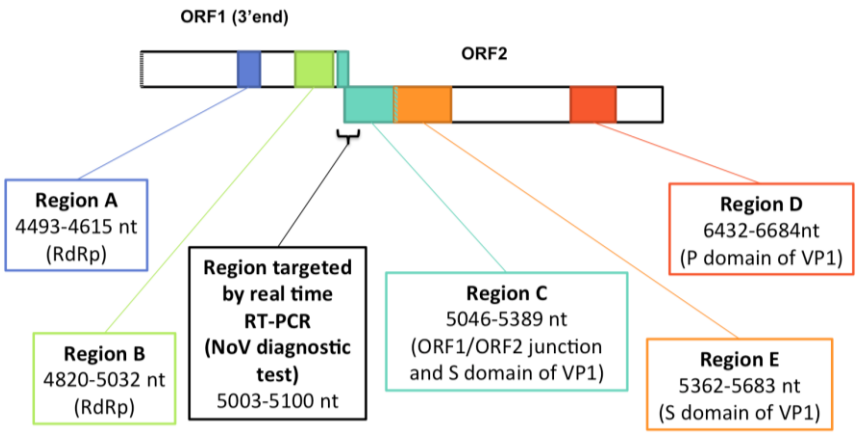


Figure 1.3. NoV regions frequently used for genotyping purposes

Positions are given based on X86557. References describing the primers used for each region are: [81] for region A, [82] for region B, [83] for region C, [84] for region D, [85] for region E and [80] for real time RT-PCR.

1.6.4. Multi-gastrointestinal-pathogen diagnostic platforms

These platforms are RT-PCR-based multiplexed assays. The xTAG Gastrointestinal Pathogen Panel (Luminex Corporation, Toronto, Canada), FilmArray GI Panel (BioFire Diagnostics Inc., Salt Lake City, UT, United States), and Verigene Enteric Pathogens Test (Nanosphere) have been approved by the FDA for diagnostic purposes [58]; while the first two have been approved by Health Canada. The major advantage of these tests is the parallel detection of multiple enteric pathogens (14, 22 and 9, respectively) including bacteria, viruses and parasites within a relatively short time. A disadvantage of these platforms is the high cost of consumables and equipment.

1.7. Molecular epidemiology

1.7.1. Outbreaks

Norovirus is the leading cause of gastroenteritis outbreaks around the world, responsible for over 50% of all gastroenteritis outbreaks [86-92]. Typical settings for NoV

outbreaks include nursing homes, retirement centres, hospitals, schools, day care centres, restaurants, cruise ships and military facilities [70, 93-95].

NoV GII.4 has been the most predominant genotype around the world since the 1990's causing over 60% of all NoV outbreaks [96]. Genotype GII.4 has showed a fast paced evolution with new intra-genotype phylogenetic clusters (variants) emerging periodically and spreading rapidly. Five GII.4 variants have emerged since 2000 and caused global pandemics in different years: Farmington Hills 2002, Hunter 2004, Den Haag 2006b, New Orleans 2009 and Sydney 2012, [86, 96, 97]. GII.4 variants Asia 2003 and Yerseke 2006a were also responsible for increased outbreak activity in Asia and Australia, respectively, but maintained limited geographical circulation [96]. More recently, starting during the winter of 2014-2015, outbreaks of genotype GII.17 became predominant in China [98] and Japan [99] and have raised concerns of a possible global NoV pandemic in a naïve population [100]. GII.17 strains have been also been detected in Europe and North America [101-104]; however, to date, there are no reports of high levels of outbreak activity due to GII.17 in these regions where GII.4 Sydney remains as the predominant strain.

While GII.4 variants predominate in outbreaks, differences in genotype distribution have been observed depending on the type of outbreak setting. Genotype I and genotype II-non-GII.4 strains appear to be more common in settings associated with foodborne or waterborne transmission (e.g. restaurants, schools, child care centers and social events) compared to settings where transmission is often person-to-person (long-term-care facilities and cruise ships) [94, 105-107].

1.7.2. Sporadic disease

Sporadic NoV disease is frequent among all ages, but children under 6 months as well as individuals between 18-64 years show lower prevalence compared to other age groups [108-110]. Moreover, after the introduction of Rotavirus vaccination programs NoV has emerged as the leading cause of gastroenteritis among children [111-114].

In a meta-analysis including data from 23 countries, Patel *et al.* estimated that NoV is responsible for 12% of all gastroenteritis cases in the community and clinics (mild and moderate diarrhoea) and 11% of all gastroenteritis cases in hospitals (severe diarrhoea) [66]. NoV sporadic disease is largely under reported in Canada; during 2001-2002, the Public Health Agency of Canada performed a population-based survey of retrospective self-reported gastroenteritis in the community and found that only 20% of gastroenteritis cases visit a physician and only 4% submit samples for analysis [115, 116]. After adjusting for underdiagnosis, Thomas *et al.* estimated that in Canada, during 2006, 4.0 million episodes of gastroenteritis illness with NoV disease occurred accounting for 104 cases per 1,000 persons-year [117]. For comparison, the United States, reported an incidence of 41 and 69.5 cases per 1,000 person-years at the community level and 5.4 and 5.6 cases per 1,000 person-years for outpatient visits during 2004-2005 and 2012-2013 respectively [118, 119]; a prospective cohort study in the Netherlands reported an incidence of 32 cases per 1,000 persons-year in 1998-1999 [108]; and a similar study in the United Kingdom reported 47 cases per 1,000 persons-year in 2008-2009 [120]. The incidence estimated by Thomas *et al.* is about 2-3 times higher than that observed in the United States, Netherlands and United Kingdom probably due to differences among populations and years of study. There is evidence that variant Den Haag 2006b emerged and circulated in Canada during 2006-2007 reaching the highest levels in NoV outbreak activity registered in Alberta to date [94, 121]. On the other

hand, it is also possible that the estimate for Canada might include noise from retrospective self-reported data, which can overestimate incidence [122].

NoV strains from genogroups I and II are the most common among sporadic cases. Genotype GII.4 is found overwhelmingly worldwide, although few temporospatial changes in strain prevalence have also been reported (e.g. genotype GII.3 in Malawi in 2005-2006 [123] and genotype GII.17 in China since 2014 [124]). Childhood sporadic NoV gastroenteritis plays a role in some of these changes, as children are more susceptible to infection by multiple strains during the first years of life [125-127].

1.7.3. Seasonality

The winter seasonality of NoV has been described since the earliest reports of the disease [1, 128]. The reason for this seasonality has not been elucidated and several factors have been proposed including environmental factors and host related factors. Each factor is highly plausible; thus, NoV seasonality could actually be result of a combination of some, if not all of these factors.

Low temperatures and increased rainfall have been positively associated with epidemic NoV disease in the Northern hemisphere; although, the association is less clear for tropical areas including South America, Africa and Australia due to scarce data which is not robust [129]. A systematic review of the global seasonality of NoV identified that, overall, about 75% of NoV cases and outbreaks occurred during in the cool months (between October and March in the Northern Hemisphere, and between April and September in the Southern Hemisphere) [129]. Other environmental factors that have been associated are low lake surface temperatures (≤4°C) and high river flows [130]. Congruently, experimental studies have shown that high temperatures reduce the infectivity and stability for MNV and FCV

(culturable surrogates of HuNoV) [131] and an absolute humidity below 0.007 Kg water/Kg air, commonly observed in winter periods, is favourable for NoV stability [132].

Host-related factors such as behaviour, health status and birth rates could also explain NoV seasonality. During winter periods, crowding and time spent indoors increase which could facilitate NoV transmission by aerosolization (see section 1.8) [133]. Another hypothesis is that variations in light/dark annual cycles (i.e. a decrease in light hours during winter periods) may trigger physiological changes making the host more susceptible to infection [134]. Herd immunity could also shape increases in NoV activity [135]; although no studies have confirmed this association by measuring the population's immunity, birth rate, an indicator of how frequently the pool of susceptible individuals within a population is replenished, has been positively associated with NoV season strength [129].

1.8. Transmission

Norovirus is extremely transmissible and highly infectious, with reported reproduction numbers (Ro) of 14 secondary cases per primary case in the absence of hygiene control practices and 2 after implementation of enhanced hygiene measures [136]. Moreover, the estimated infection dose is low [the HID50 of NV -a GI.1 strain- is 18.2 (95%CI 1.03-4,350)] [137-139] and there are several modes of transmission including foodborne, waterborne, direct (person-to-person) and indirect (person-fomite-person) transmission. Moreover, NoV can be spread by aerosolization of projectile vomiting or feces, which could potentially lead to either direct or indirect transmission.

Foodborne transmission accounts for approximately 37% of all NoV outbreaks around the world [140] and 26% of all NoV outbreaks occurring in the United States, where NoV is also responsible for 58% of all foodborne disease [141]. In Canada, about 1 million annual

cases of foodborne NoV disease were estimated to occur in 2006 [117]. Food items that act as vehicles for foodborne transmission of NoV are produce, shellfish and ready-to-eat food [106]. Most NoV foodborne outbreaks (~89%) have been associated with the health and poor hygiene practices of food workers [142].

Waterborne transmission of NoV has been documented by multiple studies [143-148]. In Europe, NoV has been identified in ~9% of recreational waters [149], representing a potential vehicle of exposure.

Person-to-person transmission is considered the primary mode of transmission for the majority (66%) of reported NoV outbreaks in the United States [92] and can occur via the oral-fecal route or via ingestion of aerosolized vomitus [150, 151]. Projectile vomiting is a very effective way of transmission: NoV has been found in 38-92% of vomitus samples at a high concentration (median concentration ranges between 5×10^3 to 9.2×10^5 virus per mL) [137, 152] and a single vomiting event could potentially contaminate an area up to $7.8 m^2$ [153]. There is evidence suggesting that NoV might survive in vomiting suspension and emesis up to 42h [154, 155], further contributing to indirect transmission of NoV via fomites. Moreover, experimental studies have shown that HuNoV can persist >61 days in water and up to 56 days in stainless steel and plastic surfaces [156]. This evidence, along with multiple compelling reports of transmission via fomites [157-162] suggests that environmental transmission could be rather common for NoV.

1.9. Clinical symptoms and pathogenesis

The incubation period of NoV lasts about 1.1-1.2 days [163]. The disease usually starts with a sudden onset of acute non-bloody diarrhoea, vomiting, nausea and abdominal cramps although some individuals could experience either vomiting only (more common in children)

or diarrhoea only (more common in adults) [164-166]. In 37-45% of patients symptoms include fever and body aches within the first 24h (hence the nickname "stomach flu" despite no involvement of influenza) [167] (Kaplan, 1982). The symptoms often self-resolve within 10-61 h in challenge studies [168] but the virus can still be detected in stools for an average of 28 days (range 13-56 days) after infection, with the highest vial load on feces being observed 2-5 days after infection [168]. Some rare complications associated with NoV include: necrotizing enterocolitis [169-172], post-infectious irritable bowel syndrome [173] exacerbation of inflammatory bowel disease with bloody diarrhea [174], benign convulsions [175, 176] and encephalopathy [177, 178].

The pathophysiologic mechanism behind NoV diarrhea is still unclear. Besides broadening and blunting of the villi, there are no signs of major structural damage of the intestinal wall. Some secretory/absorptive processes are altered, including malabsorption of D-xylose, fat and lactose [179]. The small intestine is presumed to be the site of primary infection; however, to date, there is no experimental evidence showing that HuNoV infects enterocytes. Recently, it has been discovered that HuNoV can withstand low-level replication in B-cells [3, 180], somewhat resembling MNV in mice, which can infect macrophages, dendritic cells, and B cells [180, 181].

1.10. Populations at risk

1.10.1. Children

Between 12-21% of all gastroenteritis disease in children younger than 5 years of age is attributed to NoV [66, 182, 183]. Children are largely prone to infection by NoV and present a higher prevalence of GI and GII-non-GII.4 strains compared to other groups [126]. A prospective cohort study involving children in low- and middle-income countries observed

that by 12 months of age 68% of the children already suffered at least one NoV infection, and that by two years of age the percentage increased to 89% [184]. A similar cohort study in Ecuador found that during the first 3 years of life, 66% children had ≥1 episode of NoV infection, 40% had 2 infections and 16% had >2 infections [185]. Infection in premature infants could lead to life-threatening severe disease as NoV has been implicated in several cases of necrotizing enterocolitis [169-172]. Children ≤2 years of age have a longer and more severe NoV disease compared to children 2-5 years old (average of 7 days and 3.5 days, respectively) [186]. Compared to rotavirus, the severity of NoV infection in children is similar for hospitalized and outbreaks cases [187, 188] but lower for cases from the community [184]. [189].

1.10.2. Elderly

The process of aging leads to a reduction in the capacity of the immune system to respond to pathogens. Innate, cell-mediated and humoral immune responses are affected; a few examples of characteristic changes that occur in the elderly are reduction in the pools of naïve T-cells and naïve B-cells, inflamm-aging (constant stimulation of the immune system causing subclinical inflammation) and decreased phagocytic activity by neutrophils and macrophages [189-191]. Correspondingly, the elderly have the highest rates of NoV mortality, approximately 200% higher than that reported for children <5 years of age [192]. The elderly present more NoV associated hospitalization in intensive care units compared to children, they are also more likely to receive antibiotics and require a longer stay at the hospital (≥3 days) [126] resulting in greater costs [192]. While diarrhea, nausea and vomiting recede within 3-4 days, non specific symptoms like anorexia, thirst and lethargy can persist for up to 19 days in these patients [193]. Moreover, the elderly are more likely to have previous underlying conditions that might increase the severity of NoV disease [126].

1.10.3. Immunocompromised individuals

NoV infections have been reported in patients suffering congenital immunodeficiency, solid organ transplantation, hematopoietic stem-cell transplantation (HSCT), cancer or cancer chemotherapy and human immunodeficiency virus (HIV) infection [194, 195]. In immunocompromised patients NoV disease can be life threatening; a systematic review of norovirus-associated deaths identified chemotheraphy and HSCT as the most common underlying condition, observed in 59% of NoV deaths of subjects with an underlying condition [196]. Different studies suggest that the burden of NoV disease in immunocompromised patients can be considerable. Frange *et al.* identified 11 NoV-positive patients in a group of 62 children with primary immunodeficiencies [197]. In a retrospective study involving 55 HSCT recipients <21 years of age, Robles *et al.*, reported a cumulative incidence of 12% at two years after transplantation [198]. Roddie *et al.* observed 2 cases of NoV among 12 HSCT adult patients within a period of 6 months [199] and Schorn *et al.* observed 9 NoV cases among 78 adult kidney allograft recipients within a 2 year period [200]. In most immunocompromised patients NoV infection can become chronic, lasting months or even years [52, 197-204] causing severe weight loss, dehydration and malnutrition. Furthermore, it has been hypothesized that immunocompromised patients could be reservoirs where new NoV strains could evolve by accumulating ample mutations leading to antigenic change [202, 205].

1.11. Genetic susceptibility to Norovirus

The histo-blood goup antigen (HBGA) phenotype profile of an individual has been identified as a factor determining the susceptibility to infection by NoV. HBGAs are comprised by the ABH and Lewis families of antigens, a highly polymorphic group of oligosaccharide epitopes that are present on the surface of red blood cells, epithelial cells of

the gastrointestinal, respiratory and genito-urinary tracts and also as free antigens in biological fluids including saliva, milk and intestinal secretions [206]. The biosynthesis of ABH and Lewis antigens is governed by several fucosyl transferases (Figure 1.4). The interaction between NoV and HBGAs is a prerequisite for NoV attachment into host cells and human saliva and follows the lock and key principle [54, 207]. Noroviruses are highly diverse at the HBGA binding interface and thus, display different specificities for ABH and Lewis antigens. Eight binding patterns have been described so far based on binding assays with synthetic oligosaccharides and human saliva [208] (Table 1.2). Importantly, these patterns should not be considered genogroup/genotype specific in all instances, as a simple amino acid mutation can result in a change of HBGA specificity [53].

Patterns 1 and 2 have been supported with data from challenge studies and observational studies on patients suffering natural NoV infection. For instance, individuals with blood group B and non-secretor phenotype are resistant to infection by NV, a GI.1 strain (Pattern 2) [209-212]. Non-secretors are also resistant to infection by GII.4 strains (Pattern 1) [185, 213, 214]. Non-secretors individuals do not have a functional FUT2 enzyme (Figure 1.4) and thus do not present ABH antigens in saliva or intestinal epithelial cells [215]. The proportion of non-secretors varies between 3 to 54% among humans [216, 217] and represent ~20% of the Caucasian population [218].

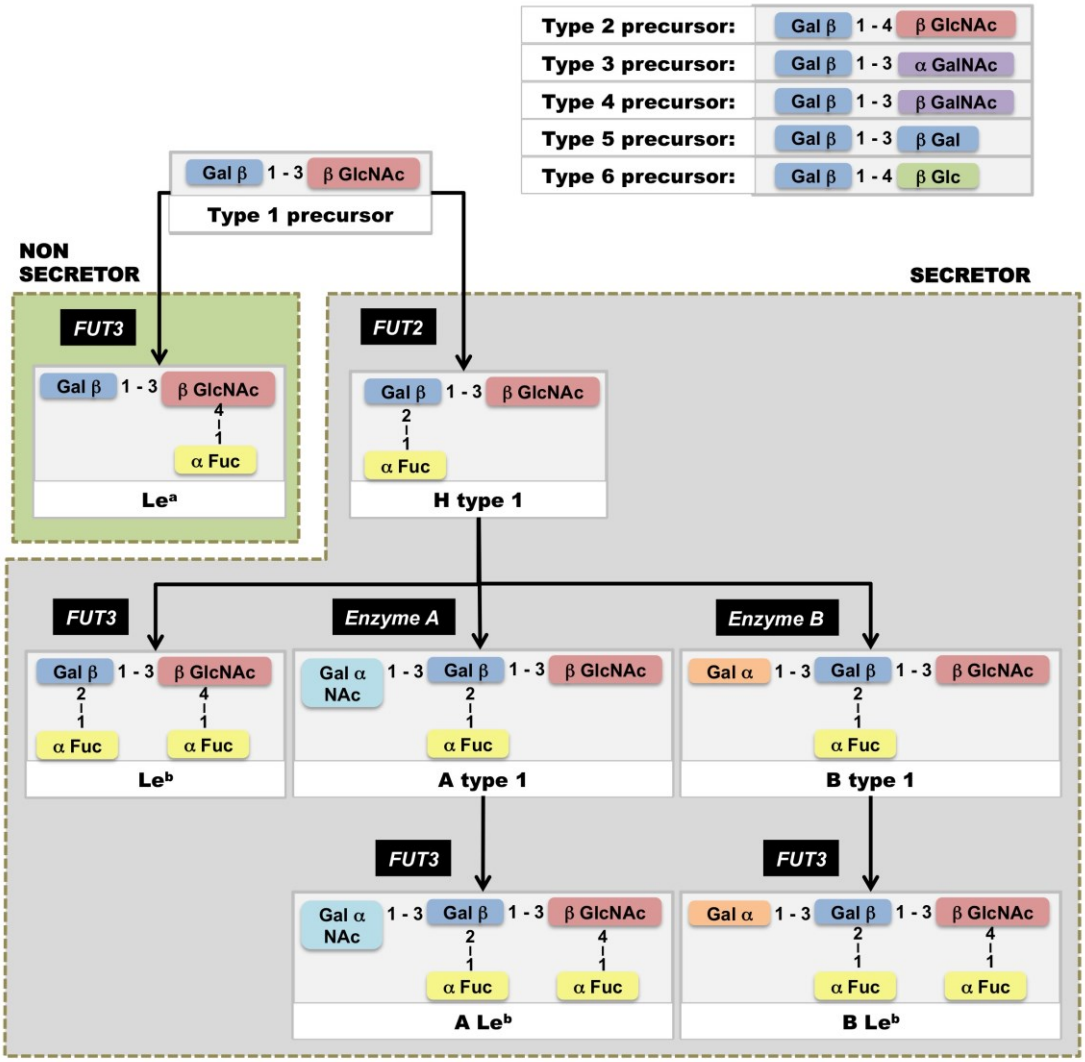


Figure 1.4. Biosynthetic pathways of HBGAs type 1

The HBGAs are synthesized from 6 different types of precursor molecules. Types 1-4 have been associated with NoV binding; type 5 has not been found in humans and type 6 has been found only in exocrine secretions like milk and urine [206]. The biosynthetic pathway of type 1 antigens is shown in the graph. FUT2 is a α-1,2 fucosyl transferase that adds a fucose moiety to the precursor molecule producing H antigens in saliva and other mucosal secretions [FUT1 enzyme (not shown), is responsible for the same activity on erythrocytes]. The A and/or B antigens can be synthesized depending on the genetic background of an individual (i.e. if the corresponding enzymes are functional). FUT3 is a α-1,3/4 fucosyl transferase that synthesizes the Lewis (Le) antigens by transferring a fucose moiety to H, A or B antigens. Individuals homozygous for inactivating mutations in FUT2 are referred as "non-secretors" because they do not present ABH antigens in saliva or other body secretions. Non-secretor individuals can only synthesize Le^a antigens. The pathway for type 2

antigens is similar to type 1 antigens, except that the molecules produced by FUT3 from the type 2 precursor and H molecules are denoted Le^x and Le^y, respectively. The pathways of type 2-6 precursors are described in more detail in reference [215] and [206].

The pattern of NoV biding to HBGAs can therefore affect the prevalence of certain NoV strains. For example, NoV strains that can bind A or B epitopes but no H epitopes should be less common than those that can bind H epitopes only, given that prevalence of individuals carrying H epitopes (i.e. subjects with blood type O) is high. NoV GII.4, the most predominant genotype in humans, can bind to a broad group of HBGAs, that include H, A, B, Le^b and Le^y epitopes (Table 1.2). Moreover, there is evidence suggesting that there could be additional NoV receptor molecules yet to be identified [219-221].

Table 1.2. Patterns of NoV binding to synthetic oligosaccharide conjugates and human saliva

Binding pattern	Strain (Genotype)	Saliva binding assay				Oligosaccharide-based binding assay	
		Non Secretor	Secretor			Binding	Not binding
		O	O	A	B		
1	VA387 (GII.4)	-	√	√	√	Le^b, Le^y, H1, H3, A, B	Le^a, Le^x, H2
2	NV (GI.1)	-	√	√	-	Le^b, Le^y, H1, H3, A	Le^a, Le^x, H2, B
3	MxV (GII.3)	-	√	√	√	Le^b, A, B	Le^a, Le^x, Le^y, H1, H2, H3
4	MOH (GII.5)	-	-	√	√	A, B	Le^a, Le^x, Le^b, Le^y, H1, H2, H3
5	SMV (GII.2)	-	-	-	√	H3, B	Le^a, Le^b, H1, H2, A
6	VA207 (GII.9)	√	√	√	√	Le^x, Le^y	Le^a, Le^b, H1, H2, H3, A, B
7	Boxer (GI.8)	√	√	√	√	Le^b, Le^y,	Le^a, Le^x, H1, H2, H3, A, B
8	OIF (GII.21)	√	√	-	-	Le^a	Le^x, Le^b, Le^y, H1, H2, H3, A, B

NOTE: There are NoV strains not contemplated in these patterns that do not bind any known HBGA and thus might have other receptor(s). The table has been adapted from reference [208].

1.12. Norovirus evolution

1.12.1. Antigenic drift

Compared to other genotypes, GII.4 has the largest number of sequences in public databases and has provided significant insights into the mechanisms of NoV evolution. GII.4 variants that have circulated since 1987 present time-ordered amino acid changes, mostly at the P2 domain (the outermost part of the virus), which affect the antigenic properties of the virus [53, 222, 223]. Compared to other regions of VP1, the P2 domain presents more sites under positive selection, most likely to allow escape from herd immunity built against previous variants [53]. In support of this premise, epitopes that can be targeted to neutralize the binding of NoV GII.4 to HBGAs have been identified and mapped into 5 regions (A to F) that evolve between pandemic variants [224-228]. Moreover, convalescent sera from individuals infected with past GII.4 variants are less reactive against later variants compared to the homologous strain [53]. Based on these observations, it has been proposed that NoV GII.4 evolves by a mechanism similar to influenza's epochal evolution, in which periods of stasis in phenotype are followed by evolutionary jumps resulting in new phenotypes [53].

In view that some of the sites under positive selection in VP1 are located in close proximity to the HBGA binding sites and that a single amino acid replacement can lead to changes in HBGA specificity, it has been suggested that NoV might use antigenic drift to expand its pool of susceptible individuals within a population [229]. On the other hand, it has also been demonstrated that GII.4 strains have maintained their HBGA binding profile relatively conserved at least since 1987 [230, 231]. Thus, although changes in HBGA binding specificity might potentially occur, these might not be common due to negative selection by HBGAs [232].

The number of non-GII.4 sequences available in public databases is considerably lower compared to GII.4 [61]. Analyses on the evolution of non-GII.4 genotypes are scarce and have shown different evolutionary patterns compared to GII.4. In general, genogroup I (GI) appears to evolve slower than genogroup II (GII) as GI presents shorter intergenotype phylogenetic distances compared to GII [233]. Protein modeling analyses also suggest that there is a functional restriction for GI strains to accommodate and tolerate mutations. Specifically, the P2 subdomain of VP1 of GI strains is a few amino acids smaller and has less plasticity compared to that of GII viruses [233].

NoV GII.2 strains seem to evolve a lower rate compared to GII.4. Iritani et *al.* observed a modest change in the capsid sequences of GII.2 strains over a period of 30 years (≥83.6% in nucleotide and ≥93.5% in amino acid identities from 1976 to 2005) [234]. For comparison, while a 5% amino acid diversity has been reported between GII.4 variants that emerge every two years [60], GII.2 strains showed 2.6% amino acid diversity in a 12-year period [234]. The low evolutionary rate of this genotype might be a consequence of the little circulation these strains have among humans yet, GII.2 strains are widely spread around the world [234]. NoV GII.3 represents another peculiar pattern of evolution. Similarly to GII.4, GII.3 is also a prevalent genotype among children and displays a large genetic diversity, presenting at least 3 intra-genotype clusters [235]. Moreover, the evolutionary rate of the capsid gene in GII.3 viruses is 4.16×10^{-3} to 7.39×10^{-3} nucleotide substitutions/site/year, which is comparable to that of GII.4 strains (4.3×10^{-3} to 6.5×10^{-3} nucleotide substitutions/site/year) [235]. However, GII.3 strains seem to revert some amino acid changes in time, resulting in a low overall difference in amino acid sequences. Noticeably, the third GII.3 cluster (1998-2006) is actually more similar to the first cluster (1975-1979) and more distant to the second cluster (1987-1995) [235].

1.12.2. Recombination

NoV can undergo a process of homologous recombination, where two viruses co-infecting a cell exchange genomic sequences [236]. There are reports of NoV recombinants with breakpoints within ORF2, at the ORF1/ORF2 junction and at the ORF2/ORF3 junction [237-240]. The ORF1/ORF2 junction represents the most common hot spot for NoV recombination. A copy-choice mechanism has been proposed for the virus, where the RNA polymerase momentarily stalls at the secondary structures present at the ORF1/ORF2 overlap increasing its chances to switch templates during RNA synthesis [241] (Figure 1.5).

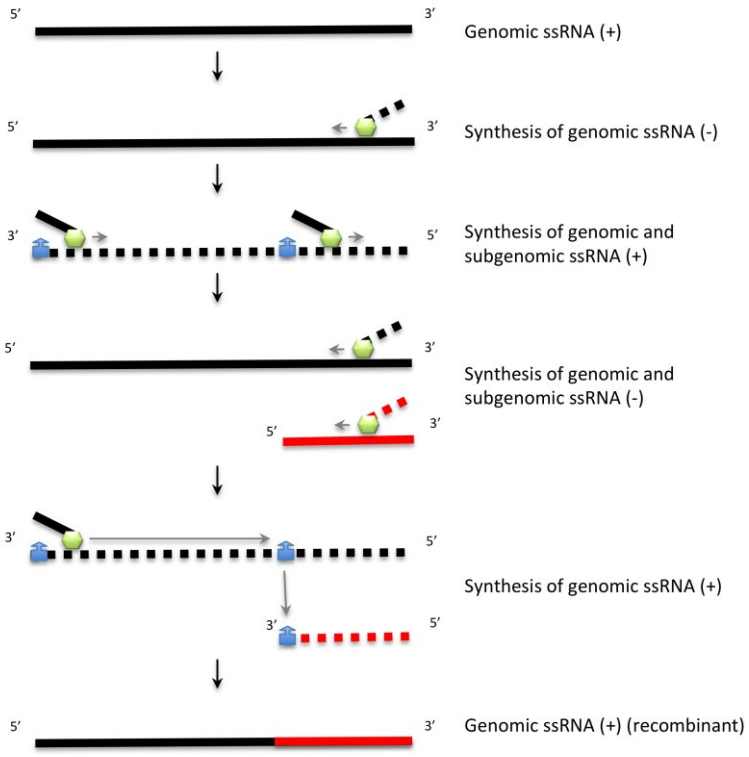


Figure 1.5. Mechanism of recombination in NoV

The mechanism is depicted as proposed by Bull *et al.* [241]. Recombination occurs between two viruses that are co-infecting a cell. During the process of viral replication, genomic and subgenomic RNA are synthesized for each virus. During synthesis of genomic RNA (+), the viral polymerase

pauses at the secondary structures (blue) present at the ORF1/ORF2 junction, promoting a switch of template from the genomic RNA of a virus (black lines) to the subgenomic RNA of the other co-infecting virus (red lines).

Recombination is an important mechanism of evolution used by NoV that further expands the large genetic diversity of the virus. A large number of intragenotype [242, 243] and intergenotype [238, 244] recombinants have been described. In addition, the first ORF1-ORF2 intergenogroup recombinant (GII.9/GI.7) was recently reported [244]. Recombination among GII.4 variants is not unusual [237] and has been associated with the generation of pandemic variants Den Haag 2006b, New Orleans 2009 and Sydney 2012 [242].

1.13. Immunity to Norovirus

The immune responses against HuNoV are not well understood due to the host-specific nature of the virus. Without animal models it is practically impossible to control for factors such as previous exposure history. Our limited knowledge in this area has been obtained through challenge studies in human volunteers, cohort studies and outbreak patients. As a result, the data reported by some studies are conflicting. For instance, early challenge studies suggest that the protection conferred after NoV infection can last 2-6 months for the homologous strain [245], whereas population-based incidence data and mathematical modeling suggests that NoV protective responses might last 4-9 years [246]. While innate immune responses most likely play an important role limiting HuNoV replication, as observed for MNV [247-250], the fact that immunocompromised patients suffer prolonged infections, suggests that adaptive immune responses are crucial for viral clearance [248, 249].

1.13.1. Humoral responses

The seroprevalence of NoV in human populations is high, ranging between 54% and 98% across different populations [251-258]. The protection conferred by serum antibodies after NoV infection is controversial. Studies have reported an association between the presence of anti-NoV IgG and protection against infection in children <5 years old [259, 260] but not in adults, in whom high pre-existing titers of anti-NoV IgG or IgA -either in serum or feces- are not correlates of protection [245, 261-265]. Furthermore, multiple exposures appear to be required to confer protection against NoV infection [266].

The HBGA blocking antibodies, surrogates of neutralizing antibodies, constitute the best correlate of protection against NoV infection identified to date [264, 267, 268]. Importantly for vaccine development, protective antibody responses appear to be genogroup specific and cross reactive within genogroups. A lack of cross protection between GI.1 and GII strains has been reported by several challenge studies [167, 269] and a study of a children cohort [266]. Cross-reactivity within NoV GII has been extensively documented [85, 269-272] and has been observed between GII.4 variants [231, 273].

A recent clinical trial of a vaccine candidate has revealed that immune responses to NoV appear to follow a pattern of “original antigenic sin” or “antigenic seniority”, where a current NoV strain can activate pre-existing B-cell memory responses against an older, antigenically different, strain [233, 274]. The premise of antigenic seniority is that individuals are likely to favour antibody responses of higher affinity, which are shaped in childhood [274].

1.13.2. Cellular mediated immunity

In comparison with humoral immunity, cellular mediated responses to HuNoV are less characterized. There are only a few studies in the area, but they all agree in that immunity seems to be skewed towards a Th1 response with some involvement of Th2 responses. In a SMV (GII.2) challenge study, post-challenge peripheral blood mononuclear cells (PBMCs) displayed a significant increase in IFN-γ (Th1), IL-2 (Th1) and IL-5 (Th2) cytokine secretion upon VLP stimulation, compared to pre-challenge PBMCs [269]. Congruent with a predominant Th1 response, post-challenge sera also displayed a significantly higher concentration of IFN-γ, IL-2 as well as IgG1 and the primary source of IFN-γ were $CD4^+$ T cells [269] [275]. Another challenge study with NV (GI.1) also reported an increased secretion of IFN-γ by CD^+4 T cells upon stimulation with the homologous strain [276]. Interestingly, this study also compared the post-challenge IFN-γ secretion responses in PBMCs stimulated with a panel of different GI genotypes and identified that 66% of the subjects had responses preferentially targeted to strains different than the challenge strain. These results suggest that cell mediated immunity, like humoral mediated immunity, could be subject to 'antigenic seniority' (i.e. affected by past exposures, see section 1.13.1).

Another study involving hospitalized children with acute gastroenteritis, identified that compared with bacterial infections, NoV and rotavirus infections induce less IL-6 production during the acute phase of the disease and that during NoV infection there is less production of IL-8 compared to rotavirus [188]. IL-6 and IL-8 cytokines, which are involved in the process of inflammation [277, 278], were also positively correlated with disease severity and duration of fever, respectively [188].

1.14. Norovirus in Alberta

In Canada, NoV is considered a notifiable disease under the Public Health Act. NoV outbreak surveillance was implemented in Alberta for the Northern and Central region of the province since 1999 and was later expanded to cover the entire province in March 2002 [94]. During outbreak investigations, samples are collected and submitted for laboratory investigations. An outbreak is defined as "two or more cases of clinical illness compatible with NoV that can be epidemiologically linked to one another" (i.e. with onsets within a 1-3 day period) and all outbreaks require laboratory confirmation of at least one sample in order to be identified as a NoV outbreak [279].

In the period between July 1999 and June 2008 there were a total of 1,057 NoV outbreaks [94]. The annual distribution of these outbreaks displayed a clear biennial pattern where years of high outbreak activity alternated with years of low outbreak activity (Figure 1.6a). NoV genogroup II represented 677 (96%) out of 707 NoV outbreaks that were genotyped. GII.4 was by far the most predominant genotype, identified in 617 (87%) of all genotyped outbreaks. The monthly distribution of GII.4 outbreaks followed the characteristic winter seasonality of NoV. Particularly, the data showed that every year with high levels of NoV outbreaks was associated with the emergence of a novel GII.4 variant in the province. During the eight period of NoV surveillance, a total of four epidemic GII.4 variants emerged during the spring, summer or fall seasons and became predominant in the immediate winter season of the same year: US95-96 (winter 2000-2001), Farmington Hills 2002 (winter 2002-2003), Hunter 2004 (2004-2005), Den Haag 2006b (2006-2007) (Figure 1.6b); variants Yerseke 2006 and Osaka 2007 were also observed but with less prevalence. Other GII genotypes also identified in the same period were GII.1, GII.2, GII.3, GII.4, GII.5, GII.6, GII.9 and GII.13.

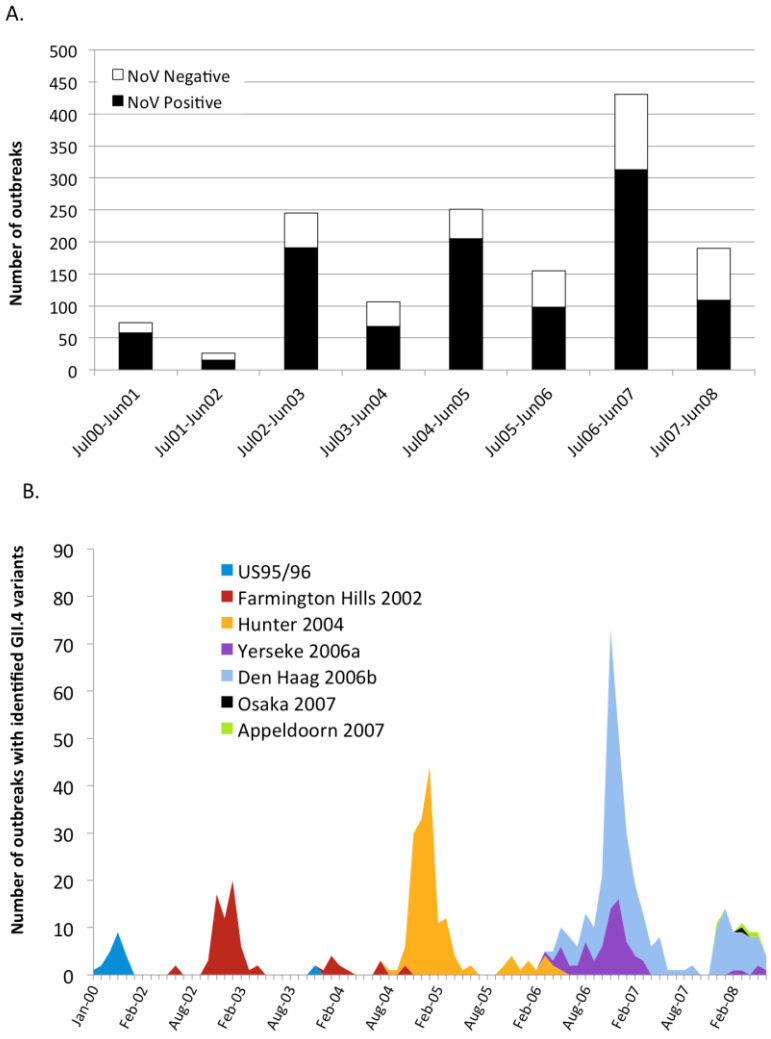


Figure 1.6. NoV outbreaks in Alberta between July 2000 and June 2008

A) Annual distribution of NoV positive outbreaks. Years were considered to start every month of June in order to capture complete NoV seasons. B) Monthly distribution of GII.4 outbreaks per month and per variant. Figure modified from Pang *et al.* [94].

The three most common outbreak settings for NoV GII.4 outbreaks were community-based group residences (69.2%), hospital acute care (14.9%) and hospital long term care (7.1%); the common settings of non-GII.4 outbreaks were community-based group residences (62.2%), schools (8.9%) and social events (7.8%) [94].

In addition, a group of cases with sporadic NoV disease occurring between July 2008 and July 2009 were analyzed [109]. The monthly distribution of cases exhibited winter seasonality similar to outbreak cases. The age groups with highest incidence of sporadic disease by NoV GII were children between >6 months and 2 years of age and individuals >60 years of age. In the case of NoV GI, the highest incidence occurred in children >5 to 10 years of age.

1.15. Current caveats

The development of effective vaccines, anti-viral drugs and effective control measures to stop HuNoV spread have been largely affected by the absence of *in vitro* and *in vivo* systems to study HuNoV. A recently developed cell culture system for HuNoV [180] has opened opportunities for advancement in the area but it still faces problems of reproducibility and low yield [3]. Cultivable surrogates for HuNoV, including FVC, Tulane Virus, MNV have been used extensively to provide insights on aspects of stability [131, 280-283], viral replication [37, 284, 285] and host immune responses [250, 286-288], but their relevance is unknown.

Gnotobiotic pigs and calves as well as chimpanzees can be inoculated via oral route with HuNoV and allow viral replication; however, the clinical outcomes in these hosts differ from those observed in humans. Gnotobiotic pigs and calves infected develop mild diarrhea only and chimpanzees present asymptomatic infection [289-292]. Balb/c RAG/$\gamma c^{-/-}$ (immunocompromised) mice can also be infected by HuNoV but are asymptomatic [293]. Wild-type mice infected with MNV are also asymptomatic and can transmit the virus by the fecal-oral route. Although MNV presents a large genetic distance compared to HuNoV, the wild-type/MNV animal model has proved to be useful in providing important clues for the development of the HuNoV cell culture system [180].

In absence of an efficient culture system, the only source of HuNoV virions is human stools. Studies analyzing the antigenic properties of HuNoV make use of virus-like particles (VLPs) produced with eukaryotic expression systems [e.g. Venezuelan Equine Encephalitis (VEE) replicon, baculovirus or yeast based systems] [53, 267, 294]. The P-particles constitute an alternate source of NoV antigens. The P-particles are formed by the P domain of VP1, are smaller in size than VLPs and share similar antigenic characteristics with VLPs [295, 296]. P-particles have the advantage that they can be expressed in soluble form by *E. coli* expression systems with reduced time and costs. VLP and P-particle based vaccines are currently being tested as potential vaccine candidates [297]. No vaccine is commercially available yet; an oral vaccine by Vaxart is in a phase one clinical trial [298] and a bivalent (GI.1 and GII.4) VLP based vaccine developed by Takeda Vaccines is currently in a 'proof-of-concept' phase clinical trial [297, 299].

1.16. Rationale of the project

In Alberta, Canada NoV outbreak surveillance data for the period 2000-2008 showed a characteristic biennial pattern of outbreak activity, where each peak of outbreak activity coincided with the emergence of a new GII.4 variant. The aim of this thesis was to continue monitoring the outbreak activity of NoV in Alberta from 2008 onwards and to identify factors that could be associated with increases in epidemic activity.

Chapter two describes the molecular epidemiology of Norovirus in Alberta between 2008 and 2015. During this 7-year period, GII.4 continued to be the most predominant genotype. Deviations from the original biennial pattern were observed, which suggested that other factors in addition to the emergence of novel GII.4 variants affected the levels of NoV outbreak activity in those years. However, since the genotype of NoV strains was allocated based on a small (322 bp) portion of the VP1 gene potentially missing genetic diversity, near-

full length genome analyses were performed to better characterize the circulating GII.4 strains which are described in **chapter three**. Phylogenetic analysis as well as recombination detection analysis confirmed that the Albertan strains were comparable to those described in global pandemics. **Chapter four** describes the validation of a next-generation sequencing-based method developed to obtain near full-length sequences of any NoV genogroup, genotype or GII.4 variant. The method also provided enough data to allow an analysis of the intra-host genetic diversity of NoV in samples from acute and chronic cases of NoV infection. Lastly, **chapter five** describes a temporal analysis of the immune responses in the population. The objective of this study was to determine if immune responses in two sub-populations in Alberta fluctuate over time as is observed for NoV outbreak levels. The results from this study showed a high seroprevalence of NoV GII.4 but no significant changes in seroprevalence or antibody titers over time.

1.17. References (chapter 1)

1. Zahorsky, J., *Hyperemesis heimis or the winter vomiting disease.* Arch. Pediatr., 1929. **46**.
2. Kapikian, A.Z., et al., *Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis.* Journal of virology, 1972. **10**(5): p. 1075-1081.
3. Jones, M.K., et al., *Human norovirus culture in B cells.* Nat Protoc, 2015. **10**(12): p. 1939-47.
4. Kapikian, A.Z., *The discovery of the 27-nm Norwalk virus: an historic perspective.* The Journal of infectious diseases, 2000. **181 Suppl 2**: p. S295-302.
5. Mayo, M.A., *Virus taxonomy - Houston 2002.* Arch Virol, 2002. **147**(5): p. 1071-6.
6. *Fields virology*, B.N. Fields, D.M. Knipe, and P.M. Howley, Editors. 2013, Wolters Kluwer/Lippincott Williams & Wilkins Health: Philadelphia :.
7. Prasad, B.V., et al., *Three-dimensional structure of baculovirus-expressed Norwalk virus capsids.* J Virol, 1994. **68**(8): p. 5117-25.
8. Prasad, B.V., et al., *X-ray crystallographic structure of the Norwalk virus capsid.* Science (New York, N.Y.), 1999. **286**(5438): p. 287-290.
9. Jiang, X., et al., *Sequence and genomic organization of Norwalk virus.* Virology, 1993. **195**(1): p. 51-61.

10. Thackray, L.B., et al., *Murine noroviruses comprising a single genogroup exhibit biological diversity despite limited sequence divergence.* Journal of virology, 2007. **81**(19): p. 10460-10473.
11. Herbert, T.P., I. Brierley, and T.D. Brown, *Identification of a protein linked to the genomic and subgenomic mRNAs of feline calicivirus and its role in translation.* J Gen Virol, 1997. **78 (Pt 5)**: p. 1033-40.
12. Lochridge, V.P. and M.E. Hardy, *Snow Mountain virus genome sequence and virus-like particle assembly.* Virus Genes, 2003. **26**(1): p. 71-82.
13. Hardy, M.E. and M.K. Estes, *Completion of the Norwalk virus genome sequence.* Virus Genes, 1996. **12**(3): p. 287-90.
14. Simmonds, P., et al., *Bioinformatic and functional analysis of RNA secondary structure elements among different genera of human and animal caliciviruses.* Nucleic Acids Res, 2008. **36**(8): p. 2530-46.
15. Bailey, D., et al., *Functional analysis of RNA structures present at the 3' extremity of the murine norovirus genome: the variable polypyrimidine tract plays a role in viral virulence.* Journal of virology, 2010. **84**(6): p. 2859-2870.
16. Vashist, S., et al., *Identification of RNA-protein interaction networks involved in the norovirus life cycle.* J Virol, 2012. **86**(22): p. 11977-90.
17. Yunus, M.A., et al., *The murine norovirus core subgenomic RNA promoter consists of a stable stem-loop that can direct accurate initiation of RNA synthesis.* J Virol, 2015. **89**(2): p. 1218-29.
18. Gutierrez-Escolano, A.L., et al., *Interaction of cellular proteins with the 5' end of Norwalk virus genomic RNA.* J Virol, 2000. **74**(18): p. 8558-62.
19. Cancio-Lonches, C., et al., *Nucleolin interacts with the feline calicivirus 3' untranslated region and the protease-polymerase NS6 and NS7 proteins, playing a role in virus replication.* J Virol, 2011. **85**(16): p. 8056-68.
20. Lopez-Manriquez, E., et al., *Norovirus genome circularization and efficient replication are facilitated by binding of PCBP2 and hnRNP A1.* J Virol, 2013. **87**(21): p. 11371-87.
21. Hardy, M.E., *Norovirus protein structure and function.* FEMS microbiology letters, 2005. **253**(1): p. 1-8.
22. Neill, J.D., *Nucleotide sequence of a region of the feline calicivirus genome which encodes picornavirus-like RNA-dependent RNA polymerase, cysteine protease and 2C polypeptides.* Virus Res, 1990. **17**(3): p. 145-60.
23. Jiang, X., et al., *Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein.* Journal of virology, 1992. **66**(11): p. 6527-6532.
24. Ettayebi, K. and M.E. Hardy, *Norwalk virus nonstructural protein p48 forms a complex with the SNARE regulator VAP-A and prevents cell surface expression of vesicular stomatitis virus G protein.* Journal of virology, 2003. **77**(21): p. 11790-11797.
25. Liu, B.L., et al., *Identification of further proteolytic cleavage sites in the Southampton calicivirus polyprotein by expression of the viral protease in E. coli.* The Journal of general virology, 1999. **80 (Pt 2)**(Pt 2): p. 291-296.
26. Baker, E.S., et al., *Inherent structural disorder and dimerisation of murine norovirus NS1-2 protein.* PLoS One, 2012. **7**(2): p. e30534.

27. Hughes, P.J. and G. Stanway, *The 2A proteins of three diverse picornaviruses are related to each other and to the H-rev107 family of proteins involved in the control of cell proliferation.* J Gen Virol, 2000. **81**(Pt 1): p. 201-7.
28. Nice, T.J., et al., *A single-amino-acid change in murine norovirus NS1/2 is sufficient for colonic tropism and persistence.* J Virol, 2013. **87**(1): p. 327-34.
29. Pfister, T. and E. Wimmer, *Polypeptide p41 of a Norwalk-like virus is a nucleic acid-independent nucleoside triphosphatase.* J Virol, 2001. **75**(4): p. 1611-9.
30. Sharp, T.M., et al., *Secretory pathway antagonism by calicivirus homologues of Norwalk virus nonstructural protein p22 is restricted to noroviruses.* Virol J, 2012. **9**: p. 181.
31. Sharp, T.M., et al., *Inhibition of cellular protein secretion by norwalk virus nonstructural protein p22 requires a mimic of an endoplasmic reticulum export signal.* PLoS One, 2010. **5**(10): p. e13130.
32. Cotten, M., et al., *Deep sequencing of norovirus genomes defines evolutionary patterns in an urban tropical setting.* Journal of virology, 2014. **88**(19): p. 11056-11069.
33. Sosnovtsev, S.V. and K.Y. Green, *Identification and genomic mapping of the ORF3 and VPg proteins in feline calicivirus virions.* Virology, 2000. **277**(1): p. 193-203.
34. Burroughs, J.N. and F. Brown, *Presence of a covalently linked protein on calicivirus RNA.* J Gen Virol, 1978. **41**(2): p. 443-6.
35. Daughenbaugh, K.F., C.E. Wobus, and M.E. Hardy, *VPg of murine norovirus binds translation initiation factors in infected cells.* Virol J, 2006. **3**: p. 33.
36. Daughenbaugh, K.F., et al., *The genome-linked protein VPg of the Norwalk virus binds eIF3, suggesting its role in translation initiation complex recruitment.* EMBO J, 2003. **22**(11): p. 2852-9.
37. Leen, E.N., et al., *A Conserved Interaction between a C-Terminal Motif in Norovirus VPg and the HEAT-1 Domain of eIF4G Is Essential for Translation Initiation.* PLoS Pathog, 2016. **12**(1): p. e1005379.
38. Goodfellow, I., et al., *Calicivirus translation initiation requires an interaction between VPg and eIF 4 E.* EMBO Rep, 2005. **6**(10): p. 968-72.
39. Zeitler, C.E., M.K. Estes, and B.V. Venkataram Prasad, *X-ray crystallographic structure of the Norwalk virus protease at 1.5-A resolution.* J Virol, 2006. **80**(10): p. 5050-8.
40. May, J., et al., *The p4-p2' amino acids surrounding human norovirus polyprotein cleavage sites define the core sequence regulating self-processing order.* J Virol, 2014. **88**(18): p. 10738-47.
41. Belliot, G., et al., *Norovirus proteinase-polymerase and polymerase are both active forms of RNA-dependent RNA polymerase.* J Virol, 2005. **79**(4): p. 2393-403.
42. Belliot, G., et al., *Nucleotidylylation of the VPg protein of a human norovirus by its proteinase-polymerase precursor protein.* Virology, 2008. **374**(1): p. 33-49.
43. Weerasekara, S., A.M. Prior, and D.H. Hua, *Current tools for norovirus drug discovery.* Expert Opin Drug Discov, 2016. **11**(6): p. 529-41.
44. Eltahla, A.A., et al., *Nonnucleoside inhibitors of norovirus RNA polymerase: scaffolds for rational drug design.* Antimicrob Agents Chemother, 2014. **58**(6): p. 3115-23.

45. Hogbom, M., et al., *The active form of the norovirus RNA-dependent RNA polymerase is a homodimer with cooperative activity.* J Gen Virol, 2009. **90**(Pt 2): p. 281-91.
46. Ng, K.K., et al., *Crystal structure of norwalk virus polymerase reveals the carboxyl terminus in the active site cleft.* J Biol Chem, 2004. **279**(16): p. 16638-45.
47. Zamyatkin, D.F., et al., *Binding of 2'-amino-2'-deoxycytidine-5'-triphosphate to norovirus polymerase induces rearrangement of the active site.* J Mol Biol, 2009. **390**(1): p. 10-6.
48. Hyde, J.L. and J.M. Mackenzie, *Subcellular localization of the MNV-1 ORF1 proteins and their potential roles in the formation of the MNV-1 replication complex.* Virology, 2010. **406**(1): p. 138-48.
49. Asanaka, M., et al., *Replication and packaging of Norwalk virus RNA in cultured mammalian cells.* Proc Natl Acad Sci U S A, 2005. **102**(29): p. 10327-32.
50. Glass, P.J., et al., *Norwalk virus open reading frame 3 encodes a minor structural protein.* Journal of virology, 2000. **74**(14): p. 6581-6591.
51. Luttermann, C. and G. Meyers, *A bipartite sequence motif induces translation reinitiation in feline calicivirus RNA.* J Biol Chem, 2007. **282**(10): p. 7056-65.
52. Nilsson, M., et al., *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.* Journal of virology, 2003. **77**(24): p. 13117-13124.
53. Lindesmith, L.C., et al., *Mechanisms of GII.4 norovirus persistence in human populations.* PLoS medicine, 2008. **5**(2): p. e31.
54. White, L.J., et al., *Attachment and entry of recombinant Norwalk virus capsids to cultured human and animal cell lines.* J Virol, 1996. **70**(10): p. 6589-97.
55. Vongpunsawad, S., B.V. Venkataram Prasad, and M.K. Estes, *Norwalk Virus Minor Capsid Protein VP2 Associates within the VP1 Shell Domain.* Journal of virology, 2013. **87**(9): p. 4818-4825.
56. Bertolotti-Ciarlet, A., et al., *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.* Journal of virology, 2003. **77**(21): p. 11603-11615.
57. Sosnovtsev, S.V., et al., *Feline calicivirus VP2 is essential for the production of infectious virions.* J Virol, 2005. **79**(7): p. 4012-24.
58. Vinje, J., *Advances in laboratory methods for detection and typing of norovirus.* Journal of clinical microbiology, 2015. **53**(2): p. 373-381.
59. Zheng, D.P., et al., *Norovirus classification and proposed strain nomenclature.* Virology, 2006. **346**(2): p. 312-323.
60. Zheng, D.P., et al., *Molecular epidemiology of genogroup II-genotype 4 noroviruses in the United States between 1994 and 2006.* Journal of clinical microbiology, 2010. **48**(1): p. 168-177.
61. Kroneman, A., et al., *Proposal for a unified norovirus nomenclature and genotyping.* Arch Virol, 2013. **158**(10): p. 2059-68.
62. Hasing, M.E., et al., *Detection and analysis of recombination in GII.4 norovirus strains causing gastroenteritis outbreaks in Alberta.* Infection, genetics and

evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2014.
63. Ahmed, S.M., et al., *Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis.* The Lancet infectious diseases, 2014. **14**(8): p. 725-730.
64. Pires, S.M., et al., *Aetiology-Specific Estimates of the Global and Regional Incidence and Mortality of Diarrhoeal Diseases Commonly Transmitted through Food.* PLoS One, 2015. **10**(12): p. e0142927.
65. Kirk, M.D., et al., *World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis.* PLoS Med, 2015. **12**(12): p. e1001921.
66. Patel, M.M., et al., *Systematic literature review of role of noroviruses in sporadic gastroenteritis.* Emerging infectious diseases, 2008. **14**(8): p. 1224-1231.
67. Gastanaduy, P.A., et al., *Burden of norovirus gastroenteritis in the ambulatory setting--United States, 2001-2009.* J Infect Dis, 2013. **207**(7): p. 1058-65.
68. Scharff, R.L., *Economic burden from health losses due to foodborne illness in the United States.* J Food Prot, 2012. **75**(1): p. 123-31.
69. Tam, C.C. and S.J. O'Brien, *Economic Cost of Campylobacter, Norovirus and Rotavirus Disease in the United Kingdom.* PLoS One, 2016. **11**(2): p. e0138526.
70. Freeland, A.L., G.H. Vaughan, Jr., and S.N. Banerjee, *Acute Gastroenteritis on Cruise Ships - United States, 2008-2014.* MMWR Morb Mortal Wkly Rep, 2016. **65**(1): p. 1-5.
71. Advisors, B.R.E., *The Contribution of the International Cruise Industry to the U.S. Economy in 2014*. 2014: Exton, PA.
72. Chen, C.H. and C.P. Yu, *Comparison of enzyme-linked immunosorbent assay and reverse transcription-polymerase chain reaction for the detection of norovirus in fecal samples.* Research Journal of Chemistry and Environment, 2014. **18**(3): p. 70-75.
73. Kele, B., G. Lengyel, and J. Deak, *Comparison of an ELISA and two reverse transcription polymerase chain reaction methods for norovirus detection.* Diagn Microbiol Infect Dis, 2011. **70**(4): p. 475-8.
74. Morillo, S.G., et al., *Norovirus 3rd Generation kit: an improvement for rapid diagnosis of sporadic gastroenteritis cases and valuable for outbreak detection.* J Virol Methods, 2011. **173**(1): p. 13-6.
75. Kirby, A., et al., *An evaluation of the RIDASCREEN and IDEIA enzyme immunoassays and the RIDAQUICK immunochromatographic test for the detection of norovirus in faecal specimens.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2010. **49**(4): p. 254-257.
76. Khamrin, P., et al., *Evaluation of immunochromatography and commercial enzyme-linked immunosorbent assay for rapid detection of norovirus antigen in stool samples.* J Virol Methods, 2008. **147**(2): p. 360-3.
77. Ambert-Balay, K. and P. Pothier, *Evaluation of 4 immunochromatographic tests for rapid detection of norovirus in faecal samples.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2013. **56**(3): p. 194-198.

78. de Bruin, E., et al., *Diagnosis of Norovirus outbreaks by commercial ELISA or RT-PCR.* J Virol Methods, 2006. **137**(2): p. 259-64.
79. Burton-MacLeod, J.A., et al., *Evaluation and comparison of two commercial enzyme-linked immunosorbent assay kits for detection of antigenically diverse human noroviruses in stool samples.* J Clin Microbiol, 2004. **42**(6): p. 2587-95.
80. Pang, X.L., J.K. Preiksaitis, and B. Lee, *Multiplex real time RT-PCR for the detection and quantitation of norovirus genogroups I and II in patients with acute gastroenteritis.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2005. **33**(2): p. 168-171.
81. Ando, T., et al., *Detection and differentiation of antigenically distinct small round-structured viruses (Norwalk-like viruses) by reverse transcription-PCR and southern hybridization.* Journal of clinical microbiology, 1995. **33**(1): p. 64-71.
82. Beuret, C., A. Baumgartner, and J. Schluep, *Virus-contaminated oysters: a three-month monitoring of oysters imported to Switzerland.* Appl Environ Microbiol, 2003. **69**(4): p. 2292-7.
83. Kojima, S., et al., *Genogroup-specific PCR primers for detection of Norwalk-like viruses.* Journal of virological methods, 2002. **100**(1-2): p. 107-114.
84. Vinje, J., R.A. Hamidjaja, and M.D. Sobsey, *Development and application of a capsid VP1 (region D) based reverse transcription PCR assay for genotyping of genogroup I and II noroviruses.* Journal of virological methods, 2004. **116**(2): p. 109-117.
85. Noel, J.S., et al., *Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995.* Journal of medical virology, 1997. **53**(4): p. 372-383.
86. van Beek, J., et al., *Indications for worldwide increased norovirus activity associated with emergence of a new variant of genotype II.4, late 2012.* Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin, 2013. **18**(1): p. 8-9.
87. Lopman, B.A., et al., *Clinical manifestation of norovirus gastroenteritis in health care settings.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2004. **39**(3): p. 318-324.
88. Kroneman, A., et al., *Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the Foodborne Viruses in Europe network from 1 July 2001 to 30 June 2006.* Journal of clinical microbiology, 2008. **46**(9): p. 2959-2965.
89. Vega, E., et al., *Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013.* J Clin Microbiol, 2014. **52**(1): p. 147-55.
90. Lau, C.S., et al., *High rate and changing molecular epidemiology pattern of norovirus infections in sporadic cases and outbreaks of gastroenteritis in Hong Kong.* Journal of medical virology, 2004. **73**(1): p. 113-117.
91. Kim, H.S., et al., *Emergence of GII.4 Sydney norovirus in South Korea during the winter of 2012-2013.* J Microbiol Biotechnol, 2013. **23**(11): p. 1641-3.
92. Hall, A.J., et al., *Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States.* Emerg Infect Dis, 2013. **19**(8): p. 1305-9.

93. Fankhauser, R.L., et al., *Epidemiologic and molecular trends of "Norwalk-like viruses" associated with outbreaks of gastroenteritis in the United States.* Journal of Infectious Diseases, 2002. **186**(1): p. 1-7.
94. Pang, X.L., et al., *Influence of novel norovirus GII.4 variants on gastroenteritis outbreak dynamics in Alberta and the Northern Territories, Canada between 2000 and 2008.* PloS one, 2010. **5**(7): p. e11599.
95. Delacour, H., P. Dubrous, and J.L. Koeck, *Noroviruses: a challenge for military forces.* J R Army Med Corps, 2010. **156**(4): p. 251-4.
96. Siebenga, J.J., et al., *Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007.* Journal of Infectious Diseases, 2009. **200**(5): p. 802-812.
97. Vega, E., et al., *Novel surveillance network for norovirus gastroenteritis outbreaks, United States.* Emerging infectious diseases, 2011. **17**(8): p. 1389-1395.
98. Lu, J., et al., *Gastroenteritis Outbreaks Caused by Norovirus GII.17, Guangdong Province, China, 2014-2015.* Emerg Infect Dis, 2015. **21**(7): p. 1240-2.
99. Matsushima, Y., et al., *Genetic analyses of GII.17 norovirus strains in diarrheal disease outbreaks from December 2014 to March 2015 in Japan reveal a novel polymerase sequence and amino acid substitutions in the capsid region.* Euro Surveill, 2015. **20**(26).
100. de Graaf, M., et al., *Emergence of a novel GII.17 norovirus - End of the GII.4 era?* Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 2015. **20**(26): p. 21178.
101. Medici, M.C., et al., *Identification of the novel Kawasaki 2014 GII.17 human norovirus strain in Italy, 2015.* Euro Surveill, 2015. **20**(35): p. 30010.
102. Dinu, S., et al., *Molecular identification of emergent GII.P17-GII.17 norovirus genotype, Romania, 2015.* Euro Surveill, 2016. **21**(7).
103. Parra, G.I. and K.Y. Green, *Genome of Emerging Norovirus GII.17, United States, 2014.* Emerging infectious diseases, 2015. **21**(8): p. 1477-1479.
104. LeBlanc, J.J., et al., *Outbreak of Norovirus GII.P17-GII.17 in the Canadian Province of Nova Scotia.* Canadian Journal of Infectious Diseases and Medical Microbiology, 2016. **2016**: p. 6.
105. Division of Viral Diseases, N.C.f.I., C.f.D.C. Respiratory Diseases, and Prevention, *Updated norovirus outbreak management and disease prevention guidelines.* MMWR.Recommendations and reports : Morbidity and mortality weekly report.Recommendations and reports / Centers for Disease Control, 2011. **60**(RR-3): p. 1-18.
106. Bitler, E.J., et al., *Norovirus outbreaks: a systematic review of commonly implicated transmission routes and vehicles.* Epidemiol Infect, 2013. **141**(8): p. 1563-71.
107. Verhoef, L., et al., Emerging Infectious Diseases, 2010. **16**(4): p. 617-624.
108. de Wit, M.A., et al., *Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology.* Am J Epidemiol, 2001. **154**(7): p. 666-74.
109. Pang, X.L., J.K. Preiksaitis, and B.E. Lee, *Enhanced enteric virus detection in sporadic gastroenteritis using a multi-target real-time PCR panel: A one-year study.* Journal of medical virology, 2014. **86**(9): p. 1594-1601.

110. Rockx, B., et al., *Natural history of human calicivirus infection: a prospective cohort study.* Clin Infect Dis, 2002. **35**(3): p. 246-53.
111. Koo, H.L., et al., *Noroviruses: The Most Common Pediatric Viral Enteric Pathogen at a Large University Hospital After Introduction of Rotavirus Vaccination.* J Pediatric Infect Dis Soc, 2013. **2**(1): p. 57-60.
112. McAtee, C.L., et al., *Burden of Norovirus and Rotavirus in Children After Rotavirus Vaccine Introduction, Cochabamba, Bolivia.* Am J Trop Med Hyg, 2016. **94**(1): p. 212-7.
113. Platts-Mills, J.A., et al., *Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED).* Lancet Glob Health, 2015. **3**(9): p. e564-75.
114. Hemming, M., et al., *Major reduction of rotavirus, but not norovirus, gastroenteritis in children seen in hospital after the introduction of RotaTeq vaccine into the National Immunization Programme in Finland.* Eur J Pediatr, 2013. **172**(6): p. 739-46.
115. Majowicz, S.E., et al., *Estimating the under-reporting rate for infectious gastrointestinal illness in Ontario.* Can J Public Health, 2005. **96**(3): p. 178-81.
116. Majowicz, S.E., et al., *Magnitude and distribution of acute, self-reported gastrointestinal illness in a Canadian community.* Epidemiol Infect, 2004. **132**(4): p. 607-17.
117. Thomas, M.K., et al., *Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006.* Foodborne pathogens and disease, 2013. **10**(7): p. 639-648.
118. Grytdal, S.P., et al., *Incidence of Norovirus and Other Viral Pathogens That Cause Acute Gastroenteritis (AGE) among Kaiser Permanente Member Populations in the United States, 2012-2013.* PLoS One, 2016. **11**(4): p. e0148395.
119. Hall, A.J., et al., *Incidence of acute gastroenteritis and role of norovirus, Georgia, USA, 2004-2005.* Emerg Infect Dis, 2011. **17**(8): p. 1381-8.
120. Tam, C.C., et al., *Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice.* Gut, 2012. **61**(1): p. 69-77.
121. Hasing, M.E., et al., *Emergence of a new norovirus GII.4 variant and changes in the historical biennial pattern of norovirus outbreak activity in Alberta, Canada, from 2008 to 2013.* Journal of clinical microbiology, 2013. **51**(7): p. 2204-2211.
122. Wheeler, J.G., et al., *Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive.* BMJ, 1999. **318**(7190): p. 1046-50.
123. Phan, T.G., et al., *Genetic heterogeneity, evolution, and recombination in noroviruses.* Journal of medical virology, 2007. **79**(9): p. 1388-1400.
124. Fu, J., et al., *Emergence of a new GII.17 norovirus variant in patients with acute gastroenteritis in Jiangsu, China, September 2014 to March 2015.* Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 2015. **20**(24): p. 21157.

125. Blazevic, V., et al., *Multiple consecutive norovirus infections in the first 2 years of life.* Eur J Pediatr, 2015. **174**(12): p. 1679-83.
126. Yi, J., et al., *Molecular epidemiology of norovirus in children and the elderly in Atlanta, Georgia, United States.* J Med Virol, 2016. **88**(6): p. 961-70.
127. Lee, B.E., et al., *Genetic relatedness of noroviruses identified in sporadic gastroenteritis in children and gastroenteritis outbreaks in northern Alberta.* Journal of medical virology, 2008. **80**(2): p. 330-337.
128. Adler, J.L. and R. Zickl, *Winter vomiting disease.* J Infect Dis, 1969. **119**(6): p. 668-73.
129. Ahmed, S.M., B.A. Lopman, and K. Levy, *A systematic review and meta-analysis of the global seasonality of norovirus.* PLoS One, 2013. **8**(10): p. e75922.
130. Greer, A.L., S.J. Drews, and D.N. Fisman, *Why "winter" vomiting disease? Seasonality, hydrology, and Norovirus epidemiology in Toronto, Canada.* Ecohealth, 2009. **6**(2): p. 192-9.
131. Cannon, J.L., et al., *Surrogates for the study of norovirus stability and inactivation in the environment: aA comparison of murine norovirus and feline calicivirus.* J Food Prot, 2006. **69**(11): p. 2761-5.
132. Colas de la Noue, A., et al., *Absolute Humidity Influences the Seasonal Persistence and Infectivity of Human Norovirus.* Appl Environ Microbiol, 2014. **80**(23): p. 7196-205.
133. Lofgren, E., et al., *Influenza seasonality: underlying causes and modeling theories.* J Virol, 2007. **81**(11): p. 5429-36.
134. Dowell, S.F., *Seasonal variation in host susceptibility and cycles of certain infectious diseases.* Emerg Infect Dis, 2001. **7**(3): p. 369-74.
135. Lopman, B., et al., *Host, weather and virological factors drive norovirus epidemiology: time-series analysis of laboratory surveillance data in England and Wales.* PLoS One, 2009. **4**(8): p. e6671.
136. Heijne, J.C., et al., *Enhanced hygiene measures and norovirus transmission during an outbreak.* Emerg Infect Dis, 2009. **15**(1): p. 24-30.
137. Atmar, R.L., et al., *Determination of the 50% human infectious dose for Norwalk virus.* J Infect Dis, 2014. **209**(7): p. 1016-22.
138. Teunis, P.F., et al., *Norwalk virus: how infectious is it?* Journal of medical virology, 2008. **80**(8): p. 1468-1476.
139. Kirby, A.E., P.F. Teunis, and C.L. Moe, *Two human challenge studies confirm high infectivity of Norwalk virus.* J Infect Dis, 2015. **211**(1): p. 166-7.
140. Verhoef, L., et al., *Norovirus genotype profiles associated with foodborne transmission, 1999-2012.* Emerg Infect Dis, 2015. **21**(4): p. 592-9.
141. Scallan, E., et al., *Foodborne illness acquired in the United States--major pathogens.* Emerging infectious diseases, 2011. **17**(1): p. 7-15.
142. Gould, L.H., et al., *Contributing factors in restaurant-associated foodborne disease outbreaks, FoodNet sites, 2006 and 2007.* J Food Prot, 2013. **76**(11): p. 1824-8.
143. Qin, M., et al., *A Waterborne Gastroenteritis Outbreak Caused by Norovirus GII.17 in a Hotel, Hebei, China, December 2014.* Food Environ Virol, 2016.
144. Zhou, N., et al., *A waterborne norovirus gastroenteritis outbreak in a school, eastern China.* Epidemiol Infect, 2016. **144**(6): p. 1212-9.

145. Mouly, D., et al., *Description of two waterborne disease outbreaks in France: a comparative study with data from cohort studies and from health administrative databases.* Epidemiol Infect, 2016. **144**(3): p. 591-601.
146. Di Bartolo, I., et al., *Waterborne norovirus outbreak during a summer excursion in Northern Italy.* New Microbiol, 2015. **38**(1): p. 109-12.
147. Giammanco, G.M., et al., *Investigation and control of a Norovirus outbreak of probable waterborne transmission through a municipal groundwater system.* J Water Health, 2014. **12**(3): p. 452-64.
148. Laine, J., et al., *The duration of gastrointestinal and joint symptoms after a large waterborne outbreak of gastroenteritis in Finland in 2007--a questionnaire-based 15-month follow-up study.* PLoS One, 2014. **9**(1): p. e85457.
149. Wyn-Jones, A.P., et al., *Surveillance of adenoviruses and noroviruses in European recreational waters.* Water Res, 2011. **45**(3): p. 1025-38.
150. Marks, P.J., et al., *A school outbreak of Norwalk-like virus: evidence for airborne transmission.* Epidemiol Infect, 2003. **131**(1): p. 727-36.
151. Marks, P.J., et al., *Evidence for airborne transmission of Norwalk-like virus (NLV) in a hotel restaurant.* Epidemiol Infect, 2000. **124**(3): p. 481-7.
152. Kirby, A.E., A. Streby, and C.L. Moe, *Vomiting as a Symptom and Transmission Risk in Norovirus Illness: Evidence from Human Challenge Studies.* PLoS One, 2016. **11**(4): p. e0143759.
153. Makison Booth, C., *Vomiting Larry: a simulated vomiting system for assessing environmental contamination from projectile vomiting related to norovirus infection.* J Infect Prev, 2014. **15**(5): p. 176-180.
154. Tung-Thompson, G., et al., *Persistence of human norovirus RT-qPCR signals in simulated gastric fluid.* Food Environ Virol, 2015. **7**(1): p. 32-40.
155. Cheesbrough, J.S., L. Barkess-Jones, and D.W. Brown, *Possible prolonged environmental survival of small round structured viruses.* J Hosp Infect, 1997. **35**(4): p. 325-6.
156. Kotwal, G. and J.L. Cannon, *Environmental persistence and transfer of enteric viruses.* Curr Opin Virol, 2014. **4**: p. 37-43.
157. Centers for Disease, C. and Prevention, *Norovirus outbreak in an elementary school--District of Columbia, February 2007.* MMWR Morb Mortal Wkly Rep, 2008. **56**(51-52): p. 1340-3.
158. Thornley, C.N., et al., *Recurring norovirus transmission on an airplane.* Clin Infect Dis, 2011. **53**(6): p. 515-20.
159. Evans, M.R., et al., *An outbreak of viral gastroenteritis following environmental contamination at a concert hall.* Epidemiol Infect, 2002. **129**(2): p. 355-60.
160. Gallimore, C.I., et al., *Contamination of the hospital environment with gastroenteric viruses: comparison of two pediatric wards over a winter season.* J Clin Microbiol, 2008. **46**(9): p. 3112-5.
161. Gallimore, C.I., et al., *Environmental monitoring for gastroenteric viruses in a pediatric primary immunodeficiency unit.* J Clin Microbiol, 2006. **44**(2): p. 395-9.
162. Kimura, H., et al., *A norovirus outbreak associated with environmental contamination at a hotel.* Epidemiol Infect, 2011. **139**(2): p. 317-25.

163. Lee, R.M., et al., *Incubation periods of viral gastroenteritis: a systematic review.* BMC Infect Dis, 2013. **13**: p. 446.
164. Division of Viral Diseases, N.C.f.I., C.f.D.C. Respiratory Diseases, and Prevention, *Updated norovirus outbreak management and disease prevention guidelines.* MMWR Recomm Rep, 2011. **60**(RR-3): p. 1-18.
165. Glass, R.I., U.D. Parashar, and M.K. Estes, *Norovirus gastroenteritis.* New England Journal of Medicine, 2009. **361**(18): p. 1776-1785.
166. Robilotti, E., S. Deresinski, and B.A. Pinsky, *Norovirus.* Clin Microbiol Rev, 2015. **28**(1): p. 134-64.
167. Wyatt, R.G., et al., *Comparison of three agents of acute infectious nonbacterial gastroenteritis by cross-challenge in volunteers.* J Infect Dis, 1974. **129**(6): p. 709-14.
168. Atmar, R.L., et al., *Norwalk virus shedding after experimental human infection.* Emerging infectious diseases, 2008. **14**(10): p. 1553-1557.
169. Turcios-Ruiz, R.M., et al., *Outbreak of necrotizing enterocolitis caused by norovirus in a neonatal intensive care unit.* J Pediatr, 2008. **153**(3): p. 339-44.
170. Stuart, R.L., et al., *An outbreak of necrotizing enterocolitis associated with norovirus genotype GII.3.* Pediatric Infectious Disease Journal, 2010. **29**(7): p. 644-647.
171. Pelizzo, G., et al., *Isolated colon ischemia with norovirus infection in preterm babies: a case series.* J Med Case Rep, 2013. **7**: p. 108.
172. Armbrust, S., et al., *Norovirus infections in preterm infants: wide variety of clinical courses.* BMC Res Notes, 2009. **2**: p. 96.
173. Marshall, J.K., et al., *Postinfectious irritable bowel syndrome after a food-borne outbreak of acute gastroenteritis attributed to a viral pathogen.* Clin Gastroenterol Hepatol, 2007. **5**(4): p. 457-60.
174. Khan, R.R., et al., *Gastrointestinal norovirus infection associated with exacerbation of inflammatory bowel disease.* J Pediatr Gastroenterol Nutr, 2009. **48**(3): p. 328-33.
175. Kawano, G., et al., *Benign infantile convulsions associated with mild gastroenteritis: a retrospective study of 39 cases including virological tests and efficacy of anticonvulsants.* Brain Dev, 2007. **29**(10): p. 617-22.
176. Chen, S.Y., et al., *Norovirus infection as a cause of diarrhea-associated benign infantile seizures.* Clinical Infectious Diseases, 2009. **48**(7): p. 849-855.
177. Ito, S., et al., *Norovirus-associated encephalopathy.* Pediatric Infectious Disease Journal, 2006. **25**(7): p. 651-652.
178. Obinata, K., et al., *Norovirus encephalopathy in a previously healthy child.* Pediatr Infect Dis J, 2010. **29**(11): p. 1057-9.
179. Karst, S.M., S. Zhu, and I.G. Goodfellow, *The molecular pathology of noroviruses.* J Pathol, 2015. **235**(2): p. 206-16.
180. Jones, M.K., et al., *Enteric bacteria promote human and mouse norovirus infection of B cells.* Science (New York, N.Y.), 2014. **346**(6210): p. 755-759.
181. Wobus, C.E., et al., *Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages.* PLoS Biol, 2004. **2**(12): p. e432.
182. Payne, D.C., et al., *Norovirus and medically attended gastroenteritis in U.S. children.* N Engl J Med, 2013. **368**(12): p. 1121-30.

183. Borrows, C.L. and P.C. Turner, *Seasonal screening for viral gastroenteritis in young children and elderly hospitalized patients: is it worthwhile?* J Hosp Infect, 2014. **87**(2): p. 98-102.
184. Rouhani, S., et al., *Norovirus Infection and Acquired Immunity in 8 Countries: Results From the MAL-ED Study.* Clin Infect Dis, 2016. **62**(10): p. 1210-7.
185. Lopman, B.A., et al., *Norovirus Infection and Disease in an Ecuadorian Birth Cohort: Association of Certain Norovirus Genotypes With Host FUT2 Secretor Status.* J Infect Dis, 2015. **211**(11): p. 1813-21.
186. Murata, T., et al., *Prolonged norovirus shedding in infants <or=6 months of age with gastroenteritis.* Pediatr Infect Dis J, 2007. **26**(1): p. 46-9.
187. Sakai, Y., et al., *Clinical severity of Norwalk virus and Sapporo virus gastroenteritis in children in Hokkaido, Japan.* Pediatr Infect Dis J, 2001. **20**(9): p. 849-53.
188. Chen, S.M., et al., *The significance of serum and fecal levels of interleukin-6 and interleukin-8 in hospitalized children with acute rotavirus and norovirus gastroenteritis.* Pediatr Neonatol, 2014. **55**(2): p. 120-6.
189. Weiskopf, D., B. Weinberger, and B. Grubeck-Loebenstein, *The aging of the immune system.* Transpl Int, 2009. **22**(11): p. 1041-50.
190. Castle, S.C., *Clinical relevance of age-related immune dysfunction.* Clin Infect Dis, 2000. **31**(2): p. 578-85.
191. Aw, D., A.B. Silva, and D.B. Palmer, *Immunosenescence: emerging challenges for an ageing population.* Immunology, 2007. **120**(4): p. 435-46.
192. Lindsay, L., et al., *A decade of norovirus disease risk among older adults in upper-middle and high income countries: a systematic review.* BMC Infect Dis, 2015. **15**: p. 425.
193. Goller, J.L., et al., *Long-term features of norovirus gastroenteritis in the elderly.* J Hosp Infect, 2004. **58**(4): p. 286-91.
194. Bok, K. and K.Y. Green, *Norovirus gastroenteritis in immunocompromised patients.* N Engl J Med, 2012. **367**(22): p. 2126-32.
195. Green, K.Y., *Norovirus infection in immunocompromised hosts.* Clin Microbiol Infect, 2014. **20**(8): p. 717-23.
196. Trivedi, T.K., et al., *Clinical characteristics of norovirus-associated deaths: a systematic literature review.* American Journal of Infection Control, 2013. **41**(7): p. 654-657.
197. Frange, P., et al., *Prevalence and clinical impact of norovirus fecal shedding in children with inherited immune deficiencies.* The Journal of infectious diseases, 2012. **206**(8): p. 1269-1274.
198. Robles, J.D., et al., *Norovirus infection in pediatric hematopoietic stem cell transplantation recipients: incidence, risk factors, and outcome.* Biol Blood Marrow Transplant, 2012. **18**(12): p. 1883-9.
199. Roddie, C., et al., *Allogeneic hematopoietic stem cell transplantation and norovirus gastroenteritis: a previously unrecognized cause of morbidity.* Clin Infect Dis, 2009. **49**(7): p. 1061-8.
200. Schorn, R., et al., *Chronic norovirus infection after kidney transplantation: molecular evidence for immune-driven viral evolution.* Clinical infectious diseases :

an official publication of the Infectious Diseases Society of America, 2010. **51**(3): p. 307-314.
201. Carlsson, B., et al., *Quasispecies dynamics and molecular evolution of human norovirus capsid P region during chronic infection.* The Journal of general virology, 2009. **90**(Pt 2): p. 432-441.
202. Siebenga, J.J., et al., *High prevalence of prolonged norovirus shedding and illness among hospitalized patients: a model for in vivo molecular evolution.* The Journal of infectious diseases, 2008. **198**(7): p. 994-1001.
203. Ludwig, A., et al., *Quantitative detection of norovirus excretion in pediatric patients with cancer and prolonged gastroenteritis and shedding of norovirus.* Journal of medical virology, 2008. **80**(8): p. 1461-1467.
204. Sukhrie, F.H., et al., *Chronic shedders as reservoir for nosocomial transmission of norovirus.* Journal of clinical microbiology, 2010. **48**(11): p. 4303-4305.
205. Karst, S.M. and R.S. Baric, *What Is the Reservoir of Emergent Human Norovirus Strains?* Journal of virology, 2015. **89**(11): p. 5756-5759.
206. Oriol, R., *Genetic control of the fucosylation of ABH precursor chains. Evidence for new epistatic interactions in different cells and tissues.* J Immunogenet, 1990. **17**(4-5): p. 235-45.
207. Marionneau, S., et al., *Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals.* Gastroenterology, 2002. **122**(7): p. 1967-77.
208. Tan, M. and X. Jiang, *Norovirus-host interaction: implications for disease control and prevention.* Expert Rev Mol Med, 2007. **9**(19): p. 1-22.
209. Lindesmith, L., et al., *Human susceptibility and resistance to Norwalk virus infection.* Nature medicine, 2003. **9**(5): p. 548-553.
210. Hutson, A.M., et al., *Norwalk virus infection associates with secretor status genotyped from sera.* Journal of medical virology, 2005. **77**(1): p. 116-120.
211. Hutson, A.M., et al., *Norwalk virus infection and disease is associated with ABO histo-blood group type.* The Journal of infectious diseases, 2002. **185**(9): p. 1335-1337.
212. Rockx, B.H., et al., *Association of histo-blood group antigens and susceptibility to norovirus infections.* Journal of Infectious Diseases, 2005. **191**(5): p. 749-754.
213. Frenck, R., et al., *Predicting susceptibility to norovirus GII.4 by use of a challenge model involving humans.* The Journal of infectious diseases, 2012. **206**(9): p. 1386-1393.
214. Van Trang, N., et al., *Association between norovirus and rotavirus infection and histo-blood group antigen types in Vietnamese children.* J Clin Microbiol, 2014. **52**(5): p. 1366-74.
215. Marionneau, S., et al., *ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world.* Biochimie, 2001. **83**(7): p. 565-573.
216. Bucardo, F., et al., *Genetic susceptibility to symptomatic norovirus infection in Nicaragua.* J Med Virol, 2009. **81**(4): p. 728-35.
217. Erney, R.M., et al., *Variability of human milk neutral oligosaccharides in a diverse population.* J Pediatr Gastroenterol Nutr, 2000. **30**(2): p. 181-92.

218. Le Pendu, J., et al., *Mendelian resistance to human norovirus infections.* Semin Immunol, 2006. **18**(6): p. 375-86.
219. Carmona-Vicente, N., et al., *Antibodies against Lewis antigens inhibit the binding of human norovirus GII.4 virus-like particles to saliva but not to intestinal Caco-2 cells.* Virol J, 2016. **13**(1): p. 82.
220. Tamura, M., et al., *Interaction of recombinant norwalk virus particles with the 105-kilodalton cellular binding protein, a candidate receptor molecule for virus attachment.* J Virol, 2000. **74**(24): p. 11589-97.
221. Tamura, M., et al., *Genogroup II noroviruses efficiently bind to heparan sulfate proteoglycan associated with the cellular membrane.* J Virol, 2004. **78**(8): p. 3817-26.
222. Siebenga, J.J., et al., *Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006.* Journal of virology, 2007. **81**(18): p. 9932-9941.
223. Allen, D.J., et al., *Analysis of amino acid variation in the P2 domain of the GII-4 norovirus VP1 protein reveals putative variant-specific epitopes.* PloS one, 2008. **3**(1): p. e1485.
224. Lindesmith, L.C., et al., *Monoclonal antibody-based antigenic mapping of norovirus GII.4-2002.* J Virol, 2012. **86**(2): p. 873-83.
225. Debbink, K., et al., *Emergence of new pandemic GII.4 Sydney norovirus strain correlates with escape from herd immunity.* The Journal of infectious diseases, 2013. **208**(11): p. 1877-1887.
226. Lindesmith, L.C., et al., *Emergence of a norovirus GII.4 strain correlates with changes in evolving blockade epitopes.* Journal of virology, 2013. **87**(5): p. 2803-2813.
227. Debbink, K., et al., *Genetic mapping of a highly variable norovirus GII.4 blockade epitope: potential role in escape from human herd immunity.* Journal of virology, 2012. **86**(2): p. 1214-1226.
228. Lindesmith, L.C., et al., *Particle conformation regulates antibody access to a conserved GII.4 norovirus blockade epitope.* J Virol, 2014. **88**(16): p. 8826-42.
229. Donaldson, E.F., et al., *Norovirus pathogenesis: mechanisms of persistence and immune evasion in human populations.* Immunological reviews, 2008. **225**: p. 190-211.
230. Yang, Y., et al., *Genetic and phenotypic characterization of GII-4 noroviruses that circulated during 1987 to 2008.* J Virol, 2010. **84**(18): p. 9595-607.
231. Dai, Y.C., et al., *Antigenic Relatedness of Norovirus GII.4 Variants Determined by Human Challenge Sera.* PloS one, 2015. **10**(4): p. e0124945.
232. Tan, M. and X. Jiang, *Norovirus-host interaction: multi-selections by human histo-blood group antigens.* Trends Microbiol, 2011. **19**(8): p. 382-8.
233. Donaldson, E.F., et al., *Viral shape-shifting: norovirus evasion of the human immune system.* Nature reviews.Microbiology, 2010. **8**(3): p. 231-241.
234. Iritani, N., et al., *Genetic analysis of the capsid gene of genotype GII.2 noroviruses.* Journal of virology, 2008. **82**(15): p. 7336-7345.
235. Boon, D., et al., *Comparative evolution of GII.3 and GII.4 norovirus over a 31-year period.* Journal of virology, 2011. **85**(17): p. 8656-8666.

236. Lai, M.M., *RNA recombination in animal and plant viruses.* Microbiological reviews, 1992. **56**(1): p. 61-79.
237. Eden, J.S., et al., *Recombination within the pandemic norovirus GII.4 lineage.* Journal of virology, 2013. **87**(11): p. 6270-6282.
238. Bull, R.A., M.M. Tanaka, and P.A. White, *Norovirus recombination.* The Journal of general virology, 2007. **88**(Pt 12): p. 3347-3359.
239. Motomura, K., et al., *Divergent evolution of norovirus GII/4 by genome recombination from May 2006 to February 2009 in Japan.* Journal of virology, 2010. **84**(16): p. 8085-8097.
240. Motomura, K., et al., *Identification of monomorphic and divergent haplotypes in the 2006-2007 norovirus GII/4 epidemic population by genomewide tracing of evolutionary history.* Journal of virology, 2008. **82**(22): p. 11247-11262.
241. Bull, R.A., et al., *Norovirus recombination in ORF1/ORF2 overlap.* Emerging Infectious Diseases, 2005. **11**(7): p. 1079-1085.
242. Eden, J.S., et al., *The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012.* Virology, 2014. **450-451**: p. 106-113.
243. Lam, T.T., et al., *The recombinant origin of emerging human norovirus GII.4/2008: intra-genotypic exchange of the capsid P2 domain.* The Journal of general virology, 2012. **93**(Pt 4): p. 817-822.
244. Ruether, I.G., et al., *Characterization of novel intergenogroup and intergenotype recombinant noroviruses from central Greece.* Mol Cell Probes, 2014. **28**(4): p. 204-10.
245. Parrino, T.A., et al., *Clinical immunity in acute gastroenteritis caused by Norwalk agent.* The New England journal of medicine, 1977. **297**(2): p. 86-89.
246. Simmons, K., et al., *Duration of immunity to norovirus gastroenteritis.* Emerging infectious diseases, 2013. **19**(8): p. 1260-1267.
247. McCartney, S.A., et al., *MDA-5 recognition of a murine norovirus.* PLoS Pathog, 2008. **4**(7): p. e1000108.
248. Sarvestani, S.T., et al., *Norovirus Infection: Replication, Manipulation of Host, and Interaction with the Host Immune Response.* J Interferon Cytokine Res, 2016. **36**(4): p. 215-25.
249. Newman, K.L. and J.S. Leon, *Norovirus immunology: Of mice and mechanisms.* Eur J Immunol, 2015. **45**(10): p. 2742-57.
250. Karst, S.M., et al., *STAT1-dependent innate immunity to a Norwalk-like virus.* Science, 2003. **299**(5612): p. 1575-1578.
251. Greenberg, H.B., et al., *Prevalence of antibody to the Norwalk virus in various countries.* Infect Immun, 1979. **26**(1): p. 270-3.
252. Nurminen, K., et al., *Prevalence of norovirus GII-4 antibodies in Finnish children.* Journal of medical virology, 2011. **83**(3): p. 525-531.
253. Pelosi, E., et al., *The seroepidemiology of genogroup 1 and genogroup 2 Norwalk-like viruses in Italy.* Journal of medical virology, 1999. **58**(1): p. 93-99.
254. Nicollier-Jamot, B., et al., *Molecular cloning, expression, self-assembly, antigenicity, and seroepidemiology of a genogroup II norovirus isolated in France.* Journal of clinical microbiology, 2003. **41**(8): p. 3901-3904.

255. Dai, Y.C., et al., *Seroprevalence of antibodies against noroviruses among students in a Chinese military medical university.* Journal of clinical microbiology, 2004. **42**(10): p. 4615-4619.
256. Kobayashi, S., et al., *Seroepidemiological study of norovirus infection in Aichi Prefecture, Japan.* Microbiology and immunology, 2009. **53**(6): p. 356-359.
257. Ajami, N.J., et al., *Antibody responses to norovirus genogroup GI.1 and GII.4 proteases in volunteers administered Norwalk virus.* Clinical and vaccine immunology : CVI, 2012. **19**(12): p. 1980-1983.
258. Son, H., et al., *Seroepidemiology of predominant norovirus strains circulating in Korea by using recombinant virus-like particle antigens.* Foodborne pathogens and disease, 2013. **10**(5): p. 461-466.
259. Ryder, R.W., et al., *Evidence of immunity induced by naturally acquired rotavirus and Norwalk virus infection on two remote Panamanian islands.* The Journal of infectious diseases, 1985. **151**(1): p. 99-105.
260. Lew, J.F., et al., *Detection of Norwalk virus or Norwalk-like virus infections in Finnish infants and young children.* The Journal of infectious diseases, 1994. **169**(6): p. 1364-1367.
261. Johnson, P.C., et al., *Multiple-challenge study of host susceptibility to Norwalk gastroenteritis in US adults.* The Journal of infectious diseases, 1990. **161**(1): p. 18-21.
262. Gray, J.J., et al., *Detection of immunoglobulin M (IgM), IgA, and IgG Norwalk virus-specific antibodies by indirect enzyme-linked immunosorbent assay with baculovirus-expressed Norwalk virus capsid antigen in adult volunteers challenged with Norwalk virus.* J Clin Microbiol, 1994. **32**(12): p. 3059-63.
263. Okhuysen, P.C., et al., *Viral shedding and fecal IgA response after Norwalk virus infection.* The Journal of infectious diseases, 1995. **171**(3): p. 566-569.
264. Reeck, A., et al., *Serological correlate of protection against norovirus-induced gastroenteritis.* The Journal of infectious diseases, 2010. **202**(8): p. 1212-1218.
265. Atmar, R.L., et al., *Serological Correlates of Protection against a GII.4 Norovirus.* Clin Vaccine Immunol, 2015. **22**(8): p. 923-9.
266. Saito, M., et al., *Multiple norovirus infections in a birth cohort in a Peruvian Periurban community.* Clin Infect Dis, 2014. **58**(4): p. 483-91.
267. Atmar, R.L., et al., *Norovirus vaccine against experimental human Norwalk Virus illness.* The New England journal of medicine, 2011. **365**(23): p. 2178-2187.
268. Bok, K., et al., *Chimpanzees as an animal model for human norovirus infection and vaccine development.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(1): p. 325-330.
269. Lindesmith, L., et al., *Cellular and humoral immunity following Snow Mountain virus challenge.* Journal of virology, 2005. **79**(5): p. 2900-2909.
270. Rockx, B., et al., *Characterization of the homo- and heterotypic immune responses after natural norovirus infection.* Journal of medical virology, 2005. **77**(3): p. 439-446.
271. Treanor, J.J., et al., *Subclass-specific serum antibody responses to recombinant Norwalk virus capsid antigen (rNV) in adults infected with Norwalk, Snow Mountain, or Hawaii virus.* J Clin Microbiol, 1993. **31**(6): p. 1630-4.

272. Hale, A.D., et al., *Homotypic and heterotypic IgG and IgM antibody responses in adults infected with small round structured viruses.* J Med Virol, 1998. **54**(4): p. 305-12.
273. Blazevic, V., M. Malm, and T. Vesikari, *Induction of homologous and cross-reactive GII.4-specific blocking antibodies in children after GII.4 New Orleans norovirus infection.* J Med Virol, 2015. **87**(10): p. 1656-61.
274. Lindesmith, L.C., et al., *Broad blockade antibody responses in human volunteers after immunization with a multivalent norovirus VLP candidate vaccine: immunological analyses from a phase I clinical trial.* PLoS Med, 2015. **12**(3): p. e1001807.
275. Abbas, A.K., K.M. Murphy, and A. Sher, *Functional diversity of helper T lymphocytes.* Nature, 1996. **383**(6603): p. 787-93.
276. Lindesmith, L.C., et al., *Heterotypic humoral and cellular immune responses following Norwalk virus infection.* J Virol, 2010. **84**(4): p. 1800-15.
277. Kaplanski, G., et al., *IL-6: a regulator of the transition from neutrophil to monocyte recruitment during inflammation.* Trends Immunol, 2003. **24**(1): p. 25-9.
278. Harada, A., et al., *Essential involvement of interleukin-8 (IL-8) in acute inflammation.* J Leukoc Biol, 1994. **56**(5): p. 559-64.
279. Alberta, G.o., *Public Health Notifiable Disease Management Guidelines*, A. Health, Editor. 2013: Alberta.
280. Arthur, S.E. and K.E. Gibson, *Environmental persistence of Tulane virus - a surrogate for human norovirus.* Can J Microbiol, 2016. **62**(5): p. 449-54.
281. Esseili, M.A., et al., *Feline Calicivirus, Murine Norovirus, Porcine Sapovirus, and Tulane Virus Survival on Postharvest Lettuce.* Appl Environ Microbiol, 2015. **81**(15): p. 5085-92.
282. Bae, J. and K.J. Schwab, *Evaluation of murine norovirus, feline calicivirus, poliovirus, and MS2 as surrogates for human norovirus in a model of viral persistence in surface water and groundwater.* Appl Environ Microbiol, 2008. **74**(2): p. 477-84.
283. D'Souza, D.H., et al., *Persistence of caliciviruses on environmental surfaces and their transfer to food.* Int J Food Microbiol, 2006. **108**(1): p. 84-91.
284. O'Donnell, T.B., et al., *Mouse Norovirus infection promotes autophagy induction to facilitate replication but prevents final autophagosome maturation.* Virology, 2016. **492**: p. 130-9.
285. Lin, X., et al., *Subgenomic promoter recognition by the norovirus RNA-dependent RNA polymerases.* Nucleic Acids Res, 2015. **43**(1): p. 446-60.
286. Waugh, E., et al., *Characterization of the chemokine response of RAW264.7 cells to infection by murine norovirus.* Virus Res, 2014. **181**: p. 27-34.
287. Herod, M.R., et al., *Expression of the murine norovirus (MNV) ORF1 polyprotein is sufficient to induce apoptosis in a virus-free cell model.* PLoS One, 2014. **9**(3): p. e90679.
288. Baldridge, M.T., et al., *Commensal microbes and interferon-lambda determine persistence of enteric murine norovirus infection.* Science, 2015. **347**(6219): p. 266-9.
289. Cheetham, S., et al., *Pathogenesis of a genogroup II human norovirus in gnotobiotic pigs.* Journal of virology, 2006. **80**(21): p. 10372-10381.

290. Bui, T., et al., *Median infectious dose of human norovirus GII.4 in gnotobiotic pigs is decreased by simvastatin treatment and increased by age.* J Gen Virol, 2013. **94**(Pt 9): p. 2005-16.
291. Souza, M., et al., *Pathogenesis and immune responses in gnotobiotic calves after infection with the genogroup II.4-HS66 strain of human norovirus.* J Virol, 2008. **82**(4): p. 1777-86.
292. Wyatt, R.G., et al., *Experimental infection of chimpanzees with the Norwalk agent of epidemic viral gastroenteritis.* J Med Virol, 1978. **2**(2): p. 89-96.
293. Taube, S., et al., *A mouse model for human norovirus.* MBio, 2013. **4**(4).
294. Xia, M., T. Farkas, and X. Jiang, *Norovirus capsid protein expressed in yeast forms virus-like particles and stimulates systemic and mucosal immunity in mice following an oral administration of raw yeast extracts.* J Med Virol, 2007. **79**(1): p. 74-83.
295. Fang, H., et al., *Norovirus P particle efficiently elicits innate, humoral and cellular immunity.* PLoS One, 2013. **8**(4): p. e63269.
296. Tan, M. and X. Jiang, *The p domain of norovirus capsid protein forms a subviral particle that binds to histo-blood group antigen receptors.* Journal of virology, 2005. **79**(22): p. 14017-14030.
297. Riddle, M.S. and R.I. Walker, *Status of vaccine research and development for norovirus.* Vaccine, 2016. **34**(26): p. 2895-9.
298. Wasserman, E. *Vaxart launches PhI clinical trial for norovirus tablet vaccine*. Vaccines August 28, 2016; Available from: http://www.fiercepharma.com/vaccines/vaxart-launches-phi-clinical-trial-for-norovirus-tablet-vaccine.
299. Treanor, J.J., et al., *A Novel Intramuscular Bivalent Norovirus Virus-Like Particle Vaccine Candidate-Reactogenicity, Safety, and Immunogenicity in a Phase 1 Trial in Healthy Adults.* The Journal of infectious diseases, 2014.

Chapter II: Molecular Epidemiology of Norovirus Outbreaks in Alberta*

* This chapter has been previously published as: M.E. Hasing, B.E. Lee, J.K. Preiksaitis, R. Tellier, L. Honish, A. Senthilselvan, X.L. Pang. "Emergence of a new norovirus GII.4 variant and changes in the historical biennial pattern of norovirus outbreak activity in Alberta, Canada, from 2008 to 2013". *J Clin Microbiol*. 2013 Jul;51(7):2204-11. Permission to reuse contents of the publication was granted by the publisher, the American Society for Microbiology. This chapter also includes data not previously published.

2.1. Introduction

Norovirus (NoV) is the leading cause of gastroenteritis outbreaks worldwide. Based on genetic variability, NoV strains are classified into genogroups, genotypes and genotype variants. The viral RNA genome has three open reading fames (ORFs) and most of the genetic variability resides in the capsid gene, which is encoded by ORF2 and ORF3. Specifically, ORF2 encodes the major capsid gene, VP1, which contains the hypervariable region and receptor-binding site [1].

Seven NoV genogroups (GI-GVII) have been identified, GI, GII and GIV infect humans with most infections caused by GI and GII [1-3]. Based on sequence variations in ORF2, 9 GI and at least 22 GII genotypes have been described using numeric designation for the various strains, and based on variations in ORF1, 14 GI and at least 29 GII genotypes have been described using numeric or alphabetic designations [3-5]. The majority of global gastroenteritis outbreaks are caused by GII.4 strains [6]. GII.4 has demonstrated a faster evolution than other strains [7] and new GII.4 clusters or variants emerge every 2 to 5 years [8].

Our understanding of NoV GI genetic evolution and outbreak activity are limited due to their overall low prevalence. GI strains appear to be more common in outbreak settings other than health-care institutions and different GI genotypes predominate over time [9-12].

A unique biennial pattern of NoV outbreak activity was observed in Alberta, Canada between July 2000 and June 2008 [13]. Increases in NoV outbreak activity to epidemic levels during this period were always associated with the emergence of new GII.4 variants every two years. In this study, we extend our analysis and describe the NoV outbreak activity and characterize the circulating NoV outbreak strains between July 2008 and June 2015. We

also perform an updated analysis of the periodic variations of NoV outbreak activity in the province of Alberta over a 13-year period, from July 2002 to June 2015.

2.2. Materials and methods

2.2.1. Outbreak investigations

In Alberta, outbreaks are reportable to the Medical Officer of Health, and investigated by public health officials (e.g. environmental health officers); ProvLab provides laboratory testing for all communicable disease outbreaks in the province. Laboratory testing of gastroenteritis outbreak specimens may include testing for enteric bacteria, enteric viruses and/or parasitic agents. Public health officials request the testing of stool specimens for NoV when outbreak illnesses are clinically consistent with this pathogen. Since February 2002, laboratory investigations of gastroenteritis outbreaks were standardized across the province and basic demographic data and results of laboratory investigations are retained in a centralized database within the Provincial Laboratory for Public Health (ProvLab) [13, 14]. The Health Research Ethics Board of the University of Alberta approved the use of samples for this study.

2.2.2. Testing and characterization of Norovirus in stool specimens

Stool specimens submitted to ProvLab for suspected NoV gastroenteritis outbreaks were tested for NoV genogroup I (GI) and II (GII) by multiplex real time reverse transcription (RT)-PCR [15]. For NoV strain characterization cDNA of one positive stool specimen from each NoV outbreak was amplified using primers Mon381 and Mon383 (region E) [16] or G2SKF and G2SKR (region C) [17] for NoV GII, and using primers G1SKF and G1SKR (region C) [17] or CapA, CapB1 and CapB2 (region D) for NoV GI [18]. The thermocycler conditions used were as follows: 95°C for 3 min for the initial denaturation step followed by 40 cycles of

95°C for 30s, 53°C for 30s and 72°C for 40s, and a final extension step of 7 min at 72°C. PCR products were sequenced using Sanger sequencing. Genotyping was performed using the norovirus genotyping tool [4] (available at http://www.rivm.nl/mpf/norovirus/typingtool) and in-house phylogenetic analysis. Specimen and reference sequences (listed on Table 2.1) were aligned with MAFFT 6 and Neighbor-joining trees were constructed in Mega 5.05 [19] using the Kimura 2-parameter model with a branch support of 1000 bootstrap replicates.

Table 2.1. Reference sequences used for NoV genotyping.

ACCESSION NUMBER	NAME	ORF2 GENOTYPE
L23828.1	SRSV-KY-89/89/JP	GI.1
M87661.2	CVXRNA	GI.1
AJ277610.1	Whiterose/96/UK	GI.2
L07418.1	Southampton/1991/UK	GI.2
AB187514.1	Otofuke/1979/JP	GI.3
AJ277612.1	Birmingham/1993/UK	GI.3
AB042808.1	Chiba 407/1987/JP	GI.4
AJ313030.1	Queen's Arms/Leeds/1992/UK	GI.4
AJ277614.1	Musgrove/1989/UK	GI.5
AB081723.1	WUG1	GI.6
AJ277615.1	Sindlesham/1995/UK	GI.6
AJ277609.1	Winchester/1994/UK	GI.7
AF538679.1	Boxer/2001/US	GI.8
AJ277606.1	Girlington/1993/UK	GII.1
AY919139.1	Picton/2003/AU	GII.1
U07611.2	Hawaii/1971/US	GII.1
DQ366347.1	OsakaNI/2004/JP	GII.2
U70059.1	Snow_Mountain	GII.2
X81879.1	Melksham	GII.2
AF190817.1	Arg320	GII.3
EU187437.1	5017.34/2003/JP	GII.3
U02030.1	Toronto/1991/CA	GII.3
X76716.1	Bristol/1993/UK	GII.4 Bristol 1990
X86557.1	Lordsdale	GII.4 Bristol 1990
AY032605.1	MD145-12/1987/US	GII.4 Camberwell 1994
AY030098.1	MD134-7/1987/US	GII.4 Camberwell 1994
AB303923.1	DenHaag015/2000/NL	GII.4 Grimsby 1996
AJ004864.1	Grimsby	GII.4 Grimsby 1996
DQ078829.2	Sydney348/97O/AU	GII.4 Grimsby 1996

ACCESSION NUMBER	NAME	ORF2 GENOTYPE
AY502023.1	Farmington Hills/2002/US	GII.4 Farmington Hills 2002
AY587985.1	Oxford/B4S6/2002/UK	GII.4 Farmington Hills 2002
JX445152.1	AlbertaEI131/2004/CA	GII.4 Farmington Hills 2002
AB294779.1	Kaiso/030556/2003/JP	GII.4 Kaiso 2003
AB303929.1	EmmenE006/2002/NL	GII.4 Kaiso 2003
DQ078794.2	Hunter_284E/04O/AU	GII.4 Hunter 2004
EF126961.1	Dongen46/2006/NL	GII.4 Hunter 2004
JX445153.1	AlbertaEI142/2006/CA	GII.4 2004 Hunter 2004
AB220921.1	Chiba/04-1050/2005/JP	GII.4 Chiba 2005
AB447448.1	Sakai2/2006/JP	GII.4 Chiba 2005
DQ369797.1	Guangzhou/NVgz01/CN	GII.4 Chiba 2005
GU445325.2	New Orleans1805/2009/US	GII.4 New Orleans 2009
JN595867.1	New Orleans/2010/US	GII.4 New Orleans 2009
JX445168.1	AlbertaEI388/2011/CA	GII.4 New Orleans 2009
JX459907.1	Woonona/SWW3309/2012/AU	GII.4 Sydney 2012
JX459908.1	Sydney/NSW0514/2012/AU	GII.4 Sydney 2012
AB447432.1	Aomori1/2006/JP	GII.4 Yerseke 2006a
JX445157.1	AlbertaEI513/2006	GII.4Yerseke 2006a
EF126966.1	Nijmegen115/2006/NL	GII.4 Den Haag 2006b
EU078417.1	OSD-CS/2006/US	GII.4 Den Haag 2006b
JX445159.1	AlbertaEI009/2008/CA	GII.4 Den Haag 2006b
AB434770.1	OC07138/07/JP	GII.4 Osaka 2007
AB541319.1	Osaka1/2007/JP	GII.4 Osaka 2007
GQ845369.3	Armidale/NSW390I/2008/AU	GII.4 Osaka 2007
AB445395.1	Apeldoorn317/2007/NL	GII.4 Apeldoorn 2007
HQ009513.1	JB15/KOR/2008	GII.4 Apeldoorn 2007
AB212306.1	Hokkaido/133/2003/JP	GII.5
AF414423.1	White_River/290/1994/US	GII.5
AY682550.1	S63/1999/FR	GII.5
AB039778.1	Saitama_U16	GII.6
AF414410.1	Miami/292/1994/US	GII.6
AJ277620.1	Seacroft/90/UK	GII.6
AB258331.1	Osaka/F140/2006/JP	GII.7
DQ078846.2	Sydney4477/02S/AU	GII.7
GQ849129.1	NSW088L/2007/AU	GII.7
AB039780.1	Saitama_U25	GII.8
AB067543.1	U25GII	GII.8
AF195848.1	Amsterdam/98-18/1998/NL	GII.8
AY038599.2	VA97207/1997	GII.9

ACCESSION NUMBER	NAME	ORF2 GENOTYPE
DQ379715.1	Goulburn_Valley_G5175_C/1983/AU	GII.9
AF427118.1	Erfurt/546/00/DE	GII.10
AY237415.2	Mc37	GII.10
AB074893.1	Sw918/1997/JP	GII.11
AB126320.1	swine43/JP	GII.11
AB039775.1	Saitama_U1	GII.12
AF414420.1	Honolulu/314/1994/US	GII.12
AJ277618.1	Wortley/1990/UK	GII.12
AB078334.1	Kashiwa47	GII.13
AY113106.1	Fayetteville/1998/US	GII.13
DQ379714.1	Goulburn_Valley_G5175_B/1983/AU	GII.13
AY130761.1	M7/1999/US	GII.14
EF547404.1	Maizuru/010426/2001/JP	GII.14
AF542090.1	Mex7076/1999	GII.15
AY130762.1	J23/1999/US	GII.15
AY502010.1	Tiffin/1999/US	GII.16
AY682551.1	VannesL23/1999/FR	GII.16
AY772730.1	Neustrelitz260/2000/DE	GII.16
AY502009.1	CS-E1/2002/US	GII.17
DQ438972.1	Katrina-17/2005/US	GII.17
AY823304.1	OH-QW101/2003/US	GII.18
AY823305.1	OH-QW125/2003/US	GII.18
AY823306.1	OH-QW170/2003/US	GII.19
AY823307.1	OH-QW218/2003/US	GII.19
AB083780.1	YURI	GII.NA2
AY675554.1	IF1998/2003/IQ	GII.NA3
AJ844470.1	Chiba/04-0502/2004/JP	GII.NA4
EU373815.1	Luckenwalde591/2002/DE	GIINA1
EU424333.1	Leverkusen267/2005/DE	GIINA1

2.2.3. Amplification and phylogenetic analysis of GII.4-2012 capsid sequences

The two earliest NoV GII.4-2012 strains, AlbertaEI337 and AlbertaEI63, were further characterized by sequencing the complete ORF2. Nucleic acid extracts were amplified with Taq DNA Polymerase (Invitrogen, CA, USA) according to the manufacturer's instructions, using three overlapping sets of primers to span the entire major capsid gene (VP1): 1) NV4611

and Mon381, 2) Mon383 and NVR7 and 3) NVF8 and NVR8 [16, 20, 21]. PCR products were sequenced using Sanger sequencing.

The phylogenetic analysis was conducted in Mega 5.05. Specimen and reference sequences were aligned using multiple sequence comparison by log-expectation (MUSCLE). Pairwise distances and maximum likelihood trees were calculated using the best nucleotide and amino acid substitution models based on Bayesian information criterion scores, the Jones-Taylor-Thornton model, and Kimura 2-parameter model, respectively, with gamma-distributed rate variation across sites and 1000 bootstrap replicates.

2.2.4. Data inclusion and statistical analysis

All NoV-positive outbreaks, defined by the detection of NoV in one or more of the outbreak stool specimens submitted for testing, and NoV negative outbreaks, defined by all outbreak stool specimens testing negative for NoV, were included in this study. Outbreaks in community-based and hospital-based long-term care facilities, senior lodges and group homes were grouped together as LTC/SL/GH for the analysis of outbreak settings. Outbreaks in LTC/SL/GH and acute care hospitals were further combined as healthcare-related institutional outbreaks vs. non-healthcare-related outbreaks which included outbreaks in daycares, school, camps, dormitories, community functions, conference and other settings for the comparison of outbreak settings between GI and GII using Chi square test. The distribution of gender per setting was analyzed using Chi square test or Fisher's exact test, depending on the sample size. NoV strains causing outbreaks between July 2008 and June 2015 in Alberta were sequenced and characterized as described. NoV outbreak activity was analyzed using annual observational periods from July 1 to June 30 of the following year because of the winter seasonality of NoV gastroenteritis. Temporal variation in the number of NoV positive outbreaks in Alberta between July 2002 and June 2012 was analyzed using the

Edwards test [22], correlogram [23] and periodogram [24]. The Edwards test considers the null hypothesis of no seasonality versus the alternative that the time series is a sinusoidal curve with one peak and one trough. Autocorrelation functions are plotted against the time in the correlogram. The periodogram is a sample estimate of spectral density function used in the frequency-domain time series methods. The seasonality in the proportion of positive NoV was analyzed using the Walter & Elwood test, which allows for the adjustment of varied numbers of the total tests done for each year [25, 26].

2.3. Results

2.3.1. Norovirus outbreaks by GII and GI

Between July 1st, 2008 and June 30th, 2015, 2,236 gastroenteritis outbreak investigations were initiated at ProvLab by public health officials. NoV was listed as the suspect agent for 1,963 outbreaks and specimens were received for NoV testing for 1,312 (66.8%) of these outbreaks. NoV testing was also performed for 62 of 87 (71.3%) outbreaks where no suspect agent was indicated in the ProvLab database and 47 of 186 (25.3%%) outbreaks where NoV was not initially listed as a suspected agent. Of the 1,421 outbreaks with laboratory testing for NoV, 944 (66.4%) were NoV positive: 839 (88.9%) were GII, 94 (10%) were GI, and 11 (1.2%) were mixed GI and GII outbreaks (Table 2.2). Examination of the monthly NoV genogroup-specific outbreaks revealed fluctuations in the proportions of NoV GI outbreaks with the highest proportion of GI outbreaks during the period Jul 2012-Jun 2013 (Figure 2.1). GI caused 30.4% of all NoV outbreaks during Jul 2012-Jun 2013 compared to 7.6% for Jul 2014-Jun 2015, 5.9% for Jul 2013-Jun 2014, 4.3% for Jul 2011-Jun 2012, 2.5% for Jul 2010-Jun 2011, 15.5% for Jul 2009-Jun 2010, and 11.7% for Jul 2008-Jun 2009.

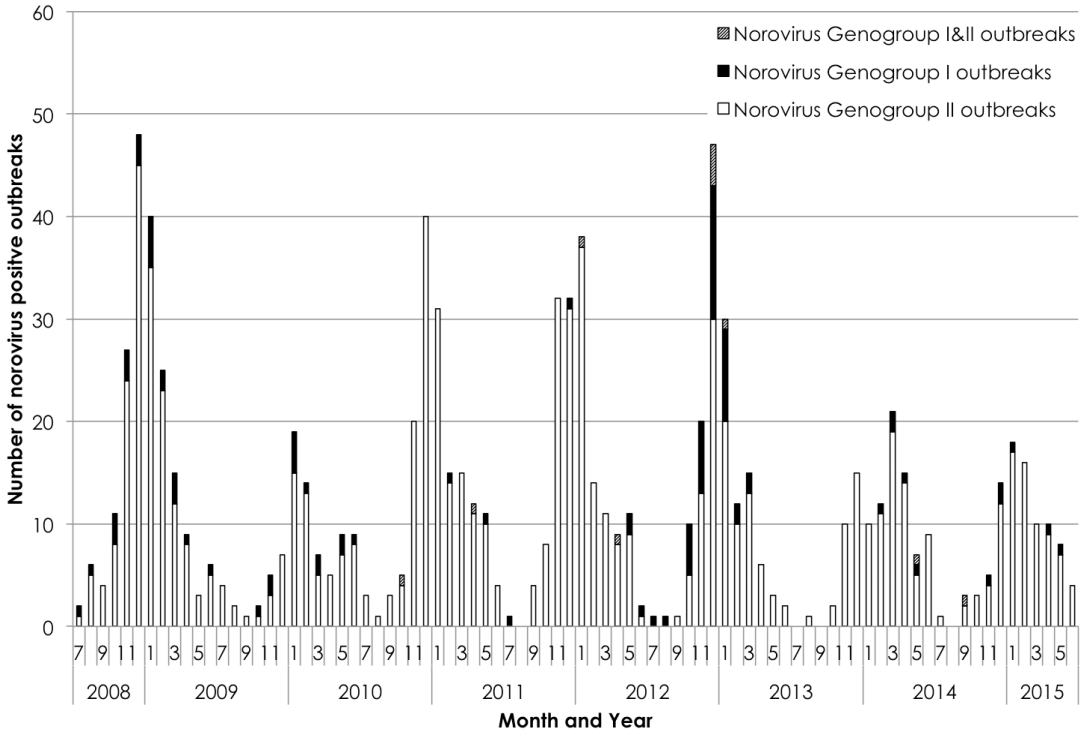
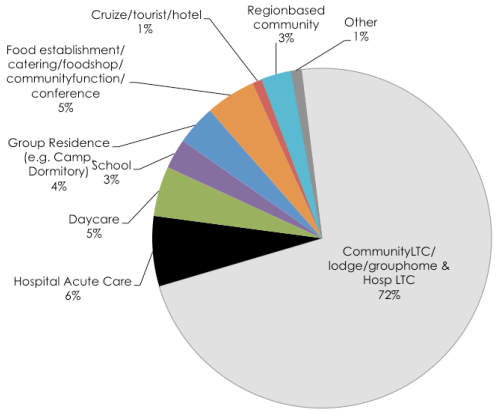


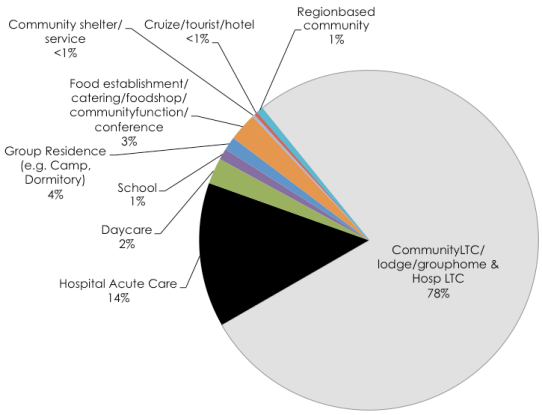
Figure 2.1. Monthly distribution of Norovirus outbreaks by genogroup in Alberta, Canada.

Data shown corresponds to the period July 1, 2008 to June 30, 2015.

The four most common settings for GI outbreaks were LTC/SL/GH (72.4% [76/105]), followed by acute care hospital (6.7% [7/105]), food establishment/catering events (4.8% [5/105]) and day care centers (4.8% [5/105]) (Figure 2.2). GII outbreaks had a similar pattern with most of the outbreaks in LTC/SL/GH (77.5% [659/850]), followed by acute care hospital (13.8% [117/850]), food establishment/catering events (2.7% [23/850]) and day care centers (2.5% [21/850]). A total of 21.0% (22/105) of GI outbreaks occurred in non-healthcare-related setting versus 8.7% (74/850) for GII, (p=0.000082, Chi square test).



NoV GI outbreaks



NoV GII outbreaks

2.3.2. Norovirus GII diversity

Sequences were obtained for 764 (89.9%) out of the 850 outbreaks with NoV GII. Thirteen different GII genotypes were identified and GII.4 (86.6%) was the most prevalent to be followed by GII.6 (2.4%), GII.12 (2.1%), GII.13 (2.0%), GII.1 (1.7%) and GII.3 (1.6%) (Table 2.2). The emergence and circulation of the various GII.4 variants since 2002 was shown in Figure 2.2. Two of the five GII.4 variants, GII.4 Osaka 2007 and GII.4 Apeldoorn 2007 (previously known as 2008a and 2008b, respectively [13]), caused only 11 outbreaks since their emergence in the winter of 2007. GII.4 Den Haag 2006b was the predominant strain in 2008/2009, causing 95.0% of all GII.4 outbreaks. GII.4 New Orleans 2009 first emerged in June 2009 and became predominant in 2009/2010 causing 67.5% of GII.4 outbreaks with GII.4 Den Haag 2006b at 32.5% during the same period. In 2010/2011, GII.4 New Orleans 2009 and GII.4 Den Haag 2006b variants were identified in 91.7% and 7.6% of GII.4 outbreaks, respectively. GII.4 New Orleans 2009 remained the most prevalent variant in 2011/2012 causing 97.7% of GII.4 outbreaks. A new variant, GII.4 Sydney 2012, first appeared in September 2011 and was identified in only 2 outbreaks in 2011/2012. However, GII.4 Sydney 2012 became the predominant variant in 2012/2013, causing 89.6% of GII.4

outbreaks and remained predominant during 2013/2014 and 2014/2015 representing 81.8% and 100% of all GII.4 outbreaks.

Table 2.2.Number of NoV outbreaks by genogroups and genotypes, from July 1 2008 to June 30 2015

Genotype or Genogroup	No. (%) outbreaks per observation period							Total
	July 2008 – June 2009	July 2009 - June 2010	July 2010 – June 2011	July 2011 - June 2012	July 2012 - June 2013	July 2013- June 2014	July 2014 - June 2015	
GI	23	13	2	5	40	5	6	94
Mixed GI and GII	0	0	2	2	5	1	1	11
GII	173	71	156	155	103	96	85	839
GII Sequenced	152(87.9)	62(87.3)	143(90.5)	150(95.5)	100(92.6)	85(87.6)	72(83.7)	764(89.9)
GII.1	0	0	0	13	0	0	0	13
GII.2	0	0	1	2	1	0	0	4
GII.3	0	10	1	1	0	0	0	12
GII.4	141	40	132	130	96	66	57	662
GII.4Den Haag 2006 (2006b)	134	13	10	1	4	1	0	163
GII.4 Osaka 2007 (2008a)	2	0	0	0	0	0	0	2
GII.4 Apeldoorn (2008b)	4	0	0	0	0	0	0	4
GII.4 New Orleans 2009	1	27	121	127	5	6	0	287
GII.4 Sydney 2012	0	0	0	2	86	54	57	199
GII.4 unidentified variant	0	0	1	0	1	5	0	7
GII.5	0	0	0	0	0	7	1	8
GII.6	0	0	4	0	2	7	5	18
GII.7	1	0	0	0	1	0	1	3
GII.8	0	0	0	0	0	1	1	2
GII.12	1	11	4	0	0	0	0	16
GII.13	5	0	1	1	0	4	4	15
GII.14	3	1	0	1	0	0	0	5
GII.15	1	0	0	2	0	0	0	3
GII.17	0	0	0	0	0	0	3	3
GI sequenced	21(91.3)	11(84.6)	3(75)	4(57.1)	34(75.6)	5(83.3)	4(57.1)	82(78.1)
GI.1	0	0	1	2	0	0	1	4
GI.2	0	0	0	0	0	2	0	2
GI.3	20	1	1	0	0	1	1	24
GI.4	1	2	1	1	0	1	0	6
GI.5	0	0	0	0	0	0	1	1
GI.6	0	7	0	1	20	0	1	29
GI.7	0	1	0	0	14	0	0	15
GI.9	0	0	0	0	0	1	0	1

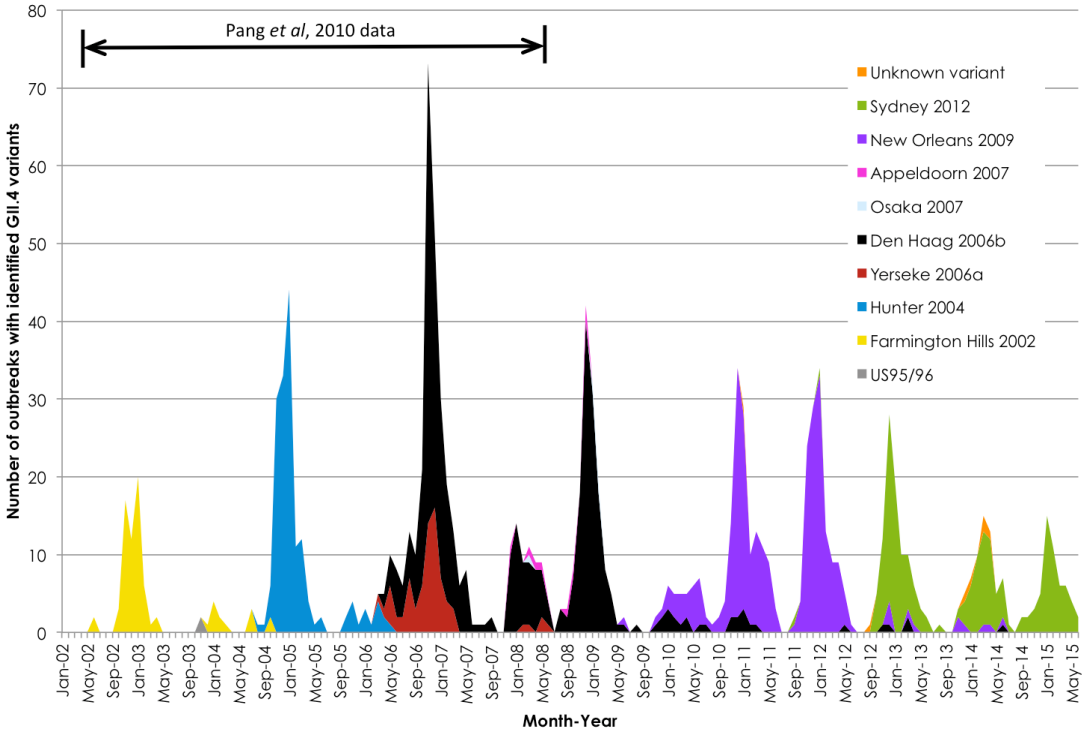


Figure 2.2. GII.4 variants identified in Norovirus outbreaks in Alberta, Canada, between July 1st, 2002 and June 30, 2015.

Phylogenetic analysis based on complete VP1 sequences of the first two GII.4 Sydney 2012 strains (Figure 2.3), showed that AlbertaEI337, circulated during September 2011, had a 2.9% nt and 2.0% aa difference when compared with Sydney 2012 (JX459908.1) whereas AlbertaEI63, circulated during January 2011, had a 0.5% nt and 0.4% aa difference.

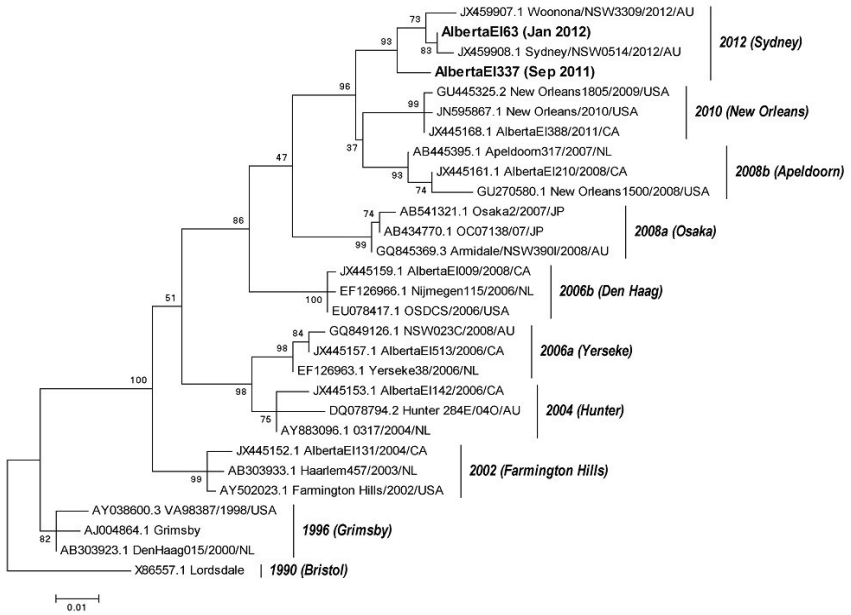


Figure 2.3. Phylogenetic analysis of complete VP1 amino acid sequences of GII.4 variants identified in Alberta

The names used by the Norovirus Genotyping Tool are shown in parentheses. The scale bars represent the number of substitutions per site.

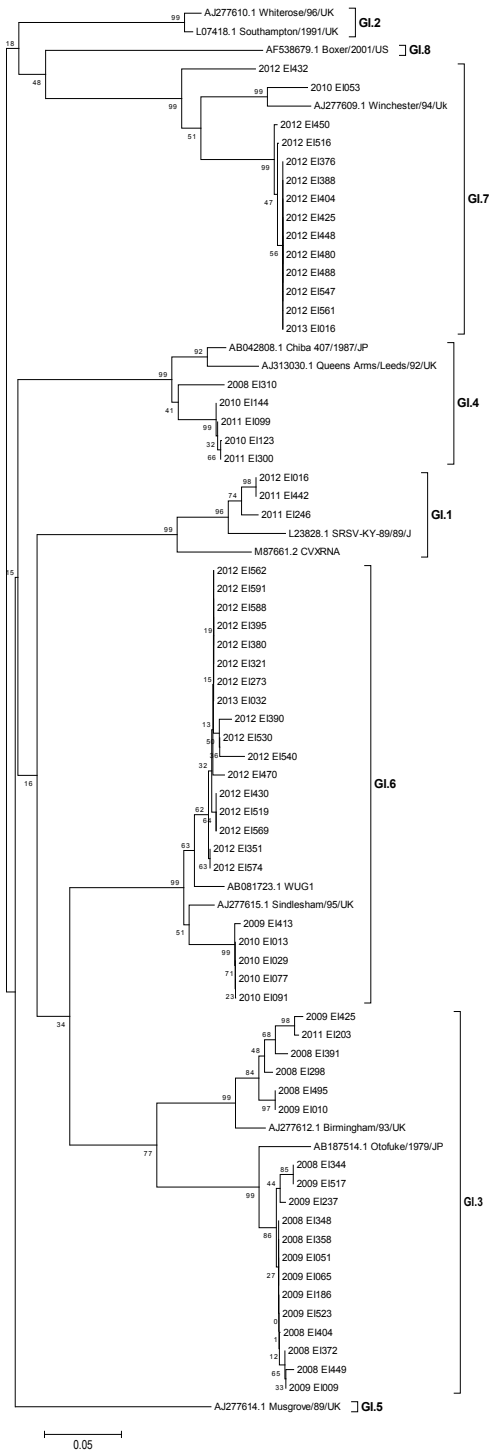


Figure 2.4. Neighbor-joining tree of NoV GI sequences (region C) from samples identified in Alberta

The phylogenetic analysis was performed using Kimura 2 parameter model with gamma-distribution of rate variation among sites and 1000 bootstrap replicates. The scale bar represents the number of substitutions per site.

2.3.3. Norovirus GI diversity

Eighty-two of the 105 (78.1%) outbreaks with NoV GI were successfully sequenced and five GI genotypes were identified (Table 2.1). The three most common genotypes were GI.6 causing 35.4% of sequenced GI outbreaks, followed by GI.3 causing 29.3%, and GI.7 18.3% of outbreaks. GI.3 was predominant in 2008/2009 causing 95.2% of GI outbreaks, whereas GI.6 was predominant in 2009/2010 and in 2012/2013 causing 63.6% and 58.8% of GI outbreaks, respectively. GI.7 also had increased activity during 2012/2013 and was responsible for 41.2% of GI outbreaks.

Phylogenetic analysis based on region C (320 nt) in the capsid gene for the GI genotypes (n=63) is shown in Figure 2.4. The latest GI.7 sequences (n=13), circulated in 2012 and 2013, clustered separately from the earlier GI.7 strain in 2010. A distinct group, formed by 12 of the 13 latest GI.7 sequences, showed a 12% nt difference with the older strain. Similarly, the most recent GI.6 strains, circulated in 2012 and 2013 (n=5), clustered separately from the earlier strains circulated in 2009 and 2010 (n=17), with a 5.5% nt difference between the two groups. Two clusters were also observed within GI.3 strains but there was no grouping by the time of circulation.

2.3.4 Demographic data of individuals affected in Norovirus outbreaks

The age and sex of the individuals affected in NoV outbreaks is shown on Table 2.3. The age of the subjects ranged between 1 and 90 years old, but NoV outbreaks affected mostly the elderly. A higher number of males was observed for GI outbreaks occurring at region based community setting compared to that observed in GII and NoV negative outbreaks (Fisher's exact test, P=0.0008).

Table 2.3. Age and sex of subjects involved in NoV outbreak investigations by outbreak setting

Setting	NoV Negative outbreaks			NoV GI outbreaks			NoV GII outbreaks		
	N	Age IQR (Q1-Q3)	Sex F:M ratio	N	Age IQR (Q1-Q3)	Sex F:M ratio	N	Age IQR (Q1-Q3)	Sex F:M ratio
Community long term care/lodge/group home	727	78.4-90.5	2.3	118	79.9-90.2	4.2	820	79.6-90.1	2.5
Hospital long term care	113	74.1-92.2	1.8	10	34-87.2	0.4	109	69.3-87.3	1.8
Hospital acute care	262	53.2-84.2	1.0	7	10.9-53.7	1.3	158	57.3-85.1	1.3
Daycare	64	1.5-6.3	1.0	6	1.4-30.0	4.0	17	1.5-30.3	1.9
School	1	5.1-5.1	-	4	10.6-23.0	1.0	7	14.9-20.8	0.9
Group residence (e.g. Camp, Dormitory)	13	17.0-81.8	0.6	5	13.5-40.6	0.3	2	20.2-43.5	0.1
Food establishment/ Catering/foodshop/community function/conference	27	27.5-49.0	1.2	10	35.4-57.7	0.7	25	30.5-56.3	1.5
Community shelter/service	6	2.4-39.0	0.2	-	-	-		9.9-38.7	0.0
Cruize/tourist/hotel	7	25.0-27.8	2.5	1	31.3-31.3	0.0	5	22.5-45.5	2.5
Region based community	32	3.1-35.1	3.3	13	1.4-13.5	0.2	7	4.4-37.6	1.4
Other	1	2.7-2.7	-	2	82.6-83.0	-		-	

N=number of individuals

IQR= interquartile range

F:M= female to male ratio

2.3.4. Temporal variations in norovirus outbreak levels

Significant seasonality (p = <0.01, Edwards test) with a temporal pattern consisting of a periodicity of 2 years and highest peaks at 2 years was observed when the number of NoV positive outbreaks from observation period 2002/2003 to 2011/2012 was analyzed using correlogram and periodogram analysis (Figure 2.5). A clear change from the biennial pattern was observed for 2010/2011 and 2011/2012 as these two intervals had a similar number of NoV positive outbreaks, 160 and 162, respectively . When the data was analyzed in terms of the proportion of NoV positive outbreaks for the first 10 observation periods, no significant seasonality (p=0.23, Walter & Elwood test) and no biennial pattern was detected. The

statistical analysis was not performed for the last three periods, 2012/2013, 2013/2014 and 2014/2015 which showed a clear decrease in NoV outbreak numbers. (Figure 2.5).

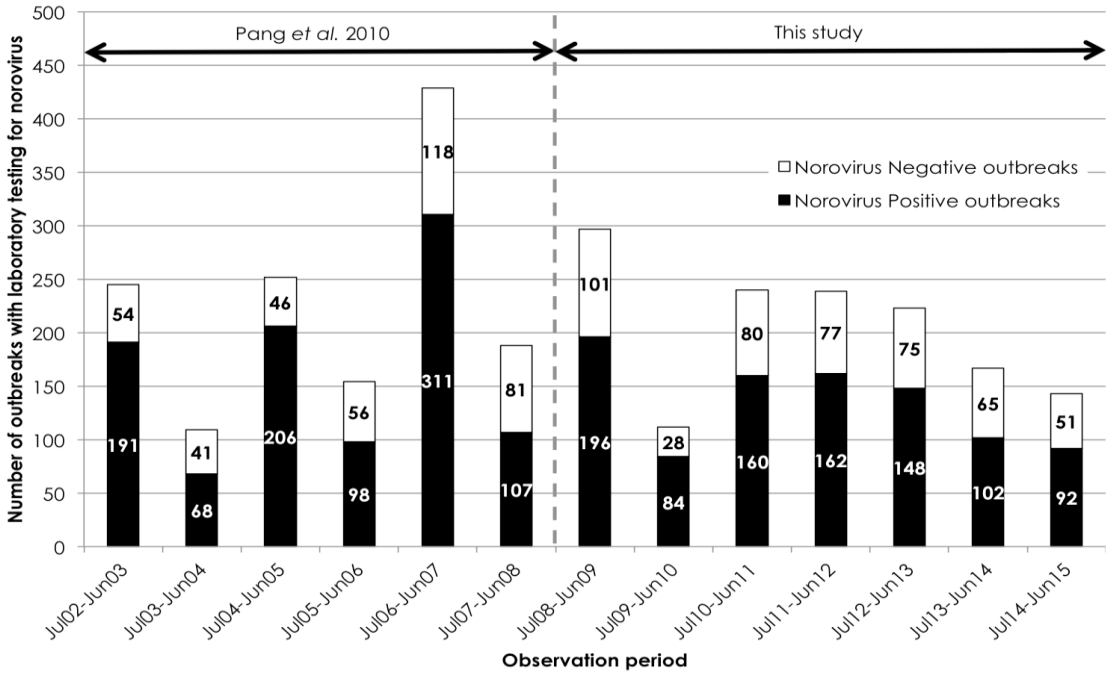


Figure 2.5. Numbers of Norovirus positive and negative outbreaks in Alberta, Canada by observation period.

2.4. Discussion

In our current study, we observed that the levels of NoV outbreak activity in 2010/2011 through 2014/2015 deviated from the historical biennial pattern of alternative years of quiescent and epidemic activity levels previously described [13]. Several novel observations were made when the data from July 2008 to June 2015 was added to the historical observations. NoV GII.4 Den Haag 2006b persisted and predominated over three consecutive observational periods between 2006/2007 and 2008/2009 while other new variants such as GII.4 Osaka 2007 and Apeldoorn 2007 never became predominant. The

persistence of a biennial pattern despite the predominance of a single variant, GII.4 Den Haag 2006, for three years demonstrates that changes in the level of NoV outbreak activity can occur in the absence of new variants. We postulate that the high levels of outbreak activity caused by GII.4 Den Haag 2006b in 2008/2009 was related to the waning of short-lived immunity from infection with the same variant in 2006/2007 interplaying with other factors such as virulence of GII.4 Den Haag 2006b, circulation of NoV in the environment and human reservoirs and, exposure of susceptible populations.

In contrast to previous observations, the emergence and predominance of GII.4 New Orleans 2009 was not associated with high levels of NoV outbreak activity in 2009/2010. Similar observations related to this variant were made in the United States during the 2009/2010 winter season [27]. The low outbreak activity related to GII.4 New Orleans 2009 may be due to multiple factors. Pandemic H1N1 influenza was rampant in Alberta between May 2009 and December 2009, limiting laboratory-based gastroenteritis outbreak investigations during the winter of 2009 because of resource constraints [28], possibly reducing the number of NoV outbreaks documented during this period. On the other hand, the number of suspected NoV outbreaks reported to the Medical Officers of Health in Alberta during this period was also lower than anticipated (unpublished data, Alberta Health Services). We postulate that the widespread and aggressive public hand hygiene and infection control educational campaigns targeting pandemic influenza might have decreased the transmission of GII.4 New Orleans 2009 within the population-at-risk for NoV outbreaks in 2009/2010.

Similar levels of NoV outbreak activity were observed for the second and third observation periods during which GII.4 New Orleans 2009 remained predominant (2010/2011 and 2011/2012). This was different from the biennial pattern observed for GII.4-2006 and other variants. Lower than normal levels of herd immunity resulting from less

transmission within the general population in 2009/2010, combined with short-lived or incomplete immunity in those exposed may have resulted in increased susceptibility to NoV infection in the Alberta population during both 2010/2011 and 2011/2012.

Differences in the virulence of GII.4 variants may also have contributed to the NoV outbreak burden. Genotyping was based on partial sequences of either the S domain (region E) or the N-terminal/shell (N/S) domain of VP1 (region C). Sequence alterations or antigenic variation occurring in other regions within the capsid, such as the P2 subdomain, were not analyzed. GII.4 Den Haag 2006b was unusual in its predominance over three consecutive observation periods and continued to represent ~30% of circulating strains after the emergence of GII.4 New Orleans 2009. Complete capsid sequence analysis of GII.4 Den Haag 2006b and GII.4 New Orleans 2009 strains from different observation periods would allow a more comprehensive analysis of genetic variability, drift and virulence of NoV GII.4 strains.

The newest variant, GII.4 Sydney 2012, was first reported in Australia in March 2012, thereafter in Europe and New Zealand [29], and has emerged as the predominant strain causing NoV outbreaks globally [30]. In Alberta, the first GII.4 Sydney 2012-like strain was identified in September 2011 and the second in January 2012. However, this variant did not become predominant until October 2012, almost one year after its first appearance. Although it has been suggested that GII.4 Sydney 2012 variants were the result of recombination of two earlier GII.4 variants [29], the amino acid distance between the first GII.4 Sydney 2012-like strain and later epidemic ones (Sydney 2012) suggests that the present GII.4 Sydney 2012 strains might also have evolved by antigenic drift. Analysis of region C sequences of GII.4 Sydney 2012 strains identified in Alberta since October 2012 (representing 44 outbreaks, see Figure 2.6), showed base changes in two nucleotide positions (5327 and 5339, relative to JX459908.1) compared to the initially identified GII.4 Sydney 2012 strain. The first GII.4 Sydney 2012-like strain in Alberta seemed to be a “pre-Sydney 2012 variant” with unique

intermediate sequence features. This implies that new emergent NoV GII.4 variants could be monitored and detected by phylogenic analysis as transitional strains before they evolve into new variants with higher virulence or which escape herd immunity.

Strains of genotype GII.17 have been reported increasingly around the world, but especially in Asia, where GII.17 strains have displaced GII.4 and have become the most predominant genotype. GII.17 emerged in China in October 2014 and, in Japan in December 2014 [31-34]. However, at least until June 2015, GII.17 did not become predominant in Alberta, causing only three GII.17 outbreaks in the province, with the first outbreak occurring in November 2014. Similar to Canada, Italy, Romania and United States also identified GII.17 during the same period, without observing major increases of GII.17 outbreak activity [35-37].

There are few published studies of GI genotypes in NoV outbreaks, however, the updated Alberta NoV outbreak data demonstrated GI prevalence similar to that of other countries. The overall percentage of NoV outbreaks over a decade of NoV investigation (from July 2008 to June 2015) caused by GI in Alberta was 11.1% (range: 2.5% to 30.4%) [13], comparable to that reported by Australia (4.3% in 2002-2010) [9] and New Zealand (8.7% in 2002/2009) [38]. However, we have observed episodic high levels of GI outbreak activity with peaks in 2003/2004 (29.4%) [13] and during the period 2012/2013 period (30.4%). Data from Noronet, an international NoV surveillance network that collects outbreak data from fourteen different countries around the globe, showed that GI.4, GI.6 and GI.3 were the ORF2-based GI genotypes most commonly reported between 2008 and 2014 [39, 40]. In Alberta, the three most predominant GI genotypes were GI.3, GI.6 and GI.7. GI.3 had high circulation in 2008/2009, while GI.6 and GI.7 were predominant in 2012/2013. In contrast with the trend reported by Noronet, GI.4 was seen only occasionally in Alberta throughout the entire 7-year study period. The reasons for these changes in GI genotype diversity and

prevalence are unknown; although differences in nucleotide sequences were observed in the late GI.6 and GI.7 strains, additional research is required to determine if antigenic drift, route of transmission or type of outbreak setting are factors associated with the increases of GI outbreak activity. Our observations related to GI outbreak activity will need to be interpreted in the context of ongoing global NoV surveillance.

In Alberta, GI and GII outbreaks occurred mostly in long-term care facilities, senior lodges and healthcare facilities as described in Europe [41] and affected mostly the elderly. The higher proportion of GI in non-healthcare related settings was also reported in Australia [9]. There were some variations in the predominant outbreak setting for GI over different observation periods but the number of outbreaks per category was too small for trending analysis.

In conclusion, our updated NoV outbreak activity data from the period July 2008 to June 2015 demonstrated a deviation from the historical biennial pattern of NoV outbreak activity previously reported in Alberta (July 2002- June 2008). Increased NoV outbreak activity was not always associated with new GII.4 variants. The recent divergence from the biennial pattern may be due to multiple factors, which need further study such as host susceptibility, levels of herd immunity, the virulence of specific NoV strains and public health interventions that might impact transmission.

Figure 2.6. Neighbor-joining tree of NoV GII.4 sequences (region C) including GII.4 Sydney 2012 strains identified in Alberta.

The phylogenetic analysis was performed using Kimura 2 parameter model with gamma-distribution of rate variation among sites and 1000 bootstrap replicates

2.5. Acknowledgements

This work was supported by the Provincial Laboratory for Public Health, Alberta, Canada; and Alberta Health, Canada [RES0015904]. We thank Michael Janke, Fatema Dhalla, Andy Wang and Jay Gamma for their assistance with norovirus sequencing and Dr. Bart Hazes for his expert opinion regarding sequence analysis. We also thank the Provincial Laboratory for Public Health (ProvLab) staff, especially Dr. Marie Louie and Dr. Greg Tyrrell, the collaborative teams participating in outbreak investigations in Alberta, Northwest Territories, Yukon and Nunavut, and the members of the Public Health (ProvLab) Outbreak Investigation Committee (OINC).

2.6. References (chapter 2)

1. Donaldson, E.F., et al., *Viral shape-shifting: norovirus evasion of the human immune system.* Nature reviews.Microbiology, 2010. **8**(3): p. 231-241.
2. Glass, R.I., U.D. Parashar, and M.K. Estes, *Norovirus gastroenteritis.* New England Journal of Medicine, 2009. **361**(18): p. 1776-1785.
3. Vinje, J., *Advances in laboratory methods for detection and typing of norovirus.* Journal of clinical microbiology, 2015. **53**(2): p. 373-381.
4. Kroneman, A., et al., *An automated genotyping tool for enteroviruses and noroviruses.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2011. **51**(2): p. 121-125.
5. Kroneman, A., et al., *Proposal for a unified norovirus nomenclature and genotyping.* Arch Virol, 2013. **158**(10): p. 2059-68.
6. Siebenga, J.J., et al., *Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007.* Journal of Infectious Diseases, 2009. **200**(5): p. 802-812.
7. Bull, R.A., et al., *Rapid evolution of pandemic noroviruses of the GII.4 lineage.* PLoS pathogens, 2010. **6**(3): p. e1000831.
8. Zheng, D.P., et al., *Molecular epidemiology of genogroup II-genotype 4 noroviruses in the United States between 1994 and 2006.* Journal of clinical microbiology, 2010. **48**(1): p. 168-177.
9. Bruggink, L.D., et al., *Molecular and epidemiological features of gastroenteritis outbreaks involving genogroup I norovirus in Victoria, Australia, 2002-2010.* Journal of medical virology, 2012. **84**(9): p. 1437-1448.
10. Division of Viral Diseases, N.C.f.I., C.f.D.C. Respiratory Diseases, and Prevention, *Updated norovirus outbreak management and disease prevention guidelines.*

MMWR.Recommendations and reports : Morbidity and mortality weekly report.Recommendations and reports / Centers for Disease Control, 2011. **60**(RR-3): p. 1-18.
11. Bitler, E.J., et al., *Norovirus outbreaks: a systematic review of commonly implicated transmission routes and vehicles.* Epidemiol Infect, 2013. **141**(8): p. 1563-71.
12. Kroneman, A., et al., *Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the Foodborne Viruses in Europe network from 1 July 2001 to 30 June 2006.* Journal of clinical microbiology, 2008. **46**(9): p. 2959-2965.
13. Pang, X.L., et al., *Influence of novel norovirus GII.4 variants on gastroenteritis outbreak dynamics in Alberta and the Northern Territories, Canada between 2000 and 2008.* PloS one, 2010. **5**(7): p. e11599.
14. Mukhi, S.N., et al., *DIAL: A Platform for real-time Laboratory Surveillance.* Online journal of public health informatics, 2010. **2**(3): p. 10.5210/ojphi.v2i3.3041. Epub 2010 Dec 23.
15. Pang, X.L., J.K. Preiksaitis, and B. Lee, *Multiplex real time RT-PCR for the detection and quantitation of norovirus genogroups I and II in patients with acute gastroenteritis.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2005. **33**(2): p. 168-171.
16. Noel, J.S., et al., *Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995.* Journal of medical virology, 1997. **53**(4): p. 372-383.
17. Kojima, S., et al., *Genogroup-specific PCR primers for detection of Norwalk-like viruses.* Journal of virological methods, 2002. **100**(1-2): p. 107-114.
18. Vinje, J., R.A. Hamidjaja, and M.D. Sobsey, *Development and application of a capsid VP1 (region D) based reverse transcription PCR assay for genotyping of genogroup I and II noroviruses.* Journal of virological methods, 2004. **116**(2): p. 109-117.
19. Tamura, K., et al., *MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.* Mol Biol Evol, 2011. **28**(10): p. 2731-9.
20. Yuen, L.K., et al., *Heminested multiplex reverse transcription-PCR for detection and differentiation of Norwalk-like virus genogroups 1 and 2 in fecal samples.* Journal of clinical microbiology, 2001. **39**(7): p. 2690-2694.
21. Chhabra, P., A.M. Walimbe, and S.D. Chitambar, *Complete genome characterization of Genogroup II norovirus strains from India: Evidence of recombination in ORF2/3 overlap.* Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2010. **10**(7): p. 1101-1109.
22. Fernandez-Duran, J.J. and M.M. Gregorio-Dominguez, *Testing for seasonality using circular distributions based on non-negative trigonometric sums as alternative hypotheses.* Statistical methods in medical research, 2011.
23. Cleophas, T.J. and A.H. Zwinderman, *Assessing seasonality in clinical research.* Clinical chemistry and laboratory medicine : CCLM / FESCC, 2012.
24. Zeger, S.L., R. Irizarry, and R.D. Peng, *On time series analysis of public health and biomedical data.* Annual Review of Public Health, 2006. **27**: p. 57-79.

25. Walter, S.D. and J.M. Elwood, *A test for seasonality of events with a variable population at risk.* British journal of preventive & social medicine, 1975. **29**(1): p. 18-21.
26. Siffel, C., C.J. Alverson, and A. Correa, *Analysis of seasonal variation of birth defects in Atlanta.* Birth defects research.Part A, Clinical and molecular teratology, 2005. **73**(10): p. 655-662.
27. Yen, C., et al., *Impact of an emergent norovirus variant in 2009 on norovirus outbreak activity in the United States.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2011. **53**(6): p. 568-571.
28. Lee, B.E., et al., *Determination of the relative economic impact of different molecular-based laboratory algorithms for respiratory viral pathogen detection, including Pandemic (H1N1), using a secure web based platform.* Virology journal, 2011. **8**: p. 277-422X-8-277.
29. van Beek, J., et al., *Indications for worldwide increased norovirus activity associated with emergence of a new variant of genotype II.4, late 2012.* Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin, 2013. **18**(1): p. 8-9.
30. Centers for Disease, C. and Prevention, *Notes from the Field: Emergence of New Norovirus Strain GII.4 Sydney - United States, 2012.* MMWR.Morbidity and mortality weekly report, 2013. **62**: p. 55.
31. Lu, J., et al., *Gastroenteritis Outbreaks Caused by Norovirus GII.17, Guangdong Province, China, 2014-2015.* Emerg Infect Dis, 2015. **21**(7): p. 1240-2.
32. Han, J., et al., *Emergence and predominance of norovirus GII.17 in Huzhou, China, 2014-2015.* Virol J, 2015. **12**: p. 139.
33. Fu, J., et al., *Emergence of a new GII.17 norovirus variant in patients with acute gastroenteritis in Jiangsu, China, September 2014 to March 2015.* Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 2015. **20**(24): p. 21157.
34. Matsushima, Y., et al., *Genetic analyses of GII.17 norovirus strains in diarrheal disease outbreaks from December 2014 to March 2015 in Japan reveal a novel polymerase sequence and amino acid substitutions in the capsid region.* Euro Surveill, 2015. **20**(26).
35. Parra, G.I. and K.Y. Green, *Genome of Emerging Norovirus GII.17, United States, 2014.* Emerging infectious diseases, 2015. **21**(8): p. 1477-1479.
36. Medici, M.C., et al., *Identification of the novel Kawasaki 2014 GII.17 human norovirus strain in Italy, 2015.* Euro Surveill, 2015. **20**(35): p. 30010.
37. Dinu, S., et al., *Molecular identification of emergent GII.P17-GII.17 norovirus genotype, Romania, 2015.* Euro Surveill, 2016. **21**(7).
38. Greening, G.E., et al., *Molecular epidemiology of norovirus gastroenteritis outbreaks in New Zealand from 2002-2009.* Journal of medical virology, 2012. **84**(9): p. 1449-1458.
39. van Beek, J.B., D.; Kroneman, A.; Vennema, H.; Koopmans, M., *Noronet report, April 2012*. 2012, National Institute for Public Health and the Environment: Bilthoven, The Netherlands.
40. van Beek, J.K., A.; Vennema, H.; Koopmans, M., *Noronet report, April 2014.* 2014.

41. Koopmans, M., *Noroviruses in healthcare settings: a challenging problem.* J Hosp Infect, 2009. **73**(4): p. 331-7.

Chapter III: Detection and analysis of recombination in GII.4 norovirus strains causing gastroenteritis outbreaks in Alberta[†]

[†] This chapter has been published as: ME Hasing, B Hazes, BE Lee, JK Preiksaitis, XL Pang. "Detection and analysis of recombination in GII.4 norovirus strains causing gastroenteritis outbreaks in Alberta". *Infect Genet Evol.* 2014 Oct;27:181-92. Permission to reuse this publication in my thesis was granted by the publisher, Elsevier.

3.1. Introduction

Norovirus (NoV) is responsible for more than half of all gastroenteritis outbreaks globally [1]. It has been estimated that every year NoV causes ~200,000 deaths in children <5 y of age in developing countries [2] and 21 million illnesses, 71,000 hospitalizations and 800 deaths in the United States [3]. About 90% of all NoV outbreaks in Alberta occur in health care facilities, which include acute care hospitals, group homes, long term care and supportive living facilities [chapter 2]. To date, there is no specific drug therapy or licensed vaccine available for the treatment or prevention of NoV infection. Progress in understanding the pathogenesis of human NoV has been hampered by the absence of an efficient cell culture system and an animal infection model. Our current knowledge regarding the transmissibility of NoV within populations has been inferred from epidemiological and clinical data, *in vitro* studies with heterologous proteins and culture surrogates, and a few challenge studies in human volunteers.

Norovirus has a single-stranded, positive sense, RNA genome of ~7.6 Kb that contains 3 open reading frames (ORFs). ORF1 encodes a ~1700 amino acid polyprotein that is cleaved into 6 non-structural proteins: p48, a N-terminal protein of unknown function; a 2C-like helicase; a 3A-like protein; VPg, a viral genome-linked protein; a 3C-like protease; and a RNA-dependent RNA polymerase (RdRp) [4]. ORF2 and ORF3 encode the VP1 and VP2 structural proteins, respectively. VP1 is the major capsid protein and is formed by two domains: the shell (S) domain and the protruding (P) domain. The P domain of VP1 is further divided into two subdomains: P1 and P2. P2 contains binding sites for histoblood group antigens (HBGAs), the putative receptor molecules for human NoV [5]. A total of 90 VP1 dimers form the final viral capsid with low copy numbers of VP2 [6, 7]. VP2 is not

required for virus-like particle (VLP) formation [8]; however, the conservation of ORF3 in the Caliciviridae family suggests that VP2 has an important role in the viral life cycle [7].

Noroviruses are classified into 7 genogroups (I -VII) and only genogroups I, II and IV have been found to infect humans [9, 10]. At least 41 ORF2-based genotypes have been described so far [9, 10]. Norovirus genogroup II genotype 4 (GII.4) is of high relevance in public health as ~60% of NoV outbreaks are caused by this genotype globally [11]. Norovirus GII.4 evolves rapidly resulting in new genetic clusters or variants every 2-5 years that quickly replace previous circulating GII.4 strains [11]. While some GII.4 variants such as Cairo 2007, Asia 2003, Japan 2008 only circulated in limited geographic regions [12-14], variants US95/96 1995, Farmington Hills 2002, Hunter 2004, Den Haag 2006b, New Orleans 2009 and Sydney 2012 spread globally causing pandemic NoV outbreaks [11, 15-19].

Norovirus evolves by mechanisms of antigenic drift [20-22] and recombination [23-25] leading to gradual and sudden antigenic changes, respectively. Recombination in the case of NoV requires the presence of co-infecting strains in the same host cell and generates chimeric genomes with stretches from each parental sequence [26, 27]. The most commonly identified hotspot for NoV recombination is located at the ORF1/2 junction (ORF1/2 intergenic recombinants), although ORF2/3 intergenic recombinants and ORF2-intragenic recombinants with a breakpoint ~150bp upstream the S/P domain junction have also been described [24, 25, 28]. Recombination in RNA viruses has been associated with expansions in host range, increased virulence, evasion of host immunity and resistance to antivirals [29]. A trait associated with increased fitness in NoV is the ability to generate genetic diversity relatively fast. Recombinant NoV might benefit by acquiring non-structural genes that promote increased genetic diversity, e.g. an RdRp with increased error rate [30]or a new capsid gene to increase immune evasion.

Several studies have identified multiple NoV GII.4 strains including CHDC 1970, Asia 2003, Cairo 2007, Den Haag 2006b, Japan 2008b, Apeldoorn 2007, Osaka 2007, New Orleans 2009 and Sydney 2012 as recombinants [23-25] and more recently, recombinants of New Orleans and Sydney 2012 have been reported in Europe [31, 32]. The NoV outbreak surveillance program established in Alberta, Canada, by the Provincial Laboratory of Public Health, currently collects NoV genotyping data based on partial ORF2 sequences that do not have sufficient length to detect recombinant NoV. In this study we analyzed near full-length virus sequences of selected isolates to detect and analyze recombination in NoV GII.4 strains causing gastroenteritis outbreaks in Alberta.

3.2. Materials and methods

3.2.1. Stool specimens

The Health Research Ethics Board of the University of Alberta granted ethics approval for the use of samples. The identification of GII.4 variants from NoV outbreaks occurring in Alberta and the Northern Territories between July 2002 and February 2013 was described previously [33][chapter 2]. Twenty outbreaks of NoV were selected to represent NoV GII.4 variants Farmington Hills 2002 (n=1), Hunter 2004 (n=2), Yerseke 2006a (n=2), Apeldoorn 2007 (n=1), Den Haag 2006b (n=6), New Orleans 2009 (n=6) and Sydney 2012 (n=2). Variants Den Haag 2006b and New Orleans 2009 were sampled each year of circulation, as these two variants were responsible for increased epidemic activity in more than one winter [chapter 2]. The Sydney 2012 outbreaks correspond to the first two outbreaks reported in Alberta for this variant. One stool sample from each outbreak was selected for whole genome sequencing.

3.2.2. Genomic amplification and DNA sequencing

RNA extraction was performed on 20% stool dilutions with the MagaZorb® Total RNA mini prep kit (Promega). Viral cDNA was obtained using Superscript™ II Reverse Transcriptase (Invitrogen) with random primers. PCR reactions were performed using *Taq* DNA polymerase (Invitrogen) and the primers described by [34-36] summarized in Table 3.1, to obtain overlapping fragments spanning the entire NoV genome. Thermocycler conditions were 94°C for 5 min for initial denaturation, followed by 40 cycles of 94°C for 1 min, 50°C for 1 min and 68°C for 3 min and final extension at 70°C for 7 min. PCR fragments were gel purified using the QIAquick gel extraction kit (Qiagen), and sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) in the ABI PRISM ® 3100-Avant Genetic Analyzer (Applied Biosystems).

Table 3.1. List of primers used for near full-length norovirus genome sequencing

Primer	Sequence (5'→3")	Position*	Orientation	Reference
NVF1	GTGAATGAAGATGGCGTCTA	1-20	+	[36]
NVR1	AAACTCCAAAGAGCTCTGCCA	912-932	-	[36]
NVF2	GGACTTTCGCAGGCATAGTGGAGT	876-899	+	[36]
NVR2	TCCTGTGAGGGAGGCTGCGAT	1538-1558	-	[36]
NVF3	GAGAATCGCTGCTGCACGTT	1402-1421	+	[36]
NVR3	CGTCAGACTCAAGTGTGTAGGT	2291-2312	-	[36]
NVF4	GATCGCAACAAAGTGCTTGCCT	2120-2141	+	[36]
NVR4	CAACCTGAACCAAAGTTGACTA	3054-3075	-	[36]
NVF5	GCGACCGAGGAAGACTTCTGTGA	2813-2835	+	[36]
NVR5	TGGGCTCTGTAAATGGTTTCA	3770-3791	-	[36]
NVF6	AGCACCAAGACGAAATTCTGGAG	3680-3660	+	[36]
NVR6	ATGGAGTTCCATTGGGAGGTGCA	4481-4503	-	[36]
NV4611-F	CWGCAGCMCTDGAAATCATGG	4338-4358	+	[34]
MON383	CAAGAGACTGTGAAGACATCATC	5661-5683	-	[35]
MON381	CCAGAATGTACAATGGTTATGC	5362–5383	+	[35]
NVR7	GAAGTGCTGCACCCATTCCT	6475-6495	-	[36]
NVF8	GGTGAGCAACTTCTTTTCTT	6394-6413	+	[36]
NVR8	TTTGTCATGGGGGCGTTGATT	7033-7053	-	[36]
NVF9	GCACAAATTGAGGCCACTAA	6929-6948	+	[36]
NVR9	AAAGACACTAAAGAAAAGGAAAGA	7563-7585	-	[36]
NVF669	TYAGRCCNTAYCARGAYTGGAA	699-720	+	This study

*Positions given relative to X86557.1

3.2.3. Preparation of reference dataset

In order to provide a wide diversity of reference sequences for the identification of recombination events in the Albertan sequences, a comprehensive dataset was created based on all full-length and near full-length NoV GII.4 sequences available at Genbank as of September 4th, 2013. Sequences were renamed with their corresponding accession numbers and duplicates were removed using REFGEN [37]. Sequences with >10 internal unresolved nucleotide positions were removed, and sequences of variants Den Haag 2006b, Apeldoorn 2007 and New Orleans 2009 that clustered in the same clade in Neighbor Joining trees constructed for each ORF were reduced to 10 sequences each. Unique recombinant sequences that have been previously reported and characterized by others [24, 25] were excluded from the dataset. A total of 100 reference sequences, representing 16 out of 19 ORF2-GII.4 variants described to date were present in the final dataset. The three exceptions were Japan 2001, Kaiso 2003 and the South American cluster as no full-length sequences were available in Genbank for these variants. Sequences of three possible new GII.4 variants tentatively named as "Maryland 1", "Maryland 2" and "Randwick 2011" were also included in the analysis.

3.2.4. Phylogenetic and recombination detection analyses

Sample and reference nucleotide sequences were aligned with MAFFT, using the E-INS-i strategy. The alignment was visually inspected and manually adjusted in MEGA5.10 [38]. Recombination detection was performed with the software package RDP4.13 [39] and Simplot 3.5.1 [40]. RDP4.13 analysis was set up with the automated option including the following seven methods: RDP, GENECONV, Chimaera, MaxChi, and 3Seq for primary scans and BootScan and SiScan for secondary scans. For this study, only those sequences with events that were significant ($p \leq 0.05$) by at least five methods - including 3Seq unless only one

parent is present in the dataset - were reported as recombinants. 3Seq was selected to confirm recombination events because of its power, robustness against rate variation and low false positive rate [27, 41], but similarly to Chimaera, 3Seq can miss recombination events when only one parent is present [39]. Phylogenetic trees were constructed with MEGA 5.10 for each region of the alignment limited by recombination breakpoints, using the best nucleotide substitution model determined by Bayesian Information Criterion and branch support calculated based on 1000 bootstrap replicates.

3.3. Results

3.3.1. Norovirus epidemic activity, outbreak settings and patient demographics associated with samples selected for near full-length genome sequencing

Twenty near full-length sequences were obtained representing seven NoV GII.4 variants defined by partial ORF2 sequences (ORF2-GII.4 variants) from outbreaks in Alberta: Farmington Hills 2002 (n=1), Hunter 2004 (n=2), Yerseke 2006a (n=2), Den Haag 2006b (n=6), Apeldoorn 2007 (n=1), New Orleans (n=6) and Sydney (n=2) (Table 3.2). Six of the seven variants were selected isolates of the predominant strains associated with outbreak activity of epidemic proportions that had occurred in the province during one or more winter seasons between July 2002 and January 2013. The exception was variant Apeldoorn 2007 that had circulated at low levels causing only 8 outbreaks between July 2007 and June 2009 [33][chapter 2].

Table 3.2. Near full-length NoV GII.4 sequences obtained in this study

ORF2-GII.4 Variant	Strain	Accession Number	Period	Outbreak setting
Farmington Hills 2002	AlbertaEI131/2004	JX445152.1	Jul 2004-Jun2005	Community LTC/SL
Hunter 2004	AlbertaEI142/2006	JX445153.1	Jul 2005-Jun 2006	Community LTC/SL
	AlbertaEI190/2006	JX445154.1	Jul 2005-Jun 2006	Hospital LTC
Yerseke 2006a	AlbertaEI498/2006	JX445156.1	Jul 2006-Jun 2007	Community LTC/SL
	AlbertaEI513/2006	JX445157.1	Jul 2006-Jun 2007	Community LTC/SL
Den Haag 2006b	AlbertaEI438/2006	JX445155.1	Jul 2006-Jun 2007	Acute care hospital
	AlbertaEI603/2006	JX445158.1	Jul 2006-Jun 2007	Community LTC/SL
	AlbertaEI009/2008	JX445159.1	Jul 2007-Jun 2008	Community LTC/SL
	AlbertaEI102/2008	JX445160.1	Jul 2007-Jun 2008	Community LTC/SL
	AlbertaEI425/2008	JX445162.1	Jul 2008-Jun 2009	Community LTC/SL
	AlbertaEI109/2009	JX445163.1	Jul 2008-Jun 2009	Community LTC/SL
Apeldoorn 2007	AlbertaEI210/2008	JX445161.1	Jul 2007-Jun 2008	Community LTC/SL
New Orleans 2009	AlbertaEI045/2010	JX445164.1	Jul 2009-Jun 2010	Acute care hospital
	AlbertaEI119/2010	JX445165.1	Jul 2009-Jun 2010	Community LTC/SL
	AlbertaEI204/2010	JX445166.1	Jul 2010-Jun 2011	Community LTC/SL
	AlbertaEI065/2011	JX445167.1	Jul 2010-Jun 2011	Community LTC/SL
	AlbertaEI388/2011	JX445168.1	Jul 2011-Jun 2012	Community LTC/SL
	AlbertaEI003/2012	JX445169.1	Jul 2011-Jun 2012	Acute care hospital
Sydney 2012	AlbertaEI337/2011	KF509947.2	Jul 2011-Jun 2012	Community LTC/SL
	AlbertaEI063/2012	KF509946.2	Jul 2011-Jun 2012	Community LTC/SL

Variant Osaka 2007 was also identified in Alberta but was not included in this study as no sample was available for whole genome sequencing.

TC=long-term care

SL=senior lodge

The patient samples were collected during 20 NoV outbreaks that occurred in health care facilities, with 16 (80%) occurring in community-based long-term care facilities or senior lodge (LTC/SL), one (5%) in a hospital-based LTC and three (15%) in acute care hospitals. The median age of the 41 sampled patients who tested positive for NoV at the 17 community- or hospital-based LTC/SL outbreaks was 84 years (range: 48-93 years) and the median age of the 6 patients infected with NoV at acute care hospital outbreaks was 89 years (range: 79-95 years).

3.3.2. Characterization of recombinant NoV reference sequences from GenBank

The sequence dataset used to detect and analyze recombinants included 100 reference sequences from GenBank representing 21 ORF2-GII.4 variants. Analysis using seven different methods (RDP, GENECONV, Chimaera, MaxChi, 3Seq, BootScan and SiScan) and

RDPv.4.13 software identified variants Asia 2003, Osaka 2007, Apeldoorn 2007, Japan 2008, Randwick 2011, and Sydney 2012 as recombinants. The breakpoints and putative parental sequences for the six recombinant clusters matched those described in previous studies [24, 25] and are summarized in Table 3.3.

Table 3.3. Recombinant NoV GII.4 sequences identified in this study

	GII.4 variants						New recombinants	
	Asia 2003	**Osaka 2007**	**Japan 2008**	**Randwick 2011**	**Apeldoorn 2007 (includes AlbertaEI210/ 2008)**	**Sydney 2012 (includes AlbertaEI063/ 2012)**	**Seoul/1071 /2010/KR (JX448566)**	**LC31912 /2012/US (KF429777)**
Breakpoint(s) † (RDP/Simplot)	5106/5043	ND/5083 6640/6697	5020/5094	5638/5692	5028/5026	5030/5067	5596/5639	5114/5058
ORF1 (5-5104)	GII.P12*	GII.Pe*	Apeldoorn 2007	Den Haag 2006b	Hunter2004	Osaka 2007	New Orleans 2009	New Orleans 2009
ORF2a (5085-5600)	Unknown	Unknown	Den Haag 2006b	Den Haag 2006b	Pre-Apeldoorn‡	Alberta 2011§	New Orleans 2009	Den Haag 2006b
ORF2b (5601-6707)	Unknown	Unknown	Den Haag 2006b	New Orleans 2009	Pre-Apeldoorn‡	Alberta 2011§	Apeldoorn 2007	Den Haag 2006b
ORF3 (6707-7516)	Unknown	Den Haag 2006b	Den Haag 2006b	New Orleans 2009	Pre-Apeldoorn‡	Alberta 2011§	Apeldoorn 2007	Den Haag 2006b
p-Values:								
RDP	1.0×10^{-10}	9.0×10^{-15}	1.5×10^{-19}	1.0×10^{-9}	2.6×10^{-11}	3.5×10^{-10}	7.0×10^{-9}	1.2×10^{-32}
GENECONV	NS	1.5×10^{-13}	1.2×10^{-10}	NS	8.0×10^{-9}	NS	2.6×10^{-3}	3.4×10^{-28}
BOOTSCAN	6.8×10^{-11}	2.9×10^{-16}	1.3×10^{-18}	4.3×10^{-4}	9.3×10^{-12}	3.9×10^{-10}	2.7×10^{-9}	2.9×10^{-32}
MAXCHI	3.2×10^{-9}	4.9×10^{-8}	2.8×10^{-21}	1.4×10^{-17}	3.3×10^{-7}	4.5×10^{-15}	4.0×10^{-16}	2.6×10^{-24}
CHIMAERA	2.0×10^{-10}	5.0×10^{-5}	1.9×10^{-23}	9.9×10^{-10}	1.0×10^{-8}	2.3×10^{-8}	6.2×10^{-3}	3.1×10^{-26}
SISCAN	1.7×10^{-17}	1.4×10^{-10}	9.7×10^{-34}	5.8×10^{-31}	3.3×10^{-10}	5.4×10^{-37}	9.7×10^{-32}	5.3×10^{-40}
3SEQ	1.5×10^{-10}	NS	4.3×10^{-61}	1.6×10^{-47}	9.0×10^{-14}	2.0×10^{-57}	9.6×10^{-45}	4.8×10^{-84}

GC=GENECONV, BS=Bootscan, MC=MaxChi, CH=Chimaera, SS=SiScan, 3S=3Seq

NS=Not significant; ND=Not detected

*Genotype names according the Norovirus Genotyping Tool [42].

†Breakpoint positions are given relative to the alignment.

‡Representatives of this strain have not been reported yet; its possible origin has been discussed in [24, 25]

§AlbertaEI337/2011, identified in this study, is the representative of this cluster.

In contrast with previous findings, New Orleans 2009 which was described as a recombinant of Yerseke 2006a and Apeldoorn 2007 by Eden *et al.* [25] was not detected as a recombinant by our analysis with RDPv4.13. However, the low divergence between New Orleans 2009 and Yerseke 2006a and their relative position in ML phylogenetic trees for different genomic regions (Fig.3.1-3.3) may have prevented identification of New Orleans 2009 using RDPv4.13.

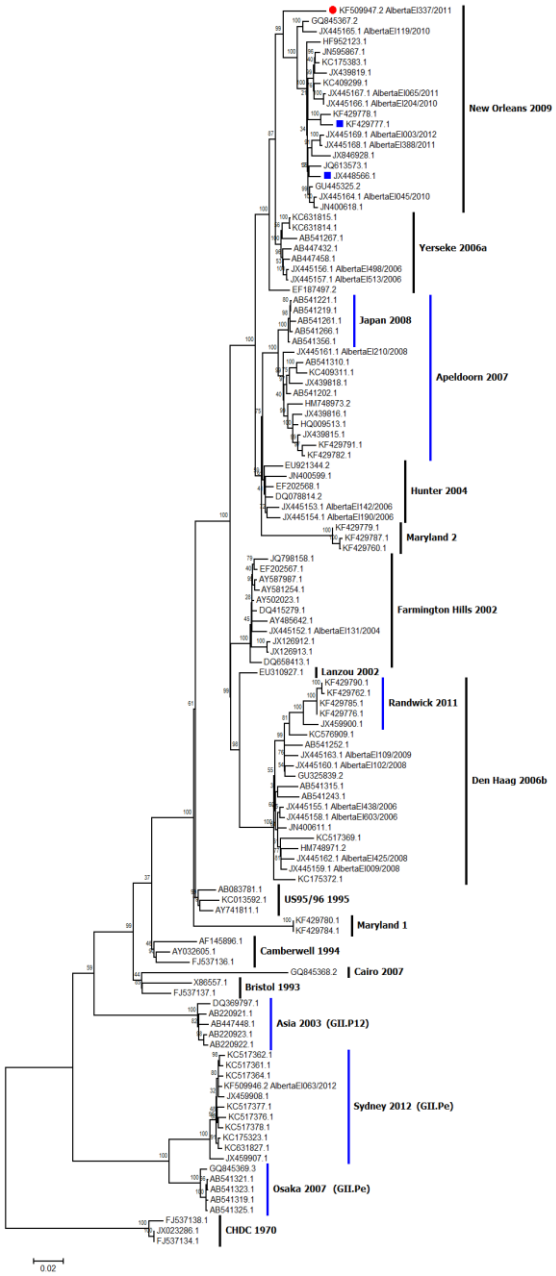


Figure 3.1. Maximum likelihood phylogenies of NoV GII.4 sequences in ORF1

The analysis included positions 5-5104nt, 5100 bp. The ML tree was inferred using the General Time Reversible model, assuming gamma-distributed substitution rates with 5 categories and invariant sites (G+I) and a 1000 bootstrap replicates. Red circle: Alberta 2011; blue squares: unique recombinant sequences; blue bars: variants with recombinant origin.

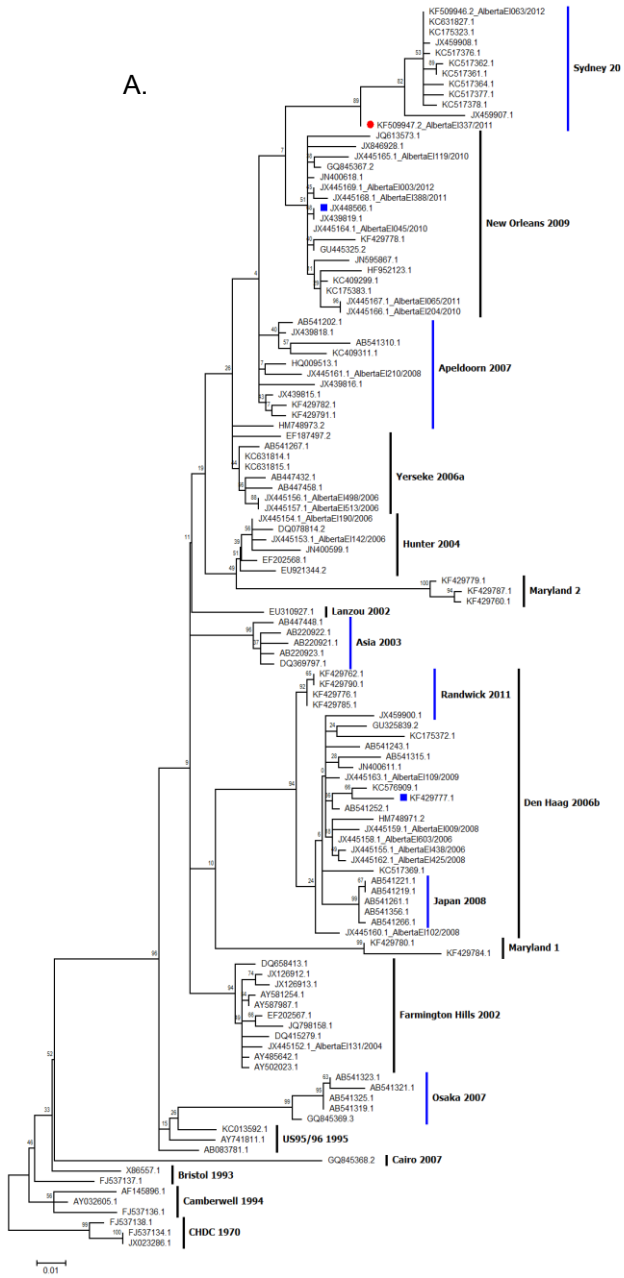
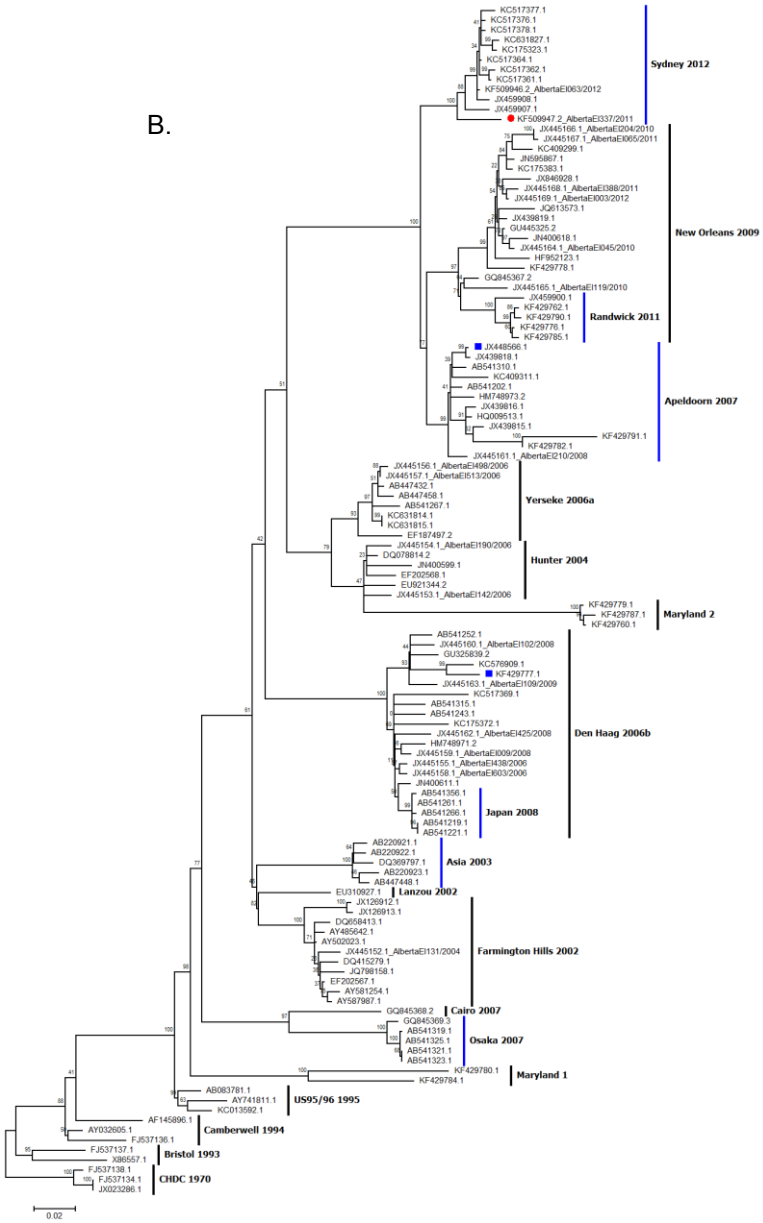


Figure 3.2. Maximum likelihood phylogenies of NoV GII.4 sequences in ORF2

The analysis included: (a) ORF2a (positions 5085-5600nt, 516bp) and (b) ORF2b (positions 5601-6707nt, 1107bp). Both ML trees were inferred using the General Time Reversible model, assuming gamma-distributed substitution rates with 5 categories and invariant sites (G+I), and 1000 bootstrap replicates. Red circle: Alberta 2011; blue squares: unique recombinant sequences; blue bars: variants with recombinant origin.

B.


KC517377.1
KC517376.1
KC517378.1
KC631827.1
KC175323.1
KC517364.1
KC517362.1
KC517361.1
KF509946.2_AlbertaEI063/2012
JX459908.1
JX459907.1
KF509947.2_AlbertaEI337/2011
Sydney 2012
JX445166.1_AlbertaEI204/2010
JX445167.1_AlbertaEI065/2011
KC409299.1
JN595867.1
KC175383.1
JX846928.1
JX445168.1_AlbertaEI388/2011
JX445169.1_AlbertaEI003/2012
JQ613573.1
JX439819.1
GU445325.2
JN400618.1
JX445164.1_AlbertaEI045/2010
HF952123.1
KF429778.1
GQ845367.2
JX445165.1_AlbertaEI119/2010
JX459900.1
KF429782.1
KF429790.1
KF429776.1
KF429785.1
Randwick 2011
New Orleans 2009
JX448566.1
JX439818.1
AB541310.1
KC409311.1
AB541202.1
HM748973.2
JX439816.1
HQ009513.1
JX439815.1
KF429792.1
KF429791.1
JX445161.1_AlbertaEI210/2008
Apeldoorn 2007
JX445156.1_AlbertaEI498/2006
JX445157.1_AlbertaEI513/2006
AB447432.1
AB447458.1
AB541267.1
KC631814.1
KC631815.1
Yerseke 2006a
EF187497.2
JX445154.1_AlbertaEI190/2006
DQ078814.2
JN400599.1
EF202568.1
EU821344.2
JX445153.1_AlbertaEI142/2006
Hunter 2004
KF429779.1
KF429787.1
KF429760.1
Maryland 2
AB541252.1
JX445160.1_AlbertaEI102/2008
GU325839.2
KC576909.1
KF429777.1
JX445163.1_AlbertaEI109/2009
KC517369.1
AB541315.1
AB541243.1
KC175372.1
JX445162.1_AlbertaEI425/2008
HM748971.2
JX445159.1_AlbertaEI009/2008
JX445155.1_AlbertaEI438/2006
JX445158.1_AlbertaEI603/2006
JN400611.1
AB541356.1
AB541281.1
AB541266.1
AB541219.1
AB541221.1
Japan 2008
Den Haag 2006b
AB220921.1
AB220922.1
DQ369797.1
AB220923.1
AB447448.1
Asia 2003
EU310927.1
Lanzou 2002
JX126912.1
JX126913.1
DQ658413.1
AY485642.1
AY502023.1
JX445152.1_AlbertaEI131/2004
DQ415279.1
JQ798158.1
EF202567.1
AY581254.1
AY587987.1
Farmington Hills 2002
GQ845368.2
Cairo 2007
GQ845369.3
AB541319.1
AB541325.1
AB541321.1
AB541323.1
Osaka 2007
KF429780.1
KF429784.1
Maryland 1
AB083781.1
AY741811.1
KC013592.1
US95/96 1995
AF145896.1
AY032605.1
FJ537136.1
Camberwell 1994
FJ537137.1
X86557.1
Bristol 1993
FJ537138.1
FJ537134.1
JX023286.1
CHDC 1970
0.02

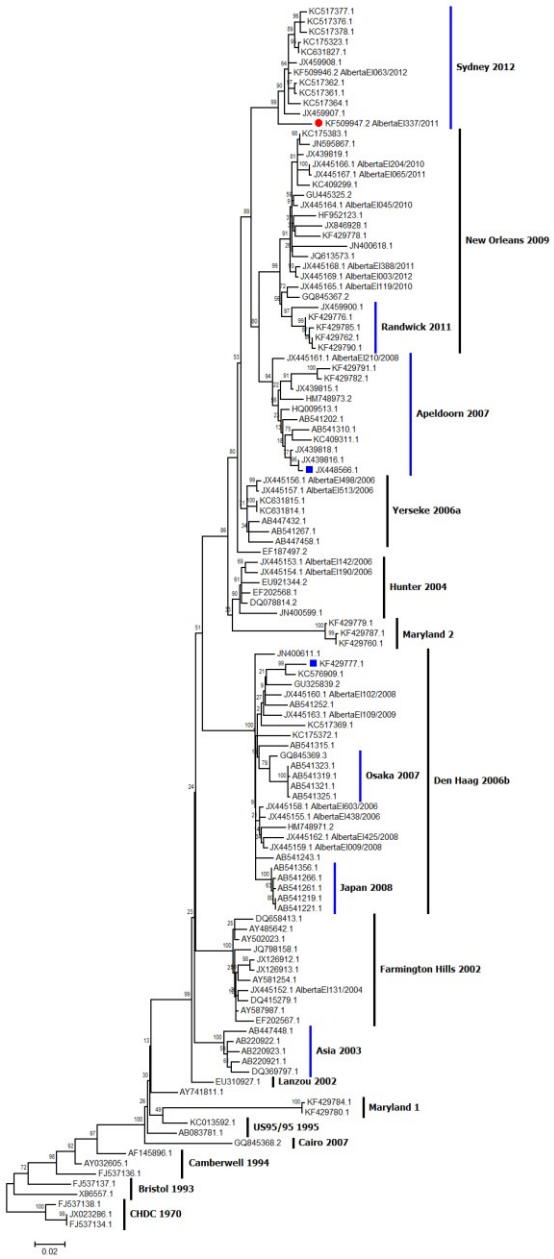


Figure 3.3. Maximum likelihood phylogenies of NoV GII.4 sequences in ORF3

The analysis included positions 6707-7516nt, 810bp. The ML tree was inferred with the Kimura 2 parameter model and assuming a gamma distribution of substitution rates with 5 categories and 1000 bootstrap replicates. Red circle: Alberta 2011; blue squares: unique recombinant sequences; blue bars: variants with recombinant origin.

Two previously sequenced genomes were identified as recombinants for the first time. The first strain, Seoul/1071/2010/KR (JX448566.1) was identified as an ORF2-intragenic recombinant of New Orleans 2009 and Apeldoorn 2007 as shown in Table 3.3 and Fig. 3.4(a). A second strain, LC31912/2012/US (KF429777.1) was identified as an ORF1/2 recombinant of New Orleans 2009 and Den Haag 2006b as shown in Table 3.3 and Fig. 3.4(b).

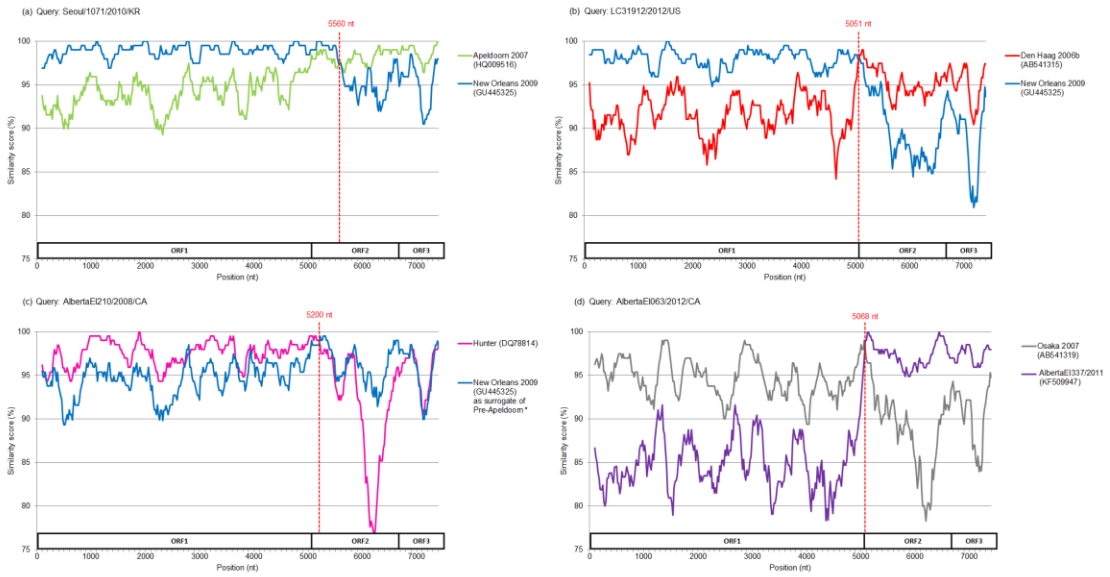


Figure 3.4. Simplot analysis of recombinant sequences identified in this study

Plots are shown for (a) Seoul1071/2010/KR, (b) LC31912/2012/US, (c) AlbertaEI210/2008/CA, and (d) AlbertaEI063/2012/CA. The breakpoints positions for each recombinant are shown in the top section of each plot and are marked with dashed lines. All plots were obtained using a 200bp sliding window, steps of 20bp, and K-2-p model. Representatives of pre-Apeldoorn have not been reported yet; its possible origin has been discussed in [24, 25].

3.3.3. Norovirus GII.4 recombinants from outbreaks in Alberta

Two out of 20 sequences, AlbertaEI210/2008 and AlbertaEI063/2012, were identified by RDPv4.13 as recombinants. AlbertaEI210/2008 is a member of the Apeldoorn 2007 clade,

with a breakpoint at position 5151nt, Fig. 3.4(c), whereas Alberta/EI063/2012 is a Sydney 2012 strain, with a breakpoint at position 5030nt, Fig. 3.4(d). Because our strategy of sequencing near full-length genomes through ten separate PCR reactions can potentially lead to artificial recombinants, we amplified the ORF1/2 junction via a single PCR reaction using primers NV4611 and Mon383. Maximum likelihood trees constructed for the regions flanked by these two primers before and after the breakpoints confirmed that both recombinants were not artifacts.

3.3.4. Comparison of previous ORF2-genotyping results vs. whole genome analysis

In total, 19 out of 20 NoV sequences from Alberta, grouped with a known parental or recombinant reference strains across all three ORFs (Fig. 3.1-3.3) matching our previous ORF2- genotyping results. The exception was AlbertaEI337/2011, which branched at the base of New Orleans 2009 with high bootstrap support in all maximum likelihood trees and maintained the same relative position independently of the genomic region used for analysis, a pattern consistent with non-recombinant strains that have evolved by antigenic drift. RDPv4.13 identified AlbertaEI337/2011 as the parental strain of Sydney 2012 with Osaka 2007 as the other parent. The event was highly significant for six out of the seven analytic methods (Table 3.3) and the breakpoint was detected at the ORF1/2 junction, around positions 5030-5067nt of the alignment, Fig. 3.4(d).

However, in the absence of similar reference sequences and due to its short distance from the New Orleans 2009 clade, AlbertaEI337/2011 was inaccurately interpreted as an ORF1/2 recombinant of New Orleans 2009 and Sydney 2012 by the Norovirus Genotyping Tool [42]. Recently, ORF1/2 recombinants of New Orleans 2009 and Sydney 2012 have been reported in Denmark and Italy [31, 32] but no full-length sequence of these recombinants was available in GenBank at the time of preparing this manuscript for inclusion in our dataset for

detection and analysis of recombination. Maximum likelihood trees constructed with partial sequences of ORF1 and ORF2, Fig. 3.5, showed that, in contrast to AlbertaEI337/2011, the European sequences branched off further from the root and were closer to more recent sequences of both genes suggesting a recombinant origin for these strains.

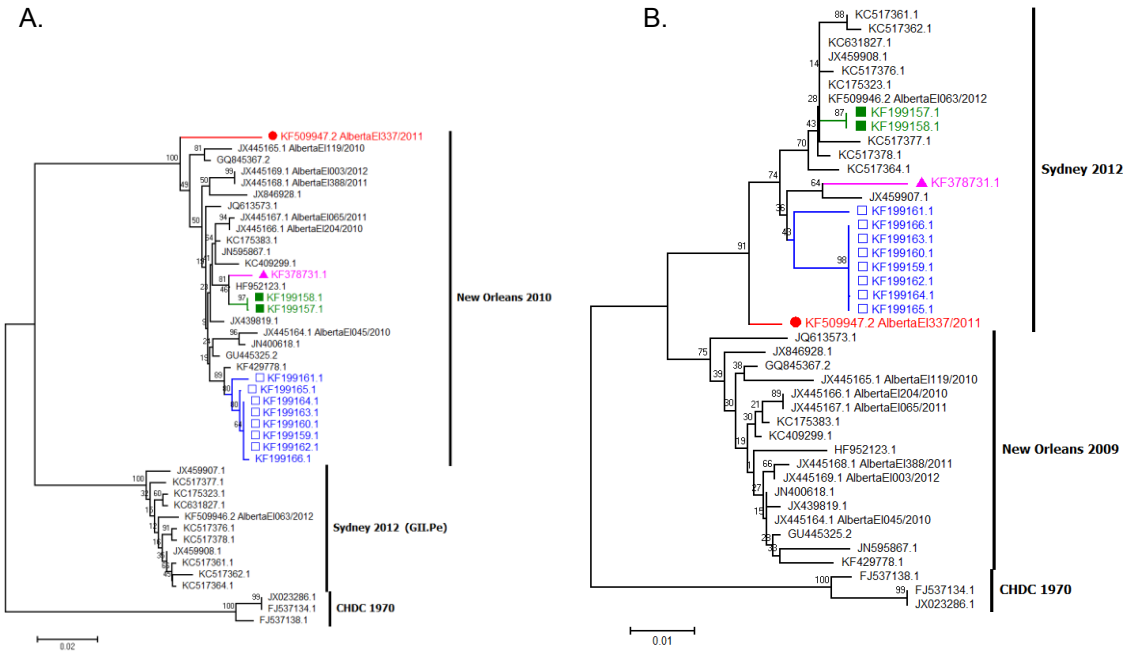


Figure 3.5. Comparison of AlbertaEI337/2011 with recombinant sequences from Europe

(a) Neighbor joining tree for the 3'end of ORF1 (648bp) and (b) the 5'end of ORF2 (478bp) AlbertaEI337/2011 is identified by a red filled circle, ORF1/2 recombinants of New Orleans 2009 and Sydney 2012 from three independent recombination events are shown with a pink triangle (isolate from Denmark), green filled squares and blue open squares (isolates from Italy). Trees were inferred with the Kimura 2 parameter model assuming a gamma distribution of substitution rates with 5 categories and branch support was calculated based on 1000 replicates.

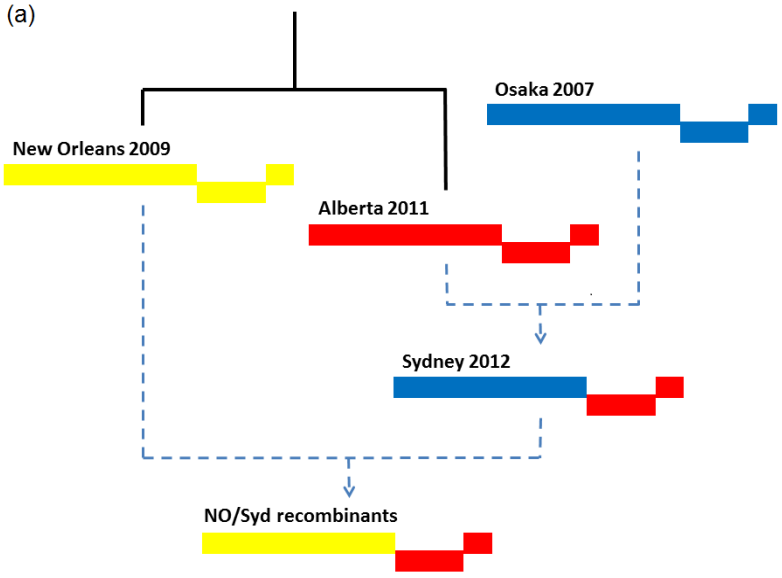
3.4. Discussion

Detection and analysis of recombination performed on near full-length NoV GII.4 sequences from 20 outbreaks in Alberta between July 2002 and January 2013 identified two recombinant isolates. The first recombinant corresponded to Apeldoorn 2007, a variant that circulated at low levels in the province for two years [33][chapter 2]. The second recombinant was characterized as Sydney 2012, a predominant outbreak strain that has circulated in Alberta since January 2012 that also caused global epidemic activity. Other previously reported GII.4 recombinants either circulated only locally (e.g. Asia 2003, Japan 2008) [12, 23] or at low levels globally (e.g. Apeldoorn 2007 and Osaka 2007) [17, 19, 43]. Even though recombination could be increasing NoV genetic fitness by combining different non-structural genes with capsid genes, antigenic drift might be more effective in creating new pandemic strains by allowing evasion of host herd protective immunity acquired against previously circulating NoV. We observed that many new NoV GII.4 variants branched at the base, i.e., from ancestors of previous clusters and postulate that this phylogenetic tree pattern could be used for early prediction or detection of new pandemic NoV GII.4 variants.

Our results indicate that most Alberta full-length sequences were homologous to reference sequences of NoV GII.4 described in global epidemics matching our previous ORF2- genotyping results except for AlbertaEI337/2011. This strain, previously identified based on ORF2 as a Sydney 2012, was characterized instead as a non-recombinant strain representing the most basal example of the New Orleans 2009 clade. Interestingly, it also forms the most basal branch of Sydney 2012 in the ORF2/3-based tree, indicating it most closely resembles one of the parents of the recombinant cluster Sydney 2012. Such sequences have not been reported previously, although their existence and time of emergence has been inferred from phylogenetic analysis by Eden *et al.* [44]. In this study we have shown that

AlbertaEI337/2011-like sequences, hereafter referred as Alberta 2011 likely represented a transient intermediate strain in the evolution of Sydney 2012. We believe that Alberta 2011 evolved from an early ancestor of New Orleans 2009 by antigenic drift and then recombined with Osaka 2007 to generate Sydney 2012, as proposed in Fig. 3.6(a). Unfortunately, we have only AlbertaEI337/2011 to support this evolution pathway of Sydney 2012.

(a)



(b)

	Epitope A						Epitope D				Epitope E		
GII.4 Variant	294	296	297	298	368	372	393	394	395	396	407	412	413
New Orleans 2009	P	S	R	N	A	D	S	T	T	P	S	N	I
Alberta 2011	T	S	R	N	E	D	S	T	T	H	S	N	T
Sydney 2012	T	S	R	N	E	D	G	T	T	H	S	N	T

Figure 3.6. Proposed origin of NoV GII.4 Alberta 2011 and Sydney 2012 strains

(a) Alberta 2011 evolved by antigenic drift from an ancestor of New Orleans 2009 and underwent recombination with Osaka 2007 generating Sydney 2012. Ulterior recombination between New Orleans 2009 and Sydney 2012 produced NO/Syd recombinants. Solid lines represent evolution by antigenic drift and broken lines represent recombination events. Representatives of Alberta 2011, Sydney 2012 and NO/Syd recombinants are KF509947.2, GU445325.2 and KF378731.1, respectively. (b) Changes in blockade epitopes for Sydney 2012 and Alberta 2011 vs. New Orleans 2009. Alberta 2011 displays 4 out of 5 changes associated with evasion of protective immunity by Sydney 2012.

The timing of AlbertaEI337/2011 and Sydney 2012's circulation in Alberta also supports our observations. AlbertaEI337/2011 caused an outbreak in September 2011 whereas the first Sydney 2012 strain was detected four months later in January 2012. Interestingly, AlbertaEI337/2011 already displayed most of the changes in blockade epitopes, i.e. potential neutralizing epitopes, that have been associated with evasion of herd immunity by Sydney 2012 [45] as shown in Fig. 3.6(b). The transmission success of Sydney 2012, in contrast to Alberta 2011, suggests that in addition to antigenic drift, additional factors such as increased fitness related to differences in non-structural genes made Sydney 2012 the predominant strain seen in the following years. Of note, five out of six NoV GII.4 recombinant clusters detected in this study (Asia 2003, Osaka 2007, Randwick 2011, Apeldoorn 2007 and Sydney 2012) retained contemporary GII.4 sequences at ORF2/3 whereas their ORF1 sequences were obtained from older GII.4 variants or different genotypes. Thus, there might be some genetic advantage for NoV to behave in this manner.

Our study has limitations that are common to all studies of recombination events. The first limitation was the absence of some full-length genome parental sequences, which impaired the identification of some recombinant strains, e.g., Cairo 2007. A second limitation was the presence of multiple recombinant reference sequences potentially leading to misinterpretations of the order of recombination events by RDPv.4.13 [39]. New Orleans 2009 and Apeldoorn 2007 might be examples of this; however, in support of our results, Apeldoorn 2007 was still identified as a recombinant after removing all New Orleans 2009 sequences from the dataset and our findings also agree with data published by Lam *et al.* [24]. Finally, in an effort to avoid false positives, we may have underestimated some recombination signals that were obscured by antigenic drift using our criteria for identifying true recombination events. Further studies using different sequence datasets and

recombination detection methods are necessary to establish the best approach to detect recombination events in NoV.

Here we demonstrated that Alberta 2011 was incorrectly identified as an ORF1/2-recombinant of New Orleans 2009 and Sydney 2012 by the Norovirus Genotyping Tool. However, real recombinants with this description have originated on at least three independent occasions and have circulated in Italy and Denmark as early as December 2012 [31, 32]. Therefore, a genotyping process covering both ORFs and cautious observation of clade order within each ORF are necessary to differentiate Alberta 2011, Sydney 2012 and Sydney 2012 recombinants. Future studies on the molecular epidemiology of these three strains could help understand the impact of changes in non-structural genes on NoV outbreak activity levels.

In summary, our study result provides detailed analysis of the genetic diversity of NoV GII.4 in Alberta, and underscores the need to incorporate genomic regions that are suitable for the identification of recombinant NoV in surveillance programs. Because sequencing more than one gene target of each outbreak strain might not be feasible for surveillance purposes, based on our results we recommend to use the ORF1/2 junction for genotyping via a single PCR reaction to detect antigenic drift versus recombination. We hypothesize that sequences branching at the base of previous variants may be candidates for future new NoV GII.4 pandemic strains and suggest to monitor for and further characterize (e.g. by whole genome analysis) such sequences.

3.5. Acknowledgements

This work was supported by Provincial Laboratory for Public Health, Alberta, Canada, and research grant funded by Alberta Health, Canada (RES0015904). We thank the

Provincial Laboratory for Public Health (ProvLab) staff, the collaborative teams in Alberta and the Northern Territories and the members of the Public Health (ProvLab) Outbreak Investigation Committee (OINC) for their work in outbreak investigations. We also thank Dr. David Evans for his scientific comments on this study.

3.6. References (chapter 3)

1. Atmar, R.L. and M.K. Estes, *The epidemiologic and clinical importance of norovirus infection.* Gastroenterology clinics of North America, 2006. **35**(2): p. 275-90, viii.
2. Patel, M.M., et al., *Systematic literature review of role of noroviruses in sporadic gastroenteritis.* Emerging infectious diseases, 2008. **14**(8): p. 1224-1231.
3. Hall, A.J., et al., *Norovirus disease in the United States.* Emerging infectious diseases, 2013. **19**(8): p. 1198-1205.
4. Liu, B.L., et al., *Identification of further proteolytic cleavage sites in the Southampton calicivirus polyprotein by expression of the viral protease in E. coli.* The Journal of general virology, 1999. **80 (Pt 2)**(Pt 2): p. 291-296.
5. Donaldson, E.F., et al., *Viral shape-shifting: norovirus evasion of the human immune system.* Nature reviews.Microbiology, 2010. **8**(3): p. 231-241.
6. Vongpunsawad, S., B.V. Venkataram Prasad, and M.K. Estes, *Norwalk Virus Minor Capsid Protein VP2 Associates within the VP1 Shell Domain.* Journal of virology, 2013. **87**(9): p. 4818-4825.
7. Glass, P.J., et al., *Norwalk virus open reading frame 3 encodes a minor structural protein.* Journal of virology, 2000. **74**(14): p. 6581-6591.
8. Bertolotti-Ciarlet, A., et al., *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.* Journal of virology, 2003. **77**(21): p. 11603-11615.
9. *Fields virology*, B.N. Fields, D.M. Knipe, and P.M. Howley, Editors. 2013, Wolters Kluwer/Lippincott Williams & Wilkins Health: Philadelphia :.
10. Vinje, J., *Advances in laboratory methods for detection and typing of norovirus.* Journal of clinical microbiology, 2015. **53**(2): p. 373-381.
11. Siebenga, J.J., et al., *Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007.* Journal of Infectious Diseases, 2009. **200**(5): p. 802-812.
12. Okada, M., et al., *Genetic analysis of noroviruses associated with fatalities in healthcare facilities.* Archives of Virology, 2006. **151**(8): p. 1635-1641.
13. Motomura, K., et al., *Divergent evolution of norovirus GII/4 by genome recombination from May 2006 to February 2009 in Japan.* Journal of virology, 2010. **84**(16): p. 8085-8097.
14. Kamel, A.H., et al., *Predominance and circulation of enteric viruses in the region of Greater Cairo, Egypt.* Journal of clinical microbiology, 2009. **47**(4): p. 1037-1045.

15. Zintz, C., et al., *Prevalence and genetic characterization of caliciviruses among children hospitalized for acute gastroenteritis in the United States.* Infection, Genetics & Evolution, 2005. **5**(3): p. 281-290.
16. Yen, C., et al., *Impact of an emergent norovirus variant in 2009 on norovirus outbreak activity in the United States.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2011. **53**(6): p. 568-571.
17. van Beek, J., et al., *Indications for worldwide increased norovirus activity associated with emergence of a new variant of genotype II.4, late 2012.* Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin, 2013. **18**(1): p. 8-9.
18. Hasing, M.E., et al., *Emergence of a new norovirus GII.4 variant and changes in the historical biennial pattern of norovirus outbreak activity in Alberta, Canada, from 2008 to 2013.* Journal of clinical microbiology, 2013. **51**(7): p. 2204-2211.
19. van Beek, J., et al., *Noronet report, April 2013*. 2013, National Institute for Public Health and the Environment: Bilthoven, The Netherlands.
20. Siebenga, J.J., et al., *Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006.* Journal of virology, 2007. **81**(18): p. 9932-9941.
21. Lindesmith, L.C., E.F. Donaldson, and R.S. Baric, *Norovirus GII.4 strain antigenic variation.* Journal of virology, 2011. **85**(1): p. 231-242.
22. Lindesmith, L.C., et al., *Emergence of a norovirus GII.4 strain correlates with changes in evolving blockade epitopes.* Journal of virology, 2013. **87**(5): p. 2803-2813.
23. Motomura, K., et al., *Divergent Evolution of Norovirus GII/4 by Genome Recombination from May 2006 to February 2009 in Japan.* The Journal of Virology, 2010. **84**(16): p. 8085-8097.
24. Lam, T.T., et al., *The recombinant origin of emerging human norovirus GII.4/2008: intra-genotypic exchange of the capsid P2 domain.* The Journal of general virology, 2012. **93**(Pt 4): p. 817-822.
25. Eden, J.S., et al., *Recombination within the pandemic norovirus GII.4 lineage.* Journal of virology, 2013. **87**(11): p. 6270-6282.
26. Bull, R.A., et al., *Norovirus recombination in ORF1/ORF2 overlap.* Emerging Infectious Diseases, 2005. **11**(7): p. 1079-1085.
27. Boni, M.F., D. Posada, and M.W. Feldman, *An exact nonparametric method for inferring mosaic structure in sequence triplets.* Genetics, 2007. **176**(2): p. 1035-1047.
28. Bull, R.A., M.M. Tanaka, and P.A. White, *Norovirus recombination.* The Journal of general virology, 2007. **88**(Pt 12): p. 3347-3359.
29. Simon-Loriere, E. and E.C. Holmes, *Why do RNA viruses recombine?* Nature reviews.Microbiology, 2011. **9**(8): p. 617-626.
30. Bull, R.A., et al., *Comparison of the replication properties of murine and human calicivirus RNA-dependent RNA polymerases.* Virus genes, 2010.
31. Martella, V., et al., *Evidence for recombination between the pandemic GII.4 norovirus strains New Orleans 2009 and Sydney 2012.* Journal of clinical microbiology, 2013.

32. Fonager, J., S. Barzinci, and T.K. Fischer, *Emergence of a new recombinant Sydney 2012 norovirus variant in Denmark, 26 December 2012 to 22 March 2013.* Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 2013. **18**(25): p. 20506.
33. Pang, X.L., et al., *Influence of novel norovirus GII.4 variants on gastroenteritis outbreak dynamics in Alberta and the Northern Territories, Canada between 2000 and 2008.* PloS one, 2010. **5**(7): p. e11599.
34. Yuen, L.K., et al., *Heminested multiplex reverse transcription-PCR for detection and differentiation of Norwalk-like virus genogroups 1 and 2 in fecal samples.* Journal of clinical microbiology, 2001. **39**(7): p. 2690-2694.
35. Noel, J.S., et al., *Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995.* Journal of medical virology, 1997. **53**(4): p. 372-383.
36. Chhabra, P., A.M. Walimbe, and S.D. Chitambar, *Complete genome characterization of Genogroup II norovirus strains from India: Evidence of recombination in ORF2/3 overlap.* Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2010. **10**(7): p. 1101-1109.
37. Leonard, G., J.R. Stevens, and T.A. Richards, *REFGEN and TREENAMER: automated sequence data handling for phylogenetic analysis in the genomic era.* Evolutionary bioinformatics online, 2009. **5**: p. 1-4.
38. Tamura, K., et al., *MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.* Mol Biol Evol, 2011. **28**(10): p. 2731-9.
39. Martin, D.P., et al., *RDP3: a flexible and fast computer program for analyzing recombination.* Bioinformatics (Oxford, England), 2010. **26**(19): p. 2462-2463.
40. Lole, K.S., et al., *Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination.* J Virol, 1999. **73**(1): p. 152-60.
41. Boni, M.F., et al., *Guidelines for identifying homologous recombination events in influenza A virus.* PloS one, 2010. **5**(5): p. e10434.
42. Kroneman, A., et al., *An automated genotyping tool for enteroviruses and noroviruses.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2011. **51**(2): p. 121-125.
43. Siebenga, J.J., et al., *Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007.* The Journal of infectious diseases, 2009. **200**(5): p. 802-812.
44. Eden, J.S., et al., *The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012.* Virology, 2014. **450-451**: p. 106-113.
45. Debbink, K., et al., *Emergence of new pandemic GII.4 Sydney norovirus strain correlates with escape from herd immunity.* The Journal of infectious diseases, 2013. **208**(11): p. 1877-1887.

Chapter IV: A NGS based method to study the intra-host population of Norovirus[‡]

[‡] This chapter has been published as: M.E. Hasing, B. Hazes, B.E. Lee, J.K. Preiksaitis, X.L. Pang. "A next generation sequencing-based method to study the intra-host genetic diversity of norovirus in patients with acute and chronic infection". *BMC Genomics*. 2016 Jul 1; 17(1):480. Permission to reuse this publication in my thesis was granted by the publisher, BioMed Central.

4.1. Introduction

Norovirus (NoV) is recognized as a leading cause of epidemic and sporadic gastroenteritis around the world [1]. The viral RNA genome is about 7500nt long and contains three ORFs. ORF1 encodes for a polyprotein that is cleaved into 6 non-structural proteins [2]. ORF2 and ORF3 encode the major (VP1) and minor (VP2) capsid proteins, respectively. NoVs are classified, based on VP1 amino acid sequences, into seven genogroups (GI-GVII) that are further divided into genotypes. To date, 41 NoV genotypes have been reported of which at least 29 have been found in humans [3], however, genotype GII.4 alone is responsible for over 60% of all NoV outbreaks worldwide [4]. New genetic clusters of GII.4, commonly referred to as GII.4 variants, arise every 2 to 5 years and spread rapidly often causing global pandemics [4]. Novel GII.4 variants evolve by antigenic drift and display changes in VP1 epitopes that can avert immune responses mounted against previous variants [5-7]. Homologous recombination is another mechanism responsible for the genetic diversity among NoV GII.4 and it commonly occurs at the ORF1/ORF2 junction allowing the virus to exchange structural and non-structural genes between different GII.4 variants or even different genotypes [8][chapter 3].

Norovirus acute gastroenteritis is a self-limited illness typically lasting 2 to 3 days while viral shedding can range from 13 to 56 days [9]. In immunocompromised patients, however, NoV shedding is usually prolonged [10-14] and cases of chronic NoV infection with shedding over 1-2 years have been reported [12, 13, 15, 16]. Due to their long shedding periods and weak immune responses it has been hypothesized that immunocompromised patients with chronic NoV infection might be reservoirs where new GII.4 variants emerge [11, 17]. Indeed, ORF2 sequence analysis in these individuals has shown that the virus can accumulate mutations in VP1 and develop an intra-host NoV population with large genetic

diversity [11, 16, 18-23]. However, it is still unclear how other regions of the viral genome evolve in immunocompromised hosts. More recently, the elderly and malnourished host have also been proposed as NoV reservoirs where NoV might accumulate mutations [17] but no studies have yet been performed in humans to confirm this hypothesis.

The gold standard method to study intra-host NoV populations involves cloning viral RT-PCR amplicons into plasmids followed by Sanger sequencing [16, 19, 22, 24], a labour intensive method that requires a relatively large number of clones to be processed in order to obtain an accurate assessment of the viral diversity. An alternative approach uses next generation sequencing (NGS) technologies, which process millions of fragments of nucleic acid in a single experiment. NGS is highly cost-effective in terms of cost per base, however, since NoV RNA represents a very small fraction of all stool RNA, as little as 0.01% [25], the cost per viral base can be considerable. Fortunately, the efficiency can be improved by viral RNA enrichment, and strategies previously used for NoV enrichment include polyA tail selection [26], RT-PCR amplification using NoV-specific primers [20, 23, 27], VIDISCA [28, 29] and virus purification [18]. Depletion of rRNA is another possible enrichment strategy [30]. For most cells, rRNA is the most abundant species of RNA and its removal with commercial kits can substantially increase the prevalence of non-rRNA [31, 32]. The method poses an advantage over others in that it maintains the original representation of the viral population present in the sample and is less likely to be affected by NoV RNA fragmentation.

In this study we establish a next generation sequencing-based method for NoV using bacterial rRNA depletion as an enrichment strategy for NoV RNA then use the resulting data to examine the intra-host viral population in samples from patients involved in NoV outbreaks (mostly elderly) and in longitudinal samples collected from an immunocompromised host with chronic NoV infection.

4.2. Materials and methods

4.2.1. Patient samples

All stool specimens were collected in Alberta between 2012 and 2014 and stored at -80°C until analysis. A total of six stool samples previously genotyped within our routine program of NoV surveillance in Alberta were included in this study [33][chapter 2]. The patient's description and NoV genotype associated to each sample are listed in Table 4.1. Samples OU1, OU2, OU3 and OU4 were collected from outbreak patients with acute NoV infection. The near full-length norovirus genome for sample OU1 was determined in a previous study using Sanger sequencing [chapter 3] and was included for comparison purposes. Samples SP1 and SP2 were collected four months apart from a pediatric hematopoietic stem cell transplant patient who first tested positive for NoV 6 months before sample SP1 was collected. The Biomedical Health Research Ethics Board of the University of Alberta approved this study.

Table 4.1. Samples analyzed by NGS

Sample	NoV genotype	Collection date	Patient description, age group
OU1	GII.4 Sydney 2012	January 2012	Outbreak patient (senior residence), 70-90Y
OU2	GII.4 Sydney 2012	September 2014	Outbreak patient (supportive living), >70-90Y
OU3	GII.5	November 2013	Outbreak patient (hospital acute care), 30-50Y
OU4	GI.7	November 2012	Outbreak patient (senior residence), 70-90Y
SP1	GII.4 Den Haag 2006b	December 2012	Bone marrow transplant patient, <18Y
SP2	GII.4 Den Haag 2006b	April 2013	

4.2.2. RNA extraction

Samples OU1, OU2, OU3, SP1 and SP2 were processed as follows: 50 to 75mg of stool were mixed with 20µL of proteinase K and 200µL of lysis buffer from Magazorb®RNA mini-prep kit (Promega, Madison, WI). Nucleic acids were extracted with 1mL of Trizol® (Life

Technologies, Carlsbad, CA) and according to the manufacturer's instructions. The yield of RNA per extraction was estimated using NanoDrop 1000 and the process was repeated to obtain at least 50ug of RNA per sample. The extracts were then treated with DNase (Promega, Madison, WI) and purified by phenol chloroform extraction and ethanol precipitation. RNA extracts were eluted through OneStep™ PCR Inhibitor Removal columns (Zymo Research, Irvine, CA). Bacterial rRNA was depleted using Ribo-Zero® bacterial kit (Epicentre, Madison, WI) according to the manufacturer's instructions using 5μg of RNA per sample and purifying by ethanol precipitation as the final step. Depletion of bacterial rRNA was performed once per sample. For comparison purposes, sample OU4 was processed using our routine enteric virus nucleic acid extraction method previously described [33][chapter 2] using Magazorb®RNA mini-prep kit. The nucleic extract was treated with DNase and purified by phenol chloroform extraction followed by ethanol precipitation. Sample OU4 was not depleted of bacterial rRNA. The presence of norovirus RNA was confirmed in all extracts by RT-qPCR as previously described [34].

4.2.3. Illumina library preparation and sequencing

Sample libraries were prepared from 1μg of RNA using the TruSeq RNA sample preparation kit v2 (Illumina, San Diego, CA) following the manufacturer's instructions with a fragmentation time of 1 min during the "elute-fragment-prime" step and unique indexed adapters for each sample. cDNA libraries were quantified with Qubit and the average fragment size was estimated using the Agilent 2100 Bioanalyzer. A control library of phage X714 was also included in each sample. All sample libraries (n=6) were sequenced once on a single Illumina MiSeq run to produce paired end reads of 250bp each, resulting in reads of 121bp each after removing adapters.

4.2.4. Identification, characterization and removal of rRNA reads

Raw sequence reads were quality-trimmed and filtered with Prinseq-lite, version 0.20.4 [35] using the following criteria: the first nucleotide at each end (5' and 3') was trimmed and the following nucleotides were also trimmed stepwise if their base quality was below 20. Sequences with an average base quality below 20 or with more than 90% of Ns were also removed. A description of the reads that were filtered is provided in Appendix A.

Ribosomal RNA reads were identified and filtered out with SortmeRNA [36] using the 23S/28S large subunit (LSU) and 16S/18S small subunit (SSU) rRNA SILVA 119 databases and the 5S and 5.8S rRNA Rfam databases for all three domains of life (Eukarya, Bacteria and Archaea). All reads failing to pass filters (i.e. non-rRNA reads) were maintained as paired-ends reads and used in downstream analyses the command used for analysis is described in Appendix B)

A subset of 150,000 single-end reads with length ≥80nt and identified as bacterial 23S rRNA, the most predominant type of rRNA found in all samples, was uploaded in the SILVAngs data analysis service (https://www.arb-silva.de/ngs/) for identification of operational taxonomic units (OTUs).

4.2.5. Assembly of NoV genomes

Non-rRNA paired-ends reads of samples OU1, OU2, OU3 and OU4 were mapped using Bowtie, version 2-2.2.5 [37] to Genbank reference sequences KF509946.2 (genotype GII.4 Sydney), KC631827.1 (genotype GII.4 2012 Sydney), KJ196277.1 (genotype GII.5) and JN899243.1 (genotype GI.7, partial genome), respectively, whereas SP1 and SP2 reads were aligned to KC576909 (genotype GII.4 2006b Den Haag). The consensus sequence of each sample was obtained from the alignments (SAM files) using Samtools.

Since no full-length or near full-length genome NoV GI.7 sequences were available in GenBank, additional steps were followed for OU4. OU4 reads were assembled *de novo* with Velvet version 1.2.10 using several hash lengths with read category set to 'short paired' and including the consensus sequence (partial ORF1, complete ORF2 and ORF3) as a long read (the command used for this analysis is described in Appendix B). Norovirus contigs were identified among all Velvet assemblies with BLAST [38] using JN899243.1, a genotype GI.9 strain, as query sequence. All NoV contigs and the partial consensus sequence were aligned using MEGA 6.0 [39] to obtain the final genome assembly.

The ends of each NoV genome were extended beyond the reference sequences by using the first and last 15 nt as query for matches among fastq sequences using the Unix "grep" command. Matching reads were aligned to the consensus sequence and any extra nucleotides (5'or 3' overhangs) were incorporated into the consensus sequence. The process was repeated until no more extra nucleotides were found at either end of the consensus sequence.

Sample reads were mapped with Bowtie 2 to their respective NoV extended consensus assembly. The final alignments (SAM files) were used to calculate the coverage per genome position using BEDTools, version 2.14.3-1 [40].

4.2.6. NoV sequencing using Sanger's method

The NoV strain from sample OU3 was sequenced using Sanger's method to compare results against those obtained with Illumina MiSeq. Nine pairs of primers (described in Table 4.2) were designed to retrieve overlapping PCR amplicons between ~600 to 1100 bp long, spanning altogether all NoV ORFs. The RT and PCR reactions were performed as previously described [chapter 3]. PCR products were obtained for six out of nine pair of primers and

were sequenced in both directions. The assembly of the sequences produced two non-overlapping contigs: a 3,448bp sequence containing a partial ORF1 (incomplete at the 5' and 3' ends) and a 1,906 bp sequence spanning ORF2 and ORF3 (incomplete at the 5' and 3' ends, respectively).

Table 4.2. Primers used for sequencing NoV strain OU3 using Sanger's method

Name	Type	Sequence	Start *	End*	Sense	Tm (°C)	Reference
NVF1	Forward	GTGAATGAAGATGGCGTCTA	-12	8	+	52.2	[41]
1GII5R	Reverse	AACTCCAAAGAGCTCTGCAAG	950	970	-	55.3	This study
2GII5F	Forward	AGGATCTCATAGGGAAGTTGAG	855	876	+	53.3	This study
2GII5R	Reverse	GCGCATCAGTCACAGGGTTGCTCATTC	1686	1712	-	55.6	This study
3GII5F	Forward	CTCAGGTGATCAGAGAGTGG	1591	1610	+	54.2	This study
3GII5R	Reverse	CAGTGGCCTCCACTTGTTTTTC	2570	2591	-	56.7	This study
4GII5F	Forward	AGTGAGATACTATGTTAAATGTGTTCAA	2446	2476	+	55	This study
4GII5R	Reverse	GCGGTGTGCACCCCTATAA	3510	3528	-	57.5	This study
5GII5F	Forward	AATGGGTACCCACGCCAC	3352	3369	+	58	This study
5GII5R	Reverse	TTCAGGTTCAGCTGAGAACCTAAC	4382	4405	-	56.6	This study
6GII5F	Forward	AGAGTTGGCATGAACATGAATG	4241	4262	+	53.6	This study
6GII5R	Reverse	ACAAAATTAGTTCTAATCCAGGGGTC	5253	5278	-	54.9	This study
G2SKF	Forward	CNTGGGAGGGCGATCGCAA	5070	5088	+	61.7	[42]
7GII5R	Reverse	CCTGTTTGCTGGATTGCTTTCAC	6133	6155	-	57.2	This study
8GII5F	Forward	GTGACAGGACAGGTCCCTAATG	5976	5997	+	57	This study
8GII5R	Reverse	GCTGGTGGTCTTCATTGAACC	7055	7075	-	56.3	This study
9GII5F	Forward	TGATATGATAGCAATCAAACAGGGAG	6925	6950	+	54.8	This study
9GII5R	Reverse	GACTCCCCCTTCTTGCGAAG	7478	7497	-	57.9	This study

4.2.7. Characterization of non-rRNA, non-NoV sequences

The sequences failing to align with Bowtie 2 to the final NoV consensus sequence were analyzed with BLAST to further characterize the major components of stool RNA. All reads from a single end were queried against the non-redundant nt database using megablast (standalone version with databases downloaded on July 24, 2015; see parameters of BLAST analysis in Appendix B). Results were analyzed with SPSS after removing duplicates, i.e. if a read had more than one BLAST hit, then only the hit with the lowest e-value was included in

the analysis. Bacterial hits belonging to the normal human gut were identified using as reference the microoganisms reported in previous studies [43, 44].

4.2.8. Analysis of single nucleotide variants

Single nucleotide variants (SNVs) were called with FreeBayes [45] and visually inspected in Tablet version 1.14.04.10 [46]. The following criteria was used for SNV calling with FreeBayes: -K (report all alleles passing filters) , --haplotype_length =1 (call haplotypes as 1nt long), –m or mapping quality =10 (chance that the read truly originated elsewhere of 1 in 10), –q or base quality =20 (chance of a wrong base call of 1 in 100), –F or alternate fraction =0.02 (call SNVs with frequencies ≥2%), --min-coverage =10 (call SNVs for positions with coverage ≥10X) and –C or min-alternate-count=5 (call SNVs with coverage ≥5X) (the command used for the analysis is described in Appendix B). The choice to set up the analysis to detect variants with frequencies ≥2% was made based on a study reporting that sequencing errors with the MiSeq platform can produce false variants that are undistinguishable from true low frequencies variants at ≤ 1% with a 1000X average coverage [47]. We also set up the analysis to call SNVs with a coverage of ≥5X based on: 1) a study that eliminated virtually all false positives by calling variants if counted ≥10 times independently [48] and 2) the lower coverage per genome position achieved with our samples compared to Van den Hoecke *et al.* [48].

4.3. Results

4.3.1. Characterization of the major components of stool RNA after bacterial rRNA depletion

To determine whether our bacterial rRNA depletion process was effective we characterized all rRNA reads present in Illumina MiSeq NGS data from total (n=1) and bacterial rRNA-depleted (n=5) stool RNA. Bacterial 23S and 16S rRNA dominated the pool of total RNA (Table 4.3) with a combined abundance in the non-enriched OU4 sample reaching 99%. In the rRNA-depleted samples (OU1, OU2, OU3, SP1 and SP2) bacterial 23S and 16S rRNA remained prominently present with a combined abundance of 63.52% to 95.23%.

Table 4.3. Percentage of rRNA reads per source

	OU1	OU2	OU3	OU4*	SP1	SP2
Total reads	7,370,512	4,703,034	9,386,008	8,536,368	5,891,654	2,909,390
Reads passing QC filters	7,362,154 (99.89%)	4,696,680 (99.86%)	9,375,772 (99.89%)	8,516,490 (99.77%)	5,781,324 (98.13%)	2,874,654 (98.81%)
Per source**:						
rRNA reads	6,4291,66 (87.33%)	3,054,402 (65.03%)	7,651,316 (81.61%)	8,461,361 (99.35%)	4,591,394 (79.42%)	2,751,629 (95.72%)
Eukarya 28S	4.75%	0.70%	5.83%	0.01%	0.54%	0.21%
Eukarya 18S	2.76%	0.66%	3.62%	0.00%	0.12%	0.23%
Bacteria 23S	70.87%	57.01%	60.87%	73.82%	75.30%	90.71%
Bacteria 16S	6.07%	6.51%	9.76%	25.50%	3.45%	4.52%
Archaea 23S	2.09%	0.01%	0.79	0.00%	0.00%	0.00%
Archaea 16S	0.18%	0.00%	0.19%	0.00%	0.00%	0.04%
5S	0.08%	0.12%	0.37%	0.00%	0.01%	0.00%
5.8S	0.52%	0.02%	0.16%	0.02%	0.00%	0.00%
non-rRNA reads	932,988 (12.67%)	1,642,278 (34.97%)	1,724,456 (18.39%)	55,129 (0.65%)	1,189,930 (20.58%)	123,025 (4.28%)

rRNA reads were identified with SortmeRNA using the SILVA 119 small (16S/18S, SSU) and large subunit (23S/28S, LSU) and the rfam 5S and 5.8S ribosomal RNA databases for all three domains of life (*Bacteria*, *Archaea* and *Eukarya*).

*Total stool (undepleted) RNA

**Percentages per source were calculated relative to reads passing QC filters.

Although it is not possible to quantify the enrichment effect of rRNA depletion without paired samples, the percentages of NoV sequences obtained for each sample (Table 4.4) suggests that the non-rRNA fraction was enriched. For the non-enriched OU4 sample, complete elimination of rRNA would have given a 120-fold enrichment of NoV sequences (2.43% vs. 0.02%). Sample OU2, had the best (lowest) experimental enrichment factor, only 3-fold below the optimal (5.36% vs. 1.88%) whereas sample SP2 had the poorest (highest) experimental enrichment factor, 23-fold below the optimal (7.55% vs. 0.33%) but still considerably better than the non-enriched sample. The best and worst experimental enrichment corresponded, respectively, to the samples with the lowest and highest remaining bacterial 23S rRNA fraction. In general, bacterial rRNA depletion seemed to perform better for 16S rRNA than for 23S rRNA. As a result, the ratio of 23S to 16S bacterial rRNA for depleted RNA samples was higher (6.4 to 21.8) compared to total RNA (2.9).

Table 4.4. Norovirus reads and coverage per sample

Sample	Reads mapping to NoV	Percentage of NoV reads		Average coverage	% of positions called *	Length of consensus sequence (bp)
		Vs. non-rRNA reads	Vs. quality filtered reads			
OU1	37,863	4.01%	0.51%	590X	100%	7,532
OU2	88,270	5.36%	1.88%	1,318X	99.55%	7,525
OU3	746	0.04%	0.01%	11X	90.59%	N.A.**
OU4	1,420	2.43%	0.02%	22X	N.A.	7,657
SP1	102,805	8.54%	1.78%	1,603X	99.91%	7,524
SP2	9,589	7.55%	0.33%	149X	99.91%	7,535

N.A.: not applicable

* Coverage was calculated compared to KF509946 (OU1), KC631827 (OU2), KJ196277 (OU3), JN899243.1 (OU4) and KC576909 (SP1 and SP2).

**The alignment of OU3 produced 9 non-overlapping contigs of 1744, 281, 518, 394, 434, 106, 119, 487 and 2760 nt long.

Further characterization of bacterial 23S rRNA reads showed a higher genetic diversity (i.e. higher number of OTUs for a given number of analysed sequences) in total RNA (OU4) vs. rRNA depleted RNA (OU1, OU2, OU3, SP1 and SP2) (Figure 4.1a). This suggests

that the rRNA depletion process was biased towards removing sequences from some bacterial species or taxonomic groups.

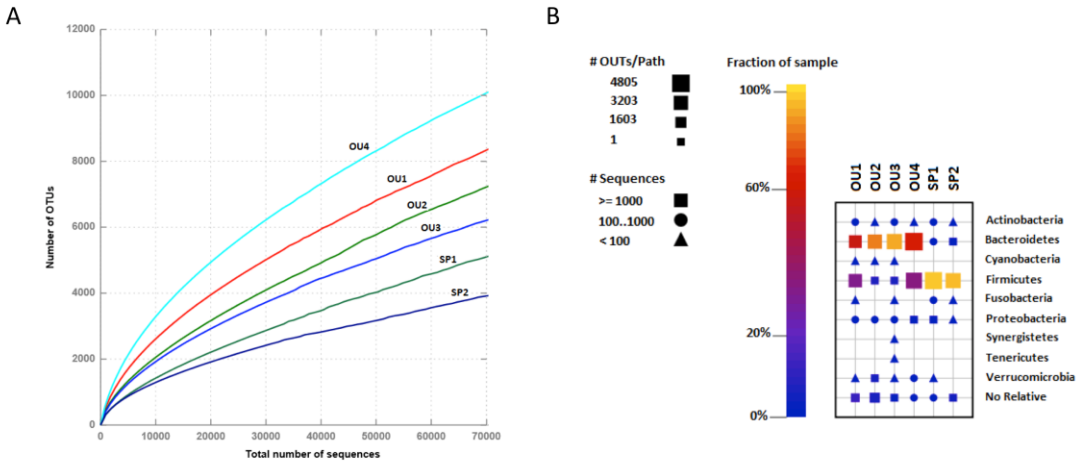


Figure 4.1. Taxonomic fingerprint of bacterial rRNA 23S sequences present in stool RNA of individuals with NoV infection

A. Rarefaction curves (number of OTUs detected per number of reads). B. Distribution of bacterial rRNA 23S sequences by phyla. The analysis was performed using the SILVA NGS data analysis service.

In all samples, the majority of bacterial 23S rRNA (94.5 to 99.9 %) was from the phyla Bacteroidetes and Firmicutes (Figure 4.1b). The lowest bacterial diversity was found in the samples from the immunocompromised patient, SP1 and SP2, which in contrast to the samples from outbreak patients, showed higher abundance of Firmicutes over Bacteroidetes. These relative abundances are in agreement with a previous study describing that microbiome of subjects with NoV infections resembles that of healthy subjects, being dominated by the phylum Bacteroidetes with fewer members of Firmicutes; however, cases showing an increased number of Firmicutes over Bacteroidetes (as observed in our immunocompromised patient) have also been reported [49, 50].

4.3.2. Non-rRNA-non-NoV RNA

In order to fully characterize other components of stool RNA, all non-rRNA-non-NoV reads (NRNNR) were analyzed with BLAST. Between 45% and 90% of NRNNR had a BLAST hit (Figure 4.2). The kingdoms with most hits varied from sample to sample and were either Bacteria (OU1, OU2, OU3 and SP1) or Eukarya (OU4 and SP2).

The majority of bacterial hits (47-91%) were from microorganisms considered normal flora of the human gut. The bacterial species with most BLAST hits varied among samples and were: *Bacteroides fragilis* (40% and 18% of NRNNR in SP1 and OU2, respectively), *Faecalibacterium prausnitzii* (8% and 7% of NRNNR in OU1 and OU3), *Anaerostipes hadrus* (28% of the NRNNR in SP2) and an uncultured bacterium (3% of the NRNNR in OU4).

Eukaryotic hits were more abundant than bacterial hits in samples SP2 and OU4. Most of the eukaryotic hits in SP2 were from animal species considered common food sources (e.g. *Sus scrofa* and *Gallus gallus*). Almost all eukaryotic hits for OU4 (99%) were from *Xanthophyllomyces dendrorhous*, a yeast not yet reported to be part of the normal eukaryotic gut flora in humans in previous studies [51, 52]. *X. dendrorhous* was also identified in the other five samples although in lower proportion (3% to 20% in of NRNNR). Human sequences were found in all six samples representing 2.5%, 0.2%, 0.2%, 0.1%, 13.7% and 8.4% of NRNNR in OU1, OU2, OU3, OU4, SP1 and SP2, respectively.

Archaeal hits represented 0.01% to 1.5% of NRNNR among the six samples while viral hits were 0.05% to 0.8% of NRNNR. Viral hits included plant viruses and bacteriophages. Pepper Mild Mottle virus (PMMV), a plant virus previously identified in human stools [53], was present in samples OU1, OU2, SP1 and SP2. OU2 had PMMV sequences in sufficient amount to yield a full-length genome (GenBank accession number KU311159) with an average

coverage of 27X. All six samples had a small proportion of NoV sequences that were missed and therefore not filtered by Bowtie (0.002% to 0.5% of NRNNR) as well as sequences from phage phiX174 (0.01% to 0.6% of NRNNR). The presence of the phiX174 phage was expected since a library of this phage was used as a spike-in control for the MiSeq sequencing run.

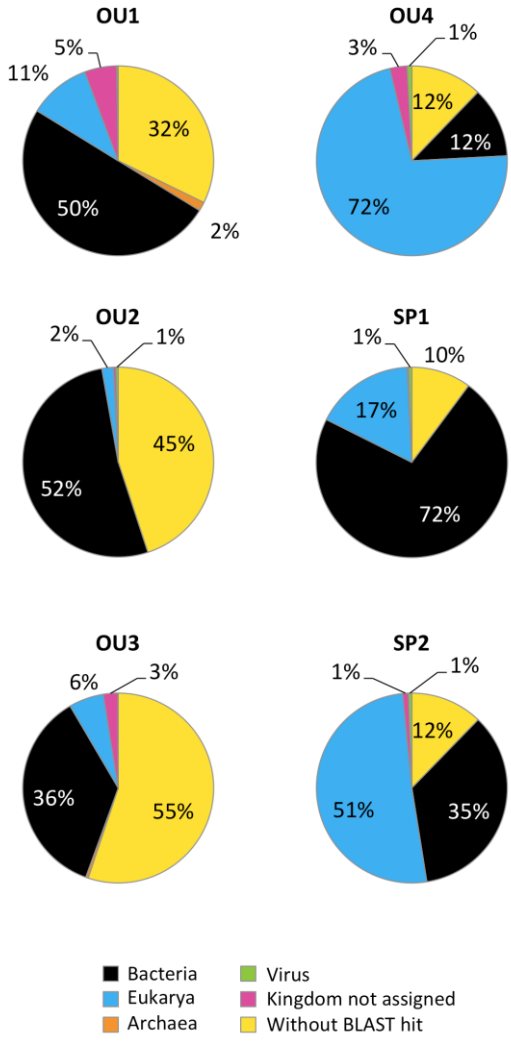


Figure 4.2. Characterization of non-rRNA-non-NoV RNA in stool RNA.

The non-rRNA-non-NoV reads from six different stool samples were characterized using BLAST. Results were summarized at the kingdom level.

4.3.1. Mapping of NoV sequencing reads

The samples included in this study (n=6) are described in Table 4.2. NoV sequences represented 0.01% to 1.88% of all quality-filtered reads and 0.04% to 8.54% of the non-rRNA reads. The average coverage of the final consensus sequences ranged between 11X and 1,603X (Table 4.3 and Figure 4.3). Five out of six samples yielded near full-length NoV sequences and called 99.91 to 100 % of the reference genome. OU3, the sample with the lowest percentage of NoV reads and lowest coverage, failed to yield a complete NoV genome sequence. The percentage of NoV sequences per sample showed a strong correlation with the Ct values of the RT-qPCR performed on RNA stool extracts (Spearman's rho correlation coefficient: 0.886, P=0.019, two-tailed). An equally strong correlation was observed between the percentage of NoV sequences and viral titer per ng of RNA (Appendix C). These results suggest that the poor yield of NoV sequences from sample OU3 was due to a low abundance of viral RNA rather than failure of the enrichment method.

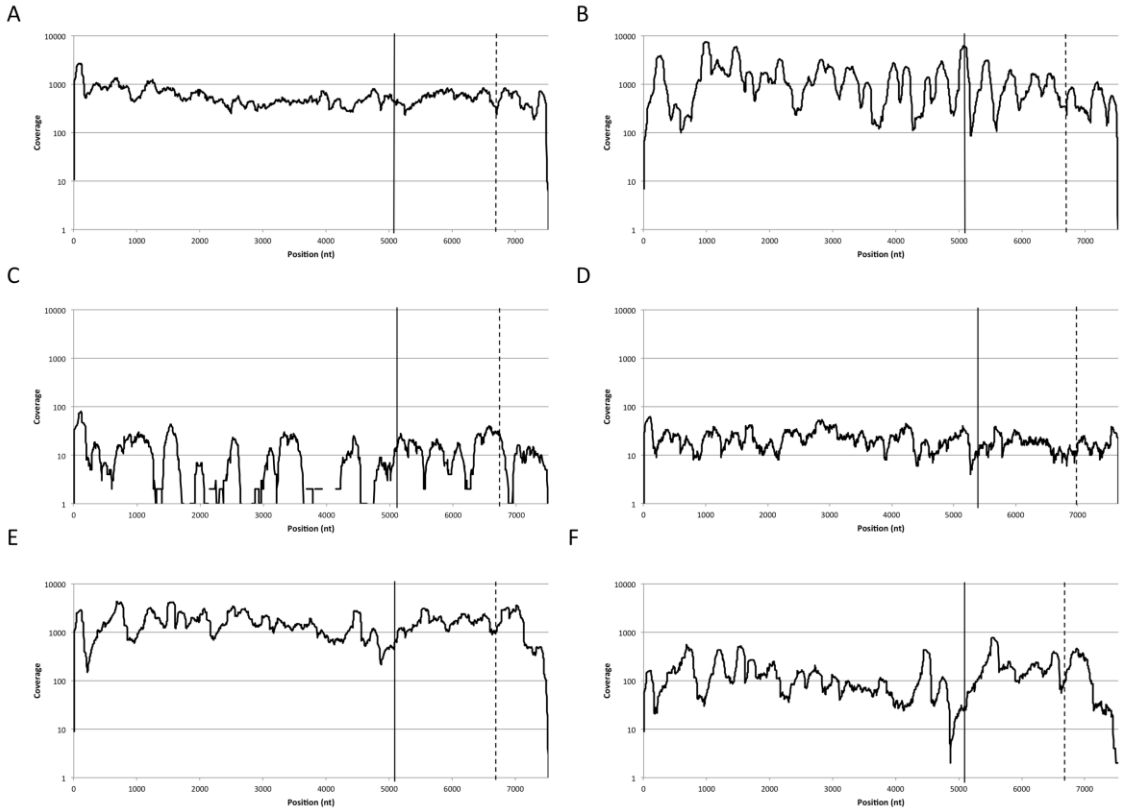


Figure 4.3. Coverage per NoV genome position.

Number of reads (in log scale) that aligned with the corresponding NoV consensus sequence of OU1 (A), OU2 (B), OU3 (C), OU4 (D), SP1 (E) and SP2 (F). Solid and broken lines indicate the start of ORF2 and ORF3, respectively.

In order to validate our NGS method, the NoV consensus sequences of samples OU1 and OU3 obtained with the MiSeq platform were compared to those obtained using Sanger sequencing. There was a large concordance between both methods. NoV OU1, which achieved high coverage with MiSeq, showed a concordance of almost 100% between the sequences from both methods except for one nucleotide. The base mismatch was located near the 3'end of the genome and was confirmed as an error in Sanger sequencing after review of the Sanger chromatogram and the coverage with MiSeq at that position (21X). Furthermore, the 5' and 3' ends of the genome could be extended by 9 and 24 nucleotides,

respectively, with MiSeq. These nucleotides were not covered with Sanger because of their close proximity to the primers used for PCR amplification and sequencing. The consensus sequence with lowest coverage, NoV OU3, showed three mismatches over a total of 5,081 nt that could be sequenced using Sanger and MiSeq. The mismatches were identified at 1732 nt, 2167 nt and 2239 nt (positions are given relative to KU311160) and had, respectively, coverages of 1X, 2X and 2X and Phred base quality scores of 16, 39 and 39. All three mismatches occurred at wobble bases resulting in synonymous mutations and the last two were located in the p22 gene. Since these positions had very low coverage, it is possible that the mismatches were created by sequencing errors with MiSeq, however, the high base quality associated with at least 2 of these three nucleotides and their relative position in the genome suggests that they could also represent true SNVs originally present in the sample.

4.3.2. Detection of Single Nucleotide Polymorphisms

Single nucleotide variants (SNV) were identified to measure the intra-host genetic population of NoV. In order to reduce false positives, we only called a SNV if it was observed at least 5 times and represented a minimum of 2% of all observations. Using these criteria we did not detect any SNVs in all acute infection samples (OU1, OU2, OU3 and OU4). OU3 and OU4 have relatively low coverage (Table 4.4) limiting the ability to detect SNVs, but OU1 and OU2 have an average coverage above 250X and still no SNVs were found. In contrast, the first sample from the immunocompromised subject with chronic NoV infection, SP1, had 51 SNVs while SP2, the second sample collected from the same individual 4 months later, had 122 SNVs, indicating that the genetic diversity of NoV in this patient was higher and also increased over time (Figure 4.4). The increased number of SNVs in SP2 vs. SP1 was confirmed even after controlling for differences in coverage. The higher number of SNVs in

the SP2 sample was not due to more sensitive detection, since SP2 had a lower average coverage than SP1, OU1, and OU2 (Table 4.4).

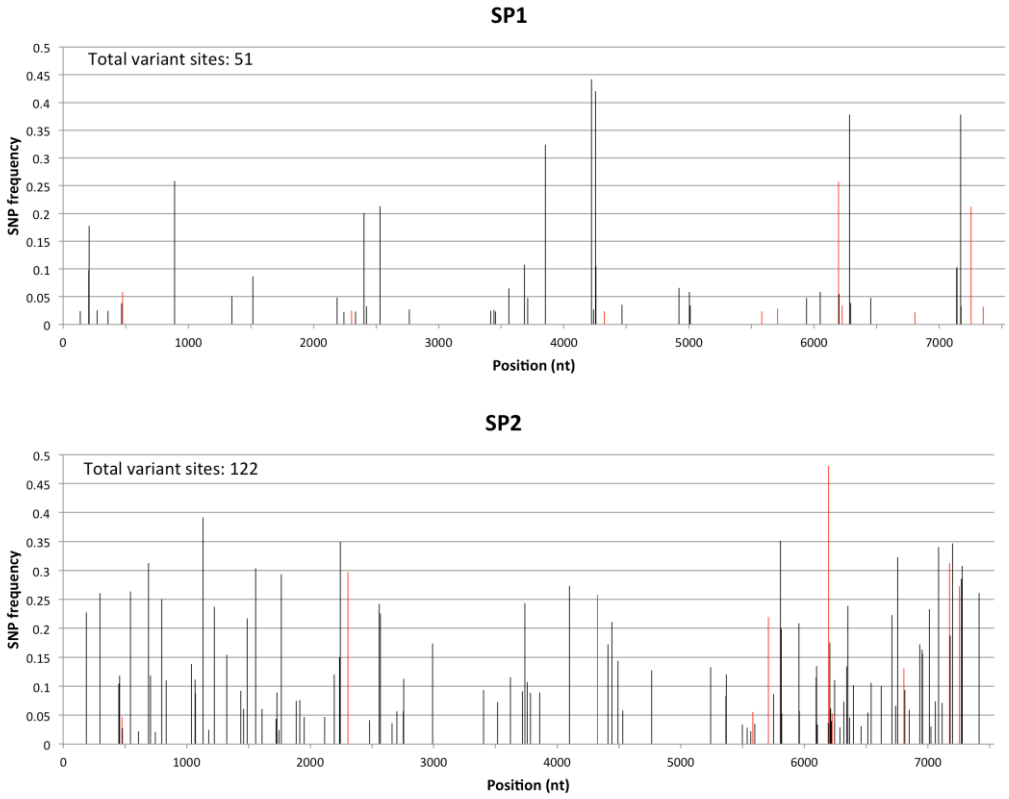


Figure 4.4. Distribution of NoV SNV frequencies across the viral genome

Samples SP1 and SP2 were collected four months apart from an immunocompromised bone marrow transplant patient with chronic NoV infection. SNV calling was performed using Freebayes. Only SNVs with frequencies ≥2% and ≥5X coverage are reported. Positions with coverage <10X were excluded from the analysis. SNVs shared in common between SP1 and SP2 are shown in red.

The percentages of non-synonymous single nucleotide variants (nsSNVs) for SP1 and SP2 were 27.5% and 20.5%, respectively. There were differences in the distribution of nsSNV across genes between SP1 and SP2. In SP1, most nsSNVs occurred in p22 (4/14), VP2 (4/14) and VP1 (3/14) whereas in SP2, the majority of nsSNVs occurred in VP2 (12/25) and VP1 (9/25) (Table 4.5). The enrichment of non-synonymous mutations in VP2, the P2 domain of

VP1 and p22 was statistically significant ($P<0.05$) after controlling for gene/domain size (Table 4.5) and is consistent with an increased rate of amino acid divergence in these proteins. The amino acid residues affected by nsSNVs (summarized in Figure 4.5) were mostly scattered across VP2 (5, 15, 80, 88, 140, 149, 150, 159, 162, 169, 187, 193, 195 and 240) whereas for VP1 all were located within the P2 domain (294, 340, 341, 344, 373, 374, 377, 378, 380, 403 and 406). Of the 11 mutations occurring at the P2 domain of VP1, two (affecting amino acid residues 294 and 340) were located at epitopes reported to be targeted by antibodies that block the binding of NoV to human blood group antigens (surrogates of neutralizing antibodies) [7].

A comparison between the consensus sequences of SP1 and SP2 revealed differences in 9 positions, of which 2 were synonymous (203nt and 3851nt), 6 were non synonymous (2188nt, 6095nt, 6193nt, 6281nt, 7141nt and 7167nt; the corresponding amino acid changes are shown in Figure 4.5) and 1 was located at the 3'UTR (7509nt). Among the 9 differences between the consensus sequences of SP1 and SP2, 6 were SNVs at SP1 only (203nt, 2188nt, 3851nt, 6281nt, 7141nt and 7167nt), 1 was a SNV in SP2 only (6095nt), 1 was a SNV in SP1 and SP2 (6193nt) and 1 was not a SNV in SP1 nor SP2 (7509nt). There were a total of 10 SNVs occurring at the same nucleotide positions of SP1 and SP2 (474nt, 2303nt, 4323nt, 5580nt, 5706nt, 6193nt, 6222nt, 6803nt, 7172nt and 7252nt; shown with red color in Figure 4.4).

Table 4.5. Distribution of single-nucleotide variants in NoV by gene

ORF	Gene (other names)	Position (nt)	SP1		SP2	
			S	NS (P-value*)	S	NS (P-value*)
1	p48 (NS1-2, N-term, p37)	1-983	6	2 (0.5)	11	2 (0.8)
	NTPase (NS3)	984-2081	2	0 (1)	18	1 (1)
	p22 (p20, NS4)	2082-2618	3	4 (0.01)	7	1 (0.8)
	VPg (NS5)	2619-3014	1	0 (1)	6	0 (1)
	3CLpro (NS6)	3015-3560	4	0 (1)	2	0 (1)
	RdRp (NS7)	3561-5090	11	1 (0.9)	13	0 (1)
2	VP1	5074-6690	7	3 (0.6)	30	9 (0.06
	S domain	5074-5736	2	0 (1)	9	0 (1)
	P1 subdomain	5737-5895, 6325-6690	1	0 (1)	13	0 (1)
	P2 subdomain	5896-6324	4	3 (0.03)	8	9 ($4x10^{-6}$)
3	VP2	6693-7496	3	4 (0.04)	10	12 ($1x10^{-6}$)
	TOTAL		37	14	97	25

Single nucleotide variants (SNVs) were identified using FreeBayes in two samples (SP1 and SP2) collected four months apart from an immunocompromised bone marrow transplant patient with chronic NoV infection. Only those SNVs found at frequencies ≥2% and with 5X coverage are reported. Positions with coverage <10X were excluded from the analysis.

* P-values were calculated using binomial distribution and indicate the probability of observing the corresponding number of non-synonymous single nucleotide variants in the specified region after controlling for gene/domain size.

Gene	p48			NTPase	p22					RdRp
Amino Acid position	48	161	184	28	36	54	85	150	158	453
SP1 Major Variant	P	F	S	I	R	I	S	N	D	T
SP1 Minor Variant (frequency)	S (2.4%)	L (5.8%)	-	-	K (4.8%)	T (2.2%)	N (2.3%)	D (21.3%)	-	A (6.5%)
SP2 Major Variant	P	F	S	I	K	I	S	N	D	T
SP2 Minor Variant (frequency)	-	L (4.7%)	A (26.4%)	V (11.1%)	-	-	-	-	E (24.2%)	-

Gene	VP1 (P2 domain)										
Amino Acid position	294	340	341	344	373	374	377	378	380	403	406
SP1 Major Variant	A	Q	G	T	N	D	P	H	S	A	C
SP1 Minor Variant (frequency)	-	-	-	-	-	N (25.7%)	-	-	-	V (37.8%)	S (3.9%)
SP2 Major Variant	A	Q	D	T	N	N	P	H	S	V	C
SP2 Minor Variant (frequency)	T (20.9%)	L (11.5%)	G (13.5%)	A (3.3%)	S (3.6%)	D (48.1%)	S (5.5%)	R (5.4%)	G (6.1%)	-	-

Gene	VP2													
AA position	5	15	80	88	140	149	150	159	162	169	187	193	195	240
SP1 Major Variant	L	S	K	E	L	P	I	S	V	H	N	S	A	D
SP1 Minor Variant (frequency)	-	-	-	-	-	S (10.0%)	T (10.3%)	P (37.8%)	-	-	S (21.2%)	-	-	-
SP2 Major Variant	L	S	K	E	L	P	T	P	V	H	N	S	A	D
SP2 Minor Variant (frequency)	F (22.2%)	G (6.6%)	R (17.2%)	Q (15.6%)	P (7.1%)	-	-	-	A (18.8%)	Y (34.6%)	S (27.3%)	L (28.6%)	V (13.0%) T (11.1%)	N (26.1%)

Figure 4.5. Distribution per gene of the amino acid residues affected by non-synonymous single nucleotide variants in NoV

Samples SP1 and SP2 were collected four months apart from an immunocompromised bone marrow transplant patient with chronic NoV infection. Residues that changed between SP1 and SP2 are highlighted in green.

4.4. Discussion

In this study we established a method to analyze the intra-host genetic diversity of NoV in samples from patients with acute and chronic NoV infection. Our next generation sequencing-based method coupled to bacterial rRNA depletion as a NoV RNA enrichment strategy produced between 0.01% and 1.9% of NoV reads. Characterization of the non-NoV reads confirmed that samples still contained 64 to 95% of bacterial rRNA reads after depletion which differs with a previous study reporting 1 to 5% of bacterial rRNA reads after using the same depletion method with stool RNA [32]. The reasons for the lower rRNA depletion efficiency observed with our samples could not be identified. Since the method we used for depletion works with a library of probes that must first hybridize the bacterial rRNA for subsequent removal, it is possible that the hybridization step was affected by sequence incompatibility. Also, rRNA depletion was carried out with the maximum amount of input RNA recommended by the manufacturer, which could have affected the efficiency of the method. Based on our data we estimate that if all 16S and 23S bacterial rRNA could be removed from stool RNA, NoV sequences would have represented between 0.04% and 8.5% of all NGS reads.

For all but one sample we were able to obtain sufficient sequence data to retrieve near full-length NoV genome sequences. Moreover, even the non-enriched sample (OU4) provided enough data for *de novo* assembly of a NoV genome, AlbertaEI404/2012/CA, which to our knowledge, is the first near full-length NoV GI.7 genome reported. Strains of genotype GI.7 were observed with increased prevalence in Alberta, Canada, between July 2012 and June 2013 [chapter 2] and also among children in Pakistan between April 2006 and March 2008 [54].

Our analysis included four cases of acute NoV infection. Two samples from these acute cases, OU1 and OU2, yielded sufficient coverage to identify SNVs at frequencies ≥2%. However, no NoV SNVs were detected in these two samples, indicating that the viral population in typical acute infections is homogeneous. In contrast, we identified numerous SNVs in an immunocompromised patient with chronic NoV infection. Interestingly, there was also an increase of SNV over time (51 SNVs with ≥2% frequency in the first sample and 122 SNVs with ≥2% frequency in the second sample collected 4 months after), revealing intra-host NoV evolution. As of the date of writing this manuscript there were two published studies comparing the intra-host populations of NoV between immunocompetent and immunocompromised patients by using NGS data. Both reported similar observations [18, 23]. Vega *et al.* found no SNVs throughout the viral genome (ORF1, ORF2 and ORF3) in an immunocompetent subject with acute NoV illness but identified multiple SNVs in three different immunocompromised bone marrow transplant patients (15, 67 and 235 SNVs with ≥10% frequency). In an analysis of just the ORF2 and partial ORF3 regions, Bull *et al.* detected multiple SNVs in immunocompetent individuals with acute NoV infection (5 to 8 SNVs with ≥2% frequency). However, an immunocompromised individual with chronic NoV infection displayed considerably higher NoV diversity that also increased over time (48, 59 and 109 SNVs with ≥2% frequency, in three samples collected longitudinally).

Interestingly, about half of the SNVs reported by Bull *et al.* were non synonymous (46 to 64% in immunocompetent and 34 to 48% in immunocompromised subjects) whereas we observed a rate of 21 and 28% in the immunocompromised patient. Our rates are similar to those reported in another study with bone marrow transplant patients receiving immunosuppressive drugs (236 non synonymous mutations out of a total of 1082 mutations across 13 samples, equivalent to an overall rate of 21.8%) [20]. It could be argued that Bull *et al.* achieved higher coverage and therefore higher resolution of SNVs with PCR, however, at

least two of our samples (one from an outbreak patient and another from the immunocompromised patient) had average coverage levels greater than 1300X, well above 950X, the average coverage reported by Bull *et al.* We can only speculate that the differences between our study and that of Bull *et al.* could be due to the processes of quality filtering and trimming of reads as well as the parameters used for SNV calling, which probably were more stringent in our study.

We observed an enrichment of non-synonymous SNVs in the P2 domain of VP1 which matches the expectation that the most exposed and possibly the most antigenic part of the virus bears the highest pressure to diverge. In fact, the majority of studies analyzing NoV intra-host genetic diversity have been restricted to identify potential changes at VP1 [11, 12, 23, 24, 55, 56]. By analyzing near-full length NoV genomes, we also observed an enrichment of non-synonymous mutations in VP2, which suggests that during chronic infection this gene can also be under pressure to diverge. Although our observations are based on samples from a single immunocompromised patient, we believe this is plausible because it agrees with previous analysis of sequence alignments of multiple GII.4 strains showing high evolutionary rates for VP2 [27, 57]. The role of VP2 in the viral life cycle is still unknown. Only a few copies of VP2 molecules are present in the final virion (the precise number is yet unclear). VP2 appears to bind the interior surface of the capsid and has been shown to enhance the expression and half-life of VP1 [58]. Since VP2 is rich in basic residues and therefore is positively charged, it has also been suggested that it interacts with the negatively charged genomic RNA, and possibly plays a role during the encapsidation process [58]. The VP2 protein of murine noroviruses regulates the maturation of antigen presenting cells and is an important determinant of NoV protective immunity [59]. It is interesting to consider that human NoV VP2 might be an important epitope of host immune responses.

By analyzing whole NoV genome sequences using traditional cloning, Chan *et al.*, found accumulation of mutations at VP1 and VP2 (38 and 15 nt resulting in 9 and 5 amino acid changes, respectively) over a period of 4 months in an immunocompromised patient with agammaglobulinemia and thymoma [22]. Conversely, Vega *et al.* reported that, in comparison to NoV strains in circulation, NoV mutations in immunocompromised bone marrow transplant patients occur mostly at random positions except for gene p22, which presents a significantly large proportion of mutations and positively selected sites [18]. We speculate that the difference in results observed in these studies and our study might be due to temporal changes in the strength and specificity of host immune responses during chronic infection. This will be dependent on the type of immunosuppressed host, the type of exogenous immunosuppression administered, the time of infection during the patient's clinical course and the time of the patient's immune reconstitution. Regarding this point, we observed that the distribution of non synonymous mutations across genes changed with time which suggests that NoV might face changes in its evolutionary trajectory during chronic infection as reported for other viruses such as HCV and HIV [60-62].

Besides whole genome sequencing, NGS methods are powerful tools for studying viral population dynamics within a host because of the high sequencing depth that can be achieved. In addition, samples can be prepared relatively quickly without the need of cloning into vectors and multiple isolates can be sequenced in parallel (in our case, six samples were analyzed in one MiSeq 2x121bp run). The number of samples that can be sequenced depends on the throughput of the sequencing technology and the abundance of the viral RNA in the sample. Different forms of enrichment can enhance the later. We demonstrated that bacterial rRNA depletion is a beneficial treatment that can be further optimized. Hybrid capture is a promising new enrichment strategy that remains to be examined in the future [63].

Finally, our study is the first to look at the intra-host diversity of NoV by analyzing near full length-genomes from acute cases of NoV infection and in longitudinally collected samples from an immunocompromised patient with chronic NoV infection. We identified a larger viral diversity in the immunocompromised patient, which increased over time and we also observed that genes VP2, p22 and VP1 accumulated the majority of non-synonymous mutations during chronic infection. Further studies are needed to observe if the accumulation of mutations at VP2 is consistent across immunocompromised patients with chronic infection and unveil the role of VP2 and p22 proteins in relationship to host immune responses. We are currently planning a larger follow-up study to assess the intra–host genetic diversity of NoV in solid organ transplant recipients.

4.5. Acknowledgements

The authors thank Dr. David Evans for his valuable suggestions on next generation sequencing analysis and Min Cao for her technical support.

4.6. References (chapter 4)

1. Ahmed, S.M., et al., *Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis.* The Lancet infectious diseases, 2014. **14**(8): p. 725-730.
2. Hardy, M.E., *Norovirus protein structure and function.* FEMS microbiology letters, 2005. **253**(1): p. 1-8.
3. Vinje, J., *Advances in laboratory methods for detection and typing of norovirus.* Journal of clinical microbiology, 2015. **53**(2): p. 373-381.
4. Siebenga, J.J., et al., *Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007.* The Journal of infectious diseases, 2009. **200**(5): p. 802-812.
5. Lindesmith, L.C., et al., *Mechanisms of GII.4 norovirus persistence in human populations.* PLoS medicine, 2008. **5**(2): p. e31.
6. Lindesmith, L.C., et al., *Emergence of a norovirus GII.4 strain correlates with changes in evolving blockade epitopes.* Journal of virology, 2013. **87**(5): p. 2803-2813.

7. Lindesmith, L.C., et al., *Immunogenetic mechanisms driving norovirus GII.4 antigenic variation.* PLoS pathogens, 2012. **8**(5): p. e1002705.
8. Eden, J.S., et al., *Recombination within the pandemic norovirus GII.4 lineage.* Journal of virology, 2013. **87**(11): p. 6270-6282.
9. Atmar, R.L., et al., *Norwalk virus shedding after experimental human infection.* Emerging infectious diseases, 2008. **14**(10): p. 1553-1557.
10. Sukhrie, F.H., et al., *Chronic shedders as reservoir for nosocomial transmission of norovirus.* Journal of clinical microbiology, 2010. **48**(11): p. 4303-4305.
11. Siebenga, J.J., et al., *High prevalence of prolonged norovirus shedding and illness among hospitalized patients: a model for in vivo molecular evolution.* The Journal of infectious diseases, 2008. **198**(7): p. 994-1001.
12. Schorn, R., et al., *Chronic norovirus infection after kidney transplantation: molecular evidence for immune-driven viral evolution.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2010. **51**(3): p. 307-314.
13. Ludwig, A., et al., *Quantitative detection of norovirus excretion in pediatric patients with cancer and prolonged gastroenteritis and shedding of norovirus.* Journal of medical virology, 2008. **80**(8): p. 1461-1467.
14. Frange, P., et al., *Prevalence and clinical impact of norovirus fecal shedding in children with inherited immune deficiencies.* The Journal of infectious diseases, 2012. **206**(8): p. 1269-1274.
15. Zintz, C., et al., *Prevalence and genetic characterization of caliciviruses among children hospitalized for acute gastroenteritis in the United States.* Infection, Genetics & Evolution, 2005. **5**(3): p. 281-290.
16. Carlsson, B., et al., *Quasispecies dynamics and molecular evolution of human norovirus capsid P region during chronic infection.* The Journal of general virology, 2009. **90**(Pt 2): p. 432-441.
17. Karst, S.M. and R.S. Baric, *What Is the Reservoir of Emergent Human Norovirus Strains?* Journal of virology, 2015. **89**(11): p. 5756-5759.
18. Vega, E., et al., *RNA populations in immunocompromised patients as reservoirs for novel norovirus variants.* Journal of virology, 2014. **88**(24): p. 14184-14196.
19. Nilsson, M., et al., *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.* Journal of virology, 2003. **77**(24): p. 13117-13124.
20. Kundu, S., et al., *Next-generation whole genome sequencing identifies the direction of norovirus transmission in linked patients.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2013. **57**(3): p. 407-414.
21. Hoffmann, D., et al., *Norovirus GII.4 and GII.7 capsid sequences undergo positive selection in chronically infected patients.* Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2012. **12**(2): p. 461-466.
22. Chan, M.C., et al., *Covariation of major and minor viral capsid proteins in norovirus genogroup II genotype 4 strains.* Journal of virology, 2012. **86**(2): p. 1227-1232.
23. Bull, R.A., et al., *Contribution of intra- and interhost dynamics to norovirus evolution.* Journal of virology, 2012. **86**(6): p. 3219-3229.

24. Miyoshi, T., et al., *Long-term viral shedding and viral genome mutation in norovirus infection.* Journal of medical virology, 2015. **87**(11): p. 1872-1880.
25. Batty, E.M., et al., *A modified RNA-Seq approach for whole genome sequencing of RNA viruses from faecal and blood samples.* PloS one, 2013. **8**(6): p. e66129.
26. Wong, T.H., et al., *Whole genome sequencing and de novo assembly identifies Sydney-like variant noroviruses and recombinants during the winter 2012/2013 outbreak in England.* Virology journal, 2013. **10**: p. 335-422X-10-335.
27. Cotten, M., et al., *Deep sequencing of norovirus genomes defines evolutionary patterns in an urban tropical setting.* Journal of virology, 2014. **88**(19): p. 11056-11069.
28. de Vries, M., et al., *Performance of VIDISCA-454 in feces-suspensions and serum.* Viruses, 2012. **4**(8): p. 1328-1334.
29. Cotten, M., et al., *Full genome virus detection in fecal samples using sensitive nucleic acid preparation, deep sequencing, and a novel iterative sequence classification algorithm.* PloS one, 2014. **9**(4): p. e93269.
30. Bavelaar, H.H., et al., *Whole genome sequencing of fecal samples as a tool for the diagnosis and genetic characterization of norovirus.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2015. **72**: p. 122-125.
31. He, S., et al., *Validation of two ribosomal RNA removal methods for microbial metatranscriptomics.* Nature methods, 2010. **7**(10): p. 807-812.
32. Giannoukos, G., et al., *Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes.* Genome biology, 2012. **13**(3): p. R23-2012-13-3-r23.
33. Pang, X.L., et al., *Influence of novel norovirus GII.4 variants on gastroenteritis outbreak dynamics in Alberta and the Northern Territories, Canada between 2000 and 2008.* PloS one, 2010. **5**(7): p. e11599.
34. Pang, X.L., J.K. Preiksaitis, and B. Lee, *Multiplex real time RT-PCR for the detection and quantitation of norovirus genogroups I and II in patients with acute gastroenteritis.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2005. **33**(2): p. 168-171.
35. Schmieder, R. and R. Edwards, *Quality control and preprocessing of metagenomic datasets.* Bioinformatics (Oxford, England), 2011. **27**(6): p. 863-864.
36. Kopylova, E., L. Noe, and H. Touzet, *SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data.* Bioinformatics (Oxford, England), 2012. **28**(24): p. 3211-3217.
37. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nature methods, 2012. **9**(4): p. 357-359.
38. Altschul, S.F., et al., *Basic local alignment search tool.* Journal of Molecular Biology, 1990. **215**(3): p. 403-410.
39. Tamura, K., et al., *MEGA6: Molecular Evolutionary Genetics Analysis version 6.0.* Molecular biology and evolution, 2013. **30**(12): p. 2725-2729.
40. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics (Oxford, England), 2010. **26**(6): p. 841-842.

41. Chhabra, P., A.M. Walimbe, and S.D. Chitambar, *Complete genome characterization of Genogroup II norovirus strains from India: Evidence of recombination in ORF2/3 overlap.* Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2010. **10**(7): p. 1101-1109.
42. Kojima, S., et al., *Genogroup-specific PCR primers for detection of Norwalk-like viruses.* Journal of virological methods, 2002. **100**(1-2): p. 107-114.
43. Walker, A.W., et al., *Dominant and diet-responsive groups of bacteria within the human colonic microbiota.* The ISME journal, 2011. **5**(2): p. 220-230.
44. Qin, J., et al., *A human gut microbial gene catalogue established by metagenomic sequencing.* Nature, 2010. **464**(7285): p. 59-65.
45. Garrison, E. and G. Marth *Haplotype-based variant detection from short-read sequencing*. 2012. 02 Nov 2015-arXiv:1207.3907v2.
46. Milne, I., et al., *Tablet--next generation sequence assembly visualization.* Bioinformatics (Oxford, England), 2010. **26**(3): p. 401-402.
47. Thys, K., et al., *Performance assessment of the Illumina massively parallel sequencing platform for deep sequencing analysis of viral minority variants.* Journal of virological methods, 2015. **221**: p. 29-38.
48. Van den Hoecke, S., et al., *Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing.* BMC genomics, 2015. **16**: p. 79-015-1284-z.
49. Youmans, B.P., et al., *Characterization of the human gut microbiome during travelers' diarrhea.* Gut Microbes, 2015. **6**(2): p. 110-9.
50. Nelson, A.M., et al., *Disruption of the human gut microbiota following Norovirus infection.* PLoS One, 2012. **7**(10): p. e48224.
51. Scanlan, P.D. and J.R. Marchesi, *Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces.* The ISME journal, 2008. **2**(12): p. 1183-1193.
52. Hamad, I., et al., *Molecular detection of eukaryotes in a single human stool sample from Senegal.* PloS one, 2012. **7**(7): p. e40888.
53. Zhang, T., et al., *RNA viral community in human feces: prevalence of plant pathogenic viruses.* PLoS biology, 2006. **4**(1): p. e3.
54. Alam, A., et al., *Genetic characterization of norovirus strains in hospitalized children from Pakistan.* Journal of medical virology, 2015.
55. Mai, H., et al., *GII.4 Sydney_2012 norovirus infection in immunocompromised patients in Beijing and its rapid evolution in vivo.* Journal of medical virology, 2015.
56. Debbink, K., et al., *Within-host evolution results in antigenically distinct GII.4 noroviruses.* Journal of virology, 2014. **88**(13): p. 7244-7255.
57. Bok, K., et al., *Evolutionary dynamics of GII.4 noroviruses over a 34-year period.* Journal of virology, 2009. **83**(22): p. 11890-11901.
58. Vongpunsawad, S., B.V. Venkataram Prasad, and M.K. Estes, *Norwalk Virus Minor Capsid Protein VP2 Associates within the VP1 Shell Domain.* Journal of virology, 2013. **87**(9): p. 4818-4825.
59. Zhu, S., et al., *Identification of immune and viral correlates of norovirus protective immunity through comparative study of intra-cluster norovirus strains.* PLoS pathogens, 2013. **9**(9): p. e1003592.

60. Shankarappa, R., et al., *Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection.* Journal of virology, 1999. **73**(12): p. 10489-10502.
61. Ramachandran, S., et al., *Temporal variations in the hepatitis C virus intrahost population during chronic infection.* Journal of virology, 2011. **85**(13): p. 6369-6380.
62. Bull, R.A., et al., *Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection.* PLoS pathogens, 2011. **7**(9): p. e1002243.
63. Mamanova, L., et al., *Target-enrichment strategies for next-generation sequencing.* Nature methods, 2010. **7**(2): p. 111-118.

Chapter V: Analysis of the antibody responses to NoV GII.4 in Alberta, Canada[§]

[§] The study presented in this chapter was led by Dr. Pang and had the collaboration of Dr. Jason Jiang at the Cincinnati Children's Hospital. No part of this chapter has been previously published.

5.1. Introduction

Norovirus (NoV) is the most common cause of gastroenteritis worldwide. Children and the elderly are the populations at higher risk of NoV disease with higher mobility and mortality rates. It has been estimated that NoV causes 677 million diarrhoeal cases per year among all ages and 213,515 deaths globally [1]. Because of its ubiquitous nature, NoV has a high economic burden with estimated annual costs of 284 million dollars in the United States (period 2001-2009) [2] and 81 million pounds in the United Kingdom (period 2008-2009) [3].

NoV GII.4 is the most predominant genotype among humans. In Alberta, NoV GII.4 strains were responsible for at least 66% of all NoV outbreaks occurring between 2008 to 2013 [chapter 2]. Periodic peaks of NoV outbreak epidemics have been observed in different regions across the world including Alberta, Canada and are often associated with the emergence of novel NoV GII.4 variants [4-6]. Hence, evasion of herd immune responses are presumed to be an important force driving NoV GII.4 evolution [7].

The protective immune responses against human NoV are still poorly understood as result of the absence of both, an efficient culture system and animal models capable to mimic human infection. Observational studies have identified a correlation between pre-existing antibodies and protection against NoV infection in children [8-10] and maternal antibodies are suspected to provide protection in infants less than 6 months old [11-13]. However, challenge studies with human volunteers have reported a lack of association in adults [14-16]. HBGA receptor blocking antibodies (surrogates of neutralizing antibodies) are the best characterized correlates of protection against NoV infection [16-18]. Protection after infection appears to be short-lived as challenge studies have demonstrated that re-infection

by the same NoV strain can occur within 2-6 months [14, 15] but mathematical models and outbreak data suggests that immunity could last 4 to 8 years [19].

NoV GII.4 evolves by antigenic drift through a process similar to influenza's epochal evolution where antigenically novel viruses escape the responses build against previous circulated variants. Nevertheless, recent studies have shown that HBGA receptor blocking (protective) antibodies are highly cross-reactive against GII.4 variants [20, 21]. Therefore, short term or incomplete herd immunity might also be an important factor affecting the epidemic activity of NoV. In this study we examined temporal fluctuations of anti-NoV GII.4 antibody responses in two populations of Alberta to identify possible associations with outbreak activity levels.

5.2. Materials and methods

5.2.1. Sample population

The Health Research Ethics Board of the University of Alberta approved the use of samples for this study. Sera samples used in this study were collected by the Provincial Laboratory from two populations: i) pregnant women (PW) receiving testing as part of the Prenatal Screening program in Alberta (the prenatal screening program in Alberta recommends screening of up to five serological markers: HIV, Syphilis, Hepatitis B, rubella and varicella immunity screen, during first trimester of pregnancy as well as repeated syphilis screen at delivery [22] but the samples included in the current study were collected from pregnant women at various trimesters depending on their access to prenatal care) and ii) individuals undergoing diagnostic viral serological testing (VS) by ProvLab for multiple reasons including: diagnostic and confirmatory testing of hepatitis C and syphilis (HIV positive samples were excluded from the pool of samples used in this study), serological

testing for organ transplant, blood and body fluid exposure investigations, diagnostic for tertiary adult and tertiary pediatric hospital and, specialized serological screening (e.g., West Nile virus).

The history of gastroenteritis disease was unknown for all subjects. The samples selected for the study were collected during 2012 to 2014 and the months of February and September to represent, respectively, the peak and trough of NoV outbreak seasonality shifted 1 month forward in time to consider IgG production. A group of VS samples collected in 2008 from Sexually Transmitted Infection Clinics were included as baseline controls for antibody responses before the circulation of GII.4 Sydney 2012. In addition, pre and post infection sera or plasma from an immunocompetent subject and 7 immunocompromised transplant patients were used to assess the ability of the ELISA test to detect seroconversion after NoV infection. All sera were kept at -20°C until testing.

5.2.2. Production of NoV GII.4 capsid antigens

NoV antigens were expressed and produced in Dr. Jason Jiang's laboratory as previously described [23]. The P-domain of a GII.4 Sydney strain from Alberta was amplified by PCR using the forward primer 5'-TCAAGAACTAAACCATTCTCT-3' (+) and the reverse primer P590, 5'-GCATGCGGCCGCTTAGCAAAAGCAATCGCCACGGCAATCGCATAATGCACGTCTGCGCCCCG C-3' (-), which includes the coding sequence of a CDCRGDCFC motif at the C-terminus to enhance p-particle formation. The resulting cDNA was cloned into the expression vector pGEX-4T-1 (Amersham Biosciences) using the BamHI and NotI restriccion sites. The P protein was expressed in BL21 *E. coli* cells induced with 0.4mM IPTG at room temperature and overnight. The P-protein GST fusion protein was purified using glutathione Sepharose 4 Fast Flow (Amersham Biosciences) and the GST tag was removed by thrombin cleavage. The

presence of P-dimers was confirmed by size exclusion chromatography and by HBGA binding assays using a panel of saliva from A, B O and non-secretor individuals [24] (Appendix D).

5.2.3. ELISA test for anti-NoV IgG

The antibody titers of anti-NoV IgG were measured by ELISA as previously described [25] with modifications. Ninety-six-well plates (Immulon 2HB, Thermo Scientific) were coated at 4°C overnight with a solution of NoV P-particles in PBS at a concentration 0.2μg/mL (100μL per well). Plates were blocked with a solution of 5% nonfat dry milk in PBS (200μL), incubated at 37°C for 1h and washed five times with 0.05% Tween 20 in PBS (PBS-T). Two-fold serial dilutions of sera (from 1:100 to 1:12,800) in 2% nonfat dry milk-PBS were added to each well and plates were incubated for 1h at 37°C. The plates were washed five times with PBS-T before adding 100μL of a 1:3000 solution of HRP-conjugated rabbit IgG anti-human IgG (MP Biomedicals) in 2% nonfat dry milk-PBS. After 1h incubation at 37°C and washing the plates five times with PBS-T, 100μL of TMB substrate (BD Biosciences) were added. The reaction was stopped by adding 100μL of 1M phosphoric acid after 10 min incubation at room temperature. An absorbance at 450nm with a value equal or greater than 0.2 was used as a cut off to indicate a positive reaction. The titer was reported as the reciprocal of the highest dilution with an optical density equal or greater than 0.2. A negative and a positive control were included in each plate for quality control purposes. The negative control was selected from the available pool of samples and showed low reactivity (OD_{450}<0.2) at the lowest dilution tested (1:100).

5.2.4. Statistical analysis

Statistical analyses were performed using SPSS 23. All P values reported are 2- sided. Data from pregnant women (PW) and data from individuals from virology serology diagnostic

(VS) were analyzed independently. Differences in seroprevalence between all time points were analyzed using the likelihood ratio chi-square test. Differences in distribution of anti-NoV IgG titers among months were analyzed using non-parametric Kruskal-Wallis test because logarithmic transformation of anti-NoV IgG titers failed to produced normal distributed data (Kolmogorov-Smirnov and Shapiro-Wilk tests of normality had p-values of 0.000 for PW and VS). Moreover, log-transformed titers from VS had equal variance (Levene Statistic=1.434, df1=5, df2=792, P=0.210), but PW transformed data did not (Levene Statistic=4.595, df1=5, df2=1791, P=0.000).

5.3. Results

5.3.1. Validation of the ELISA anti-NoV IgG test

In order to assess the ability of our ELISA test to detect increases in anti-NoV GII.4 IgG titers after NoV infection, pre and post-infection sera from an immunocompetent subject and 7 immunocompromised patients were analyzed. The results from this test is shown on Figure 5.1. Overall, all individuals showed between 1.4 to 32-fold increase in titers (Figure 5.1). Patient 7, an immunocompromised subject who suffered infection by GII.4 Sydney 2012, the same variant of the antigen used for the ELISA tests, had the largest increase (32-fold change) in titers. In contrast, the immunocompetent subject, who suffered infection by NoV GII.3, presented a 4-fold change in titers.

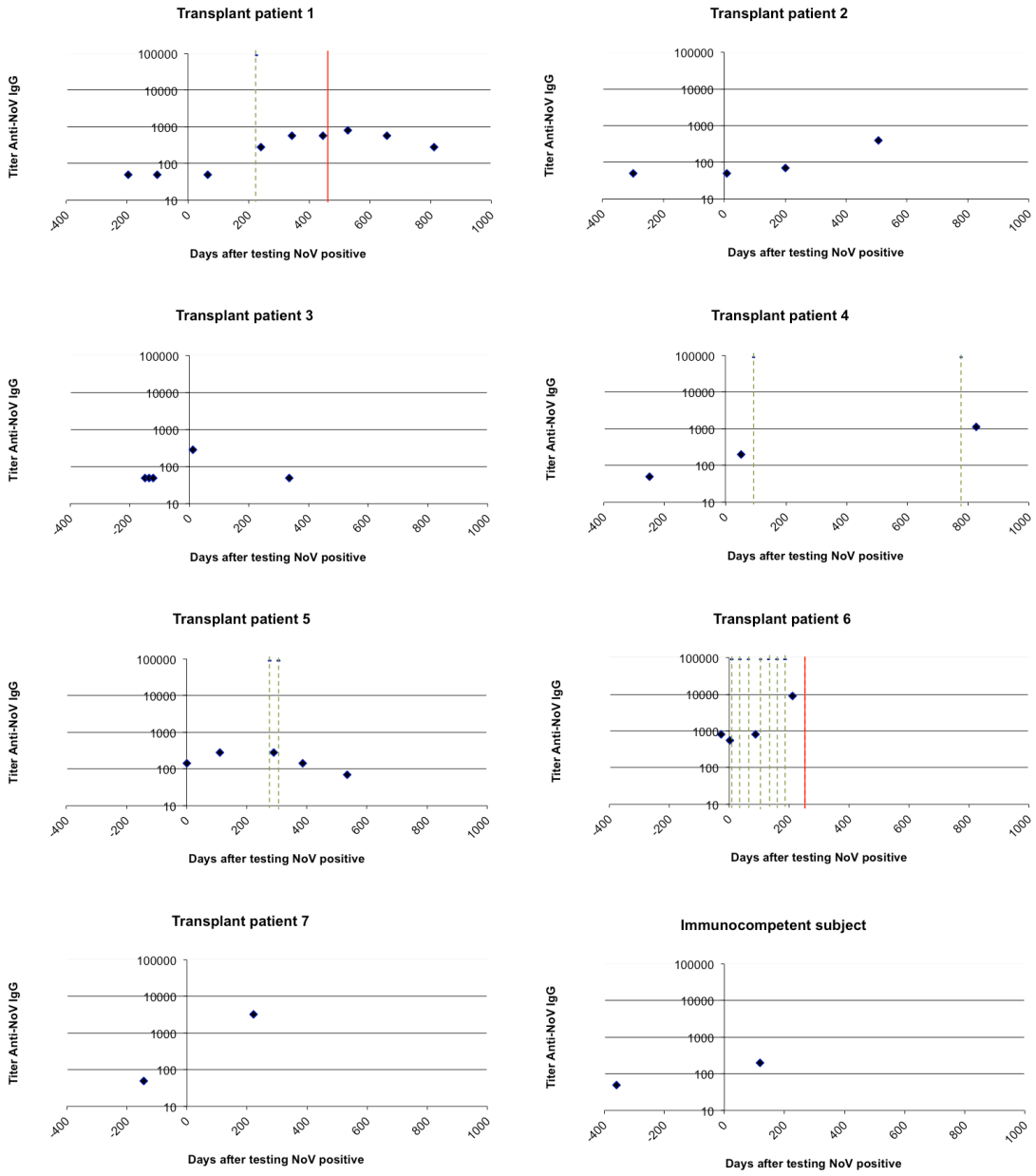


Figure 5.1. Detection of seroconversion in individuals suffering natural infection by Norovirus.

Each point in the plots represents the geometric mean titer of two tests performed in separate days. Patients 1 and 6 suffered a second infection by NoV (shown in the corresponding graphs as a solid red line). Patients 1, 4, 5, and 6 received IVIG (relative times are shown as broken green lines); patient 2 and 3 did not received IVIG during the time of sample collection; the history of IVIG (intravenous immunoglobulin) administration to patient 7 is unknown.

5.3.2. Seroprevalence of Norovirus

Sample sizes and collection times are summarized in Table 5.1. A total of 2,595 samples including 1,797 from PW and 798 from VS subjects were collected from 2008 to 2014. The overall seroprevalence of NoV GII.4 was 84.2% among pregnant women and 76.7% among VS subjects. Seroprevalence was not statistically significantly different between all time points in the two group populations (Chi-Square Likelihood Ratio=9.756, df=5, P=0.082 for pregnant women and Chi-Square Likelihood Ratio=10.128, df=5, P=0.072 for VS subjects) (Figure 5.2).

Table 5.1. Samples included in this study

Month, Year	Pregnant women (PW)	Virology serology individuals (VS)
February, 2008	N.A.	130
September, 2008	N.A.	139
February, 2012	299	N.A.
September, 2012	300	160
February, 2013	300	N.A.
September, 2013	300	131
February, 2014	299	108
September, 2014	299	130
Total	1797	798

N.A.=not available for testing

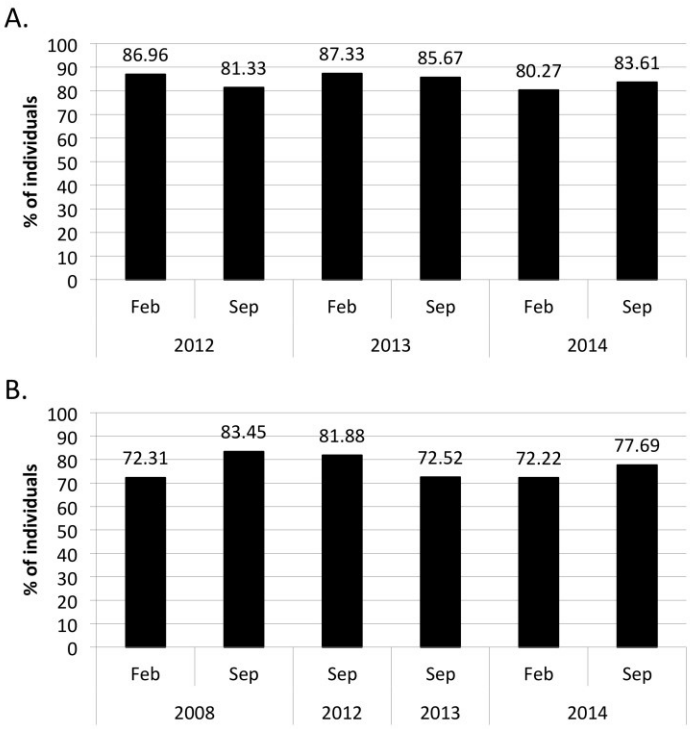


Figure 5.2: Seroprevalence of NoV in two populations in Alberta at different time points

The percentage of individuals with anti-NoV GII.4-IgG among pregnant women (A) and individuals whose samples were collected for viral serological testing of other viruses (B).

5.3.3. Analysis of anti-NoV IgG titers in time

The distribution of anti-NoV IgG titers in each time point were analyzed using reverse cumulative distribution (RCD) plots, which provide a visual assessment of possible changes including spread (variance), skewness and kurtosis in data that does not fit well with a normal distribution [26]. The curves showed a similar shape indicating similar spread and skewness of titers between time points (Figure 5.3). PW subjects in February 2013 had a slightly higher median antibody response compared to those on earlier and later time points (Table 5.2) which was also noted in the corresponding RCD plot (Figure 5.3A) as the curve situated furthest to the right. However, differences in anti-NoV titer distribution were not

statistically significant at alpha 0.05 (Kruskal-Wallis, P=0.070). The RCD plot for VS individuals (Figure 5.3B) showed no distinctive distribution for February 2008 or September 2008, the control groups for baseline antibody responses. Also, samples from February 2013 were not available for comparison as for the PW population. VS subjects in February 2008 had the lowest median antibody responses compared to those from other months, but differences were not statistically significant at alpha = 0.05 (Kruskal-Wallis, P=0.077).

Table 5.2. Anti-NoV GII.4 IgG GMT titers in PW and VS populations.

Month, year	Pregnant Women GMT (95% CI)[a]	Virology Serology GMT (95% CI)[b]
February, 2008	-	146.8 (124.8-172.7)
September, 2008	-	189.3 (160.3-223.7)
February, 2012	226.2 (199.2-256.8)	-
September, 2012	194.1 (173.7-216.8)	205.3 (173.7-242.7)
February, 2013	263.9 (231.2-301.2)	-
September, 2013	204.7 (183.4-228.5)	166.2 (139.4-198.1)
February, 2014	210.5 (185.0-239.4)	167.1 (135.0-206.9)
September, 2014	219.9 (195.6-247.3)	156.5 (133.5-183.5)
TOTAL	218.8 (208.3-229.9)	172.2 (160.5-184.8)

a. Differences were not statistically significant at alpha =0.05 (Kruskal-Wallis, P=0.070)

b. Differences were not statistically significant at alpha =0.05 (Kruskal-Wallis, P=0.077)

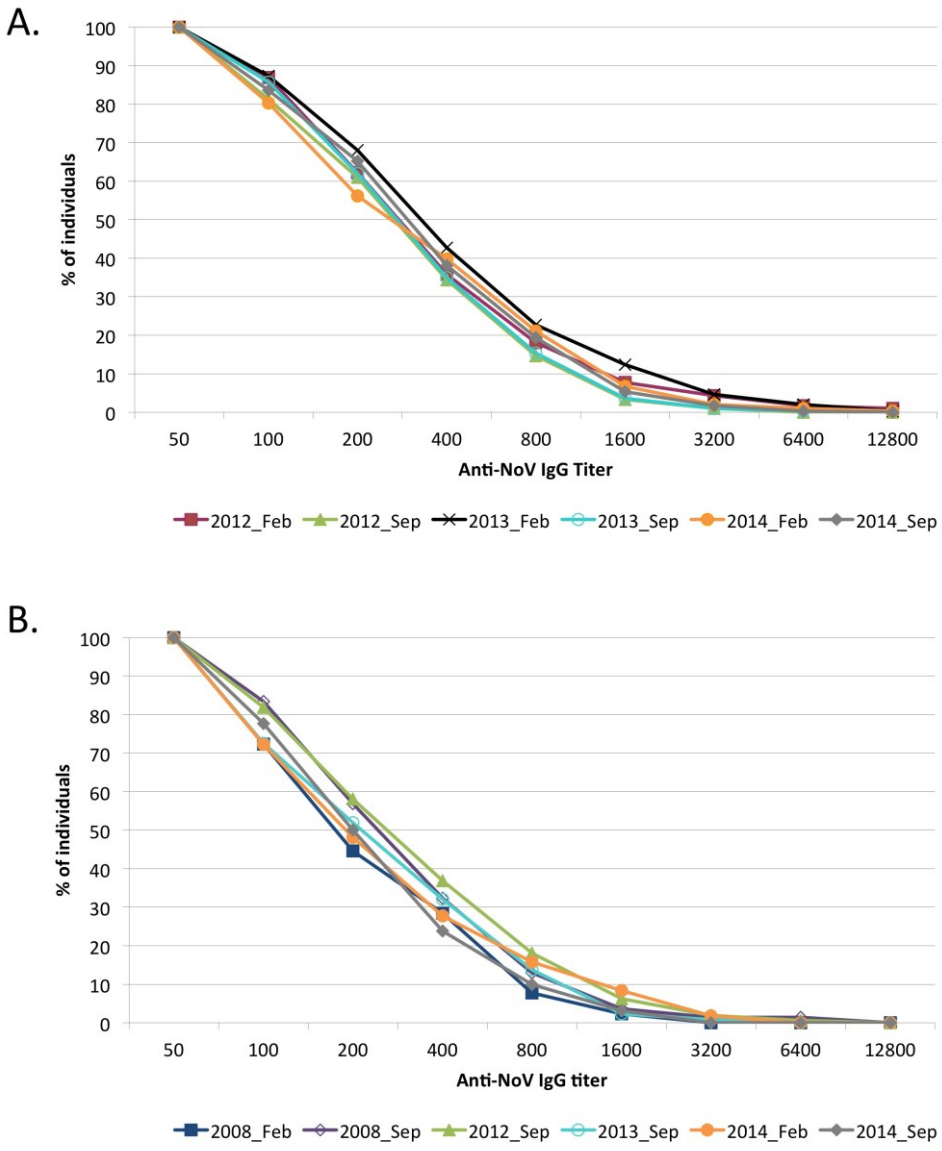


Figure 5.3. Reverse cumulative distribution plots of anti-NoV IgG titers in two adult populations in Alberta

The plots were constructed as described by Reed *et al.* [26] to provide a visual assessment of the distribution of the data. Plot (A) was constructed using data from pregnant women and plot (B) was constructed with data from individuals receiving testing for virology serology. The y-axis represents the percentage of individuals having a titer equal or greater than that displayed on the x-axis. The median antibody response for each curve corresponds to the value at 50%.

5.4. Discussion

Recent increases in outbreak activity levels of NoV in Alberta revealed that additional factors besides the antigenic evolution of NoV GII.4 might be responsible [Chapter 2]. Here, in this study, we measured the antibody responses in two populations of Alberta and investigated possible temporal variations that could explain increased NoV outbreak activity.

Overall, 84.2% of pregnant women and 76.7% of individuals receiving serology testing for other viruses had antibodies against NoV GII.4 Sydney 2012. Similar rates of seropositivity have been reported in different regions across the world (Table 3) and reflect the large predominance of NoV GII.4 in humans. About 16% (PW) and 23% (VS) of the population were seronegative to NoV GII.4 Sydney, which is similar to the proportion of non-secretor individuals among human populations who are known to be resistant to NoV infection [27-29]. However, no additional tests were performed to confirm if the subjects identified as seronegatives were also non-secretors.

The seroprevalence in pregnant women was higher than that reported in a previous study involving pregnant women also (84.2% vs 56.3%) [30]. In this study, pregnant women also had higher NoV seroprevalence and geometric mean titers than VS subjects in September 2013, February 2014 and September 2014. A possible explanation of the higher prevalence of antibody responses in pregnant women compared to VS individuals could be that pregnant women are more likely to be in contact with young children, which are known to have a high incidence of NoV infection. The present study provides novel information regarding the level of anti-NoV responses during pregnancy, a condition known to alter immune responses and render women at higher risk for severe illness by pathogens including influenza, HEV, HSV, varicella and malaria [31, 32]. During the first trimester of pregnancy, there is a change to dominant antibody-mediated (T-helper type II or Th2) immune responses that suppresses

cellular mediated (T-helper type I or Th1) immune responses and allows tolerance of foetal antigens [31, 32], whereas in the second and third trimester, both, Th1 and Th2 responses decrease while elements of innate immunity increase [31]. Maternal antibodies have an important impact in NoV epidemics among children; there is substantial evidence suggestive of a protective role: the seroprevalence in infants less than 11 months old is higher than for children 12-23 months old [10, 11, 33-35] and the incidence of NoV infection is also lower compared to that of children aged 12-23 months old [10, 36].

Table 5.3. Seroprevalence of NoV GII.4 reported by previous studies.

Country	Year of sample collection	Number of subjects tested	Age range (y)	Seroprevalence	Antigen	Reference
India	2007-2013	191	0- 5	62.2%	Den Haag 2006b	[33]
Spain	2006-2009	71	16-40*	56.3%	Den Haag 2006b	[30]
Spain	2008-2011	434	0-86	98.6%	Den Haag 2006b	[13]
China	2009-2012	362	0-5	86.2%	Den Haag 2006b	[34]
Korea	2005-2006	346	0-70	94.5%	Farmington Hills 2002	[35]
India	1999-2000	1,044	0-60	99.5%	Farmington Hills 2002	[37]
United Kingdom	2000-2005	1,034	0-60	99.4%	Farmington Hills 2002	[37]
Finland	2006-2008	492	0-14	63.4%	US95_96	[10]
Japan	2006-2007	400	1-62	62.5%	US95_96	[38]
Netherlands	1998-2001	630	20->60	94.0%	Bristol 1993	[39]
France	2000-2001	1,078	0-96	74.1%	US 95_96	[40]
Italy	1996	1,729	0-95	93.3-96.3%	Bristol 1993	[41]

*This study was performed with samples from pregnant women.

One limitation of this study is that the immune responses observed in individuals under virology serology diagnostic testing might not be representative of the general population of Alberta, as some of these individuals were likely suffering infection by another virus, which could have limited NoV infection. However, seroprevalence and geometric mean titers were consistent in time. In that regards, sera collected from these subjects during 2008, several years prior to the emergence of NoV GII.4 Sydney in 2012, showed high base line antibody levels against this strain, supporting previous reports describing high cross reactivity of immune responses to NoV GII.4 variants [21, 42].

Seroprevalence and geometric mean titers of antibodies against NoV GII.4 Sydney were not significantly different in time. However, the results from these analyses should be interpreted with caution since our sample size was not large enough to minimize the likelihood of type II errors. The scarcity of data on local NoV seroprevalence and GMT did not allow us to estimate a sample size with enough power during the planning stages of the study. In addition, log-transformed titers deviated from a normal distribution forcing the use of a less powerful non-parametric test for data analysis. We observed the largest difference in seroprevalence and GMT in prenatal women, between September 2012 and February 2013, which coincides with the emergence of Sydney 2012 in the province (Figure 5.4). However no samples were available to verify a similar change in the VS population. NoV GII.4 have been circulating among humans at least since 1980s [43], thus it is possible that the level of immune responses against NoV GII.4 has reached a plateau in the population that complicated the measurement of significant changes. Concordantly with this hypothesis, the number of GII.4 outbreaks in Alberta has shown a step-wise decrease during the three consecutive years when sera samples were collected (Figure 5.4).

Lastly, an important limitation of the study is that anti-NoV IgG levels have not been correlated with protection against NoV infection in challenge studies [14-16]. The association has not been identified despite that anti-NoV IgG is a marker of previous exposure and the association between multiple NoV exposures and protection against infection [44, 45]. However, in a recent study, temporal variations in anti-NoV IgG responses were successfully identified across cohorts of children and NoV genotypes after controlling for cross-reactive responses [46].

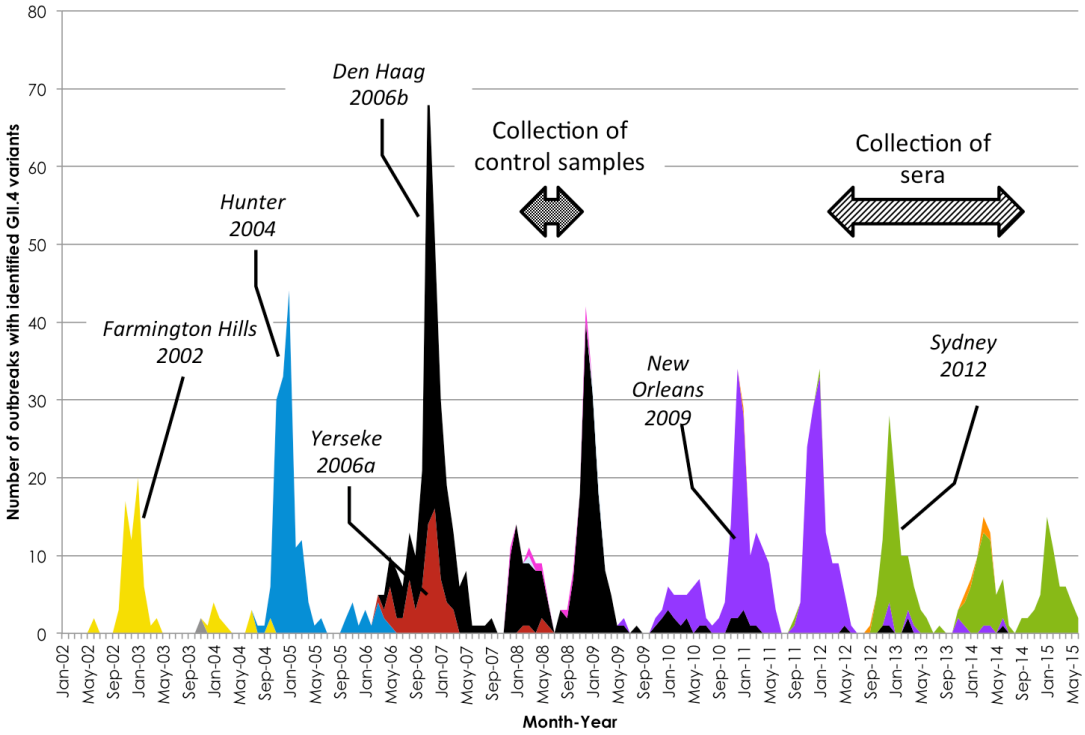


Figure 5.4. Time of sera collection in relationship to NoV GII.4 outbreak levels in the province

In summary, this study is the first to describe the seroprevalence of NoV among pregnant women in Canada and also provides preliminary data on NoV antibody responses, which could be useful for studies analyzing immune responses at the population level. Future studies should be aimed at examining correlates of protective antibody responses (HBGA-receptor blocking antibodies) in populations at highest risk of NoV illness, such as children, elderly and immunocompromised subjects.

5.5. References (chapter 5)

1. Pires, S.M., et al., *Aetiology-Specific Estimates of the Global and Regional Incidence and Mortality of Diarrhoeal Diseases Commonly Transmitted through Food.* PLoS One, 2015. **10**(12): p. e0142927.
2. Gastanaduy, P.A., et al., *Burden of norovirus gastroenteritis in the ambulatory setting--United States, 2001-2009.* J Infect Dis, 2013. **207**(7): p. 1058-65.
3. Tam, C.C. and S.J. O'Brien, *Economic Cost of Campylobacter, Norovirus and Rotavirus Disease in the United Kingdom.* PLoS One, 2016. **11**(2): p. e0138526.
4. Siebenga, J.J., et al., *Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007.* The Journal of infectious diseases, 2009. **200**(5): p. 802-812.
5. Pang, X.L., et al., *Influence of novel norovirus GII.4 variants on gastroenteritis outbreak dynamics in Alberta and the Northern Territories, Canada between 2000 and 2008.* PloS one, 2010. **5**(7): p. e11599.
6. Hasing, M.E., et al., *Emergence of a new norovirus GII.4 variant and changes in the historical biennial pattern of norovirus outbreak activity in Alberta, Canada, from 2008 to 2013.* Journal of clinical microbiology, 2013. **51**(7): p. 2204-2211.
7. Lindesmith, L.C., et al., *Immunogenetic mechanisms driving norovirus GII.4 antigenic variation.* PLoS pathogens, 2012. **8**(5): p. e1002705.
8. Ryder, R.W., et al., *Evidence of immunity induced by naturally acquired rotavirus and Norwalk virus infection on two remote Panamanian islands.* The Journal of infectious diseases, 1985. **151**(1): p. 99-105.
9. Lew, J.F., et al., *Detection of Norwalk virus or Norwalk-like virus infections in Finnish infants and young children.* The Journal of infectious diseases, 1994. **169**(6): p. 1364-1367.
10. Nurminen, K., et al., *Prevalence of norovirus GII-4 antibodies in Finnish children.* Journal of medical virology, 2011. **83**(3): p. 525-531.
11. Gray, J.J., et al., *Prevalence of antibodies to Norwalk virus in England: detection by enzyme-linked immunosorbent assay using baculovirus-expressed Norwalk virus capsid antigen.* Journal of clinical microbiology, 1993. **31**(4): p. 1022-1025.
12. Numata, K., et al., *Epidemiological study of Norwalk virus infections in Japan and Southeast Asia by enzyme-linked immunosorbent assays with Norwalk virus capsid protein produced by the baculovirus expression system.* Journal of clinical microbiology, 1994. **32**(1): p. 121-126.
13. Carmona-Vicente, N., et al., *Norovirus infections and seroprevalence of genotype GII.4-specific antibodies in a Spanish population.* J Med Virol, 2015. **87**(4): p. 675-82.
14. Parrino, T.A., et al., *Clinical immunity in acute gastroenteritis caused by Norwalk agent.* The New England journal of medicine, 1977. **297**(2): p. 86-89.
15. Johnson, P.C., et al., *Multiple-challenge study of host susceptibility to Norwalk gastroenteritis in US adults.* The Journal of infectious diseases, 1990. **161**(1): p. 18-21.

16. Reeck, A., et al., *Serological correlate of protection against norovirus-induced gastroenteritis.* The Journal of infectious diseases, 2010. **202**(8): p. 1212-1218.
17. Atmar, R.L., et al., *Norovirus vaccine against experimental human Norwalk Virus illness.* The New England journal of medicine, 2011. **365**(23): p. 2178-2187.
18. Bok, K., et al., *Chimpanzees as an animal model for human norovirus infection and vaccine development.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(1): p. 325-330.
19. Simmons, K., et al., *Duration of immunity to norovirus gastroenteritis.* Emerging infectious diseases, 2013. **19**(8): p. 1260-1267.
20. Blazevic, V., M. Malm, and T. Vesikari, *Induction of homologous and cross-reactive GII.4-specific blocking antibodies in children after GII.4 New Orleans norovirus infection.* J Med Virol, 2015. **87**(10): p. 1656-61.
21. Dai, Y.C., et al., *Antigenic Relatedness of Norovirus GII.4 Variants Determined by Human Challenge Sera.* PloS one, 2015. **10**(4): p. e0124945.
22. Plitt AA, O.M., Sahni V, Lee BE, Charlton C, Simmonds K., *Examination of a Prenatal Syphilis Screening Program, Alberta, Canada: 2010-2011.* Canadian Journal of Public Health. In press, 2016.
23. Tan, M. and X. Jiang, *The p domain of norovirus capsid protein forms a subviral particle that binds to histo-blood group antigen receptors.* Journal of virology, 2005. **79**(22): p. 14017-14030.
24. Yang, Y., et al., *Genetic and phenotypic characterization of GII-4 noroviruses that circulated during 1987 to 2008.* J Virol, 2010. **84**(18): p. 9595-607.
25. Frenck, R., et al., *Predicting susceptibility to norovirus GII.4 by use of a challenge model involving humans.* The Journal of infectious diseases, 2012. **206**(9): p. 1386-1393.
26. Reed, G.F., B.D. Meade, and M.C. Steinhoff, *The reverse cumulative distribution plot: a graphic method for exploratory analysis of antibody data.* Pediatrics, 1995. **96**(3 Pt 2): p. 600-3.
27. Thorven, M., et al., *A homozygous nonsense mutation (428G-->A) in the human secretor (FUT2) gene provides resistance to symptomatic norovirus (GGII) infections.* Journal of virology, 2005. **79**(24): p. 15351-15355.
28. Kindberg, E., et al., *A nonsense mutation (428G-->A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection.* AIDS, 2006. **20**(5): p. 685-9.
29. Bucardo, F., et al., *Genetic susceptibility to symptomatic norovirus infection in Nicaragua.* J Med Virol, 2009. **81**(4): p. 728-35.
30. Khodayar-Pardo, P., et al., *Norovirus GII.4 antibodies in breast milk and serum samples: their role preventing virus-like particles binding to their receptors.* Pediatr Infect Dis J, 2014. **33**(6): p. 554-9.
31. Kourtis, A.P., J.S. Read, and D.J. Jamieson, *Pregnancy and infection.* N Engl J Med, 2014. **370**(23): p. 2211-8.
32. Jamieson, D.J., R.N. Theiler, and S.A. Rasmussen, *Emerging infections and pregnancy.* Emerging Infectious Diseases, 2006. **12**(11): p. 1638-1643.
33. Kulkarni, R., K. Lole, and S.D. Chitambar, *Seroprevalence of antibodies against GII.4 norovirus among children in Pune, India.* J Med Virol, 2016.

34. Liu, X., et al., *Seroepidemiology of Norovirus GII.3 and GII.4 Infections in Children with Diarrhea in Xi'an, China.* Foodborne Pathog Dis, 2015. **12**(6): p. 500-5.
35. Son, H., et al., *Seroepidemiology of predominant norovirus strains circulating in Korea by using recombinant virus-like particle antigens.* Foodborne pathogens and disease, 2013. **10**(5): p. 461-466.
36. Pang, X.L., J.K. Preiksaitis, and B.E. Lee, *Enhanced enteric virus detection in sporadic gastroenteritis using a multi-target real-time PCR panel: A one-year study.* Journal of medical virology, 2014. **86**(9): p. 1594-1601.
37. Menon, V.K., et al., *Comparison of age-stratified seroprevalence of antibodies against norovirus GII in India and the United Kingdom.* PLoS One, 2013. **8**(2): p. e56239.
38. Kobayashi, S., et al., *Seroepidemiological study of norovirus infection in Aichi Prefecture, Japan.* Microbiology and immunology, 2009. **53**(6): p. 356-359.
39. Widdowson, M.A., et al., *Detection of serum antibodies to bovine norovirus in veterinarians and the general population in the Netherlands.* J Med Virol, 2005. **76**(1): p. 119-28.
40. Nicollier-Jamot, B., et al., *Molecular cloning, expression, self-assembly, antigenicity, and seroepidemiology of a genogroup II norovirus isolated in France.* Journal of clinical microbiology, 2003. **41**(8): p. 3901-3904.
41. Pelosi, E., et al., *The seroepidemiology of genogroup 1 and genogroup 2 Norwalk-like viruses in Italy.* Journal of medical virology, 1999. **58**(1): p. 93-99.
42. Lindesmith, L.C., et al., *Broad blockade antibody responses in human volunteers after immunization with a multivalent norovirus VLP candidate vaccine: immunological analyses from a phase I clinical trial.* PLoS Med, 2015. **12**(3): p. e1001807.
43. Zheng, D.P., et al., *Molecular epidemiology of genogroup II-genotype 4 noroviruses in the United States between 1994 and 2006.* Journal of clinical microbiology, 2010. **48**(1): p. 168-177.
44. Saito, M., et al., *Multiple norovirus infections in a birth cohort in a Peruvian Periurban community.* Clin Infect Dis, 2014. **58**(4): p. 483-91.
45. O'Ryan, M.L., et al., *Symptomatic and asymptomatic rotavirus and norovirus infections during infancy in a Chilean birth cohort.* Pediatr Infect Dis J, 2009. **28**(10): p. 879-84.
46. van Beek, J., et al., *Comparison of norovirus genogroup I, II, and IV seroprevalence among children in the Netherlands, 1963, 1983, and 2006.* J Gen Virol, 2016.

Chapter VI: Discussion and Future Directions

6.1. Discussion

Norovirus is the leading cause of gastroenteritis among individuals of all ages, affecting largely children, the elderly and immunocompromised individuals. The first 8 years of NoV surveillance in Alberta (July 2000 to June 2008) identified a biennial pattern of NoV outbreak activity, which appeared to be driven by the emergence of new GII.4 variants in the province. The main objectives of the thesis were to continue monitoring the outbreak activity of NoV for the following 7 years (July 2008 to June 2015) and to study novel NoV GII.4 variants as well as the level of immune responses in the population as factors contributing to the dynamics of NoV outbreak activity in the province. A discussion of the main findings of this work is presented below.

GII.4 remained as the most predominant genotype causing the majority of NoV outbreaks in Alberta

During the seven-year period, GII.4 continued to be the predominant genotype in Alberta responsible for 47.6% to 80.2% of all NoV outbreaks per year [chapter 2]. However, while seven different GII.4 variants were identified in the province in the first 8 years of NoV outbreak surveillance (US95/96, Farmington Hills 2002, Hunter 2004, Yerseke 2006a, Den Haag 2006b, Osaka 2007 and Apeldoorn 2007), only two GII.4 variants emerged in the following 7 years (New Orleans in June 2009 and Sydney in January 2012), reflecting a decrease in the rate with which novel NoV GII.4 clusters appeared. Similarly as observed in historical NoV outbreak surveillance data, the novel GII.4 variants New Orleans 2009 and Sydney 2012 were responsible for peaks of increased outbreak activity in Alberta.

Since the winter of 2014-2015, several countries have reported the emergence of GII.17 strains, particularly in Asia, where GII.17 has rapidly replaced GII.4. For instance, China identified the emergence of GII.17 in October 2014, which quickly became the most predominant NoV genotype [1-3]. Likewise, Japan had an increase of GII.17 cases since December 2014, and by March 2015, GII.17 was identified as the most predominant genotype [4]. In contrast, as of June 2015, GII.17 did not become predominant in Alberta. Only three GII.17 outbreaks were identified in the province, with the first occurring in November 2014, only one month after GII.17 emerged in China. The activity observed in Alberta for GII.17 resembles that reported in Italy, Romania and United States during the same period [5-7]. The causes behind such marked differences in GII.17 prevalence between countries despite sharing similar circulation times can only be speculated; at the moment, an increased range of host genetic susceptibility does not seem to be a cause of these differences since GII.17 recognizes the same group of HBGA molecules as GII.4 [8].

The biennial pattern of NoV outbreak activity disappeared

The biennial pattern observed in the period July 2000- June 2008 was characterized by a year of high outbreak activity followed by a quiescent year. As part of the pattern, the years with high outbreak activity were associated with the emergence of a novel GII.4 variant during the months of spring or summer preceding the winter season. Based on these observations, it was initially hypothesized that the peaks of high outbreak activity were due to novel GII.4 variants capable of evading the immune responses built against previous variants [9].

The follow-up data [chapter 2], however, showed deviations in the pattern suggesting that, in addition to NoV GII.4 antigenic change, other factor(s) might also contribute to high outbreak activity. For instance, in absence of a novel GII.4 variant during the winter of 2008-

2009, variant Den Haag 2006b caused a large number of outbreaks despite being responsible for the largest peak of NoV outbreak activity so far recorded in the province two years before, during July 2006-June 2007. Variant New Orleans 2009 also presented changes in the pattern; it emerged in the province during 2009-2010 but circulated at low levels for over a year until the winter of 2010-2011 when it caused the first of two peaks of increased outbreak activity. The atypical dynamics of Den Haag 2006b and New Orleans 2009 in Alberta demonstrated that high outbreak activity can occur with strains that have already been in circulation for over a year and suggest that a protective factor(s) could have limited the spread of Den Haag 2006b during July 2007-June 2008 and New Orleans 2009 during July 2009-June 2010. For instance, widespread hygiene practices and infection control campaigns against pandemic influenza H1N1 might have reduced the transmission of GII.4 New Orleans 2009 during July 2009-June 2010. Interestingly, similar to the observation made in Alberta, the emergence of New Orleans 2009 was not strongly associated with increased outbreak activity in the United States [10] , South Korea [11] and Japan [12] .

After an apparent stabilization of NoV outbreaks levels in 2010-2011 and 2011-2012, the number of NoV outbreaks in Alberta has been in decline (Figure 2.5). The causes behind this decrease are unknown; it is tempting to consider that the levels of herd immunity against NoV GII.4 might be reaching a plateau. Congruent with this idea, recent data suggests that HBGA blockade antibody responses, which are correlates of protective immunity, are largely cross-reactive between GII.4 variants [13, 14]. Based on time-scaled phylogenetic trees, it has been estimated that the NoV GII.4 lineage might have emerged around the year 1924 [15] and archived stool and sera specimens dating back to the 1960s provided evidence that NoV GII.4 strains have been in circulation among humans at least since the 1970s [16-19]. Therefore, it is possible that over three decades of circulation and multiple pandemics of NoV GII.4, the

population might have build up sufficient herd immunity to slow down the spread of this genotype.

NoV strains circulating in Alberta were similar to those observed in global epidemics

Since the initial assessment of the genetic diversity of NoV GII.4 was based on a small portion of ORF2 and did not allow the identification of recombinant strains, a near-full length genome characterization was performed on a group of 20 NoV strains representing the most prevalent GII.4 variants that have circulated in the province [chapter 3]. Among seven different GII.4 variants that were analyzed, only Apeldoorn 2007 and Sydney 2012 were identified as recombinants, which agreed with results from previous studies [20, 21]. Overall, phylogenetic analyses indicated that the GII.4 strains from Alberta were similar to those circulating in other regions around the world.

An exception was AlbertaEI337/2011, the first strain to present a Sydney-like ORF2 in Alberta. This strain, here named Alberta 2011, appears to have evolved by antigenic drift from an ancestor of New Orleans 2009 and resembles the parental strain that most likely underwent recombination with Osaka 2007 to create the pandemic strain Sydney 2012 (Figure 3.6a). Alberta 2011 was already in circulation in September 2011, 4 months before Sydney 2012 was first identified in the province. Despite both strains having similar amino acid sequences at VP1, and thus, share close antigenic resemblance, Alberta 2011 did not become predominant as Sydney 2012 did. This difference suggests that the ORF1 sequence acquired by Sydney 2012 via recombination might have increased the fitness and provided an evolutionary advantage to this variant over Alberta 2011.

Novel GII.4 variants tend to go back to ancestral sequences during evolution

The phylogenetic analysis of near-full length genomic sequences of NoV [chapter 3] showed that novel GII.4 variants tend to branch off from ancestors (i.e. near the base) of previous clusters (see Figure 6.1). This pattern indicates that novel GII.4 variants tend to maintain some (not all) of the changes that previous clusters acquired by antigenic drift. A similar pattern has also been reported for the non-structural (NS) and hemagglutinin (HA) genes of influenza type A (H3N2) [22-26].

In addition, it was observed that recombinants of NoV GII.4 tend to scavenge old GII.4 sequences or sequences from other GII genotypes at ORF1 while maintaining more contemporary sequences at ORF2 and ORF3 (e.g. Asia 2003, Osaka 2007, Randwick 2011, Apeldoorn 2007, Sydney 2012 adopted sequences from GII.P12, GII.Pe, GII.4 Den Haag 2006b, GII.4 Hunter 2004 and GII.4 Osaka 2007, respectively) (Table 3.3). The gain in fitness from such preference is unknown and could be further analyzed in future studies. As a ssRNA virus, NoV evolves at a relatively fast pace, with an evolutionary rate estimated as 3.8 x 10^{-3} substitutions per site per year for GII strains [15] (for comparison, this rate is a million times faster than that observed in nuclear genes of mammals [27]). The described patterns of VP1 evolution and recombination could reflect mechanisms used by NoV to circumvent the irreversible accumulation of mutations that decrease viral fitness, an effect referred as 'Muller's ratchet', that has been previously reported for a RNA virus [28].

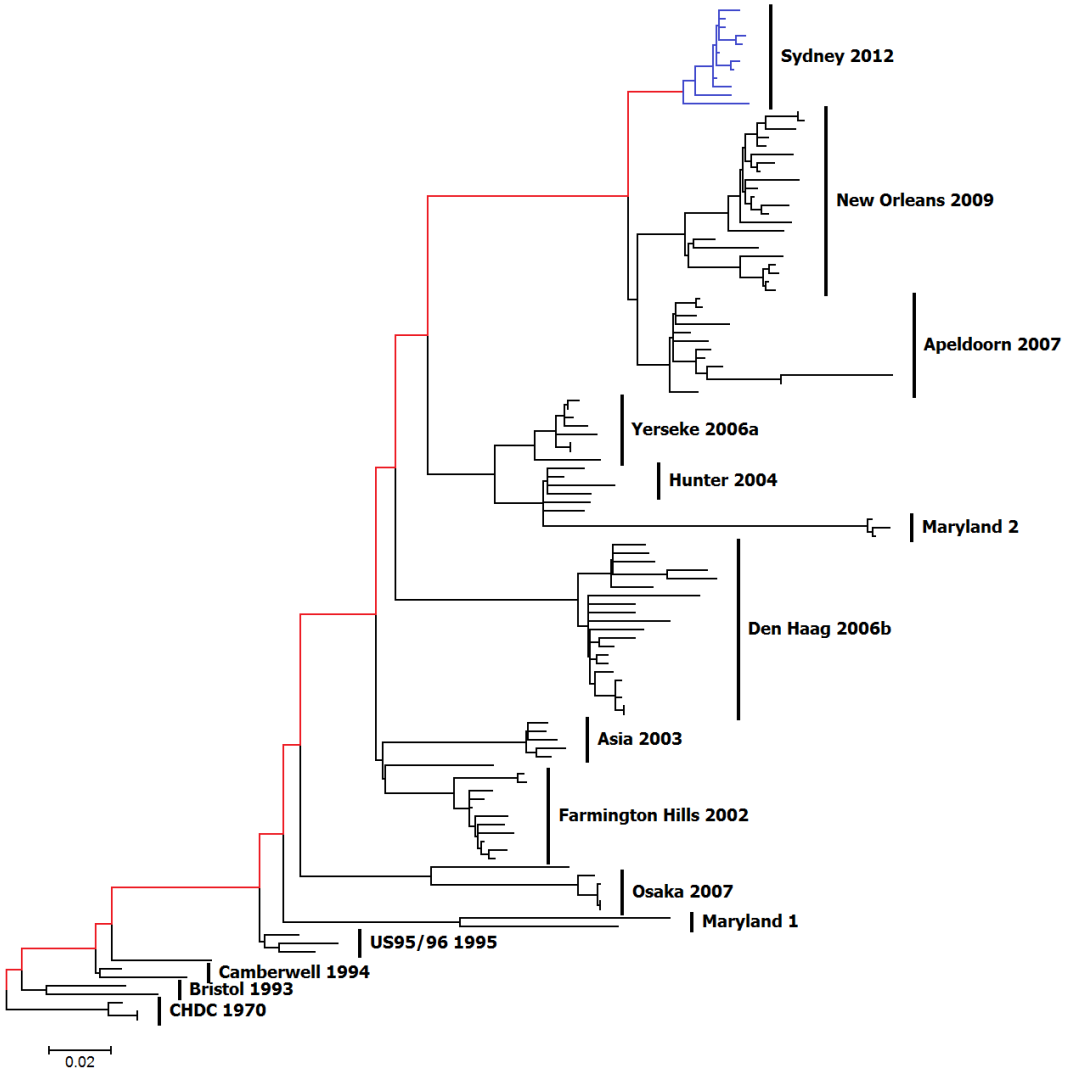


Figure 6.1. Trend in the evolution of VP1 (ORF2b) in Norovirus GII.4

The sequences, genomic region and parameters used to construct this Maximum-likelihood tree are the same as those described in Figure 3.2b. Clusters that are extinct or currently undergoing extinction are shown in black. Red lines represent evolutionary changes that have occurred in the lineage leading to the latest GII.4 variant in circulation. The latest GII.4 cluster currently in circulation is shown in blue.

Norovirus sequencing via NGS presents advantages over Sanger sequencing but is not yet cost efficient for outbreak surveillance purposes

In chapter 3, near full-length NoV genomic sequences were obtained using Sanger's method. The procedure involved running 9 separate PCR reactions and had the drawback that, due to the extensive genetic diversity of NoV, it required the use of GII.4-specific primers. As part of this thesis, a new NGS-based method was developed to obtain near-full length sequences without the use of RT-PCR amplification prior to construction of the sequencing library [chapter 4]. NoV sequences can represent as low as 0.01% of all stool RNA [29], hence, bacterial rRNA depletion was used as an alternate strategy for the enrichment of NoV sequences. A total of six samples were sequenced in a single MiSeq run. The enrichment process was not as efficient as previously reported [30]; 64% to 95% of the sequences were bacterial rRNA. Still, enough data was obtained to retrieve near full-length NoV genomic sequences for 5 of the 6 samples and the results were comparable to those obtained via Sanger sequencing.

The major advantages of the NGS-based method over Sanger sequencing are: i) it allows *de novo* near-full length sequencing of NoV strains of any genogroup or genotype and ii) it provides data free of primer bias that can be used to analyze the intra-host genetic diversity of NoV. However, there are also significant disadvantages associated to the method: i) it requires longer hands-on time for sample preparation, ii) data analysis requires the use of special software which could increase costs and turnaround times of results, iii) unless an efficient method of viral enrichment is used, the cost of sequencing is highly restrictive for outbreak surveillance purposes. Even though NGS could be used to sequence PCR amplicons to increase the efficiency and reduce costs, hands-on times would still represent a major drawback without automating library preparation.

The intra-host population of NoV is highly homogeneous during acute infection.

The NGS-based method produced enough sequence data to assess and compare the intra-host genetic diversity of NoV in two samples from outbreak patients suffering acute NoV infection and two samples collected longitudinally from an immunocompromised patient with chronic NoV infection. Both immunocompetent patients with acute infection presented a highly homogeneous NoV population (no single nucleotide variants were detected), conversely, the immunocompromised patient with chronic infection showed a large genetic diversity of NoV, which also increased in time. The poor accumulation of mutations by NoV during acute infection supports the hypothesis that NoV most likely uses special reservoirs such as immunocompromised subjects to evolve into novel strains [31].

VP2, the P2 domain of VP1 and p22 accumulated a significant number of non-synonymous mutations during chronic infection.

A significant presence of non-synonymous single nucleotide variants (nsSNVs) was identified in VP2, the P2 subdomain of VP1 and p22 during chronic NoV infection (Table 4.5). While the accumulation of nsSNV at P2 domain of VP1 is expected because this domain forms the outermost part of the capsid and is presumed to be under positive selection, the increased presence of nsSNVs at VP2 and p22 cannot be readily explained since the function of these proteins is yet unclear. VP2 might participate during NoV packaging [32, 33] and it has been demonstrated that MNV VP2 can modulate protective immune responses as well as the presentation of antigens by macrophages and in mice [34, 35]. On the other hand, p22 seems to participate in the viral cycle by highjacking the secretory pathways of the host cell and recruiting membranes into viral replication factories [34, 36, 37]. The findings from this study, although preliminary, suggest that these two proteins could be crucial for viral

adaptation during chronic infection. Interestingly, while the accumulation of nsSNV at the P2 domain and VP2 was observed at 6 and 12 months of chronic infection, p22 presented a large number of nSNVs at 6 months infection but not after 12 months, reflecting changes in evolutionary pressures during the course of chronic NoV infection. Whether this pattern of accumulation of nsSNVs at VP2 and p22 is consistent across immunocompromised patients is yet unknown and could be studied in the future to help clarify possible functions of these two proteins.

The majority of the adult population in Alberta has been exposed to NoV GII.4

A high seroprevalence of NoV GII.4 was observed in Alberta: overall, 76.7% of the subjects receiving serologic diagnostic and 84.2% of pregnant women in Alberta were positive to anti-NoV GII.4 IgG. The high seroprevalence of NoV is congruent with previous studies around the globe reporting values between 56% and 99% (Table 5.3) and highlights the ubiquitous nature of this genotype. The proportion of individuals with no detectable antibodies was similar to that of non-secretor individuals in human populations (~20%) [38-40], which are known to be genetically resistant to infection by NoV GII.4.

The high seroprevalence of NoV GII.4 and GMT titers were maintained over time.

In an effort to determine if waning of herd immune responses were implicated in the annual fluctuations of NoV outbreak levels, the titers of anti-NoV GII.4 IgG in two populations of Alberta over different years were determined and compared. No differences either in seroprevalence or geometric mean titers of anti-NoV GII.4 IgG could be identified across time points. Still, these results should be interpreted with caution for the following reasons: i) the data followed a non-normal distribution requiring the use of non-parametric tests which are known to be less powerful than parametric tests, ii) the statistics calculated to

compare the seroprevalence and GMT between time points had probabilities ranging between 0.07 and 0.08 which are very close to an alpha of 0.05, thus, the probability of having errors type II cannot be ruled out with confidence. In conclusion, the results from these analyzes indicated that anti-NoV IgG levels did not change significantly across time points, and that possible true changes in immune responses at the level of the population, if any, were too small to be observed with the sample size and statistical tests used for analysis.

An important limitation of the seroprevalence study is the use of total anti-GII.4 IgG as indicator of the immune responses in the population. Total IgG responses have been reported to be largely cross-reactive within genotypes and not good correlates of protective immunity, which in turn, might explain outbreak levels [41-43]. Nevertheless, the results from this study provide a clue of the magnitude of anti-NoV IgG seropositivity and GMT that can be observed within a population. In absence of animal models able to mimic NoV infection as it occurs in humans, cohort and human challenge studies continue to be the best approaches to study the immune responses against HuNoV. Future studies aiming to detect changes in immune responses with potential impact on NoV outbreak levels could benefit from the following recommendations. Firstly, the population under study could be narrowed to individuals at high risk of NoV (outbreak) infection (i.e. elderly, children and immunocompromised). Secondly, a cohort study design analyzing samples collected longitudinally from the same subjects would be more informative than a cross-sectional study because it would allow identifying changes in immune responses at the individual level and decrease the effect of inter-subject variation. Lastly, in addition to measuring anti-NoV total IgG, future studies should also test HBGA blocking IgG titers which have been correlated with protective immunity.

6.2. Future directions

Based on the findings of this thesis, the following areas are proposed for future studies:

6.2.1. Genetic and antigenic diversity of NoV at ORF1 and ORF3

The ongoing NoV surveillance program in Alberta could benefit from incorporating ORF1 and ORF3 regions to better identify and monitor novel NoV strains with potential epidemiological impact. Phylogenetic analyses of near full-length viral genome sequences from NoV outbreaks and characterization of the intra-host genetic diversity of NoV during chronic infection suggest that ORF1 and ORF3 might display features that increase the fitness of NoV. The large cross reactivity of HBGA blocking immune responses among GII.4 variants supports the idea that novel NoV GII.4 could be developing other traits besides antigenic diversity at the major capsid gene (ORF2) to increase their fitness and circulation among humans. Congruently, the characterization of the intra-host genetic diversity of NoV in an immunocompromised patient with chronic infection helped to identify residues at the p22 and VP2 genes that could be potential targets of positive selection. Nevertheless, these findings were based on a single patient; additional studies involving more immunocompromised individuals are necessary to identify the relevance of different genes/amino acid residues during chronic infection that could be further evaluated in NoV cell culture systems.

6.2.2. The elderly as potential reservoirs of novel NoV strains

The elderly is another group of individuals that has also been proposed as possible reservoirs of NoV [31]; however, to date, no study has been conducted in humans to support this hypothesis. The underlying premise is that elderly individuals have impaired immune responses and could shed the virus for longer periods of time compared to young adults. In

Alberta, the majority of outbreaks occur among the elderly, thus it would be relevant from the perspective of NoV evolution as well as for the purpose of NoV epidemiological surveillance to assess the intra-host evolution of NoV in these patients. In this thesis work, elderly individuals suffering NoV infection displayed a very homogenous NoV population at the acute phase of infection, but further studies with specimens collected longitudinally and at later time points of the infection are still required for a fair assessment of the elderly as potential reservoirs of novel NoV strains.

6.2.3. Immune responses to NoV and antigenic seniority

The components of the protective immune responses to NoV are still poorly understood and it is yet not clear how long protective immunity lasts after natural infection. As vaccine candidates become closer to be commercially available, it is still debatable how frequent immunization will be required to confer adequate protection. In clinical trials a VLP-based GI.1 vaccine candidate demonstrated an efficacy below expectation, with reductions of 47% in NoV symptomatic gastroenteritis and 26% on the overall rate of infection [44]. Moreover, it has been reported that the HBGA-blocking antibody responses elicited after administration of a VLP-based GI.1/GII.4 vaccine might follow a pattern of 'antigenic seniority' or 'original antigenic sin' [14], where each new strain taking a less senior position in the hierarchy of responses [14, 45, 46]. Antigenic seniority has been extensively described for other RNA viruses including influenza and dengue and has been reported to have an effect on illness severity, efficacy of vaccines and epochal trends in population immunity [14, 45-48]. Future studies on protective immune responses are needed to assess the effect of multiple variables including: waning immunity, age, antigenic seniority and cross reactivity among strains.

6.3. References (chapter 6)

1. Lu, J., et al., *Gastroenteritis Outbreaks Caused by Norovirus GII.17, Guangdong Province, China, 2014-2015.* Emerg Infect Dis, 2015. **21**(7): p. 1240-2.
2. Han, J., et al., *Emergence and predominance of norovirus GII.17 in Huzhou, China, 2014-2015.* Virol J, 2015. **12**: p. 139.
3. Fu, J., et al., *Emergence of a new GII.17 norovirus variant in patients with acute gastroenteritis in Jiangsu, China, September 2014 to March 2015.* Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 2015. **20**(24): p. 21157.
4. Matsushima, Y., et al., *Genetic analyses of GII.17 norovirus strains in diarrheal disease outbreaks from December 2014 to March 2015 in Japan reveal a novel polymerase sequence and amino acid substitutions in the capsid region.* Euro Surveill, 2015. **20**(26).
5. Parra, G.I. and K.Y. Green, *Genome of Emerging Norovirus GII.17, United States, 2014.* Emerging infectious diseases, 2015. **21**(8): p. 1477-1479.
6. Medici, M.C., et al., *Identification of the novel Kawasaki 2014 GII.17 human norovirus strain in Italy, 2015.* Euro Surveill, 2015. **20**(35): p. 30010.
7. Dinu, S., et al., *Molecular identification of emergent GII.P17-GII.17 norovirus genotype, Romania, 2015.* Euro Surveill, 2016. **21**(7).
8. Zhang, X.F., et al., *An outbreak caused by GII.17 norovirus with a wide spectrum of HBGA-associated susceptibility.* Sci Rep, 2015. **5**: p. 17687.
9. Pang, X.L., et al., *Influence of novel norovirus GII.4 variants on gastroenteritis outbreak dynamics in Alberta and the Northern Territories, Canada between 2000 and 2008.* PloS one, 2010. **5**(7): p. e11599.
10. Yen, C., et al., *Impact of an emergent norovirus variant in 2009 on norovirus outbreak activity in the United States.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2011. **53**(6): p. 568-571.
11. Cho, H.G., et al., *Emergence of Norovirus GII.4 variants in acute gastroenteritis outbreaks in South Korea between 2006 and 2013.* J Clin Virol, 2015. **72**: p. 11-5.
12. Kumazaki, M. and S. Usuku, *Genetic Analysis of Norovirus GII.4 Variant Strains Detected in Outbreaks of Gastroenteritis in Yokohama, Japan, from the 2006-2007 to the 2013-2014 Seasons.* PLoS One, 2015. **10**(11): p. e0142568.
13. Dai, Y.C., et al., *Antigenic Relatedness of Norovirus GII.4 Variants Determined by Human Challenge Sera.* PloS one, 2015. **10**(4): p. e0124945.
14. Lindesmith, L.C., et al., *Broad blockade antibody responses in human volunteers after immunization with a multivalent norovirus VLP candidate vaccine: immunological analyses from a phase I clinical trial.* PLoS Med, 2015. **12**(3): p. e1001807.
15. Kobayashi, M., et al., *Molecular evolution of the capsid gene in human norovirus genogroup II.* Sci Rep, 2016. **6**: p. 29400.
16. Mori, K., et al., *Genomic analysis of the evolutionary lineage of Norovirus GII.4 from archival specimens during 1975-1987 in Tokyo.* J Med Virol, 2016.

17. Boon, D., et al., *Comparative evolution of GII.3 and GII.4 norovirus over a 31-year period.* Journal of virology, 2011. **85**(17): p. 8656-8666.
18. Bok, K., et al., *Evolutionary dynamics of GII.4 noroviruses over a 34-year period.* Journal of virology, 2009. **83**(22): p. 11890-11901.
19. van Beek, J., et al., *Comparison of norovirus genogroup I, II, and IV seroprevalence among children in the Netherlands, 1963, 1983, and 2006.* J Gen Virol, 2016.
20. Eden, J.S., et al., *Recombination within the pandemic norovirus GII.4 lineage.* Journal of virology, 2013. **87**(11): p. 6270-6282.
21. Lam, T.T., et al., *The recombinant origin of emerging human norovirus GII.4/2008: intra-genotypic exchange of the capsid P2 domain.* The Journal of general virology, 2012. **93**(Pt 4): p. 817-822.
22. Fitch, W.M., et al., *Positive Darwinian evolution in human influenza A viruses.* Proc Natl Acad Sci U S A, 1991. **88**(10): p. 4270-4.
23. Fitch, W.M., et al., *Long term trends in the evolution of H(3) HA1 human influenza type A.* Proc Natl Acad Sci U S A, 1997. **94**(15): p. 7712-8.
24. Bush, R.M., et al., *Positive selection on the H3 hemagglutinin gene of human influenza virus A.* Mol Biol Evol, 1999. **16**(11): p. 1457-65.
25. Bush, R.M., et al., *Predicting the evolution of human influenza A.* Science (New York, N.Y.), 1999. **286**(5446): p. 1921-1925.
26. Buonagurio, D.A., et al., *Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene.* Science, 1986. **232**(4753): p. 980-2.
27. Kumar, S. and S. Subramanian, *Mutation rates in mammalian genomes.* Proc Natl Acad Sci U S A, 2002. **99**(2): p. 803-8.
28. Chao, L., *Fitness of RNA virus decreased by Muller's ratchet.* Nature, 1990. **348**(6300): p. 454-5.
29. Batty, E.M., et al., *A modified RNA-Seq approach for whole genome sequencing of RNA viruses from faecal and blood samples.* PloS one, 2013. **8**(6): p. e66129.
30. Giannoukos, G., et al., *Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes.* Genome biology, 2012. **13**(3): p. R23-2012-13-3-r23.
31. Karst, S.M. and R.S. Baric, *What Is the Reservoir of Emergent Human Norovirus Strains?* Journal of virology, 2015. **89**(11): p. 5756-5759.
32. Glass, P.J., et al., *Norwalk virus open reading frame 3 encodes a minor structural protein.* Journal of virology, 2000. **74**(14): p. 6581-6591.
33. Vongpunsawad, S., B.V. Venkataram Prasad, and M.K. Estes, *Norwalk Virus Minor Capsid Protein VP2 Associates within the VP1 Shell Domain.* Journal of virology, 2013. **87**(9): p. 4818-4825.
34. Roth, A.N. and S.M. Karst, *Norovirus mechanisms of immune antagonism.* Curr Opin Virol, 2016. **16**: p. 24-30.
35. Zhu, S., et al., *Identification of immune and viral correlates of norovirus protective immunity through comparative study of intra-cluster norovirus strains.* PLoS pathogens, 2013. **9**(9): p. e1003592.
36. Sharp, T.M., et al., *Inhibition of cellular protein secretion by norwalk virus nonstructural protein p22 requires a mimic of an endoplasmic reticulum export signal.* PLoS One, 2010. **5**(10): p. e13130.

37. Sharp, T.M., et al., *Secretory pathway antagonism by calicivirus homologues of Norwalk virus nonstructural protein p22 is restricted to noroviruses.* Virol J, 2012. **9**: p. 181.
38. Thorven, M., et al., *A homozygous nonsense mutation (428G-->A) in the human secretor (FUT2) gene provides resistance to symptomatic norovirus (GGII) infections.* Journal of virology, 2005. **79**(24): p. 15351-15355.
39. Kindberg, E., et al., *A nonsense mutation (428G-->A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection.* AIDS, 2006. **20**(5): p. 685-9.
40. Bucardo, F., et al., *Genetic susceptibility to symptomatic norovirus infection in Nicaragua.* J Med Virol, 2009. **81**(4): p. 728-35.
41. Reeck, A., et al., *Serological correlate of protection against norovirus-induced gastroenteritis.* The Journal of infectious diseases, 2010. **202**(8): p. 1212-1218.
42. Parrino, T.A., et al., *Clinical immunity in acute gastroenteritis caused by Norwalk agent.* The New England journal of medicine, 1977. **297**(2): p. 86-89.
43. Johnson, P.C., et al., *Multiple-challenge study of host susceptibility to Norwalk gastroenteritis in US adults.* The Journal of infectious diseases, 1990. **161**(1): p. 18-21.
44. Atmar, R.L., et al., *Norovirus vaccine against experimental human Norwalk Virus illness.* The New England journal of medicine, 2011. **365**(23): p. 2178-2187.
45. Lessler, J., et al., *Evidence for antigenic seniority in influenza A (H3N2) antibody responses in southern China.* PLoS Pathog, 2012. **8**(7): p. e1002802.
46. Miller, M.S., et al., *Neutralizing antibodies against previously encountered influenza virus strains increase over time: a longitudinal analysis.* Sci Transl Med, 2013. **5**(198): p. 198ra107.
47. Kim, J.H., et al., *Original antigenic sin responses to influenza viruses.* J Immunol, 2009. **183**(5): p. 3294-301.
48. Mongkolsapaya, J., et al., *Original antigenic sin and apoptosis in the pathogenesis of dengue hemorrhagic fever.* Nat Med, 2003. **9**(7): p. 921-7.

References

Abbas, A.K., K.M. Murphy, and A. Sher, *Functional diversity of helper T lymphocytes.* Nature, 1996. **383**(6603): p. 787-93.

Adler, J.L. and R. Zickl, *Winter vomiting disease.* J Infect Dis, 1969. **119**(6): p. 668-73.

Advisors, B.R.E., *The Contribution of the International Cruise Industry to the U.S. Economy in 2014*. 2014: Exton, PA.

Ahmed, S.M., et al., *Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis.* The Lancet infectious diseases, 2014. **14**(8): p. 725-730.

Ahmed, S.M., B.A. Lopman, and K. Levy, *A systematic review and meta-analysis of the global seasonality of norovirus.* PLoS One, 2013. **8**(10): p. e75922.

Ajami, N.J., et al., *Antibody responses to norovirus genogroup GI.1 and GII.4 proteases in volunteers administered Norwalk virus.* Clinical and vaccine immunology : CVI, 2012. **19**(12): p. 1980-1983.

Alam, A., et al., *Genetic characterization of norovirus strains in hospitalized children from Pakistan.* Journal of medical virology, 2015.

Alberta, G.o., *Public Health Notifiable Disease Management Guidelines*, A. Health, Editor. 2013: Alberta.

Allen, D.J., et al., *Analysis of amino acid variation in the P2 domain of the GII-4 norovirus VP1 protein reveals putative variant-specific epitopes.* PloS one, 2008. **3**(1): p. e1485.

Altschul, S.F., et al., *Basic local alignment search tool.* Journal of Molecular Biology, 1990. **215**(3): p. 403-410.

Ambert-Balay, K. and P. Pothier, *Evaluation of 4 immunochromatographic tests for rapid detection of norovirus in faecal samples.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2013. **56**(3): p. 194-198.

Ando, T., et al., *Detection and differentiation of antigenically distinct small round-structured viruses (Norwalk-like viruses) by reverse transcription-PCR and southern hybridization.* Journal of clinical microbiology, 1995. **33**(1): p. 64-71.

Armbrust, S., et al., *Norovirus infections in preterm infants: wide variety of clinical courses.* BMC Res Notes, 2009. **2**: p. 96.

Arthur, S.E. and K.E. Gibson, *Environmental persistence of Tulane virus - a surrogate for human norovirus.* Can J Microbiol, 2016. **62**(5): p. 449-54.

Asanaka, M., et al., *Replication and packaging of Norwalk virus RNA in cultured mammalian cells.* Proc Natl Acad Sci U S A, 2005. **102**(29): p. 10327-32.

Atmar, R.L., et al., *Norovirus vaccine against experimental human Norwalk Virus illness.* The New England journal of medicine, 2011. **365**(23): p. 2178-2187.

Atmar, R.L., et al., *Serological Correlates of Protection against a GII.4 Norovirus.* Clin Vaccine Immunol, 2015. **22**(8): p. 923-9.

Atmar, R.L. and M.K. Estes, *The epidemiologic and clinical importance of norovirus infection.* Gastroenterology clinics of North America, 2006. **35**(2): p. 275-90, viii.

Atmar, R.L., et al., *Norwalk virus shedding after experimental human infection.* Emerging infectious diseases, 2008. **14**(10): p. 1553-1557.

Atmar, R.L., et al., *Determination of the 50% human infectious dose for Norwalk virus.* J Infect Dis, 2014. **209**(7): p. 1016-22.

Aw, D., A.B. Silva, and D.B. Palmer, *Immunosenescence: emerging challenges for an ageing population.* Immunology, 2007. **120**(4): p. 435-46.

Bae, J. and K.J. Schwab, *Evaluation of murine norovirus, feline calicivirus, poliovirus, and MS2 as surrogates for human norovirus in a model of viral persistence in surface water and groundwater.* Appl Environ Microbiol, 2008. **74**(2): p. 477-84.

Bailey, D., et al., *Functional analysis of RNA structures present at the 3' extremity of the murine norovirus genome: the variable polypyrimidine tract plays a role in viral virulence.* Journal of virology, 2010. **84**(6): p. 2859-2870.

Baker, E.S., et al., *Inherent structural disorder and dimerisation of murine norovirus NS1-2 protein.* PLoS One, 2012. **7**(2): p. e30534.

Baldridge, M.T., et al., *Commensal microbes and interferon-lambda determine persistence of enteric murine norovirus infection.* Science, 2015. **347**(6219): p. 266-9.

Batty, E.M., et al., *A modified RNA-Seq approach for whole genome sequencing of RNA viruses from faecal and blood samples.* PloS one, 2013. **8**(6): p. e66129.

Bavelaar, H.H., et al., *Whole genome sequencing of fecal samples as a tool for the diagnosis and genetic characterization of norovirus.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2015. **72**: p. 122-125.

Belliot, G., et al., *Norovirus proteinase-polymerase and polymerase are both active forms of RNA-dependent RNA polymerase.* J Virol, 2005. **79**(4): p. 2393-403.

Belliot, G., et al., *Nucleotidylylation of the VPg protein of a human norovirus by its proteinase-polymerase precursor protein.* Virology, 2008. **374**(1): p. 33-49.

Bertolotti-Ciarlet, A., et al., *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.* Journal of virology, 2003. **77**(21): p. 11603-11615.

Beuret, C., A. Baumgartner, and J. Schluep, *Virus-contaminated oysters: a three-month monitoring of oysters imported to Switzerland.* Appl Environ Microbiol, 2003. **69**(4): p. 2292-7.

Bitler, E.J., et al., *Norovirus outbreaks: a systematic review of commonly implicated transmission routes and vehicles.* Epidemiol Infect, 2013. **141**(8): p. 1563-71.

Blazevic, V., et al., *Multiple consecutive norovirus infections in the first 2 years of life.* Eur J Pediatr, 2015. **174**(12): p. 1679-83.

Blazevic, V., M. Malm, and T. Vesikari, *Induction of homologous and cross-reactive GII.4-specific blocking antibodies in children after GII.4 New Orleans norovirus infection.* J Med Virol, 2015. **87**(10): p. 1656-61.

Bok, K., et al., *Evolutionary dynamics of GII.4 noroviruses over a 34-year period.* Journal of virology, 2009. **83**(22): p. 11890-11901.

Bok, K. and K.Y. Green, *Norovirus gastroenteritis in immunocompromised patients.* N Engl J Med, 2012. **367**(22): p. 2126-32.

Bok, K., et al., *Chimpanzees as an animal model for human norovirus infection and vaccine development.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(1): p. 325-330.

Boni, M.F., et al., *Guidelines for identifying homologous recombination events in influenza A virus.* PloS one, 2010. **5**(5): p. e10434.

Boni, M.F., D. Posada, and M.W. Feldman, *An exact nonparametric method for inferring mosaic structure in sequence triplets.* Genetics, 2007. **176**(2): p. 1035-1047.

Boon, D., et al., *Comparative evolution of GII.3 and GII.4 norovirus over a 31-year period.* Journal of virology, 2011. **85**(17): p. 8656-8666.

Borrows, C.L. and P.C. Turner, *Seasonal screening for viral gastroenteritis in young children and elderly hospitalized patients: is it worthwhile?* J Hosp Infect, 2014. **87**(2): p. 98-102.

Bruggink, L.D., et al., *Molecular and epidemiological features of gastroenteritis outbreaks involving genogroup I norovirus in Victoria, Australia, 2002-2010.* Journal of medical virology, 2012. **84**(9): p. 1437-1448.

Bucardo, F., et al., *Genetic susceptibility to symptomatic norovirus infection in Nicaragua.* J Med Virol, 2009. **81**(4): p. 728-35.

Bui, T., et al., *Median infectious dose of human norovirus GII.4 in gnotobiotic pigs is decreased by simvastatin treatment and increased by age.* J Gen Virol, 2013. **94**(Pt 9): p. 2005-16.

Bull, R.A., et al., *Contribution of intra- and interhost dynamics to norovirus evolution.* Journal of virology, 2012. **86**(6): p. 3219-3229.

Bull, R.A., et al., *Rapid evolution of pandemic noroviruses of the GII.4 lineage.* PLoS pathogens, 2010. **6**(3): p. e1000831.

Bull, R.A., et al., *Norovirus recombination in ORF1/ORF2 overlap.* Emerging Infectious Diseases, 2005. **11**(7): p. 1079-1085.

Bull, R.A., et al., *Comparison of the replication properties of murine and human calicivirus RNA-dependent RNA polymerases.* Virus genes, 2010.

Bull, R.A., et al., *Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection.* PLoS pathogens, 2011. **7**(9): p. e1002243.

Bull, R.A., M.M. Tanaka, and P.A. White, *Norovirus recombination.* The Journal of general virology, 2007. **88**(Pt 12): p. 3347-3359.

Buonagurio, D.A., et al., *Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene.* Science, 1986. **232**(4753): p. 980-2.

Burroughs, J.N. and F. Brown, *Presence of a covalently linked protein on calicivirus RNA.* J Gen Virol, 1978. **41**(2): p. 443-6.

Burton-MacLeod, J.A., et al., *Evaluation and comparison of two commercial enzyme-linked immunosorbent assay kits for detection of antigenically diverse human noroviruses in stool samples.* J Clin Microbiol, 2004. **42**(6): p. 2587-95.

Bush, R.M., et al., *Predicting the evolution of human influenza A.* Science (New York, N.Y.), 1999. **286**(5446): p. 1921-1925.

Bush, R.M., et al., *Positive selection on the H3 hemagglutinin gene of human influenza virus A.* Mol Biol Evol, 1999. **16**(11): p. 1457-65.

Cancio-Lonches, C., et al., *Nucleolin interacts with the feline calicivirus 3' untranslated region and the protease-polymerase NS6 and NS7 proteins, playing a role in virus replication.* J Virol, 2011. **85**(16): p. 8056-68.

Cannon, J.L., et al., *Surrogates for the study of norovirus stability and inactivation in the environment: aA comparison of murine norovirus and feline calicivirus.* J Food Prot, 2006. **69**(11): p. 2761-5.

Carlsson, B., et al., *Quasispecies dynamics and molecular evolution of human norovirus capsid P region during chronic infection.* The Journal of general virology, 2009. **90**(Pt 2): p. 432-441.

Carmona-Vicente, N., et al., *Antibodies against Lewis antigens inhibit the binding of human norovirus GII.4 virus-like particles to saliva but not to intestinal Caco-2 cells.* Virol J, 2016. **13**(1): p. 82.

Carmona-Vicente, N., et al., *Norovirus infections and seroprevalence of genotype GII.4-specific antibodies in a Spanish population.* J Med Virol, 2015. **87**(4): p. 675-82.

Castle, S.C., *Clinical relevance of age-related immune dysfunction.* Clin Infect Dis, 2000. **31**(2): p. 578-85.

Centers for Disease, C. and Prevention, *Norovirus outbreak in an elementary school--District of Columbia, February 2007.* MMWR Morb Mortal Wkly Rep, 2008. **56**(51-52): p. 1340-3.

Centers for Disease, C. and Prevention, *Notes from the Field: Emergence of New Norovirus Strain GII.4 Sydney - United States, 2012.* MMWR.Morbidity and mortality weekly report, 2013. **62**: p. 55.

Chan, M.C., et al., *Covariation of major and minor viral capsid proteins in norovirus genogroup II genotype 4 strains.* Journal of virology, 2012. **86**(2): p. 1227-1232.

Chao, L., *Fitness of RNA virus decreased by Muller's ratchet.* Nature, 1990. **348**(6300): p. 454-5.

Cheesbrough, J.S., L. Barkess-Jones, and D.W. Brown, *Possible prolonged environmental survival of small round structured viruses.* J Hosp Infect, 1997. **35**(4): p. 325-6.

Cheetham, S., et al., *Pathogenesis of a genogroup II human norovirus in gnotobiotic pigs.* Journal of virology, 2006. **80**(21): p. 10372-10381.

Chen, C.H. and C.P. Yu, *Comparison of enzyme-linked immunosorbent assay and reverse transcription-polymerase chain reaction for the detection of norovirus in fecal samples.* Research Journal of Chemistry and Environment, 2014. **18**(3): p. 70-75.

Chen, S.M., et al., *The significance of serum and fecal levels of interleukin-6 and interleukin-8 in hospitalized children with acute rotavirus and norovirus gastroenteritis.* Pediatr Neonatol, 2014. **55**(2): p. 120-6.

Chen, S.Y., et al., *Norovirus infection as a cause of diarrhea-associated benign infantile seizures.* Clinical Infectious Diseases, 2009. **48**(7): p. 849-855.

Chhabra, P., A.M. Walimbe, and S.D. Chitambar, *Complete genome characterization of Genogroup II norovirus strains from India: Evidence of recombination in ORF2/3 overlap.* Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2010. **10**(7): p. 1101-1109.

Cho, H.G., et al., *Emergence of Norovirus GII.4 variants in acute gastroenteritis outbreaks in South Korea between 2006 and 2013.* J Clin Virol, 2015. **72**: p. 11-5.

Cleophas, T.J. and A.H. Zwinderman, *Assessing seasonality in clinical research.* Clinical chemistry and laboratory medicine : CCLM / FESCC, 2012.

Colas de la Noue, A., et al., *Absolute Humidity Influences the Seasonal Persistence and Infectivity of Human Norovirus.* Appl Environ Microbiol, 2014. **80**(23): p. 7196-205.

Cotten, M., et al., *Full genome virus detection in fecal samples using sensitive nucleic acid preparation, deep sequencing, and a novel iterative sequence classification algorithm.* PloS one, 2014. **9**(4): p. e93269.

Cotten, M., et al., *Deep sequencing of norovirus genomes defines evolutionary patterns in an urban tropical setting.* Journal of virology, 2014. **88**(19): p. 11056-11069.

D'Souza, D.H., et al., *Persistence of caliciviruses on environmental surfaces and their transfer to food.* Int J Food Microbiol, 2006. **108**(1): p. 84-91.

Dai, Y.C., et al., *Seroprevalence of antibodies against noroviruses among students in a Chinese military medical university.* Journal of clinical microbiology, 2004. **42**(10): p. 4615-4619.

Dai, Y.C., et al., *Antigenic Relatedness of Norovirus GII.4 Variants Determined by Human Challenge Sera.* PloS one, 2015. **10**(4): p. e0124945.

Daughenbaugh, K.F., et al., *The genome-linked protein VPg of the Norwalk virus binds eIF3, suggesting its role in translation initiation complex recruitment.* EMBO J, 2003. **22**(11): p. 2852-9.

Daughenbaugh, K.F., C.E. Wobus, and M.E. Hardy, *VPg of murine norovirus binds translation initiation factors in infected cells.* Virol J, 2006. **3**: p. 33.

de Bruin, E., et al., *Diagnosis of Norovirus outbreaks by commercial ELISA or RT-PCR.* J Virol Methods, 2006. **137**(2): p. 259-64.

de Graaf, M., et al., *Emergence of a novel GII.17 norovirus - End of the GII.4 era?* Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 2015. **20**(26): p. 21178.

de Vries, M., et al., *Performance of VIDISCA-454 in feces-suspensions and serum.* Viruses, 2012. **4**(8): p. 1328-1334.

de Wit, M.A., et al., *Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology.* Am J Epidemiol, 2001. **154**(7): p. 666-74.

Debbink, K., et al., *Genetic mapping of a highly variable norovirus GII.4 blockade epitope: potential role in escape from human herd immunity.* Journal of virology, 2012. **86**(2): p. 1214-1226.

Debbink, K., et al., *Emergence of new pandemic GII.4 Sydney norovirus strain correlates with escape from herd immunity.* The Journal of infectious diseases, 2013. **208**(11): p. 1877-1887.

Debbink, K., et al., *Within-host evolution results in antigenically distinct GII.4 noroviruses.* Journal of virology, 2014. **88**(13): p. 7244-7255.

Delacour, H., P. Dubrous, and J.L. Koeck, *Noroviruses: a challenge for military forces.* J R Army Med Corps, 2010. **156**(4): p. 251-4.

Di Bartolo, I., et al., *Waterborne norovirus outbreak during a summer excursion in Northern Italy.* New Microbiol, 2015. **38**(1): p. 109-12.

Dinu, S., et al., *Molecular identification of emergent GII.P17-GII.17 norovirus genotype, Romania, 2015.* Euro Surveill, 2016. **21**(7).

Division of Viral Diseases, N.C.f.I., C.f.D.C. Respiratory Diseases, and Prevention, *Updated norovirus outbreak management and disease prevention guidelines.* MMWR Recomm Rep, 2011. **60**(RR-3): p. 1-18.

Division of Viral Diseases, N.C.f.I., C.f.D.C. Respiratory Diseases, and Prevention, *Updated norovirus outbreak management and disease prevention guidelines.* MMWR.Recommendations and reports : Morbidity and mortality weekly report.Recommendations and reports / Centers for Disease Control, 2011. **60**(RR-3): p. 1-18.

Donaldson, E.F., et al., *Norovirus pathogenesis: mechanisms of persistence and immune evasion in human populations.* Immunological reviews, 2008. **225**: p. 190-211.

Donaldson, E.F., et al., *Viral shape-shifting: norovirus evasion of the human immune system.* Nature reviews.Microbiology, 2010. **8**(3): p. 231-241.

Dowell, S.F., *Seasonal variation in host susceptibility and cycles of certain infectious diseases.* Emerg Infect Dis, 2001. **7**(3): p. 369-74.

Eden, J.S., et al., *The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012.* Virology, 2014. **450-451**: p. 106-113.

Eden, J.S., et al., *Recombination within the pandemic norovirus GII.4 lineage.* Journal of virology, 2013. **87**(11): p. 6270-6282.

Eltahla, A.A., et al., *Nonnucleoside inhibitors of norovirus RNA polymerase: scaffolds for rational drug design.* Antimicrob Agents Chemother, 2014. **58**(6): p. 3115-23.

Erney, R.M., et al., *Variability of human milk neutral oligosaccharides in a diverse population.* J Pediatr Gastroenterol Nutr, 2000. **30**(2): p. 181-92.

Esseili, M.A., et al., *Feline Calicivirus, Murine Norovirus, Porcine Sapovirus, and Tulane Virus Survival on Postharvest Lettuce.* Appl Environ Microbiol, 2015. **81**(15): p. 5085-92.

Ettayebi, K. and M.E. Hardy, *Norwalk virus nonstructural protein p48 forms a complex with the SNARE regulator VAP-A and prevents cell surface expression of vesicular stomatitis virus G protein.* Journal of virology, 2003. **77**(21): p. 11790-11797.

Evans, M.R., et al., *An outbreak of viral gastroenteritis following environmental contamination at a concert hall.* Epidemiol Infect, 2002. **129**(2): p. 355-60.

Fang, H., et al., *Norovirus P particle efficiently elicits innate, humoral and cellular immunity.* PLoS One, 2013. **8**(4): p. e63269.

Fankhauser, R.L., et al., *Epidemiologic and molecular trends of "Norwalk-like viruses" associated with outbreaks of gastroenteritis in the United States.* Journal of Infectious Diseases, 2002. **186**(1): p. 1-7.

Fernandez-Duran, J.J. and M.M. Gregorio-Dominguez, *Testing for seasonality using circular distributions based on non-negative trigonometric sums as alternative hypotheses.* Statistical methods in medical research, 2011.

Fields virology, B.N. Fields, D.M. Knipe, and P.M. Howley, Editors. 2013, Wolters Kluwer/Lippincott Williams & Wilkins Health: Philadelphia.

Fitch, W.M., et al., *Long term trends in the evolution of H(3) HA1 human influenza type A.* Proc Natl Acad Sci U S A, 1997. **94**(15): p. 7712-8.

Fitch, W.M., et al., *Positive Darwinian evolution in human influenza A viruses.* Proc Natl Acad Sci U S A, 1991. **88**(10): p. 4270-4.

Fonager, J., S. Barzinci, and T.K. Fischer, *Emergence of a new recombinant Sydney 2012 norovirus variant in Denmark, 26 December 2012 to 22 March 2013.* Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 2013. **18**(25): p. 20506.

Frange, P., et al., *Prevalence and clinical impact of norovirus fecal shedding in children with inherited immune deficiencies.* The Journal of infectious diseases, 2012. **206**(8): p. 1269-1274.

Freeland, A.L., G.H. Vaughan, Jr., and S.N. Banerjee, *Acute Gastroenteritis on Cruise Ships - United States, 2008-2014.* MMWR Morb Mortal Wkly Rep, 2016. **65**(1): p. 1-5.

Frenck, R., et al., *Predicting susceptibility to norovirus GII.4 by use of a challenge model involving humans.* The Journal of infectious diseases, 2012. **206**(9): p. 1386-1393.

Fu, J., et al., *Emergence of a new GII.17 norovirus variant in patients with acute gastroenteritis in Jiangsu, China, September 2014 to March 2015.* Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 2015. **20**(24): p. 21157.

Gallimore, C.I., et al., *Environmental monitoring for gastroenteric viruses in a pediatric primary immunodeficiency unit.* J Clin Microbiol, 2006. **44**(2): p. 395-9.

Gallimore, C.I., et al., *Contamination of the hospital environment with gastroenteric viruses: comparison of two pediatric wards over a winter season.* J Clin Microbiol, 2008. **46**(9): p. 3112-5.

Garrison, E. and G. Marth *Haplotype-based variant detection from short-read sequencing*. 2012. 02 Nov 2015-arXiv:1207.3907v2.

Gastanaduy, P.A., et al., *Burden of norovirus gastroenteritis in the ambulatory setting--United States, 2001-2009.* J Infect Dis, 2013. **207**(7): p. 1058-65.

Giammanco, G.M., et al., *Investigation and control of a Norovirus outbreak of probable waterborne transmission through a municipal groundwater system.* J Water Health, 2014. **12**(3): p. 452-64.

Giannoukos, G., et al., *Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes.* Genome biology, 2012. **13**(3): p. R23-2012-13-3-r23.

Glass, P.J., et al., *Norwalk virus open reading frame 3 encodes a minor structural protein.* Journal of virology, 2000. **74**(14): p. 6581-6591.

Glass, R.I., U.D. Parashar, and M.K. Estes, *Norovirus gastroenteritis.* New England Journal of Medicine, 2009. **361**(18): p. 1776-1785.

Goller, J.L., et al., *Long-term features of norovirus gastroenteritis in the elderly.* J Hosp Infect, 2004. **58**(4): p. 286-91.

Goodfellow, I., et al., *Calicivirus translation initiation requires an interaction between VPg and eIF 4 E.* EMBO Rep, 2005. **6**(10): p. 968-72.

Gould, L.H., et al., *Contributing factors in restaurant-associated foodborne disease outbreaks, FoodNet sites, 2006 and 2007.* J Food Prot, 2013. **76**(11): p. 1824-8.

Gray, J.J., et al., *Detection of immunoglobulin M (IgM), IgA, and IgG Norwalk virus-specific antibodies by indirect enzyme-linked immunosorbent assay with baculovirus-expressed Norwalk virus capsid antigen in adult volunteers challenged with Norwalk virus.* J Clin Microbiol, 1994. **32**(12): p. 3059-63.

Gray, J.J., et al., *Prevalence of antibodies to Norwalk virus in England: detection by enzyme-linked immunosorbent assay using baculovirus-expressed Norwalk virus capsid antigen.* Journal of clinical microbiology, 1993. **31**(4): p. 1022-1025.

Green, K.Y., *Norovirus infection in immunocompromised hosts.* Clin Microbiol Infect, 2014. **20**(8): p. 717-23.

Greenberg, H.B., et al., *Prevalence of antibody to the Norwalk virus in various countries.* Infect Immun, 1979. **26**(1): p. 270-3.

Greening, G.E., et al., *Molecular epidemiology of norovirus gastroenteritis outbreaks in New Zealand from 2002-2009.* Journal of medical virology, 2012. **84**(9): p. 1449-1458.

Greer, A.L., S.J. Drews, and D.N. Fisman, *Why "winter" vomiting disease? Seasonality, hydrology, and Norovirus epidemiology in Toronto, Canada.* Ecohealth, 2009. **6**(2): p. 192-9.

Grytdal, S.P., et al., *Incidence of Norovirus and Other Viral Pathogens That Cause Acute Gastroenteritis (AGE) among Kaiser Permanente Member Populations in the United States, 2012-2013.* PLoS One, 2016. **11**(4): p. e0148395.

Gutierrez-Escolano, A.L., et al., *Interaction of cellular proteins with the 5' end of Norwalk virus genomic RNA.* J Virol, 2000. **74**(18): p. 8558-62.

Hale, A.D., et al., *Homotypic and heterotypic IgG and IgM antibody responses in adults infected with small round structured viruses.* J Med Virol, 1998. **54**(4): p. 305-12.

Hall, A.J., et al., *Norovirus disease in the United States.* Emerging infectious diseases, 2013. **19**(8): p. 1198-1205.

Hall, A.J., et al., *Incidence of acute gastroenteritis and role of norovirus, Georgia, USA, 2004-2005.* Emerg Infect Dis, 2011. **17**(8): p. 1381-8.

Hall, A.J., et al., *Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States.* Emerg Infect Dis, 2013. **19**(8): p. 1305-9.

Hamad, I., et al., *Molecular detection of eukaryotes in a single human stool sample from Senegal.* PloS one, 2012. **7**(7): p. e40888.

Han, J., et al., *Emergence and predominance of norovirus GII.17 in Huzhou, China, 2014-2015.* Virol J, 2015. **12**: p. 139.

Harada, A., et al., *Essential involvement of interleukin-8 (IL-8) in acute inflammation.* J Leukoc Biol, 1994. **56**(5): p. 559-64.

Hardy, M.E., *Norovirus protein structure and function.* FEMS microbiology letters, 2005. **253**(1): p. 1-8.

Hardy, M.E. and M.K. Estes, *Completion of the Norwalk virus genome sequence.* Virus Genes, 1996. **12**(3): p. 287-90.

Hasing, M.E., et al., *Detection and analysis of recombination in GII.4 norovirus strains causing gastroenteritis outbreaks in Alberta.* Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2014.

Hasing, M.E., et al., *Emergence of a new norovirus GII.4 variant and changes in the historical biennial pattern of norovirus outbreak activity in Alberta, Canada, from 2008 to 2013.* Journal of clinical microbiology, 2013. **51**(7): p. 2204-2211.

He, S., et al., *Validation of two ribosomal RNA removal methods for microbial metatranscriptomics.* Nature methods, 2010. **7**(10): p. 807-812.

Heijne, J.C., et al., *Enhanced hygiene measures and norovirus transmission during an outbreak.* Emerg Infect Dis, 2009. **15**(1): p. 24-30.

Hemming, M., et al., *Major reduction of rotavirus, but not norovirus, gastroenteritis in children seen in hospital after the introduction of RotaTeq vaccine into the National Immunization Programme in Finland.* Eur J Pediatr, 2013. **172**(6): p. 739-46.

Herbert, T.P., I. Brierley, and T.D. Brown, *Identification of a protein linked to the genomic and subgenomic mRNAs of feline calicivirus and its role in translation.* J Gen Virol, 1997. **78 (Pt 5)**: p. 1033-40.

Herod, M.R., et al., *Expression of the murine norovirus (MNV) ORF1 polyprotein is sufficient to induce apoptosis in a virus-free cell model.* PLoS One, 2014. **9**(3): p. e90679.

Hoffmann, D., et al., *Norovirus GII.4 and GII.7 capsid sequences undergo positive selection in chronically infected patients.* Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2012. **12**(2): p. 461-466.

Hogbom, M., et al., *The active form of the norovirus RNA-dependent RNA polymerase is a homodimer with cooperative activity.* J Gen Virol, 2009. **90**(Pt 2): p. 281-91.

Hughes, P.J. and G. Stanway, *The 2A proteins of three diverse picornaviruses are related to each other and to the H-rev107 family of proteins involved in the control of cell proliferation.* J Gen Virol, 2000. **81**(Pt 1): p. 201-7.

Hutson, A.M., et al., *Norwalk virus infection associates with secretor status genotyped from sera.* Journal of medical virology, 2005. **77**(1): p. 116-120.

Hutson, A.M., et al., *Norwalk virus infection and disease is associated with ABO histo-blood group type.* The Journal of infectious diseases, 2002. **185**(9): p. 1335-1337.

Hyde, J.L. and J.M. Mackenzie, *Subcellular localization of the MNV-1 ORF1 proteins and their potential roles in the formation of the MNV-1 replication complex.* Virology, 2010. **406**(1): p. 138-48.

Iritani, N., et al., *Genetic analysis of the capsid gene of genotype GII.2 noroviruses.* Journal of virology, 2008. **82**(15): p. 7336-7345.

Ito, S., et al., *Norovirus-associated encephalopathy.* Pediatric Infectious Disease Journal, 2006. **25**(7): p. 651-652.

Jamieson, D.J., R.N. Theiler, and S.A. Rasmussen, *Emerging infections and pregnancy.* Emerging Infectious Diseases, 2006. **12**(11): p. 1638-1643.

Jiang, X., et al., *Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein.* Journal of virology, 1992. **66**(11): p. 6527-6532.

Jiang, X., et al., *Sequence and genomic organization of Norwalk virus.* Virology, 1993. **195**(1): p. 51-61.

Johnson, P.C., et al., *Multiple-challenge study of host susceptibility to Norwalk gastroenteritis in US adults.* The Journal of infectious diseases, 1990. **161**(1): p. 18-21.

Jones, M.K., et al., *Human norovirus culture in B cells.* Nat Protoc, 2015. **10**(12): p. 1939-47.

Jones, M.K., et al., *Enteric bacteria promote human and mouse norovirus infection of B cells.* Science (New York, N.Y.), 2014. **346**(6210): p. 755-759.

Kamel, A.H., et al., *Predominance and circulation of enteric viruses in the region of Greater Cairo, Egypt.* Journal of clinical microbiology, 2009. **47**(4): p. 1037-1045.

Kapikian, A.Z., *The discovery of the 27-nm Norwalk virus: an historic perspective.* The Journal of infectious diseases, 2000. **181 Suppl 2**: p. S295-302.

Kapikian, A.Z., et al., *Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis.* Journal of virology, 1972. **10**(5): p. 1075-1081.

Kaplanski, G., et al., *IL-6: a regulator of the transition from neutrophil to monocyte recruitment during inflammation.* Trends Immunol, 2003. **24**(1): p. 25-9.

Karst, S.M. and R.S. Baric, *What Is the Reservoir of Emergent Human Norovirus Strains?* Journal of virology, 2015. **89**(11): p. 5756-5759.

Karst, S.M., et al., *STAT1-dependent innate immunity to a Norwalk-like virus.* Science, 2003. **299**(5612): p. 1575-1578.

Karst, S.M., S. Zhu, and I.G. Goodfellow, *The molecular pathology of noroviruses.* J Pathol, 2015. **235**(2): p. 206-16.

Kawano, G., et al., *Benign infantile convulsions associated with mild gastroenteritis: a retrospective study of 39 cases including virological tests and efficacy of anticonvulsants.* Brain Dev, 2007. **29**(10): p. 617-22.

Kele, B., G. Lengyel, and J. Deak, *Comparison of an ELISA and two reverse transcription polymerase chain reaction methods for norovirus detection.* Diagn Microbiol Infect Dis, 2011. **70**(4): p. 475-8.

Khamrin, P., et al., *Evaluation of immunochromatography and commercial enzyme-linked immunosorbent assay for rapid detection of norovirus antigen in stool samples.* J Virol Methods, 2008. **147**(2): p. 360-3.

Khan, R.R., et al., *Gastrointestinal norovirus infection associated with exacerbation of inflammatory bowel disease.* J Pediatr Gastroenterol Nutr, 2009. **48**(3): p. 328-33.

Khodayar-Pardo, P., et al., *Norovirus GII.4 antibodies in breast milk and serum samples: their role preventing virus-like particles binding to their receptors.* Pediatr Infect Dis J, 2014. **33**(6): p. 554-9.

Kim, H.S., et al., *Emergence of GII.4 Sydney norovirus in South Korea during the winter of 2012-2013.* J Microbiol Biotechnol, 2013. **23**(11): p. 1641-3.

Kim, J.H., et al., *Original antigenic sin responses to influenza viruses.* J Immunol, 2009. **183**(5): p. 3294-301.

Kimura, H., et al., *A norovirus outbreak associated with environmental contamination at a hotel.* Epidemiol Infect, 2011. **139**(2): p. 317-25.

Kindberg, E., et al., *A nonsense mutation (428G-->A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection.* AIDS, 2006. **20**(5): p. 685-9.

Kirby, A., et al., *An evaluation of the RIDASCREEN and IDEIA enzyme immunoassays and the RIDAQUICK immunochromatographic test for the detection of norovirus in faecal specimens.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2010. **49**(4): p. 254-257.

Kirby, A.E., A. Streby, and C.L. Moe, *Vomiting as a Symptom and Transmission Risk in Norovirus Illness: Evidence from Human Challenge Studies.* PLoS One, 2016. **11**(4): p. e0143759.

Kirby, A.E., P.F. Teunis, and C.L. Moe, *Two human challenge studies confirm high infectivity of Norwalk virus.* J Infect Dis, 2015. **211**(1): p. 166-7.

Kirk, M.D., et al., *World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis.* PLoS Med, 2015. **12**(12): p. e1001921.

Kobayashi, M., et al., *Molecular evolution of the capsid gene in human norovirus genogroup II.* Sci Rep, 2016. **6**: p. 29400.

Kobayashi, S., et al., *Seroepidemiological study of norovirus infection in Aichi Prefecture, Japan.* Microbiology and immunology, 2009. **53**(6): p. 356-359.

Kojima, S., et al., *Genogroup-specific PCR primers for detection of Norwalk-like viruses.* Journal of virological methods, 2002. **100**(1-2): p. 107-114.

Koo, H.L., et al., *Noroviruses: The Most Common Pediatric Viral Enteric Pathogen at a Large University Hospital After Introduction of Rotavirus Vaccination.* J Pediatric Infect Dis Soc, 2013. **2**(1): p. 57-60.

Koopmans, M., *Noroviruses in healthcare settings: a challenging problem.* J Hosp Infect, 2009. **73**(4): p. 331-7.

Kopylova, E., L. Noe, and H. Touzet, *SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data.* Bioinformatics (Oxford, England), 2012. **28**(24): p. 3211-3217.

Kotwal, G. and J.L. Cannon, *Environmental persistence and transfer of enteric viruses.* Curr Opin Virol, 2014. **4**: p. 37-43.

Kourtis, A.P., J.S. Read, and D.J. Jamieson, *Pregnancy and infection.* N Engl J Med, 2014. **370**(23): p. 2211-8.

Kroneman, A., et al., *Proposal for a unified norovirus nomenclature and genotyping.* Arch Virol, 2013. **158**(10): p. 2059-68.

Kroneman, A., et al., *An automated genotyping tool for enteroviruses and noroviruses.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2011. **51**(2): p. 121-125.

Kroneman, A., et al., *Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the Foodborne Viruses in Europe network from 1 July 2001 to 30 June 2006.* Journal of clinical microbiology, 2008. **46**(9): p. 2959-2965.

Kulkarni, R., K. Lole, and S.D. Chitambar, *Seroprevalence of antibodies against GII.4 norovirus among children in Pune, India.* J Med Virol, 2016.

Kumar, S. and S. Subramanian, *Mutation rates in mammalian genomes.* Proc Natl Acad Sci U S A, 2002. **99**(2): p. 803-8.

Kumazaki, M. and S. Usuku, *Genetic Analysis of Norovirus GII.4 Variant Strains Detected in Outbreaks of Gastroenteritis in Yokohama, Japan, from the 2006-2007 to the 2013-2014 Seasons.* PLoS One, 2015. **10**(11): p. e0142568.

Kundu, S., et al., *Next-generation whole genome sequencing identifies the direction of norovirus transmission in linked patients.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2013. **57**(3): p. 407-414.

Lai, M.M., *RNA recombination in animal and plant viruses.* Microbiological reviews, 1992. **56**(1): p. 61-79.

Laine, J., et al., *The duration of gastrointestinal and joint symptoms after a large waterborne outbreak of gastroenteritis in Finland in 2007--a questionnaire-based 15-month follow-up study.* PLoS One, 2014. **9**(1): p. e85457.

Lam, T.T., et al., *The recombinant origin of emerging human norovirus GII.4/2008: intra-genotypic exchange of the capsid P2 domain.* The Journal of general virology, 2012. **93**(Pt 4): p. 817-822.

Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nature methods, 2012. **9**(4): p. 357-359.

Lau, C.S., et al., *High rate and changing molecular epidemiology pattern of norovirus infections in sporadic cases and outbreaks of gastroenteritis in Hong Kong.* Journal of medical virology, 2004. **73**(1): p. 113-117.

Le Pendu, J., et al., *Mendelian resistance to human norovirus infections.* Semin Immunol, 2006. **18**(6): p. 375-86.

LeBlanc, J.J., et al., *Outbreak of Norovirus GII.P17-GII.17 in the Canadian Province of Nova Scotia.* Canadian Journal of Infectious Diseases and Medical Microbiology, 2016. **2016**: p. 6.

Lee, B.E., et al., *Determination of the relative economic impact of different molecular-based laboratory algorithms for respiratory viral pathogen detection, including Pandemic (H1N1), using a secure web based platform.* Virology journal, 2011. **8**: p. 277-422X-8-277.

Lee, B.E., et al., *Genetic relatedness of noroviruses identified in sporadic gastroenteritis in children and gastroenteritis outbreaks in northern Alberta.* Journal of medical virology, 2008. **80**(2): p. 330-337.

Lee, R.M., et al., *Incubation periods of viral gastroenteritis: a systematic review.* BMC Infect Dis, 2013. **13**: p. 446.

Leen, E.N., et al., *A Conserved Interaction between a C-Terminal Motif in Norovirus VPg and the HEAT-1 Domain of eIF4G Is Essential for Translation Initiation.* PLoS Pathog, 2016. **12**(1): p. e1005379.

Leonard, G., J.R. Stevens, and T.A. Richards, *REFGEN and TREENAMER: automated sequence data handling for phylogenetic analysis in the genomic era.* Evolutionary bioinformatics online, 2009. **5**: p. 1-4.

Lessler, J., et al., *Evidence for antigenic seniority in influenza A (H3N2) antibody responses in southern China.* PLoS Pathog, 2012. **8**(7): p. e1002802.

Lew, J.F., et al., *Detection of Norwalk virus or Norwalk-like virus infections in Finnish infants and young children.* The Journal of infectious diseases, 1994. **169**(6): p. 1364-1367.

Lin, X., et al., *Subgenomic promoter recognition by the norovirus RNA-dependent RNA polymerases.* Nucleic Acids Res, 2015. **43**(1): p. 446-60.

Lindesmith, L., et al., *Cellular and humoral immunity following Snow Mountain virus challenge.* Journal of virology, 2005. **79**(5): p. 2900-2909.

Lindesmith, L., et al., *Human susceptibility and resistance to Norwalk virus infection.* Nature medicine, 2003. **9**(5): p. 548-553.

Lindesmith, L.C., et al., *Immunogenetic mechanisms driving norovirus GII.4 antigenic variation.* PLoS pathogens, 2012. **8**(5): p. e1002705.

Lindesmith, L.C., et al., *Emergence of a norovirus GII.4 strain correlates with changes in evolving blockade epitopes.* Journal of virology, 2013. **87**(5): p. 2803-2813.

Lindesmith, L.C., et al., *Monoclonal antibody-based antigenic mapping of norovirus GII.4-2002.* J Virol, 2012. **86**(2): p. 873-83.

Lindesmith, L.C., et al., *Heterotypic humoral and cellular immune responses following Norwalk virus infection.* J Virol, 2010. **84**(4): p. 1800-15.

Lindesmith, L.C., E.F. Donaldson, and R.S. Baric, *Norovirus GII.4 strain antigenic variation.* Journal of virology, 2011. **85**(1): p. 231-242.

Lindesmith, L.C., et al., *Particle conformation regulates antibody access to a conserved GII.4 norovirus blockade epitope.* J Virol, 2014. **88**(16): p. 8826-42.

Lindesmith, L.C., et al., *Mechanisms of GII.4 norovirus persistence in human populations.* PLoS medicine, 2008. **5**(2): p. e31.

Lindesmith, L.C., et al., *Broad blockade antibody responses in human volunteers after immunization with a multivalent norovirus VLP candidate vaccine: immunological analyses from a phase I clinical trial.* PLoS Med, 2015. **12**(3): p. e1001807.

Lindsay, L., et al., *A decade of norovirus disease risk among older adults in upper-middle and high income countries: a systematic review.* BMC Infect Dis, 2015. **15**: p. 425.

Liu, B.L., et al., *Identification of further proteolytic cleavage sites in the Southampton calicivirus polyprotein by expression of the viral protease in E. coli.* The Journal of general virology, 1999. **80 (Pt 2)**(Pt 2): p. 291-296.

Liu, X., et al., *Seroepidemiology of Norovirus GII.3 and GII.4 Infections in Children with Diarrhea in Xi'an, China.* Foodborne Pathog Dis, 2015. **12**(6): p. 500-5.

Lochridge, V.P. and M.E. Hardy, *Snow Mountain virus genome sequence and virus-like particle assembly.* Virus Genes, 2003. **26**(1): p. 71-82.

Lofgren, E., et al., *Influenza seasonality: underlying causes and modeling theories.* J Virol, 2007. **81**(11): p. 5429-36.

Lole, K.S., et al., *Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination.* J Virol, 1999. **73**(1): p. 152-60.

Lopez-Manriquez, E., et al., *Norovirus genome circularization and efficient replication are facilitated by binding of PCBP2 and hnRNP A1.* J Virol, 2013. **87**(21): p. 11371-87.

Lopman, B., et al., *Host, weather and virological factors drive norovirus epidemiology: time-series analysis of laboratory surveillance data in England and Wales.* PLoS One, 2009. **4**(8): p. e6671.

Lopman, B.A., et al., *Clinical manifestation of norovirus gastroenteritis in health care settings.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2004. **39**(3): p. 318-324.

Lopman, B.A., et al., *Norovirus Infection and Disease in an Ecuadorian Birth Cohort: Association of Certain Norovirus Genotypes With Host FUT2 Secretor Status.* J Infect Dis, 2015. **211**(11): p. 1813-21.

Lu, J., et al., *Gastroenteritis Outbreaks Caused by Norovirus GII.17, Guangdong Province, China, 2014-2015.* Emerg Infect Dis, 2015. **21**(7): p. 1240-2.

Ludwig, A., et al., *Quantitative detection of norovirus excretion in pediatric patients with cancer and prolonged gastroenteritis and shedding of norovirus.* Journal of medical virology, 2008. **80**(8): p. 1461-1467.

Luttermann, C. and G. Meyers, *A bipartite sequence motif induces translation reinitiation in feline calicivirus RNA.* J Biol Chem, 2007. **282**(10): p. 7056-65.

Mai, H., et al., *GII.4 Sydney_2012 norovirus infection in immunocompromised patients in Beijing and its rapid evolution in vivo.* Journal of medical virology, 2015.

Majowicz, S.E., et al., *Magnitude and distribution of acute, self-reported gastrointestinal illness in a Canadian community.* Epidemiol Infect, 2004. **132**(4): p. 607-17.

Majowicz, S.E., et al., *Estimating the under-reporting rate for infectious gastrointestinal illness in Ontario.* Can J Public Health, 2005. **96**(3): p. 178-81.

Makison Booth, C., *Vomiting Larry: a simulated vomiting system for assessing environmental contamination from projectile vomiting related to norovirus infection.* J Infect Prev, 2014. **15**(5): p. 176-180.

Mamanova, L., et al., *Target-enrichment strategies for next-generation sequencing.* Nature methods, 2010. **7**(2): p. 111-118.

Marionneau, S., et al., *ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world.* Biochimie, 2001. **83**(7): p. 565-573.

Marionneau, S., et al., *Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals.* Gastroenterology, 2002. **122**(7): p. 1967-77.

Marks, P.J., et al., *Evidence for airborne transmission of Norwalk-like virus (NLV) in a hotel restaurant.* Epidemiol Infect, 2000. **124**(3): p. 481-7.

Marks, P.J., et al., *A school outbreak of Norwalk-like virus: evidence for airborne transmission.* Epidemiol Infect, 2003. **131**(1): p. 727-36.

Marshall, J.K., et al., *Postinfectious irritable bowel syndrome after a food-borne outbreak of acute gastroenteritis attributed to a viral pathogen.* Clin Gastroenterol Hepatol, 2007. **5**(4): p. 457-60.

Martella, V., et al., *Evidence for recombination between the pandemic GII.4 norovirus strains New Orleans 2009 and Sydney 2012.* Journal of clinical microbiology, 2013.

Martin, D.P., et al., *RDP3: a flexible and fast computer program for analyzing recombination.* Bioinformatics (Oxford, England), 2010. **26**(19): p. 2462-2463.

Matsushima, Y., et al., *Genetic analyses of GII.17 norovirus strains in diarrheal disease outbreaks from December 2014 to March 2015 in Japan reveal a novel polymerase sequence and amino acid substitutions in the capsid region.* Euro Surveill, 2015. **20**(26).

May, J., et al., *The p4-p2' amino acids surrounding human norovirus polyprotein cleavage sites define the core sequence regulating self-processing order.* J Virol, 2014. **88**(18): p. 10738-47.

Mayo, M.A., *Virus taxonomy - Houston 2002.* Arch Virol, 2002. **147**(5): p. 1071-6.

McAtee, C.L., et al., *Burden of Norovirus and Rotavirus in Children After Rotavirus Vaccine Introduction, Cochabamba, Bolivia.* Am J Trop Med Hyg, 2016. **94**(1): p. 212-7.

McCartney, S.A., et al., *MDA-5 recognition of a murine norovirus.* PLoS Pathog, 2008. **4**(7): p. e1000108.

Medici, M.C., et al., *Identification of the novel Kawasaki 2014 GII.17 human norovirus strain in Italy, 2015.* Euro Surveill, 2015. **20**(35): p. 30010.

Menon, V.K., et al., *Comparison of age-stratified seroprevalence of antibodies against norovirus GII in India and the United Kingdom.* PLoS One, 2013. **8**(2): p. e56239.

Miller, M.S., et al., *Neutralizing antibodies against previously encountered influenza virus strains increase over time: a longitudinal analysis.* Sci Transl Med, 2013. **5**(198): p. 198ra107.

Milne, I., et al., *Tablet--next generation sequence assembly visualization.* Bioinformatics (Oxford, England), 2010. **26**(3): p. 401-402.

Miyoshi, T., et al., *Long-term viral shedding and viral genome mutation in norovirus infection.* Journal of medical virology, 2015. **87**(11): p. 1872-1880.

Mongkolsapaya, J., et al., *Original antigenic sin and apoptosis in the pathogenesis of dengue hemorrhagic fever.* Nat Med, 2003. **9**(7): p. 921-7.

Mori, K., et al., *Genomic analysis of the evolutionary lineage of Norovirus GII.4 from archival specimens during 1975-1987 in Tokyo.* J Med Virol, 2016.

Morillo, S.G., et al., *Norovirus 3rd Generation kit: an improvement for rapid diagnosis of sporadic gastroenteritis cases and valuable for outbreak detection.* J Virol Methods, 2011. **173**(1): p. 13-6.

Motomura, K., et al., *Identification of monomorphic and divergent haplotypes in the 2006-2007 norovirus GII/4 epidemic population by genomewide tracing of evolutionary history.* Journal of virology, 2008. **82**(22): p. 11247-11262.

Motomura, K., et al., *Divergent evolution of norovirus GII/4 by genome recombination from May 2006 to February 2009 in Japan.* Journal of virology, 2010. **84**(16): p. 8085-8097.

Motomura, K., et al., *Divergent Evolution of Norovirus GII/4 by Genome Recombination from May 2006 to February 2009 in Japan.* The Journal of Virology, 2010. **84**(16): p. 8085-8097.

Mouly, D., et al., *Description of two waterborne disease outbreaks in France: a comparative study with data from cohort studies and from health administrative databases.* Epidemiol Infect, 2016. **144**(3): p. 591-601.

Mukhi, S.N., et al., *DIAL: A Platform for real-time Laboratory Surveillance.* Online journal of public health informatics, 2010. **2**(3): p. 10.5210/ojphi.v2i3.3041. Epub 2010 Dec 23.

Murata, T., et al., *Prolonged norovirus shedding in infants <or=6 months of age with gastroenteritis.* Pediatr Infect Dis J, 2007. **26**(1): p. 46-9.

Neill, J.D., *Nucleotide sequence of a region of the feline calicivirus genome which encodes picornavirus-like RNA-dependent RNA polymerase, cysteine protease and 2C polypeptides.* Virus Res, 1990. **17**(3): p. 145-60.

Nelson, A.M., et al., *Disruption of the human gut microbiota following Norovirus infection.* PLoS One, 2012. **7**(10): p. e48224.

Newman, K.L. and J.S. Leon, *Norovirus immunology: Of mice and mechanisms.* Eur J Immunol, 2015. **45**(10): p. 2742-57.

Ng, K.K., et al., *Crystal structure of norwalk virus polymerase reveals the carboxyl terminus in the active site cleft.* J Biol Chem, 2004. **279**(16): p. 16638-45.

Nice, T.J., et al., *A single-amino-acid change in murine norovirus NS1/2 is sufficient for colonic tropism and persistence.* J Virol, 2013. **87**(1): p. 327-34.

Nicollier-Jamot, B., et al., *Molecular cloning, expression, self-assembly, antigenicity, and seroepidemiology of a genogroup II norovirus isolated in France.* Journal of clinical microbiology, 2003. **41**(8): p. 3901-3904.

Nilsson, M., et al., *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.* Journal of virology, 2003. **77**(24): p. 13117-13124.

Noel, J.S., et al., *Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995.* Journal of medical virology, 1997. **53**(4): p. 372-383.

Numata, K., et al., *Epidemiological study of Norwalk virus infections in Japan and Southeast Asia by enzyme-linked immunosorbent assays with Norwalk virus capsid protein produced by the baculovirus expression system.* Journal of clinical microbiology, 1994. **32**(1): p. 121-126.

Nurminen, K., et al., *Prevalence of norovirus GII-4 antibodies in Finnish children.* Journal of medical virology, 2011. **83**(3): p. 525-531.

O'Donnell, T.B., et al., *Mouse Norovirus infection promotes autophagy induction to facilitate replication but prevents final autophagosome maturation.* Virology, 2016. **492**: p. 130-9.

O'Ryan, M.L., et al., *Symptomatic and asymptomatic rotavirus and norovirus infections during infancy in a Chilean birth cohort.* Pediatr Infect Dis J, 2009. **28**(10): p. 879-84.

Obinata, K., et al., *Norovirus encephalopathy in a previously healthy child.* Pediatr Infect Dis J, 2010. **29**(11): p. 1057-9.

Okada, M., et al., *Genetic analysis of noroviruses associated with fatalities in healthcare facilities.* Archives of Virology, 2006. **151**(8): p. 1635-1641.

Okhuysen, P.C., et al., *Viral shedding and fecal IgA response after Norwalk virus infection.* The Journal of infectious diseases, 1995. **171**(3): p. 566-569.

Oriol, R., *Genetic control of the fucosylation of ABH precursor chains. Evidence for new epistatic interactions in different cells and tissues.* J Immunogenet, 1990. **17**(4-5): p. 235-45.

Pang, X.L., J.K. Preiksaitis, and B. Lee, *Multiplex real time RT-PCR for the detection and quantitation of norovirus genogroups I and II in patients with acute gastroenteritis.* Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, 2005. **33**(2): p. 168-171.

Pang, X.L., J.K. Preiksaitis, and B.E. Lee, *Enhanced enteric virus detection in sporadic gastroenteritis using a multi-target real-time PCR panel: A one-year study.* Journal of medical virology, 2014. **86**(9): p. 1594-1601.

Pang, X.L., et al., *Influence of novel norovirus GII.4 variants on gastroenteritis outbreak dynamics in Alberta and the Northern Territories, Canada between 2000 and 2008.* PloS one, 2010. **5**(7): p. e11599.

Parra, G.I. and K.Y. Green, *Genome of Emerging Norovirus GII.17, United States, 2014.* Emerging infectious diseases, 2015. **21**(8): p. 1477-1479.

Parrino, T.A., et al., *Clinical immunity in acute gastroenteritis caused by Norwalk agent.* The New England journal of medicine, 1977. **297**(2): p. 86-89.

Patel, M.M., et al., *Systematic literature review of role of noroviruses in sporadic gastroenteritis.* Emerging infectious diseases, 2008. **14**(8): p. 1224-1231.

Payne, D.C., et al., *Norovirus and medically attended gastroenteritis in U.S. children.* N Engl J Med, 2013. **368**(12): p. 1121-30.

Pelizzo, G., et al., *Isolated colon ischemia with norovirus infection in preterm babies: a case series.* J Med Case Rep, 2013. **7**: p. 108.

Pelosi, E., et al., *The seroepidemiology of genogroup 1 and genogroup 2 Norwalk-like viruses in Italy.* Journal of medical virology, 1999. **58**(1): p. 93-99.

Pfister, T. and E. Wimmer, *Polypeptide p41 of a Norwalk-like virus is a nucleic acid-independent nucleoside triphosphatase.* J Virol, 2001. **75**(4): p. 1611-9.

Phan, T.G., et al., *Genetic heterogeneity, evolution, and recombination in noroviruses.* Journal of medical virology, 2007. **79**(9): p. 1388-1400.

Pires, S.M., et al., *Aetiology-Specific Estimates of the Global and Regional Incidence and Mortality of Diarrhoeal Diseases Commonly Transmitted through Food.* PLoS One, 2015. **10**(12): p. e0142927.

Platts-Mills, J.A., et al., *Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED).* Lancet Glob Health, 2015. **3**(9): p. e564-75.

Plitt AA, O.M., Sahni V, Lee BE, Charlton C, Simmonds K., *Examination of a Prenatal Syphilis Screening Program, Alberta, Canada: 2010-2011.* Canadian Journal of Public Health. In press, 2016.

Prasad, B.V., et al., *X-ray crystallographic structure of the Norwalk virus capsid.* Science (New York, N.Y.), 1999. **286**(5438): p. 287-290.

Prasad, B.V., et al., *Three-dimensional structure of baculovirus-expressed Norwalk virus capsids.* J Virol, 1994. **68**(8): p. 5117-25.

Qin, J., et al., *A human gut microbial gene catalogue established by metagenomic sequencing.* Nature, 2010. **464**(7285): p. 59-65.

Qin, M., et al., *A Waterborne Gastroenteritis Outbreak Caused by Norovirus GII.17 in a Hotel, Hebei, China, December 2014.* Food Environ Virol, 2016.

Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics (Oxford, England), 2010. **26**(6): p. 841-842.

Ramachandran, S., et al., *Temporal variations in the hepatitis C virus intrahost population during chronic infection.* Journal of virology, 2011. **85**(13): p. 6369-6380.

Reeck, A., et al., *Serological correlate of protection against norovirus-induced gastroenteritis.* The Journal of infectious diseases, 2010. **202**(8): p. 1212-1218.

Reed, G.F., B.D. Meade, and M.C. Steinhoff, *The reverse cumulative distribution plot: a graphic method for exploratory analysis of antibody data.* Pediatrics, 1995. **96**(3 Pt 2): p. 600-3.

Riddle, M.S. and R.I. Walker, *Status of vaccine research and development for norovirus.* Vaccine, 2016. **34**(26): p. 2895-9.

Robilotti, E., S. Deresinski, and B.A. Pinsky, *Norovirus.* Clin Microbiol Rev, 2015. **28**(1): p. 134-64.

Robles, J.D., et al., *Norovirus infection in pediatric hematopoietic stem cell transplantation recipients: incidence, risk factors, and outcome.* Biol Blood Marrow Transplant, 2012. **18**(12): p. 1883-9.

Rockx, B., et al., *Characterization of the homo- and heterotypic immune responses after natural norovirus infection.* Journal of medical virology, 2005. **77**(3): p. 439-446.

Rockx, B., et al., *Natural history of human calicivirus infection: a prospective cohort study.* Clin Infect Dis, 2002. **35**(3): p. 246-53.

Rockx, B.H., et al., *Association of histo-blood group antigens and susceptibility to norovirus infections.* Journal of Infectious Diseases, 2005. **191**(5): p. 749-754.

Roddie, C., et al., *Allogeneic hematopoietic stem cell transplantation and norovirus gastroenteritis: a previously unrecognized cause of morbidity.* Clin Infect Dis, 2009. **49**(7): p. 1061-8.

Roth, A.N. and S.M. Karst, *Norovirus mechanisms of immune antagonism.* Curr Opin Virol, 2016. **16**: p. 24-30.

Rouhani, S., et al., *Norovirus Infection and Acquired Immunity in 8 Countries: Results From the MAL-ED Study.* Clin Infect Dis, 2016. **62**(10): p. 1210-7.

Ruether, I.G., et al., *Characterization of novel intergenogroup and intergenotype recombinant noroviruses from central Greece.* Mol Cell Probes, 2014. **28**(4): p. 204-10.

Ryder, R.W., et al., *Evidence of immunity induced by naturally acquired rotavirus and Norwalk virus infection on two remote Panamanian islands.* The Journal of infectious diseases, 1985. **151**(1): p. 99-105.

Saito, M., et al., *Multiple norovirus infections in a birth cohort in a Peruvian Periurban community.* Clin Infect Dis, 2014. **58**(4): p. 483-91.

Sakai, Y., et al., *Clinical severity of Norwalk virus and Sapporo virus gastroenteritis in children in Hokkaido, Japan.* Pediatr Infect Dis J, 2001. **20**(9): p. 849-53.

Sarvestani, S.T., et al., *Norovirus Infection: Replication, Manipulation of Host, and Interaction with the Host Immune Response.* J Interferon Cytokine Res, 2016. **36**(4): p. 215-25.

Scallan, E., et al., *Foodborne illness acquired in the United States--major pathogens.* Emerging infectious diseases, 2011. **17**(1): p. 7-15.

Scanlan, P.D. and J.R. Marchesi, *Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces.* The ISME journal, 2008. **2**(12): p. 1183-1193.

Scharff, R.L., *Economic burden from health losses due to foodborne illness in the United States.* J Food Prot, 2012. **75**(1): p. 123-31.

Schmieder, R. and R. Edwards, *Quality control and preprocessing of metagenomic datasets.* Bioinformatics (Oxford, England), 2011. **27**(6): p. 863-864.

Schorn, R., et al., *Chronic norovirus infection after kidney transplantation: molecular evidence for immune-driven viral evolution.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2010. **51**(3): p. 307-314.

Shankarappa, R., et al., *Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection.* Journal of virology, 1999. **73**(12): p. 10489-10502.

Sharp, T.M., et al., *Secretory pathway antagonism by calicivirus homologues of Norwalk virus nonstructural protein p22 is restricted to noroviruses.* Virol J, 2012. **9**: p. 181.

Sharp, T.M., et al., *Inhibition of cellular protein secretion by norwalk virus nonstructural protein p22 requires a mimic of an endoplasmic reticulum export signal.* PLoS One, 2010. **5**(10): p. e13130.

Siebenga, J.J., et al., *High prevalence of prolonged norovirus shedding and illness among hospitalized patients: a model for in vivo molecular evolution.* The Journal of infectious diseases, 2008. **198**(7): p. 994-1001.

Siebenga, J.J., et al., *Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006.* Journal of virology, 2007. **81**(18): p. 9932-9941.

Siebenga, J.J., et al., *Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007.* Journal of Infectious Diseases, 2009. **200**(5): p. 802-812.

Siebenga, J.J., et al., *Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007.* The Journal of infectious diseases, 2009. **200**(5): p. 802-812.

Siffel, C., C.J. Alverson, and A. Correa, *Analysis of seasonal variation of birth defects in Atlanta.* Birth defects research.Part A, Clinical and molecular teratology, 2005. **73**(10): p. 655-662.

Simmonds, P., et al., *Bioinformatic and functional analysis of RNA secondary structure elements among different genera of human and animal caliciviruses.* Nucleic Acids Res, 2008. **36**(8): p. 2530-46.

Simmons, K., et al., *Duration of immunity to norovirus gastroenteritis.* Emerging infectious diseases, 2013. **19**(8): p. 1260-1267.

Simon-Loriere, E. and E.C. Holmes, *Why do RNA viruses recombine?* Nature reviews.Microbiology, 2011. **9**(8): p. 617-626.

Son, H., et al., *Seroepidemiology of predominant norovirus strains circulating in Korea by using recombinant virus-like particle antigens.* Foodborne pathogens and disease, 2013. **10**(5): p. 461-466.

Sosnovtsev, S.V., et al., *Feline calicivirus VP2 is essential for the production of infectious virions.* J Virol, 2005. **79**(7): p. 4012-24.

Sosnovtsev, S.V. and K.Y. Green, *Identification and genomic mapping of the ORF3 and VPg proteins in feline calicivirus virions.* Virology, 2000. **277**(1): p. 193-203.

Souza, M., et al., *Pathogenesis and immune responses in gnotobiotic calves after infection with the genogroup II.4-HS66 strain of human norovirus.* J Virol, 2008. **82**(4): p. 1777-86.

Stuart, R.L., et al., *An outbreak of necrotizing enterocolitis associated with norovirus genotype GII.3.* Pediatric Infectious Disease Journal, 2010. **29**(7): p. 644-647.

Sukhrie, F.H., et al., *Chronic shedders as reservoir for nosocomial transmission of norovirus.* Journal of clinical microbiology, 2010. **48**(11): p. 4303-4305.

Tam, C.C. and S.J. O'Brien, *Economic Cost of Campylobacter, Norovirus and Rotavirus Disease in the United Kingdom.* PLoS One, 2016. **11**(2): p. e0138526.

Tam, C.C., et al., *Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice.* Gut, 2012. **61**(1): p. 69-77.

Tamura, K., et al., *MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.* Mol Biol Evol, 2011. **28**(10): p. 2731-9.

Tamura, K., et al., *MEGA6: Molecular Evolutionary Genetics Analysis version 6.0.* Molecular biology and evolution, 2013. **30**(12): p. 2725-2729.

Tamura, M., et al., *Interaction of recombinant norwalk virus particles with the 105-kilodalton cellular binding protein, a candidate receptor molecule for virus attachment.* J Virol, 2000. **74**(24): p. 11589-97.

Tamura, M., et al., *Genogroup II noroviruses efficiently bind to heparan sulfate proteoglycan associated with the cellular membrane.* J Virol, 2004. **78**(8): p. 3817-26.

Tan, M. and X. Jiang, *The p domain of norovirus capsid protein forms a subviral particle that binds to histo-blood group antigen receptors.* Journal of virology, 2005. **79**(22): p. 14017-14030.

Tan, M. and X. Jiang, *Norovirus-host interaction: implications for disease control and prevention.* Expert Rev Mol Med, 2007. **9**(19): p. 1-22.

Tan, M. and X. Jiang, *Norovirus-host interaction: multi-selections by human histo-blood group antigens.* Trends Microbiol, 2011. **19**(8): p. 382-8.

Taube, S., et al., *A mouse model for human norovirus.* MBio, 2013. **4**(4).

Teunis, P.F., et al., *Norwalk virus: how infectious is it?* Journal of medical virology, 2008. **80**(8): p. 1468-1476.

Thackray, L.B., et al., *Murine noroviruses comprising a single genogroup exhibit biological diversity despite limited sequence divergence.* Journal of virology, 2007. **81**(19): p. 10460-10473.

Thomas, M.K., et al., *Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006.* Foodborne pathogens and disease, 2013. **10**(7): p. 639-648.

Thornley, C.N., et al., *Recurring norovirus transmission on an airplane.* Clin Infect Dis, 2011. **53**(6): p. 515-20.

Thorven, M., et al., *A homozygous nonsense mutation (428G-->A) in the human secretor (FUT2) gene provides resistance to symptomatic norovirus (GGII) infections.* Journal of virology, 2005. **79**(24): p. 15351-15355.

Thys, K., et al., *Performance assessment of the Illumina massively parallel sequencing platform for deep sequencing analysis of viral minority variants.* Journal of virological methods, 2015. **221**: p. 29-38.

Treanor, J.J., et al., *A Novel Intramuscular Bivalent Norovirus Virus-Like Particle Vaccine Candidate-Reactogenicity, Safety, and Immunogenicity in a Phase 1 Trial in Healthy Adults.* The Journal of infectious diseases, 2014.

Treanor, J.J., et al., *Subclass-specific serum antibody responses to recombinant Norwalk virus capsid antigen (rNV) in adults infected with Norwalk, Snow Mountain, or Hawaii virus.* J Clin Microbiol, 1993. **31**(6): p. 1630-4.

Trivedi, T.K., et al., *Clinical characteristics of norovirus-associated deaths: a systematic literature review.* American Journal of Infection Control, 2013. **41**(7): p. 654-657.

Tung-Thompson, G., et al., *Persistence of human norovirus RT-qPCR signals in simulated gastric fluid.* Food Environ Virol, 2015. **7**(1): p. 32-40.

Turcios-Ruiz, R.M., et al., *Outbreak of necrotizing enterocolitis caused by norovirus in a neonatal intensive care unit.* J Pediatr, 2008. **153**(3): p. 339-44.

van Beek, J., et al., *Indications for worldwide increased norovirus activity associated with emergence of a new variant of genotype II.4, late 2012.* Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin, 2013. **18**(1): p. 8-9.

van Beek, J., et al., *Comparison of norovirus genogroup I, II, and IV seroprevalence among children in the Netherlands, 1963, 1983, and 2006.* J Gen Virol, 2016.

van Beek, J., et al., *Noronet report, April 2013*. 2013, National Institute for Public Health and the Environment: Bilthoven, The Netherlands.

van Beek, J.B., D.; Kroneman, A.; Vennema, H.; Koopmans, M., *Noronet report, April 2012*. 2012, National Institute for Public Health and the Environment: Bilthoven, The Netherlands.

van Beek, J.K., A.; Vennema, H.; Koopmans, M., *Noronet report, April 2014.* 2014.

Van den Hoecke, S., et al., *Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing.* BMC genomics, 2015. **16**: p. 79-015-1284-z.

Van Trang, N., et al., *Association between norovirus and rotavirus infection and histo-blood group antigen types in Vietnamese children.* J Clin Microbiol, 2014. **52**(5): p. 1366-74.

Vashist, S., et al., *Identification of RNA-protein interaction networks involved in the norovirus life cycle.* J Virol, 2012. **86**(22): p. 11977-90.

Vega, E., et al., *Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013.* J Clin Microbiol, 2014. **52**(1): p. 147-55.

Vega, E., et al., *Novel surveillance network for norovirus gastroenteritis outbreaks, United States.* Emerging infectious diseases, 2011. **17**(8): p. 1389-1395.

Vega, E., et al., *RNA populations in immunocompromised patients as reservoirs for novel norovirus variants.* Journal of virology, 2014. **88**(24): p. 14184-14196.

Verhoef, L., et al., *Norovirus genotype profiles associated with foodborne transmission, 1999-2012.* Emerg Infect Dis, 2015. **21**(4): p. 592-9.

Verhoef, L., et al., Emerging Infectious Diseases, 2010. **16**(4): p. 617-624.

Vinje, J., *Advances in laboratory methods for detection and typing of norovirus.* Journal of clinical microbiology, 2015. **53**(2): p. 373-381.

Vinje, J., R.A. Hamidjaja, and M.D. Sobsey, *Development and application of a capsid VP1 (region D) based reverse transcription PCR assay for genotyping of genogroup I and II noroviruses.* Journal of virological methods, 2004. **116**(2): p. 109-117.

Vongpunsawad, S., B.V. Venkataram Prasad, and M.K. Estes, *Norwalk Virus Minor Capsid Protein VP2 Associates within the VP1 Shell Domain.* Journal of virology, 2013. **87**(9): p. 4818-4825.

Walker, A.W., et al., *Dominant and diet-responsive groups of bacteria within the human colonic microbiota.* The ISME journal, 2011. **5**(2): p. 220-230.

Walter, S.D. and J.M. Elwood, *A test for seasonality of events with a variable population at risk.* British journal of preventive & social medicine, 1975. **29**(1): p. 18-21.

Wasserman, E. *Vaxart launches PhI clinical trial for norovirus tablet vaccine*. Vaccines August 28, 2016; Available from: http://www.fiercepharma.com/vaccines/vaxart-launches-phi-clinical-trial-for-norovirus-tablet-vaccine.

Waugh, E., et al., *Characterization of the chemokine response of RAW264.7 cells to infection by murine norovirus.* Virus Res, 2014. **181**: p. 27-34.

Weerasekara, S., A.M. Prior, and D.H. Hua, *Current tools for norovirus drug discovery.* Expert Opin Drug Discov, 2016. **11**(6): p. 529-41.

Weiskopf, D., B. Weinberger, and B. Grubeck-Loebenstein, *The aging of the immune system.* Transpl Int, 2009. **22**(11): p. 1041-50.

Wheeler, J.G., et al., *Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive.* BMJ, 1999. **318**(7190): p. 1046-50.

White, L.J., et al., *Attachment and entry of recombinant Norwalk virus capsids to cultured human and animal cell lines.* J Virol, 1996. **70**(10): p. 6589-97.

Widdowson, M.A., et al., *Detection of serum antibodies to bovine norovirus in veterinarians and the general population in the Netherlands.* J Med Virol, 2005. **76**(1): p. 119-28.

Wobus, C.E., et al., *Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages.* PLoS Biol, 2004. **2**(12): p. e432.

Wong, T.H., et al., *Whole genome sequencing and de novo assembly identifies Sydney-like variant noroviruses and recombinants during the winter 2012/2013 outbreak in England.* Virology journal, 2013. **10**: p. 335-422X-10-335.

Wyatt, R.G., et al., *Comparison of three agents of acute infectious nonbacterial gastroenteritis by cross-challenge in volunteers.* J Infect Dis, 1974. **129**(6): p. 709-14.

Wyatt, R.G., et al., *Experimental infection of chimpanzees with the Norwalk agent of epidemic viral gastroenteritis.* J Med Virol, 1978. **2**(2): p. 89-96.

Wyn-Jones, A.P., et al., *Surveillance of adenoviruses and noroviruses in European recreational waters.* Water Res, 2011. **45**(3): p. 1025-38.

Xia, M., T. Farkas, and X. Jiang, *Norovirus capsid protein expressed in yeast forms virus-like particles and stimulates systemic and mucosal immunity in mice following an oral administration of raw yeast extracts.* J Med Virol, 2007. **79**(1): p. 74-83.

Yang, Y., et al., *Genetic and phenotypic characterization of GII-4 noroviruses that circulated during 1987 to 2008.* J Virol, 2010. **84**(18): p. 9595-607.

Yen, C., et al., *Impact of an emergent norovirus variant in 2009 on norovirus outbreak activity in the United States.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2011. **53**(6): p. 568-571.

Yi, J., et al., *Molecular epidemiology of norovirus in children and the elderly in Atlanta, Georgia, United States.* J Med Virol, 2016. **88**(6): p. 961-70.

Youmans, B.P., et al., *Characterization of the human gut microbiome during travelers' diarrhea.* Gut Microbes, 2015. **6**(2): p. 110-9.

Yuen, L.K., et al., *Heminested multiplex reverse transcription-PCR for detection and differentiation of Norwalk-like virus genogroups 1 and 2 in fecal samples.* Journal of clinical microbiology, 2001. **39**(7): p. 2690-2694.

Yunus, M.A., et al., *The murine norovirus core subgenomic RNA promoter consists of a stable stem-loop that can direct accurate initiation of RNA synthesis.* J Virol, 2015. **89**(2): p. 1218-29.

Zahorsky, J., *Hyperemesis heimis or the winter vomiting disease.* Arch. Pediatr., 1929. **46**.

Zamyatkin, D.F., et al., *Binding of 2'-amino-2'-deoxycytidine-5'-triphosphate to norovirus polymerase induces rearrangement of the active site.* J Mol Biol, 2009. **390**(1): p. 10-6.

Zeger, S.L., R. Irizarry, and R.D. Peng, *On time series analysis of public health and biomedical data.* Annual Review of Public Health, 2006. **27**: p. 57-79.

Zeitler, C.E., M.K. Estes, and B.V. Venkataram Prasad, *X-ray crystallographic structure of the Norwalk virus protease at 1.5-A resolution.* J Virol, 2006. **80**(10): p. 5050-8.

Zhang, T., et al., *RNA viral community in human feces: prevalence of plant pathogenic viruses.* PLoS biology, 2006. **4**(1): p. e3.

Zhang, X.F., et al., *An outbreak caused by GII.17 norovirus with a wide spectrum of HBGA-associated susceptibility.* Sci Rep, 2015. **5**: p. 17687.

Zheng, D.P., et al., *Norovirus classification and proposed strain nomenclature.* Virology, 2006. **346**(2): p. 312-323.

Zheng, D.P., et al., *Molecular epidemiology of genogroup II-genotype 4 noroviruses in the United States between 1994 and 2006.* Journal of clinical microbiology, 2010. **48**(1): p. 168-177.

Zhou, N., et al., *A waterborne norovirus gastroenteritis outbreak in a school, eastern China.* Epidemiol Infect, 2016. **144**(6): p. 1212-9.

Zhu, S., et al., *Identification of immune and viral correlates of norovirus protective immunity through comparative study of intra-cluster norovirus strains.* PLoS pathogens, 2013. **9**(9): p. e1003592.

Zintz, C., et al., *Prevalence and genetic characterization of caliciviruses among children hospitalized for acute gastroenteritis in the United States.* Infection, Genetics & Evolution, 2005. **5**(3): p. 281-290.

Appendix A

Summary of sequences removed and trimmed by Prinseq-lite

	OU1	OU2	OU3	OU4	SP1	SP2
Input sequences (file 1)	3,685,256	2,351,517	4,693,004	4,268,184	2,945,827	1,454,695
Input bases (file 1)	426,641,653	266,202,021	545,416,436	507,672,961	338,358,182	171,507,272
Input mean length (file 1)	115.77	113.2	116.22	118.94	114.86	117.9
Input sequences (file 2)	3,685,256	2,351,517	4,693,004	4,268,184	2,945,827	1,454,695
Input bases (file 2)	426,775,056	266,278,768	545,459,092	507,711,032	338,837,175	171,556,197
Input mean length (file 2)	115.81	113.24	116.23	118.95	115.02	117.93
Good sequences (pairs)	3,681,077	2,348,340	4,687,886	4,258,245	2,890,662	1,437,327
Good bases (pairs)	844,386,819	526,586,840	1,079,589,109	1,003,842,284	657,718,003	335,796,292
Good mean length (pairs)	229.39	224.24	230.29	235.74	227.53	233.63
Good sequences (singletons file 1)	3,759 (0.10%)	2,953 (0.13%)	4,367 (0.09%)	9,106 (0.21%)	54,719 (1.86%)	16,772 (1.15%)
Good bases (singletons file 1)	438,318	342,983	509,727	1,076,074	6,247,672	1,968,153
Good mean length (singletons file 1)	116.6	116.15	116.72	118.17	114.18	117.35
Good sequences (singletons file 2)	235 (0.01%)	96 (0.00%)	195 (0.00%)	261 (0.01%)	108 (0.00%)	127 (0.01%)
Good bases (singletons file 2)	27,051	11,100	22,608	30,488	12,389	14,675
Good mean length (singletons file 2)	115.11	115.62	115.94	116.81	114.71	115.55
Bad sequences (file 1)	420 (0.01%)	224 (0.01%)	751 (0.02%)	833 (0.02%)	446 (0.02%)	596 (0.04%)
Bad bases (file 1)	37,959	17,203	41,760	53,431	26,444	55,618
Bad mean length (file 1)	90.38	76.8	55.61	64.14	59.29	93.32
Bad sequences (file 2)	3,944 (0.11%)	3,081 (0.13%)	4,923 (0.10%)	9,678 (0.23%)	55,057 (1.87%)	17,241 (1.19%)
Bad bases (file 2)	462,920	361,209	546,244	1,119,542	6,601,351	2,061,632
Bad mean length (file 2)	117.37	117.24	110.96	115.68	119.9	119.58
Sequences filtered by trim_qual_left	276	223	1,110	1,088	623	368
Sequences filtered by min_qual_mean	4,088	3,082	4,564	9,423	54,880	17,469

Appendix B

Commands used for NGS data analysis

Quality filtering and trimming or MiSeq sequence data using Prinseq-lite:

$ perl prinseq-lite.pl -fastq file1.fastq -fastq2 file2.fastq -trim_qual_right 20 -trim_qual_left 20 -trim_qual_type min -trim_qual_rule lt -trim_qual_window 1 -trim_qual_step 1 -seq_id file_IDs -seq_id_mappings -graph_data file_data -graph_stats ld,gc,qd,ns,pt,ts -log file_log -min_qual_mean 20 -ns_max_p 90 -trim_left 1 -trim_right 1 -out_good file_good -out_bad file_bad

Filtering rRNA reads with SortMeRNA:

1) to merge paired-end reads:

    $ bash merge-paired-reads.sh file_1.fastq file_2.fastq file_1_2.fastq

2) to filter reads using all rRNA databases:

    $ sortmerna --ref silva-euk-28s-id98.fasta,silva-euk-28s-id98-db:silva-euk-18s-id95.fasta,silva-euk-18s-id95-db:silva-bac-23s-id98.fasta,silva-bac-23s-id98-db:silva-bac-16s-id90.fasta,silva-bac-16s-id90-db:silva-arc-23s-id98.fasta,silva-arc-23s-id98-db:silva-arc-16s-id95.fasta,silva-arc-16s-id95-db:rfam-5s-database-id98.fasta,rfam-5s-database-id98-db:rfam-5.8s-database-id98.fasta,rfam-5.8s-database-id98-db --reads file_1_2.fastq --aligned file_1_2.fastq_rRNA.fastq --other file_1_2.fastq_norRNA.fastq fastq --log --paired_out

3) to separate paired-end reads:

    $ bash unmerge-paired-reads.sh file_1_2.fastq_norRNA.fastq file_1_norRNA.fastq file_2_norRNA.fastq

De novo assembly of Alberta404/2012 with Velvet:

2) to run velveth for multiple hash lengths (15 to 75 with steps of 4) and a long reference sequence

    $ velveth 15_75_longcontig/ 15,75,4 -fastq -shortPaired file_1_norRNA.fastq file_2_norRNA.fastq -long reference_sequence.fastq

3) to run velvetg for each hash length specified in step 2:

    $ velvetg 15_75_longcontig/_15 -ins_length 226 -ins_length_sd 102 -exp_cov 5

Creating alignments of NoV with Bowtie:

1) to index reference sequence:

$ bowtie2-build reference_sequence.fasta reference_sequence

3) to create alignments (SAM files):

$ bowtie2 --local -x reference_sequence -1 file_1_norRNA.fastq -2 file_2_norRNA.fastq --un-conc file_norRNAnoNoV_% -S file_NoV.sam

Calculating coverage per site with Bedtools:

1) to prepare a sorted BAM file from a SAM file:

$ samtools view -Sb file_NoV.sam > file_NoV.bam

$ samtools sort file_NoV.bam file_NoV_sorted.bam

2) to prepare a genome file: created a text file with the information as follows: <chromName><TAB><chromSize>. For example: OU1 7532

3) to run beedtools:

$ bedtools genomecov -d -ibam file_NoV_sorted.bam –g genome_file > output_file_bedtools

Analysis of SNPs with FreeBayes:

$ freebayes -f reference_sequence.fasta –K --haplotype-length 1 --min-alternate-fraction 0.02 -m 10 -q 20 --min-coverage 10 -C 5 --debug file_NoV_sorted.bam > output_file_freebayes

Characterization of non-rRNA-non-NoV reads with BLAST:

$ blastn -task megablast -db nt -query file_norRNAnoNoV_1.fasta -outfmt '6 qseqid qlen saccver evalue qlen length pident staxids sscinames scomnames sskingdoms' –best_hit_overhang 0.1 –best_hit_score_edge 0.1 -max_target_seqs 1 -out blast_output_file

Appendix C

Spearman's rho correlation coefficient for % NoV reads

Parameter	Spearman's rho correlation coefficient	*P* (two-tailed)
Ct value	-0.886	0.019*
NoV copies per ng of RNA	0.886	0.019*
NoV copies per uL of RNA	0.886	0.019*
RNA concentration	0.086	0.872

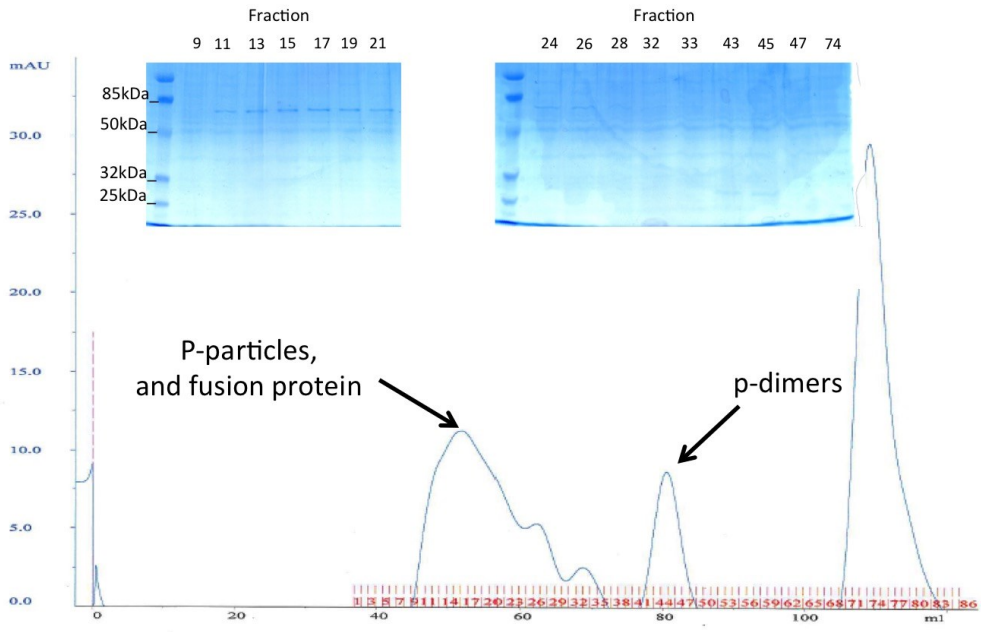
*Significant at the 0.05 level

Data used for the analysis:

Sample	% NoV reads	Ct value	NoV copies (per uL of RNA)	RNA concentration by Nanodrop (ng/uL)	NoV copies (per ng of RNA)
OU1	0.51	20.58	484,264	129.48	3,740.1
OU2	1.88	19.21	1,162,504	142.72	8,145.3
OU3	0.01	28.90	2,376	161.88	14.7
OU4	0.02	22.70	124,904	86	1,452.4
SP1	1.78	21.92	205,632	97.88	2,100.9
SP2	0.33	26.72	9,560	93.12	102.7

Appendix D

Identification of p-dimer formation by size exclusion chromatography and SDS-PAGE. The product of (GST)-P fusion protein cleaved by thrombin and purified on Glutathione Sepharose 4 Fast Flow (Amersham Biosciences) was run through a gel filtration column (Superdex 200, Amersham Biosciences) to visualize p-particle and p-dimer formation as previously described [296]. SDS-PAGE results of fractions obtained during the gel filtration process are shown for reference.



Identification of p-dimer formation by saliva-binding assay. The ability of the expressed P-protein to bind HBGA antigens was evaluated using a panel of saliva from individuals with O, A, B and non-secretor phenotypes. Serial dilutions of the protein extract produced dose response-curves indicative of HBGA binding to type A, B and O individuals. Concordantly with the HBGA binding pattern of GII.4 NoV strains, the p-protein extract did not bind saliva of a non-secretor.