Human Caliciviruses: a study of viral evolution, host genetics and disease susceptibility

Beatrice Carlsson
Cover illustration made by Rada Ellegård. Front page illustrates sapovirus particles, while the backside is an electron microscopy image of norovirus particles.

Department of Clinical and Experimental Medicine
Faculty of Health Sciences
SE- 581 85 Linköping, Sweden
Printed by Liu-tryck, Linköping, Sweden 2012
Published articles have been reprinted with the permission from respective copyright holder
ISBN: 978-91-7519-922-1
ISSN: 0345-0082
“Nothing in biology makes sense except in the light of evolution”
Theodosius Dobzhansky, 1973

“Discovery consists of looking at the same thing as everyone else and thinking something different”
Albert Szent-Gyorgy, Nobel Prize in Physiology or Medicine in 1937

To Wilhelm and Andreas,

to my beloved dad,
IN MEMORIAM
Åke Karlsson1938-2011
Supervisor:
Lennart Svensson, Professor
Division of Molecular Virology
Department of Clinical and Experimental Medicine
Linköping University, Linköping, Sweden.

Opponent:
Jan Albert, Professor
Department of Microbiology, Cell and tumor biology,
Karolinska Institute (KI), Stockholm, Sweden.

Committee Board:
Pia Forsberg, Professor
Division of Clinical Immunology,
Department of Clinical and Experimental Medicine,
Linköping University, Linköping, Sweden.

Sigvard Olofsson, Professor
Department of Clinical Virology,
University of Gothenburg, Göteborg, Sweden

Kristina Broliden, Professor
Department of Medicine,
Division of Infectious Diseases,
Center for Molecular Medicine,
Karolinska Institute (KI), Karolinska University Hospital,
Stockholm, Sweden
Articles and manuscripts included in this thesis:

I. Quasispecies dynamics and molecular evolution of human norovirus capsid P region during chronic infection.  
   J Gen Virol. 2009 Feb;90(Pt 2):432-41

II. The G428A nonsense mutation in FUT2 provides strong but not absolute protection against symptomatic GII.4 norovirus infection.  
   Carlsson B*, Kindberg E*, Buesa J, Rydell GE, Lidón MF, Montava R, Abu Mallouh R,  
   *These authors contributed equally to this work.  

III. Pediatric sapovirus infections in Nicaragua and host genetic susceptibility.  
   *These authors contributed equally to this work.  
   Submitted.

IV. Susceptibility to symptomatic sapovirus infection in Denmark is not associated with secretor or Lewis status.  
   Carlsson B, Larsson G, Böttiger B, Svensson L.  
   Submitted.

Related publications not included in this thesis:

Asymptomatic norovirus infections in Nicaraguan children and its association with viral properties and histo-blood group antigens.  

Pediatric norovirus diarrhea in Nicaragua.  
Bucardo F, Nordgren J, Carlsson B, Paniagua M, Lindgren PE, Espinoza F, Svensson L.  

Bucardo F*, Karlsson B*, Nordgren J, Paniagua M, González A, Amador JJ, Espinoza F,  
Svensson L.  
*These authors contributed equally to this work.  
Table of contents

Populärvetenskaplig sammanfattning på svenska .............................................. 1
Abstract ...................................................................................................................... 3
Abbreviations ............................................................................................................ 4
Introduction .............................................................................................................. 7
  The global burden of gastroenteritis ...................................................................... 7
  History of viral gastroenteritis ................................................................................ 8
Caliciviruses ............................................................................................................. 8
  Classification .......................................................................................................... 8
  Norovirus ................................................................................................................ 10
    Genome and capsid structure ............................................................................... 10
    Antigenicity .......................................................................................................... 13
  Sapoviruses ............................................................................................................. 14
    Genome and capsid structure ............................................................................... 14
    Antigenicity .......................................................................................................... 16
Symptoms ................................................................................................................. 17
  Diagnosis .............................................................................................................. 17
  Treatment .............................................................................................................. 18
Pathogenesis and pathology .................................................................................... 19
  Animal and cell culture model .............................................................................. 19
  Viral entry and primary replication ...................................................................... 20
  Cell tropism and release ...................................................................................... 20
Histological effects .................................................................................................. 22
  Physiology of the small intestine ......................................................................... 22
  Regulation of intestinal function .......................................................................... 23
  Diarrhea ................................................................................................................ 23
  Vomiting ............................................................................................................... 24
Epidemiology ............................................................................................................ 25
  Prevalence ............................................................................................................ 25
Susceptibility ............................................................................................................ 27
  Volunteer studies .................................................................................................. 27
Histo blood group antigens and host genetics ......................................................... 28
  PUT2 and secretor status ...................................................................................... 28
  A and B antigens .................................................................................................. 30
  PUT3 and Lewis status .......................................................................................... 30
  Norovirus .............................................................................................................. 31
  Sapovirus ............................................................................................................. 32
Evolution of caliciviruses ........................................................................................ 34
  Quasi species ........................................................................................................ 34
  Survival of the fittest ............................................................................................. 36
  Error threshold ..................................................................................................... 36
  Herd immunity and epochal evolution of NoV ..................................................... 39
  Norovirus vaccine ............................................................................................... 42
Aim of thesis ............................................................................................................ 44
Materials .................................................................................................................... 45
Methods ..................................................................................................................... 47
  Virus detection and capsid genotyping (Paper I, II, III and IV) ......................... 47
  Lewis and ABO phenotyping of saliva and blood (Paper II, III and IV) .......... 50
Populärvetenskaplig sammanfattning på svenska

I mina doktorsstudier har jag undersökt infektionsmönstret hos sapovirus och norovirus, två virus som tillhör gruppen humana calicivirus vilka orsakar gastroenterit ("magsjuka") hos människor. Gastroenterit är ett globalt hälsoproblem och orsakar häpnadsväckande nog mer än 5 gånger flera dödsfall jämfört med HIV/AIDS hos barn. Enbart norovirus som ger "vinterkräksjuka", orsakar över 200 000 dödsfall/år hos barn under 5 års ålder.

Humana calicivirus kan infektera människor i alla åldrar men är speciellt bekymmersamt för spådbarn, äldre och kroniskt sjuka vilka löper ökad risk att drabbas av svår uttorkning. Norovirusets mkt låga smitt dos (ett tiotal virusspartiklar) sammanaget med dess motståndsskick mot rengöringsmedel gör att vinterkräksjuka orsakar stora problem på sjukhus och utgör en stor merkostnad för samhället. Norovirus är dessutom en mycket vanlig orsak till livsmedelsutbrott. Livsmedel som är associerade med norovirus är vatten, ostron, bär och sallader. Inte sällan förekommer utbrotten efter besök på restauranger, kryssningsfartyg, skolor etc. Sveriges smittskyddsinsitut (SMI) beräknar att mellan 500 000 till 1 milj svenskar har drabbats av vinterkräksjuka varje vintersäsong under de senaste åren.

Oturligt nog så vet man väldigt lite om vilka faktorer som påverkar en människas mottaglighet för calicivirusinfektion. Frågor som "vem blir infekterad och varför?" är ännu inte helt klara. Dessa frågor är väldigt viktiga att besvara eftersom de påverkar både förståelsen av sjukdomsförlopp och därmed också vilka individer det är som mest mottagar risk för norovirusinfektion.

Tidigare studier har visat att ca 20% av en befolkningen i Europa och Nordamerika aldrig eller sällan drabbas av vinterkräksjuka. För att förklara detta fenomen brukar man dela in en population i två grupper; sekretorpositiva och sekretornegativa. Man brukar på gennivå sekvensera FUT2 genen på kromosom 19 och se om man har en mutation i position 428. Grovt förenklat kan man säga att personer som har en mutation kallas icke-sekretorer eller non secretors, eftersom de har ett ickefungerande FUT2 enzym och därmed inte kan binda (eller "secrete"=eng. "utsändra) de sockerstrukturer som man tror att norovirus använder som receptor. Dessa individer (ca 20% av Sveriges befolkning) är därför immun eller mindre mottagliga för norovirusinfektion. De individer som istället är sekretorer eller secretors (resterande 80% av Sveriges befolkning), kan med hjälp av sitt FUT2 enzym binda (utsändra) de sockerstrukturer som norovirus använder som receptorer och därmed är dessa individer mottagliga för infektion.

Man vet väldigt lite om vilka receptorer som humana calicivirus använder vid infektion, men man har dock sett vissa mönster när det gäller vilka personer som infekteras. Ett exempel är Norwalk virus, en GI1 stam av norovirus, som man vet helst infekterar personer som har blodgrupp A och O, men av någon anledning inte infekterar individer av blodgrupp B. Liknande kopplingar har gjorts i andra studier och det har gjort att man...
börjat misstänka att norovirus troligen kan binda till histoblodgruppsantigener, dvs i detta fall kohydraterna som definierar vilken blodgrupp man tillhör.

Dessa blodgruppsantigener kan utsöndras i tunntarmens mukosa (slemhinna) och det är också här man tror att norovirus infekterar. I fallet med Norwalk virus så tror man att viruset tar sig till mukosan och binder till sockerstrukturer (tex A antigenet) och därmed kan ta sig in i cellerna.

Stora delar av mina studier som är inkluderade i denna avhandling (Paper I, II, III och IV) är fokuserade på att karakterisera viruets ytterhölje eftersom det är starkt kopplat till receptorinteraktion och ofta är målstruktur för immunförsvar. I den första studien (Paper I) har jag undersökt hur norovirusets kapsid (ytterhölje) genomgår evolutionära förändringar hos en individ som är kroniskt norovirusinfekterad. Studien visar att dessa evolutionära förändringar ger upphov till en mängd nya virusvarianter, vilka i sin tur kan ge upphov till nya norovirusepidemier. I den andra studien (Paper II) beskriver jag en ny norovirus stam som häpnadsväckande nog även kan infektera personer som bär på den skyddade mutationen i FUT2 genen. Studie III och IV (Paper III och IV) är de första studierna som kartlägger sambandet mellan histoblodgruppsantigener och mottagligheten för sapovirusinfektion.

Sammanfattningsvis, så är studier av infektionsmönstret hos humana calcivirusinfektioner mycket viktigt eftersom omfattande utbrott kan få stora konsekvenser och orsaka höga kostnader för samhället och sjukvården. Ökad kunskap kring infektionsmönstret kan bidra till en snabbare diagnostik och att man i framtiden lättare kan identifiera riskpatienter. Därmed finns stora och betydelsefulla möjligheter till en helt annan beredskap och kunskap för att kunna begränsa och förebygga omfattande utbrott av vinterkräksjuka.
Abstract

The viruses described in this thesis are the norovirus and sapoviruses, which belong to the family of human caliciviruses and are known to cause gastroenteritis in humans. Gastroenteritis has emerged as a global health problem and is based on the large number of infected considered as one of the most common diseases today. According to estimates of the World Health Organization (WHO), diarrheal infections cause more than five times more deaths in children compared to HIV/AIDS worldwide. Norovirus the cause of the famous "winter vomiting disease", alone, cause more than 200 000 deaths each year in children less than 5 years of age.

The mechanism for emergence and evolution of new human calicivirus strains, as well as protective immunity in the human population is poorly understood. The main focus for this thesis was to elucidate the possible correlation between human calicivirus evolution, host genetics and disease susceptibility. One of the main findings presented in this thesis is the documentation of in vivo capsid gene evolution and quasispecies dynamics during chronic NoV GI.3 infection (Paper 1). In paper II, we reported that the G428A nonsense mutation in the FUT2 gene provides strong but not absolute protection against symptomatic GII.4 NoV infection. In my last two papers (Paper III and IV), we were the first to investigate host genetic susceptibility factors during authentic SaV infection.

To summarize, the results presented in this thesis show that the success of human calicivirus infection probably is determined by a delicate interplay between virus evolution and susceptibility of the host, both genetically and immunologically.
Abbreviations

aa          amino acid
(A)ₙ        poly A tail
cDNA        complimentary DNA
Cl⁻         chloride ion
3CLPro      3C-like protease
DNA         deoxyribonucleic acid
ELISA       enzyme-linked immunosorbent assay
EC cells    enterochromaffine cells
EM          electron microscopy
ENS         enteric nervous system
ET          evolutionary trace
EtOH        ethanol
Fuc         fucose residue
FUT1        fucosyltransferase 1
FUT2        fucosyltransferase 2
FUT3        fucosyltransferase 3
gal         galactose
GalNac       N-acetylgalactosamine
GlcNac       N-acetylglucosamine
Gn          gnotobiotic
HBGA        histo blood group antigen
5-HT        serotonin
IAHA        immune adherence hemagglutination assay
IEM         immune electron microscopy
i.m.        intramuscularly
i.n.        intranasally
KI          Karolinska institutet
Leᵃ         lewis a antigen
Leᵇ         lewis b antigen
Leˢ         Lewis x antigen
Leʸ         Lewis y antigen
LUX         light upon extension
MNV         murine norovirus
MPL         monophosphoryl lipid A
NoV         norovirus
nt          nucleotide
ORF         open reading frame
PAML        phylogenetic analysis by maximum likelihood
PBMC        peripheral blood mononuclear cells
PCR         polymerase chain reaction
PEC         porcine enteric calcivirus
pol         polymerase
RIA         radio immune assay
RdRp        RNA dependent RNA polymerase
RHDV        rabbit hemorrhagic disease virus
RT          reverse transcriptase
RT-PCR      real-time PCR
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>rVP6</td>
<td>recombinant VP6 protein</td>
</tr>
<tr>
<td>Se&lt;sup&gt;(w)&lt;/sup&gt;</td>
<td>weak secretor</td>
</tr>
<tr>
<td>SMI</td>
<td>Smittskyddsinstitutet, Swedish Institute for Communicable Disease Control</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>ssRNA</td>
<td>single stranded RNA</td>
</tr>
<tr>
<td>SPIEM</td>
<td>solid phase immune electron microscopy</td>
</tr>
<tr>
<td>TV</td>
<td>tulane virus</td>
</tr>
<tr>
<td>VIP</td>
<td>vasoactive intestinal peptide</td>
</tr>
<tr>
<td>vPg</td>
<td>viral protein genome-linked</td>
</tr>
<tr>
<td>VP1</td>
<td>viral capsid protein 1</td>
</tr>
<tr>
<td>VP2</td>
<td>viral capsid protein 2</td>
</tr>
<tr>
<td>VSV</td>
<td>vesicular stomatitis virus</td>
</tr>
</tbody>
</table>
Introduction

The global burden of gastroenteritis

Gastroenteritis is based on the large number of infected, one of the most common diseases worldwide (Green et al. 2001). The disease affects humans of all ages but elderly and young children are often more severely infected. Globally, gastroenteritis next to pulmonary infections is the most common cause of death in children under 5 years of age with 1.8 million deaths each year (Bryce et al. 2005). According to estimates of the World Health Organization (WHO), diarrheal infections cause more than five times more deaths in children compared to HIV/AIDS worldwide, Figure 1.

![Figure 1. WHO estimates of causes of death in neonates and children younger than 5 years of age during 2000 and 2003 (Bryce et al. 2005).](image)

Nutrition status and sanitary conditions are a great factor in many cases determine the disease outcome, since more than 95% of all deaths caused by gastroenteritis occur in developing countries (Patel et al. 2008). As a consequence of the sanitary conditions, the agent causing gastroenteritis differs between developed and developing countries. Bacteria (enterotoxogenic E.coli; ETEC, enteropathogenic E.coli; EPEC, Shigella species and Campylobacter) and parasites (e.g. Entamoeba and Giardia) that are easily spread by contaminated food and water are the most common causes of gastroenteritis in developing countries (Michelangeli and Ruiz 2003). On the other hand, in developed countries, these agents only cause less than 5% of the total cases while enteric viruses (mainly rotaviruses, enteric adenoviruses, human calciviruses and astroviruses) constitutes the vast majority.
History of viral gastroenteritis

Today it is known that many different agents such as parasites, chemicals and toxins but also bacteria and viruses can cause acute gastroenteritis. However, in the beginning of the last century only a few bacterial and parasitic agents were discovered, and only a small portion of these could be linked to outbreaks of gastroenteritis. In 1945, Reimann and coworkers described an outbreak of “epidemic diarrhea, nausea and vomiting of unknown cause” (Reimann et al. 1945). Clues to the fact that viruses were the causing agents came from volunteer studies done during 1940ties and 1950ties, showing that bacteria free stool filtrates from infected individuals caused gastroenteritis in healthy volunteers (Gordon et al. 1947). However, it was not until Kapikian and coworkers in 1972 discovered the Norwalk virus (a member of the calicivirus family), that a virus was confirmed as a causing agent of gastroenteritis (Kapikian et al. 1972). Following improvement of the electron microscope methodology, several other enteric viruses such as rotavirus (Bishop et al. 1973), astroviruses (Madeley and Cosgrove 1975a; Madeley and Cosgrove 1975b), enteric adenoviruses (Wadell et al. 1987), sapoviruses and noroviruses (additional members of the human calicivirus family) were discovered.

Caliciviruses

The name “calicivirus” is derived from the Latin word Calyx, meaning “cup”, and refer to the characteristic cup-shaped depressions found on the surface of all caliciviruses. The family of Caliciviridae consists of small (≈27 to 45nm) icosahedral non enveloped positive sense single stranded RNA viruses that can infect a broad range of hosts (Green et al. 2001). The genome is organized in two or three open reading frames (ORFs) ranging in size from 7.3 to 8.3kb and have a VPg protein covalently attached to the 5´end and is polyadenylated at the 3´end (Burroughs and Brown 1978; Dunham et al. 1998).

Classification

Caliciviruses can be divided into 4 genera consisting of noroviruses (NoV), sapoviruses (SaV), vesiviruses, lagoviruses plus 2 newly proposed genera nabovirus and recovirus, Figure 2. Caliciviruses cause various disorders, however, to briefly summarize, NoV and SaV cause epidemic gastroenteritis in humans, vesivirus cause respiratory disease in cats (feline calicivirus, FCV) and lagovirus cause hemorrhagic fever in rabbits (rabbit hemorrhagic disease virus, RHDV). Nabovirus is an enteric virus causing gastroenteritis in cattle (newbury1, while recovirus is a virus of yet unknown pathogenicity isolated in rhesus macaques (Tulane virus, TV) (Oliver et al. 2006; Farkas et al. 2008).
SaV are genetically diverse and Oka and coworkers have recently on the basis of full capsid sequencing classified them into five genogroups; GI, GII, GIII, GIV, and GV (Oka et al. 2012). Human SaV are now classified into GI, GII, GIV, and GV, while GIII are found only in pigs. According to this novel classification, SaV GI and GII were then subsequently each subdivided into seven genotypes, while both GIV and GV only contained a single genotype. The GII SaV strains that infect pigs are able to multiply in cultured cells (Flynn and Saif 1988; Chang et al. 2004), however the human strains are uncultivable. Recently, a novel SaV genogroup consisting of Canine SaV has been suggested by Li and coworkers (Li et al. 2011). A SaV strain that infects minks has also been reported (Phan et al. 2007; Cunha et al. 2010).

Noroviruses consists of five genogroups (GI-GV), where GIII infects cattle, GV contain murine strains and GI and GII infect humans (Zheng et al. 2006). The GIV NoV infects lions (Martella et al. 2007) but also contain a canine calicivirus (Roerink et al. 1999; Martella et al. 2008; Mesquita et al. 2010).

The genogroups can also be further divided into genotypes, based on polymerase gene or nucleotide capsid sequence (Katayama et al. 2002). Currently, 14 genotypes are found in GI and 17 in GII NoV, based on complete capsid sequences (Zheng et al. 2006).
**Norovirus**

Norovirus (NoV) is a major cause of acute gastroenteritis, and are the causative agent of the "winter vomiting disease". Zahorsky first described this disease in 1929, but it was not until 1972 that Kapikian and coworkers identified the prototype Norwalk strain, in an elementary school in Norwalk, Ohio (Kapikian et al. 1972).

NoV infections can vary from being asymptomatic (no symptoms), to the other end of the scale, causing severe vomiting with diarrhea resulting in fatal dehydration. Symptomatic NoV infection can be divided into three degrees based on severity: mild, moderate and severe. NoV infections are a big problem all over the world, and approximately 96% of all acute non bacterial gastroenteritis is caused by NoV (Fankhauser et al. 1998; Siebenga et al. 2009). A study by Patel and coworkers estimated that only in industrialized countries, diarrhea caused by NoV was responsible for 900000 clinical visits and 64000 cases with severe diarrhea requiring hospitalization each year. In developing countries the consequences of NoV infections are far worse, up to 200 000 children less than 5 years of age die each year often due to severe dehydration (Patel et al. 2008). However, also in industrialized countries deaths occur, e.g. in England and Wales, 80 deaths/year due to NoV infection are estimated to occur in elderly individuals (≥65 years of age) (Harris et al. 2008).

Information about NoV in Central America is limited, however in 2008 we published a study investigating pediatric NoV infection in Nicaragua (list of publications concerning calicivirus outside this thesis). Through a pediatric diarrhea surveillance program undertaken in community and hospital, a total of 542 stool samples between March 2005 and February 2006 in León, Nicaragua, were collected (Bucardo et al. 2008). The study concluded that NoV was an important etiologic agent of acute gastroenteritis in children less than 2 years of age in Nicaragua.

**Genome and capsid structure**

NoV have an approximately 27-38 nm non enveloped icosahedral capsid, having a typical "star of David" morphology when seen in electron microscope, Figure 3.

![Electron microscope image of NoV](https://example.com/nov_image.png)

*Figure 3. Electron microscope image of NoV (by Lennart Svensson, Linköping University).*
Understanding the molecular biology of NoV was long delayed due to the low amounts of virus in stool samples of infected individuals, which led to difficulties in isolating the viruses. It was not until 1990 that the first cDNA clone was produced, and thus characterization of the NoV sequence genome organization begun (Xi et al. 1990; Matsui et al. 1991; Jiang et al. 1992; Lambden et al. 1993; Lew et al. 1994; Prasad et al. 1994; Prasad et al. 1999a).

The NoV has a genome of ≈7.5kb single stranded positive sense RNA, and the genomic RNA function as a messenger RNA (mRNA) and is organized into 3 ORFs (Green et al. 2001), Figure 4. The genome has a genome-linked viral protein (VPg) attached at the 5’ end, and a poly A tail in the 3’ end. The gene encoded by the longest ORF (ORF1) is translated as a non structural polyprotein, which cleaved by the viral protease 3CLpro, results in at least 6 known proteins; p48: protein 48, NTPase: nucleoside triphosphatase, p22: protein 22, VPg: viral protein genome-linked, 3CLpro: 3C-like protease and RdRp: RNA dependent RNA polymerase (Jiang et al. 1993). The second ORF (ORF2) encodes a structural polyepitope with a molecular weight of 58kD, capsid viral protein 1 (VP1), known as the NoV capsid protein. The third ORF (ORF3) encodes a minor structural protein, viral capsid protein 2 (VP2), of debated function (Jiang et al. 1993; Glass et al. 2000).

The 6 proteins of the ORF1 aid in the replication process, copying the positive (+) sense RNA into a negative (-) sense copy to be used to produce the positive (+) sense subgenomic RNA of ORF2 and ORF3. This subgenomic RNA is then translated into multiple copies of VP1 and a few copies of VP2 that later can be assembled into a capsid. The NoV capsid consists of 180 copies of VP1, organized into 90 dimers (Prasad et al.
Prasad et al. 1999a; Chen et al. 2004b). The capsid has mainly two domains that are connected via a flexible hinge: the shell domain (S) and the protruding domain (P). The protruding domain that can be further subdivided into P1 and P2, have a high sequence variability between genotypes compared to the conserved S-domain. The S-domain forms the foundation of the icosahedral shell of the capsid, while the P-domain constitutes the outer protruding parts.

The NoV capsid has a T=3 icosahedral symmetry that are formed via a very refined arrangement of the capsid protein, VP1 (Prasad et al. 1996). Figure 5. An icosahedral structure can be described as a many sided, three dimensional, hexagonal shape made up of many small triangles. To obtain this shape, the capsid protein must adopt three slightly different conformations, termed quasi equivalent positions, referred to as A, B and C (Chen et al. 2004b). Three of these quasi equivalent subunits constitute an asymmetric subunit. Sixty copies of these asymmetric subunits form the icosahedron, enabled by the network of interactions formed when the A, B and C subunits of each asymmetric subunit form dimers with the neighboring asymmetric subunits (Prasad et al. 1999b; Chen et al. 2004b). One example of these interactions is the P domain of the A and B subunits that interact across the quasi twofold axes, forming the protruding regions giving the characteristic “cup-like” shape to the capsid.

The ability of the capsid protein to self-assemble into recombinant NoV virus-like particles (rNoV VLPs) when expressed in insect cells, have been widely used in various studies (Jiang et al. 1992; Mason et al. 1996; Rydell et al. 2009b; Koho et al. 2012). These VLPs are morphologically and antigenically similar to native NoV, which have made them an invaluable tool for characterization of potential viral-host interactions (Green et al. 1993; Marionneau et al. 2002; Rydell et al. 2009a; Lindesmith et al. 2011). VLPs induce a specific antibody response in infected volunteers, and can be used as antigens in enzyme-linked immunosorbent assays (ELISAs) (Green et al. 1993; Parker et al. 1993; Rydell et al. 2009b). The VLPs are safe to handle, since they are not infectious due to the absence of the RNA genome.
Antigenicity
The arrangement with a capsid built by multiple copies of a single protein is a common feature for plant viruses, but except for Nodaviridae (a family of insect viruses) caliciviruses are the only known animal viruses having such structure (Hosur et al. 1987). The viral capsid is the only part of the virus that is exposed to the environment and is able to interact with a potential host. Since the NoV capsid is composed only of a single protein, VP1, consequently, this protein not only contains all the determinants for sustaining the structure but also govern immunogenicity and infectivity. Due to this fact, studies of the capsid protein are usually the key target when investigating host susceptibility to viral infections, and the reason why I have studied the capsid of the GII.3 and GII.4 NoV strains in paper I and II, respectively.
Sapoviruses

Human SaV are mainly known to cause gastroenteritis among children although also adults and elderly may be infected (Noel et al. 1997; Chiba et al. 2000; Svraka et al. 2010). Epidemiological studies have showed that children (under the age of 5) in day-care centers and closed institutions were at the greatest risk to acquire SaV infection (Hansman et al. 2007a). However, there a very limited number of studies conducted on SaV compared to NoV, and therefore it is difficult to draw correct conclusions regarding the overall prevalence, antigenicity and binding specificities of SaV. SaV are usually the cause of sporadic but also food borne outbreaks of gastroenteritis (Okada et al. 2002; Dey et al. 2007; Hansman et al. 2007c; Iizuka et al. 2010). Previously, SaV outbreaks have been rarely reported, however, recently there have been reports of an emergence of SaV infections globally (Svraka et al. 2010). In studies carried out in United Kingdom, India, Finland, Denmark, Pakistan, Japan and Sweden in non-hospitalized individuals, SaV detection rates ranged from 2 to 10% (Pang et al. 2000; Phan et al. 2004; Akihara et al. 2005; Rachakonda et al. 2008; Johnsen et al. 2009; Cunliffe et al. 2010; Svraka et al. 2010). Another study has even suggested a detection rate of up to 29.9% (Nakanishi et al. 2011). In contrast, detection rates range from 0.5 to 1.4% in hospitalized children ≤5 years of age in United States, Thailand and Russia, while no SaV infection was observed in Japan and Tunisia (Sakai et al. 2001; Zintz et al. 2005; Khamrin et al. 2007; Sdiri-Loulizi et al. 2008; Podkolzin et al. 2009). However, compared to RV and NoV, the prevalence of SaV is rather low (Parashar et al. 2006; Patel et al. 2008). Symptoms of SaV infection usually are milder than symptoms of RV and NoV (Pang et al. 2000; Sakai et al. 2001), thus explaining why the detection rate of SaV infections is higher in non-hospitalized than in hospitalized children.

Genome and capsid structure

SaV were first described as causative agent of an outbreak of gastroenteritis 1977 in Sapporo, Japan (Chiba et al. 1979). However, it was not until 1995 that the first cDNA clone of SaV was produced enabling subsequent characterization of the genome (Matson et al. 1995). Recently, a full genomic sequence analysis of the original Sapporo virus strain (SV82) was obtained (Nakanishi et al. 2011).

SaV particles are approximately 41–46nm in diameter and have a typical Star of David appearance with cup-shaped depressions on the viral capsid, Figure 6.
The SaV have a single stranded positive sense RNA genome, consisting of two or three ORFs, Figure 5. The genomes of SaV GI, GIV, and GV are predicted to encode three ORFs, while the remaining genogroups (SaV GII and GIII) have two ORFs. The SaV genome has a VPg attached to the 5’ end, and a poly A tail at the 3’ end (Clarke and Lambden 1997; Clarke and Lambden 2000). Unlike the NoVs, the capsid protein of the SaVs is not encoded by a separate ORF; instead the capsid protein gene is within the same reading frame as the RNA polymerase gene creating one large fused 250 kDa polyprotein encoded by ORF1 (Clarke and Lambden 2000), Figure 7. Besides the capsid protein, ORF1 codes for at least 6 known proteins; p48: protein 48, NTPase: nucleoside triphosphatase, p22: protein 22, VPg: viral protein genome-linked, 3CLPro: 3C-like protease and RdRp: RNA dependent RNA polymerase. Just like NoV, the 3’ ORF (in this case ORF2) encodes a minor (165aa) structural protein. The presence of a third ORF, ORF3, encoding a small basic protein of 161 amino acids has been predicted. This protein is overlapping the capsid gene but is located within a different reading frame (Clarke and Lambden 2000). The function of the proteins encoded by ORF2 and ORF3 is unknown (Oka et al. 2006c).

The transcription and translation of the SaV genome is more complex than in the case of NoV, since both the SaV capsid and non structural genes are located within the same contiguous reading frame, ORF1 (Clarke and Lambden 2000). The process is poorly understood, but it must require at least one additional cleavage (compared to NoV) to release the mature RNA polymerase from the capsid protein (Oka et al. 2006c).
In similarity to NoV, expression of the SaV capsid protein leads to self assembly into VLPs which are morphologically and antigenically similar to native virions. So far, a total of 8 VLPs have been successfully produced; four GI strains: Sapporo (Numata et al. 1997), Houston/90 (Jiang et al. 1999), Parkville (Chen et al. 2004a) and Mc114 (Hansman et al. 2005a; Hansman et al. 2005b), two GII strain: C12 (Hansman et al. 2005b) and Mc10 (Oka et al. 2006a), one GII strain: PEC (Guo et al. 2001), and one GV strain: NK24 (Hansman et al. 2005b).

**Antigenicity**

Even though it has been many years since the first SaV was discovered, much remains to be elucidated regarding their pathogenicity and antigenicity. Such studies have been severely hampered due to the fact that SaV is uncultivable and that SaV virus like particles (VLPs) sometimes has proven difficult to obtain (Katayama et al. 2004; Hansman et al. 2007b).

In similarity to NoV, the capsid of SaV consists of multiple copies of a single capsid protein with a molecular weight of approximately 58 kDa. Thereby, in similarity to NoV, all the determinants for antigenicity and attachment most likely are within the SaV capsid protein (Clarke and Lambden 2000; Green et al. 2007). Most recently, Amin and coworkers using immunomics suggested putative T- and B-cell epitopes within the SaV capsid protein and mapped these locations to its predicted three-dimensional structure (Amin et al. 2011).

Not only antigenicity differs between the SaV genogroups, Oka and coworkers have reported that even SaV strains belonging to the same genogroup (GII SaV, but within different genetic clusters) have distinct antigenicity (Hansman et al. 2007b; Oka et al. 2009). It has been suggested that SaV might escape host immunity by using recombination within the capsid region (Phan et al. 2007). Katayama and coworkers have reported that recombination events have occurred between the Mc10 and C12, two GII SaV (Katayama et al. 2004). A recombination event between GII (Mc10) and GIV (SW278 and Ehime 1107) has been reported by Hansman and coworkers (Hansman et al. 2007b). Additional possible recombinations both between and within SaV genogroups, have recently been reported by Dos Anjos and coworkers (Dos Anjos et al. 2011).
Symptoms

The symptoms for SaV and NoV infection are similar; however, the symptoms of SaV infection usually are milder than symptoms of NoV (Pang et al. 2000; Sakai et al. 2001). After an incubation period in general 24 to 48 hours, symptoms such as projectile vomiting, abdominal pain and diarrhea occurs. Other symptoms include low grade fever, chills, headache and myalgias. Some recent reports also suggest that NoV infection in rare cases can be associated to acute renal failure (Kanai et al. 2010), encephalopathy (Ito et al. 2006), pneumatosis intestinalis (Kim et al. 2011), disseminated intravascular coagulation and even photophobia (CDC 2002).

Usually the disease is self-limiting after a few days, but chronic infections of immune compromised individuals have been reported (Siebenga et al. 2007a; Hoffmann et al. 2012), (paper I of this thesis).

Diagnosis

Diagnosis of NoV and SaV infection is primarily based on clinical features, but can be more accurately determined by molecular methods using PCR of viral RNA purified from stool samples of infected individuals. Diagnosis of NoV was initially made via the “Kaplan criteria”, Table 1 (Kaplan et al. 1982a; Kaplan et al. 1982b). These criteria were developed by Kaplan and coworkers in 1982, in a time with no molecular detection methods to identify outbreaks of non bacterial gastroenteritis. However, recent evaluation of these criteria has shown them still applicable (Turcios et al. 2006).

Table 1.

<table>
<thead>
<tr>
<th>Diagnosis of gastroenteritis causes by NoV: “Kaplan criteria”</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) No bacteria or parasites detected</td>
</tr>
<tr>
<td>b) Vomiting in more than 50% of cases</td>
</tr>
<tr>
<td>c) Mean duration of illness: 12 to 60 hours</td>
</tr>
<tr>
<td>d) Incubation period: 24 to 48 hours</td>
</tr>
</tbody>
</table>

Electron microscopy (EM) has long been a widely used diagnostic tool to identify virus particles in feces. The advantage of this method is that it is easy, quick to perform and may enable an instant determination of viral morphology. However, when using direct electron microscopy, a viral concentration of at least 10^6 per ml of stool is required, something that may be difficult to obtain even after concentration of the stool samples (Atmar and Estes 2001). Therefore, due to the limited amount of viral particles in stool samples from infected individuals, EM may be a somewhat uncertain diagnostic method. The sensitivity of the EM technique may be improved by using immune electron microscopy (IEM) where the viral particles are aggregated with antibodies and thereby concentrated. A further optimization of the method by higher sensitivity is solid-phase IEM (SPIEM) utilizing protein A where the viral particles are captured directly on to the EM grid (Svensson and von Bonsdorff 1982; Svensson et al. 1983).

From the 1970ies until present, a wide number of different diagnostic assays were designed to detect human calicivirus. These assays e. g. include immune adherence hemagglutination assay (IAHA) (Greenberg and Kapikian 1978), radio immune assay (RIA) and various enzyme linked immune sorbent assay (ELISAs) including western blot.
The immune assays used for detection have been optimized over the years to become more sensitive, e.g. by the use of monoclonal antibodies that are able to discriminate between different NoV genogroups.

Today, by far the most used molecular detection method for NoV and SaV is conventional PCR and real-time PCR (RT-PCR) performed on viral RNA purified from stool samples of infected individuals (Atmar and Estes 2001; Logan et al. 2007; van Maarseveen et al. 2010; Shigemoto et al. 2011; Wolfs et al. 2011). Several primer pairs have been designed to target conserved regions on both the polymerase and capsid genes. This is a very accurate and highly sensitive detection method compared to the ELISA based methods. Nordgren and coauthors in 2007 developed a novel light-upon-extension (LUX) RT-PCR assay for not only detection and quantification but also determining genogroup I and II NoV in a single run (Nordgren et al. 2008).

**Treatment**

At present there are no specific treatment available for human calicivirus infection, and in otherwise healthy well nourished individuals the infection is self-limiting. The key for an optimized recovery is to replace liquid loss, which is especially important for infants and young children to avoid dehydration (Anderson 2010). Oral rehydration is usually sufficient, but during prolonged periods of vomiting and severe diarrhea intravenous rehydration may be needed. A study by Steinhoff and coworkers, showed that oral administration of bismuth subsalicylate reduced the symptoms of gastroenteritis in adult volunteers infected with NoV (Steinhoff et al. 1980). A recent study by Siddiq and coworkers reports that the use of nitazoxanide, a broad-spectrum antimicrobial agent, has cleared NoV infection within an immune suppressed host (Siddiq et al. 2011). This drug may be an alternative to use when immune suppression must be sustained, for instance in organ transplant recipients.
Pathogenesis and pathology

Animal and cell culture model
Despite the fact that it has been known since the 1970ties that SaV and NoV cause gastroenteritis in humans, the actual pathogenesis and pathophysiology of a human calicivirus infection is still relatively unknown. Thus, lack of knowledge is mainly due to the fact that at present there are no convenient cell culture or small animal model available for human NoV and SaV, something that has hampered this field severely (Duizer et al. 2004). So far the only information available on human calicivirus is based on volunteer studies or derived from outbreak studies (Dolin et al. 1971; Agus et al. 1973; Schreiber et al. 1974; Wyatt et al. 1974; Widerlite et al. 1975; Parrino et al. 1977; Meeroff et al. 1980; Johnson et al. 1990).

Based on the knowledge that related calicivirus strains infect cattle, pigs and mice, many studies have focused on developing an experimental model using these animals. Recently, interest has also been focused on canine calicivirus (Roerink et al. 1999; Martella et al. 2008; Mesquita et al. 2010). The close genetic relationship of NoV and SaV found in pet animals and humans has caused debate concerning the risk of a possible zoonotic transmission (Bank-Wolf et al. 2010).

Wobus and coauthors reported a successful experimental system for murine NoV (MNV), where the MNV strain replicated well both in mice and in cell culture (Wobus et al. 2006). However, since MNV does not cause gastroenteritis in mice, this model system does not give many clues to the pathophysiology of the disease in humans. Studies on feline calicivirus and porcine enteric calicivirus (PEC) have, respectively, given clues to the in vitro replication and histological effects in the intestine, however, none of these is a human calicivirus.

In 2006, Cheetham and coworkers evaluated the use of gnotobiotic (Gn) pigs as a model to study the replication and pathogenesis of human NoV (Cheetham et al. 2006). They challenged the Gn pigs with fecal filtrates of the NoV/GI/4/HS66/2001/US strain, causing mild clinical disease and viral shedding. This study concludes that the tested human NoV strain least partially adapted to replication in the Gn pig host and that further adaptation may increase its pathogenicity in the pig.

Recently a study by Bok and coauthors, evaluated the use of chimpanzees as an animal model for NoV infection (Bok et al. 2011). They found that human NoV indeed can infect chimpanzees, resulting in a similar duration of shedding and level of replication as observed in humans. However, unfortunately one major drawback with this model is that the chimpanzees do not develop a clinical disease.

Several studies have been performed on development of cell culture models in mammalian cells (Katayama et al. 2006; Guix et al. 2007), and Asanaka et al reported in 2006 on a successful model system for replication and packaging of human NoV (Norwalk strain) viral RNA into virus particles (Asanaka et al. 2005). Unfortunately, none of the cell culture models so far have been able to support a complete replication system for human calicivirus. Yet another approach using tissue culture systems have also been evaluated (Vashist et al. 2009), but so far only Straub and coworkers have presented a model where they have cultivated human NoV in Caco-2 cells grown as a
three-dimensional tissue structure (Straub et al. 2007; Straub et al. 2011). However, this technique is very challenging and is no alternative for routine use.

**Viral entry and primary replication**
Both SaV and NoV are mainly spread via the fecal-oral route, and thus an infectious particle is entered via the mouth and then further transported to the digestive tract. To survive both the journey to the gastrointestinal tract, and the acidic environment of the stomach these enteric viruses are highly stable. Studies have shown that the NoV particle is stable during a variety of conditions such as freezing, temperature up to 60°C, pH in a range of 3<pH< 9 and 70% of ethanol (EtOH) (Cannon et al. 2006). Cannon and coworkers report that chlorine based solutions are the most effective disinfectant for inactivation of NoV.

Some epidemiologic data suggests an indirect spread of viruses via inhaled aerosol from projectile vomits during a gastroenteritis outbreak (Caul 1994; Chadwick and McCann 1994). Data from volunteer studies and natural infections, has shed some light into the mechanisms behind diarrhea caused by NoV (Dolin et al. 1971; Agus et al. 1973; Schreiber et al. 1974; Wyatt et al. 1974; Widerlite et al. 1975; Parrino et al. 1977; Meerroff et al. 1980; Johnson et al. 1990; Marionneau et al. 2002; Troeger et al. 2009). However, yet no such studies are available for SaV and it can only be assumed that SaV has a pathogenicity similar to NoV since they are structurally similar and both cause gastroenteritis.

The exact location for primary replication is still unknown, but it is often speculated to be located in the upper gastrointestinal tract; in duodenum or jejunum of the small intestine. However, a few studies have suggested a potential extra-intestinal spread and some cases of NoV RNAemia have been reported (Takanashi et al. 2009; Medici et al. 2010). Ito and coworkers have also reported on a case of putative NoV-associated encephalopathy (Ito et al. 2006).

Recently, Bok and coworkers reported that in chimpanzees infected with human NoV, the NoV infected cells were located exclusively in the lamina propria of the small intestine (Bok et al. 2011). Both the jejenum and duodenum were infected, however, the duration of infection was longer in the jejunum. It was also noted that in the jejunum and duodenum tissue sections, NoV was present in cells from the dendritic cell lineage. However, this finding is in contrast to a study by Lay and coworkers. They found that NoV does not replicate in human macrophages or dendritic cells (Lay et al. 2010). However, this does not exclude the possibility that NoV might replicate in a subset of intestinal DC-SIGN-positive cells. Therefore the possible correlation between NoV replication and human peripheral blood mononuclear cells (PBMCs) needs to be investigated further.

**Cell tropism and release**
Today the knowledge about NoV and SaV cell and tissue tropism is limited. White and coworkers used in 1996 radiolabelled NoV rVLPs to investigate virus-cell interactions (White et al. 1996). They found that the VLPs bound to a various amount of cell lines, but significantly more VLPs bound to differentiated human intestine Caco-2 cells compared...
to the other cell lines tested. In some of the cell lines the VLPs were also internalized, however, this was in all cases obtained to a very low extent.

Both SaV and NoV are released from the enteric tract of an infected host, via feces or vomits (Greenberg et al. 1979; Green et al. 1998). Viral shedding usually coincide with the onset and duration of symptoms in volunteers, however, viral shedding has also been observed for more than 7 days after the first symptoms (Thornhill et al. 1975; Graham et al. 1994; Schorn et al. 2010; Roos-Weil et al. 2011; Hoffmann et al. 2012).
Histological effects

Physiology of the small intestine

The term “small intestine” refers to the section of the gastrointestinal tract located between the stomach and the large intestine. The small intestine is divided into three structural parts: duodenum, jejunum, and ileum. The small intestine is the location where a large part of the digestion and absorption of nutrients occurs. However, all segments of intestine from duodenum to distal colon have mechanisms for both absorbing and secreting water and electrolytes (Field 2003). The surface area of the small intestine used for absorption and secretion is increased (about 600 fold) by the presence of villi (Michelangeli and Ruiz 2003). The villi are “folds” composed of columnar epithelial cells, mainly villus and crypt enterocytes and enteroendocrine cells. Each of the villi is lined with even smaller microvilli in order to increase surface area further, Figure 8. The enterocytes govern the secretion and absorption; secretion is performed by crypt cells and absorption occurs via the villus cells (Hodges and Gill 2010).

Figure 8. Structure of small intestine. The intestinal walls are lined with villus. Each of these villus are covered with microvilli, in order to increase the area used for absorption and secretion. Enterocytes are represented in green, enterochromaffine (EC) cells are in red while goblet cells are displayed in yellow. The figure was kindly provided by Rada Ellegård, Division of molecular virology, Linköping University.
Regulation of intestinal function

The regulation of intestinal functions are very delicate and is based on the interacting stimuli from enteroendocrine, immune and enteric nervous systems (ENS) (Lundgren and Svensson 2003; Michelangeli and Ruiz 2003). These three systems interact to “fine tune” the enterocyte secretion/absorption, gut motility and microcirculation. The ENS is often referred to as “the brain in the gut”, and is located in sheets of tissue lining the esophagus, stomach, small intestine and colon (Michelangeli and Ruiz 2003). These sheets are structurally divided in to the myenteric and the submucosal plexus. The myenteric plexus exerts control primarily over digestive tract motility, while the submucosal plexus has a role in sensing the environment within the lumen, regulating gastrointestinal blood flow, secretion and controlling epithelial cell function (Lundgren and Svensson 2003).

Diarrhea

Diarrhea and fluid loss may be caused by a number of different mechanisms (Field 2003). However to generalize, diarrhea is usually a problem of intestinal water and electrolyte balance caused by an excessive secretion and poor uptake (Hodges and Gill 2010). Osmotic diarrhea may occur when too much water is drawn into the bowel by osmotic active reagents. Water is excreted towards an osmotic gradient, caused by i.e malabsorption where nutrients are left in the lumen to pull in water (pancreatic disease or Coeliac disease) (Michelangeli and Ruiz 2003). Other substances may be vitamin C (access amounts), lactose, magnesium or sorbitol (used in sugar-free foods, chewing gum).

A second type of diarrhea is secretory diarrhea, caused by an overstimulation of the intestinal tract’s secretory system. This may be caused by an activation of the ENS leading to the release of various effector substances. These substances initiate signaling pathways causing e.g. opening of ion channels, release of chloride ions (Cl-) or vasoactive intestinal peptide (VIP) that all stimulate secretion and in the end cause diarrhea (Hodges and Gill 2010). VIP directly stimulates water- and electrolyte secretion by binding to crypt cell receptors.

A third type of diarrhea is exudative diarrhea. Exudative diarrhea may occur when the intestinal epithelial barrier function is damaged, and the hydrostatic pressure exerted in the blood or lymph vessels drives fluid loss over the epithelium.

Finally, the fourth type of diarrhea is caused by motility disturbances that change microcirculation, where both increased and decreased motility may cause an intestinal fluid loss. Many bacteria produce enterotoxins that via an activation of the ENS cause effects of the intestinal motility (Hodges and Gill 2010).

However, it is important to remark that in most cases of diarrhea caused by pathogens, more than one of the described pathophysiological mechanisms are involved. For example, cholera toxin stimulates both intestinal fluid secretion and motility (Lundgren and Jodal 1997; Fung et al. 2010).

Mechanisms by which virus and especially human calicivirus induce diarrhea are ill defined. Early volunteer studies from the 1970-ties revealed that the histological effects of a NoV infection occurred in the small intestine with an almost instant blunting and
broadening of the villi. This condition was however completely reversible, since biopsies taken after recovery were completely normal. Other typical features were also an infiltration of mono-nuclear cells and a distinct shortening of microvilli in the mucosa (Agus et al. 1973; Schreiber et al. 1973).

A recent study by Troeger and coworkers in 2009 showed that NoV infection causes an epithelial barrier dysfunction. This dysfunction includes loss of sealing protein in the tight junction, increase of epithelium apoptosis probably due to infiltration of cytotoxic cells (CD8+ T-cells) and an increased secretion of ions. This study concludes that NoV diarrhea is both explained by passive leaking and also an active transportation of ions over the tight junction (Troeger et al. 2009). Marionneau and coworkers showed by using recombinant virus like particles (rVLP) of Norwalk virus and tissue sections of the gastroduodenal junction that NoV binds to the villus and less to the crypts of the small intestine (Marionneau et al. 2002). An early study has shown that NoV infection only cause lesions in the small intestinal mucosa, while leaving the gastric mucosa intact. The same study also reported that the secretion of hydrochloric acid, pepsin and intrinsic factor remain at normal levels while the gastric motor function is altered during NoV infection (Meeroff et al. 1980).

**Vomiting**

In human and most of animals, vomiting is a protective mechanism that serves to remove toxins from the gastrointestinal tract prior to absorption, but in some species it is also used to eliminate the gut of indigestible material (Andrews and Horn 2006). Emesis might also be a response generated by cognitive, visual, or flavor stimuli and several categories of stimuli including chemotherapy, stress, intracranial pressure, and motion can produce incidental activation of the emetic reflex. The vagus nerve is an important pathway in the detection of emetic stimuli and generation of vomiting (Andrews and Horn 2006).

Projectile vomiting is one of the classical symptoms of NoV infection, but vomiting also occurs during SaV infection. The underlying mechanism for how human caliciviruses cause vomiting is however still unknown. It has been proposed that it is the abnormal gastric motor function with a delayed gastric emptying seen during NoV infection, that cause the characteristic nausea and vomiting (Meeroff et al. 1980). Another plausible mechanism for vomiting and nausea, is activation of the ENS. This is seen during rotavirus infection, and a recent study by Hagbom and coworkers reported that rotavirus stimulates release of serotonin (5-HT) from human enterochromaffin cells that activates brain structures involved in nausea and vomiting (Hagbom et al. 2011).
Epidemiology

The name "winter vomiting disease" refers to the epidemiology noted for NoV disease. Infection occurs all year round, but there is a marked increase of NoV cases during the winter months (Hedlund et al. 2000; Mounts et al. 2000; Siebenga et al. 2008a). Hedlund and coworkers performed a longitudinal survey of NoV outbreaks in Sweden between 1994 and 1998. This survey showed that NoV outbreaks were reported mainly during the winter and spring months (Hedlund et al. 2000), Figure 9. A recent study by Dey and coworkers have observed a seasonal pattern also for SaV infections (Dey et al. 2012). They reported that there was a peak of SaV infection in the cold season (November to March) in Japan during 2003-2009.

Due to short lived immunity, high viral diversity and multiple routes of transmission, NoV epidemiology is complex. Unfortunately, due to the lack of animal and cell culture models, studies of viral transmission under different environmental conditions cannot be performed yet on human caliciviruses. However, relative humidity and temperature (i.e. cool and dry conditions) have been identified to promote transmission of respiratory viruses in the laboratory as well as in human populations. Winter seasonal transmission has been reported for rotavirus, another common pathogen causing gastroenteritis, but also respiratory viruses and measles show this pattern (Mounts et al. 2000; Olofsson et al. 2011; Brittain-Long et al. 2012). Recently Lopman and coworkers performed a time series analysis of NoV laboratory surveillance data in England and Wales. Their data suggests that cool and dry weather, population immunity and viral evolution are the key players determining the complicated patterns of NoV epidemiology (Lopman et al. 2009).

Prevalence

The increased recognition of NoV as the cause of outbreaks and sporadic disease is probably due to the recent availability of improved NoV-specific diagnostics. Although strains can circulate simultaneously and different strains can circulate in distinct geographic regions at different times, genogroup GII.4 NoV are the most common cause of viral gastroenteritis worldwide (Fankhauser et al. 1998). This genotype has been the
most commonly detected since the end of the 1990ties, and cause outbreaks in nursing homes, daycare centers, cruise ships and hospitals (Green et al. 2007; Siebenga et al. 2007a; Tu et al. 2007; Siebenga et al. 2008a; Bok et al. 2009; Bull et al. 2010; Eden et al. 2010; Hoffmann et al. 2010; Boon et al. 2011).
Susceptibility

Transmission of NoV is facilitated by their high prevalence in the community, shedding of infectious virus particles from asymptomatic individuals and the high stability of the virus in the environment.

Volunteer studies

The very first volunteer study on gastroenteritis was performed by Gordon and coworkers in 1947 (Gordon et al. 1947). This study was performed subsequent to the occurrence of numerous gastroenteritis outbreaks of unknown cause that had taken place in New York State during the fall and winter of 1946-1947.

In order to learn the etiology of acute infectious nonbacterial gastroenteritis many volunteer studies were performed in the 1970-ties (Dolin et al. 1971; Agus et al. 1973; Schreiber et al. 1974; Wyatt et al. 1974; Widerlite et al. 1975; Parrino et al. 1977). One early volunteer study was performed by Dolin coworkers in 1971. In this study healthy male prisoners were challenged with inoculum prepared form stools or rectal swabs obtained from secondary cases of acute gastroenteritis obtained from four separate outbreaks (Dolin et al. 1971). The inoculum obtained from an outbreak in a primary school in Norwalk, Ohio, was the only inoculum that caused clinical disease. This strain was one year later identified as the prototype NoV, denoted as Norwalk strain (Kapikian et al. 1972). The Norwalk inoculum induced gastroenteritis in two out of three men. Later, oral administration of a stool filtrate from one of the sick volunteers induced gastroenteritis in seven of nine additional volunteers (Dolin et al. 1971). Subsequently, several other of these early volunteer studies confirmed that challenge with Norwalk virus conferred short-term but not long-term (>2 y) immunity to reinfection with the same virus (Dolin et al. 1972; Parrino et al. 1977).

In 1974, Wyatt and coworkers performed a cross-challenge study in volunteers, where they compared the infection pattern of three agents known to cause acute infectious nonbacterial gastroenteritis (Wyatt et al. 1974). These agents were later identified as the serotypically distinct NV and Hawaii viruses, which are prototypes of the now-recognized genogroup I (GI) and genogroup II (GII) NoV. Using these strains, Wyatt and coworkers were the first to document the absence of heterotypic immunity in a human population towards serotypically distinct NoV strains.

Several early volunteer studies also reported of selective NoV infections; even though challenge with high doses of the rather contagious NoV, not all individuals were infected (Parrino et al. 1977; Gary et al. 1987). It was also later noted that resistance to NoV infection often cluster with families. In 2003, Lindesmith and coworkers performed a challenge study, where they found a correlation between a 428 non sense mutation in the fucosyl transferase gene 2 (FUT2) and resistance to symptomatic NoV infection (Lindesmith et al. 2003b). Further evidence of a correlation between host genetics and susceptibility to NoV infection, was reported in 2002 by Hutson and coworkers(Hutson et al. 2002). They found a correlation between an individual’s ABO histo blood type and the risk of developing symptomatic NoV infection. They concluded that individuals with
blood type O were more likely to be infected with NoV (the Norwalk strain), compared to individuals having blood type B that had a decreased risk of infection.

**Histo blood group antigens and host genetics**

The expression of histo blood group antigens (HBGAs) and other glycosylation patterns, change during different periods of life such as during embryonic development, cell maturation, and also during cancer. To simplify, there are mainly three types of carbohydrates that determine the majority of human blood types: the secretor, Lewis and ABO(H) antigens (Le Pendu et al. 2006). Every individual can be divided into blood group A, B, AB or O (neither A or B), depending on which antigen(s) is displayed on the red blood cells. The ABO(H) antigens have also been found on epithelial cells on all organs that are in contact with the external environment, e.g. upper respiratory tract, lower genitourinary tract and the gastrointestinal tract. However, the expression of the ABH antigens is tightly regulated by several glycosyltransferases, that via a stepwise addition of monosaccharides to a precursor molecule constructs the carbohydrates (Oriol et al. 1986). The Lewis system is closely related to the ABO blood group system, since they both have a similar structure and synthesis pathway and even may be present on the same carbohydrate chains. The Lewis and ABH antigens are mainly located on four different carbohydrate chains, termed H type 1-4. However, in this thesis only the carbohydrates related to H type 1 and type 2 will be discussed. The precursor structures for the H antigens are often attached to the external part of O- or N glycoproteins, but also lacto glycolipids. The term O- and N-glycoproteins, refers to the positions of the attached glycans. The glycans are usually linked either to the side-chain nitrogen of an asparagine (N-linked) or to the side-chain oxygen of a serine or threonine residue (O-linked).

The biological role of the ABH antigens has been widely speculated, but due to their external location on epithelial cells it has been suggested that they may serve as receptors for microorganisms (Karlsson 1998). Due to the highly polymorphic nature of the ABO, secretor and Lewis genes, the human species has a great diversity of different cell surface receptors on the population level. This is an advantage in terms of human survival as a species, since no single pathogen with preferences for a specific carbohydrate receptor is able to infect all humans.

**FUT2 and secretor status**

The H antigen is an important precursor structure for the histo blood group A and B antigens, and is regulated by the α(1,2)fucosyltransferase. The α(1,2)fucosyltransferase play a key role in tissue expression of the ABH antigens, and at least two are present in human tissues; fucosyltransferase 1 (FUT1) and FUT2 often referred to as the “secretor enzyme”. FUT1 govern the expression of H antigen on red blood cells (erythrocytes) while FUT2 regulates the expression of H antigen in mainly epithelial cells and secretions (i.e. body fluids such as saliva) (Le Pendu et al. 1985; Oriol et al. 1986). The FUT1 and FUT2 adds a fucose in an α(1,2) position to the terminal galactose of the type1 precursor (Galβ1-3GlcNAcb1-R) and type 2 precursor (Galβ1-4GlcNAcb1-R) to form the H type 1 and type 2 epitopes, Figure 10.
Figure 10. Biosynthetic pathways showing the stepwise addition of monosacharides creating the ABO(H) and Lewis antigens on the type 1 and 2 chains. Antigen names are displayed in bold, glycosyltransferases are written in italic. The linkages of the different chains are represented by yellow or blue circle (Gal; galactose and Glc; glucose), yellow or blue square (GalNAc; N-acetylgalactosamine and GlcNAc; N-acetylglucosamine) and finally a red triangle representing a Fuc (fucose residue). The figure was kindly provided by Gustaf E Rydell, Institut Curie, France.

A number of fully or partially inactivating mutations, polymorphisms, have been observed on the FUT2 and FUT1 genes. Individuals carrying at least one functional allele of the FUT2 gene can express ABO antigens on most epithelial cells and in mucosal secretions and are thus termed "secretors" (SeSe; homozygous secretor or Sese; heterozygous secretor). Those that are homozygous for a non functional FUT2 allele cannot express or secrete these antigens and are referred to as "se" or "non secretors" (sese). These individuals therefore cannot have their ABO blood group determined from saliva, which is discussed in paper III and IV of this thesis.

On the other hand, an individual that lack a functional FUT1 allele are referred to as "Bombay" phenotype. This is a very rare condition (allele frequency 10^{-5} to 10^{-6}) leading to a total deficiency of ABO antigens on red blood cells (Marionneau et al. 2001). Studies have shown that many of the FUT1 mutations are linked to inactivating mutations in FUT2, due to their close proximity on chromosome 19 (Koda et al. 1997; Fernandez-Mateos et al. 1998).
The FUT2 gene is located within chromosome 19 in the human genome. In total, the length of FUT2 is 9,980 bp, and it is composed of two exons; the first containing an untranslated coding region is 118bp long and the second has a length of 2,995bp and codes for the 343 amino acids long enzyme, α(1,2)fucosyltransferase. The two exons are separated by a 6,865-bp intron.

There are many allelic variants with Se phenotype, and in total 19 different single nucleotide polymorphisms (SNPs) have been identified on the FUT2 gene (Koda et al. 2001). Studies have shown that FUT2 polymorphism has an ethnic group specific pattern, where different allele variants are located to specific geographic regions (Koda et al. 1996). However, although many of these polymorphisms are population specific, the se phenotypes are present in most populations. Two polymorphisms, are the most common cause of this phenotype: the se428 and se385. The non-functioning se428 nul allele coding for a stop codon at position 143 (Trp–Ter) is responsible for the se phenotype in approximately 20% of the European, Iranian, and African population (Kelly et al. 1995), while the se385 codes for a missense mutation at codon 129 (Ile–Phe) causing the se phenotype (or sometimes referred to as “weak secretor” (Sev)) in South East and East Asian populations (Koda et al. 2001). Two other se alleles are mainly found in smaller geographical regions; se302 found in Thai and Bangladeshi populations (Birney et al. 2007) and se571 found in Samoan populations (Soejima et al. 2007). Finally, even more rare causes of se phenotype in the FUT2 gene are: one deletion (se778), two complete deletions of the coding region (sdel, sdel2), and one fusion gene (sefus) (Soejima et al. 2007).

A and B antigens
As soon as the H antigens are formed by the FUT1 and FUT2 enzymes, the sequential addition of an α1,3 linkage with either an N-acetylgalactosamine (α1,3-GalNAc) to give the A antigen or an galactose (α1,3-Gal) to give the B antigen can occur. These events are catalyzed by the A or B enzymes, which are gene products derived from the A and B alleles, respectively, on the ABO locus. The O allele is highly similar to the A and B allele, with the exception of certain mutations leading to inactivation of the enzyme and consequently, the O phenotype in homozygous individuals (Hakomori 1999). Studies have shown that the ABO gene is highly polymorphic, and more than 160 alleles are now identified (Blumenfeld and Patnaik 2004). One example of a polymorphism is the blood group A individuals. They are divided into A1 and A2 phenotypes, where 80% are A1 and the remaining 20% are A2 phenotypes.

FUT3 and Lewis status
The FUT3 gene codes for the Lewis α(1,3/4)fucosyltransferase that synthesizes the Lewis a (Lea) and Lewis b (Leb) epitopes. Both H type 1 and H type 2 antigens can be used as a substrate for the FUT3 enzyme, however, the enzyme has a marked preference for H type 1. The FUT3 enzyme adds a fucose on the N-acetylglucosamine residue of the H-type 1 precursor in an α1,4 position to form the Lea and leb antigen, or in an α1,3 position on the H-type 2 precursor to form the Lewis x (LeX) or Lewis y (LeY) antigen. The Leb antigen is found only on secretor individuals, while non-secretors only display Lea. In contrast to the α4-fucose of the type 1 chain Lewis structures, several fucosyltransferases in addition to FUT3 (FUT4, FUT5, FUT6, FUT7 and FUT9) may
catalyze the addition of the $\alpha$1,3-fucose, characterizing the type 2 chain Lewis structures (Cailleau-Thomas et al. 2000).

The FUT3 enzyme can also act on the B and A type 1, adding a fucose residue in either the $\alpha$1,4 or $\alpha$1,3 position, respectively. This action will generate A or BLe$^b$ or A or BLe$^b$ antigens. In laboratories, antibodies towards Le$^b$ and Le$^b$ antigens are routinely used to phenotype for secretor status, since the expression of Le$^b$, but not Le$^a$, requires a functional FUT2 gene.

Homozygote carriers of inactive FUT3 alleles essentially lack Le$^b$ and Le$^b$ structures. Such individuals are denoted Lewis negatives and constitute nearly 10% of the caucasian population (Marionneau et al. 2001). It should be noted that in humans there are besides FUT3, five additional transferases (FUT4, FUT5, FUT6, FUT7 and FUT9) that are able to catalyze the addition of a fucose in $\alpha$1,4 or $\alpha$1,3 positions. These transferases in contrast to FUT3 mainly use H-type 2 structures as a substrate (Cailleau-Thomas et al. 2000).

**Norovirus**

Some viruses of Caliciviridae bind to HBGA expressed on cells in the gastrointestinal tract (Tan and Jiang 2010). Calicivirus infecting animals such as, RHDV and TV recognize H-type 2 and type A and B antigens, respectively, (Ruvoen-Clouet et al. 2000; Parkas et al. 2010). Also, Taube and co-workers have suggested a role of sialic acid moieties in murine NoV attachment to murine macrophages (Taube et al. 2009). Likewise, the bovine Newbury2 strain recognize a carbohydrate epitope of the HBGA family, which is absent in human tissues (Zakhour et al. 2009). Other studies have demonstrated a correlation between susceptibility to NoV infection and an individual’s blood type (Hutson et al. 2002; Lindesmith et al. 2003a; Rydell et al. 2011b). Thereby the ABO status was also identified as a susceptibility marker for NoV infection, and thus HBGAs are currently considered key candidates for receptors or possibly co-receptors for NoV. Besides binding of NoV to ABO(H) and Lewis antigens, a study by Tamura and coworkers report that NoV VLPs also can bind to heparan sulfate (Tamura et al. 2004). Recently, Rydell and co-workers showed that G1.3 and G1.4 VLPs could bind to Sialyl-Lewis x antigens on neoglycoproteins, suggesting that this also may be a possible determinant of NoV tropism (Rydell et al. 2009b). Together these studies indicate a role of sialylated structures or heparin sulfate in NoV cell tropism.

NoV infects individuals of all ages, however, approximately 20% of the Caucasian population are highly resistant to NoV infection due to the nonsense mutations at position 428 in the FUT2 gene. This immunity caused by the nonsense mutation, make individuals unable to display ABO histo-blood group antigens on mucosal surfaces probably necessary for NoV infection. This mutation leads to the non secretor genotype, and has been found to provide almost complete protection from experimental and authentic infection with NoV (Hutson et al. 2005; Thorven et al. 2005a; Kindberg et al. 2007b; Bucardo et al. 2009a; Rydell et al. 2011b). Tan and coworkers reported that outbreak studies of a G1.3 and a G1.4 NoV revealed an association between HBGA phenotypes and viral infection (Tan et al. 2008a). A study by Marionneau and co-workers from 2002 revealed that G1.1 VLP recognized H antigens based on the type 1 and type 3 chains and did not bind to epithelial cells from non secretors, only to those from secretor individuals (Marionneau et al. 2002).
Some sero-epidemiology studies have shown that certain non secretors are NoV antibody positive, suggesting that secretor independent infections do occur (Lindesmith et al. 2005b; Larsson et al. 2006). Even though the great majority of NoV strains have a strict secretor dependent infectivity pattern, there are actually a few exceptions. Two different GI.3 strains found by Nordgren et al and Rockx et al, have a secretor independent infectivity pattern (Rockx et al. 2005; Nordgren et al. 2010). In a challenge study performed in 2005, Lindesmith and co-workers showed that non-secretors could be infected with the GI.2 (Snow Mountain) NoV strain (Lindesmith et al. 2005a).

Finally, our study (Paper II), also indicated that even a strain of the globally dominating GI.4 NoV was able to infect a non secretor individual in a NoV outbreak in an elderly home in Spain, 2005 (Carlsson et al. 2009). However, although it is clear that also non secretors can be infected by NoV, it is still unknown what receptors NoV use while infecting these individuals (Rydell et al. 2011a).

However, even though there is a strong association between HBGAs and susceptibility to disease, the exact binding patterns and interactions between the carbohydrates and the NoV strains are probably highly variable and strain-specific rather than genogroup dependent (Huang et al. 2005; Tan et al. 2008b; Rydell et al. 2011b).

A study by White and coworkers reported that Norwalk virus VLP binding to CaCo-2 cells could be blocked by using a monoclonal antibody targeting a protruding region of the capsid protein (aa 300 to 380) (White et al. 1996). This finding suggests that this domain is involved in cell binding and receptor interactions.

Cao and coworkers have previously proposed from crystal structure studies that the residues involved in binding of the A and B trisaccharide to the GI.4 virus (VA387 strain) are located at the dimer interface and includes residues 343, 344, 345 and 374 from one protomer and 441, 442 and 443 in the other (Cao et al. 2007). A putative carbohydrate-binding patch consisting of residues 329,373, 375 and 377 located in the P2 domain has also been suggested by Chakravarty and coworkers (Chakravarty et al. 2005).

Allen and co-workers have found two antigenic regions (site A and B) in the P2 domain of the VA387 crystal structure, where aa substitutions in this area have impact on the biochemical properties as well as the entire structure of the P2 domain (Allen et al. 2008). These positions, located on the external part of the capsid, are exposed loops and thus changes in site A and B may have a strong association with the emergence of novel NoV strains.

**Sapovirus**

While human NoV susceptibility is highly associated to secretor status and thus mutations in the FUT2 (Le Pendu et al. 2006), no information has until recently (paper III and IV of this thesis) been available regarding genetic factors and susceptibility to SaV infection. In paper III we describe the genetic diversity of SaV in a Central American pediatric population and investigate for the first time role of host genetic factors and susceptibility to SaV infections. Paper IV is the first report investigating the role of host genetic factors in SaV susceptibility in the Caucasian population and we found a reduced risk of infection in individuals with blood group B (and AB), but no association to the
*FUT2* G428A nonsense mutation determining secretor status nor to the Lewis status. It remains to be elucidated whether the secretor independent NoV strains and the SaV utilize the same receptor.

The finding that individuals of blood group B (and AB) are less susceptible to SaV infection is intriguing. In 2002, Hutson and coworkers performed a challenge study with the Norwalk GI.1 NoV strain were they found that individuals of blood group B were less frequently infected compared to individuals of blood group A and O (Hutson et al. 2002). Henessey and coworkers made a similar observation regarding GI.1 NoV infection in 2003. They investigated a NoV outbreak in a British military camp in Afghanistan, and observed that personnel with blood group B were significantly less susceptible to infection (Hennessey et al. 2003). A molecular correlation to the infectivity and binding pattern of the Norwalk strain was concluded when it was realized that the $\alpha_{1,2}$-Fucose of the H-epitope and the $\alpha_{1,3}$-GalNAc of the A epitope, but not the $\alpha_{1,3}$-Gal of the B epitope, interacted strongly and specifically with the same amino acids of the virus capsid (Choi et al. 2008).

In 2007, Shirato-Horikoshi and coworkers, performed a VLP binding study but did not find any correlation between SaV (genogroup I and V) and histo blood group antigens (Shirato-Horikoshi et al. 2007). However, most likely that there is probably a genotype specific infection pattern also for SaV, as have been observed for NoV (Tan and Jiang 2005; Rydel et al. 2011b; Tan and Jiang 2011).
Evolution of caliciviruses

The mechanism for emergence and evolution of new human calicivirus strains, as well as protective immunity in the human population is poorly understood. There are a number of different NoV strains, all with slightly different HBGA binding patterns. One can thus speculate whether this feature is a vital adaptation to match the divergent HBGA distribution in the human population. However, the fact that the most common NoV strain, GII.4, has a global spread may indicate a lack of adaptation. On the other hand, it is also known that only small aa variations in the capsid binding site, may drastically alter the NoV HBGA binding pattern (Huang et al. 2005; Lindesmith et al. 2008).

Quasi species
RNA viruses have a characteristic high mutation rate, due to poor proofreading (10^{-4} mutations per nucleotide copied) and post-replicative activities associated with RNA replicases. Studies have shown that the RNA virus mutation rate might be up to 1 million times higher than mutation rate during genome replication within the human host (Holland et al. 1982). The high mutation rate and thus rapid evolution of RNA viruses has become a big problem when designing vaccines or treatment for emerging viral diseases, such as influenza or HIV (Domingo 1998; Domingo et al. 1998; Domingo et al. 2006). It has been estimated that due to the big population size formed by RNA viruses during infection, every possible mutation and even some double mutations will be produced during every replication cycle (Vignuzzi et al. 2005). Due to this error prone nature of the RNA viruses they constantly may develop resistance to antiviral drugs (Gerrish and Garcia-Lerma 2003).

The behavior of RNA virus evolution goes beyond the classical models for population genetics. A model or theory describing the error prone replication and almost infinite population sizes for RNA genome evolution was first reported by Eigen and Schuster in 1977 (Eigen and Schuster 1977). Generally, a quasispecies population consists of a quantitatively dominant genome, surrounded by a cloudlke multitude of sequences differing from the majority sequence to various extents (Domingo et al. 1998; Biebricher and Eigen 2006). Within this heterogeneous cloud of genetic variants, mutants with increased capacity to adapt to various environments will be generated. One clone will be predominant under the existing conditions because it is better fit, but when conditions change other clones may be better suited for the new conditions and thus this new clone will dominate, Figure 11.
Figure 11. Schematic representation of viral quasispecies. Due to poor proofreading of the RdRp a spectra of mutant genomes (horizontal lines, black or grey) are generated during each replication resulting in a heterogeneous population of genetic variants. Each individual mutation is represented by a blue square, red circle, yellow star or green triangle. The consensus sequence represents the dominating genome in the viral population. Evolution of the quasispecies population have a direction and may go through series of “bottlenecks”, where certain point mutations are accumulated often in response to an environmental change. The overall fitness of the viral population will decrease as the homogeneity of the population increase.

When describing genome replication within a population, the term “sequence space” is often used. The sequence space refers to the maximum amount of possible sequence variants, and the physical distance is proportional to the nucleotide similarities between sequences (Biebricher and Eigen 2006). One of the key features of a quasispecies population is that the evolution has a “direction”, meaning that each mutant is coupled to a predecessor sequence further away in sequence space. Each individual mutant are connected within a functional network of variants, and a less fit mutant may survive since it is “coupled” to a higher fitness genotype in sequence space (Domingo 1998). The functional network created by a quasispecies population may be illustrated by a study of neurotropic poliovirus reported by Vignuzzi and coworkers (Vignuzzi et al. 2006). They infected naïve mice with a neurovirulent clone of poliovirus either as part of a genetically uniform or a diverse poliovirus population. The mice infected with the genetically uniform population showed no symptoms, while mice infected with the diverse population developed CNS symptoms and died. This indicates that mutants within the diverse quasispecies population cooperated with the neurotropic clone and assisted the entry into CNS.

In the classical Darwinan evolution model, the term fitness is considered as a measurement of how well an organism fits into its environment (Orr 2009). However in the dynamic world of viruses, much more aspects such as replicative capacity, infectivity, cell tropism and immune escape must be included in the word fitness. Yet
one more dimension is added to the fitness concept when it comes to the quasispecies model. A quasispecies population is often described as located within a fitness landscape, where the broadness of the landscape refers to the number of genotypes in the sequence space. A genetically diverse but less fit population has a broad peak (wide range of mutant spectra), while a population well suited to the environment will have a narrow peak (small range of mutant spectra) (Lauring and Andino 2010).

**Survival of the flattest**

During evolutional selection the quasispecies population will function as a unit, each mutant will be linked to each other through a web of interactions. The fitness must be interpreted for the whole quasispecies population, not just for individual clones as the fitness then may be underestimated. De la Torre and coworkers reported in 1990 on a quasispecies population of vesicular stomatitis virus (VSV) where individual clones of high fitness were actively suppressed to favor clones of less fitness (de la Torre and Holland 1990). This phenomenon can be put in relation to the findings that defective clones were maintained at a constantly high frequency in a quasispecies population of Dengue virus (Aaskov et al. 2006). This phenomenon may seem very odd, but in terms of quasispecies theory can be explained as “survival of the flattest”. Sometimes for instance a high replication capacity, in spite of the fact that the individual clone is very fit, might not be good for the population since fast replicators will give rise to genetically diverse offspring, many of which are significantly less fit than the parent. Therefore in terms of quasispecies theory, a slow replicator will be favored if it gives rise to offspring that are on average more fit. Such populations have a short and flat fitness landscape, and have hence given rise to the phenomena “survival of the flattest” (Wilke 2005). A flat fitness landscape is predicted to effect the pathogenicity of a viral population. A flat fitness landscape means that the viral population consists of a wide range of mutants, and thus readily can adapt to new conditions such as when transmitted to a new host.

**Error threshold**

RNA viruses have in theory high capacity to occupy a large portion of the sequence space, however, in real life viral replication will encounter many limitations (Overbaugh and Bangham 2001). In order to sustain an efficient viral population, some mutations will be sacrificed and other more beneficial will be selected. Quasispecies theory suggests that there is an upper limit for the mutation rate, an “error threshold”. Eigen described this phenomenon for the first time in 1971, were he stated that there is a limit to the amount of information a genome can store at a given mutation rate (Eigen 1971). If this threshold or limit is preceded, the structure of the quasispecies population breaks down and the population will be spread over the sequence space. The fact that many RNA viruses replicate near the error threshold, has led to strategies for antiviral drugs that enhance the viral mutation rate to deleterious levels (Anderson et al. 2004; Jonsson et al. 2005).

**Norovirus**

The NoV capsid protein has a high sequence diversity, probably due to poor proofreading and post-replicative activities associated with RNA replicases, but sequence diversity is probably also driven by immune selection mechanisms. Investigating the evolutionary dynamics of NoV is important not only in understanding
the pathogenicity, but also in terms of effective prevention and development of control strategies. Given the high mutation and evolutionary rate of NoV, there is a high probability that quasispecies populations will be formed within infected individuals (Bull et al. 2010; Bok et al. 2011). Since NoV cannot be grown in culture, other means must be used to study NoV population dynamics and evolution. However, NoV can cause persistent infections in immune compromised patients, and these patients constitute excellent settings for NoV evolution studies. Recently, in a study involving renal transplant recipients not only NoV but also SaV was suggested to be involved in chronic infections (Roos-Wei et al. 2011). Iwakiri and coworkers have observed nucleotide changes in the SaV capsid region during prolonged SaV shedding, suggesting an evolutionary process (Iwakiri et al. 2009).

In paper I, by using evolutionary trace analysis (ET) we identified evolutionary conserved aa residues and quasispecies behavior in an individual chronically infected for 3 years with GI.3 NoV. Seventeen fecal samples were collected from August 2000 to March 2003, generating in total 134 NoV capsid sequences. Phylogenetic analysis revealed that sequences isolated at the same time point (and thus belonging to the same isolate) showed a trend of clustering together in swarms typical for a quasispecies distribution (Paper 1). The capsid P1-1 and P2 domains evolved in time-dependent manner, with a distinct break point observed between autumn/winter of year 2000 and spring to autumn of year 2001, which presumably coincided with a change of clinical symptoms. Another feature supporting this finding is that 5 out of the 12 truncated capsid sequences were found during February 2001, suggesting a prominent evolution activity during this period. This finding is in concordance with other studies, suggesting that the P2 domain of GI.4 virus is evolving in a time dependent manner showing epochal evolution with periods of slow and high evolutionary rates over time (Siebenga et al. 2007b; Lindesmith et al. 2008).

Siebenga and coworkers performed a 2 year hospital survey of NoV infections. They found that approximately 8% of all nosocomial NoV infections, resulted in chronic infection (Siebenga et al. 2008b). They also found a correlation between the degree of immune suppression and the number of aa mutations found in the antigenic sites of the viral P2 domain. The study concluded that the more suppressed immune system, the less amino acid mutations in the viral quasispecies population, thus suggesting an immune-driven selection of viral clones.

The presence of immune-driven NoV evolution in immune compromised patients was later also confirmed by Schorn and coworkers (Schorn et al. 2010). They found accumulation of point mutations over time in the P2 and P1-2 region in NoV strains (GI.4, GI.7 and GI.17), isolated from kidney transplant recipients. Many mutations resulted in amino acid changes, and up to 25 amino acids mutations were accumulated over a period of almost two years.

Recently, Hoffmann and coworkers studied norovirus evolution (two genotype II.4 and one II.7) in three chronically infected patients. The results clearly suggested that the NoV quasispecies evolved under positive selection rather than incidental drift. In average 5–9 mutations were accumulated per 100 day in each strain (Hoffmann et al. 2012). A similar evolutionary pattern has also been observed in chimpanzees challenged with human GI.1 NoV (Norwalk strain) (Bok et al. 2011). During infection most amino
acid changes were located in the P2 domain of the NoV capsid, and some of these mutations evolved under positive selection.

Boon and coworkers have performed an analysis of GII.3 and GII.4 NoV capsid evolution over a perspective of 31 Years (Boon et al. 2011). The evolutionary analysis concluded that GII.3 and GII.4 NoV have an approximate evolutionary rate of $4.16 \times 10^{-3}$ nucleotide substitutions/site/year (strict clock). However, while GII.3 NoV have maintained a steady state of amino acid changes and thus probably maintained their HBGA binding pattern, the divergence within the GII.4 viruses have increased due to an increasing number of amino acid changes and consequently altered HBGA binding patterns.

To summarize, individuals chronically infected with NoV are most likely to increase in the modern society, as a consequence to the increased number of immune suppressed transplant recipients. A chronically NoV infected immune compromised individual will not be able to eliminate the virus efficiently, but will still constitute enough selective pressure to maintain a high rate of aa evolution. This condition will within time develop a subset of viral clones (viral quasispecies), perhaps constituting novel epidemic subtypes making the chronically infected individual a viral reservoir.
Herd immunity and epochal evolution of NoV

The first characterized GII.4 NoV was the MD145 strain isolated in 1987, but it was not until a few years later in the mid-1990-ties that the GII.4 NoV cluster had become globally widespread and was considered the major epidemic strain (Green et al. 2002). Following this, almost every other year over the last decade GII.4 NoV epidemics has been associated with the emergence of novel genetic variants: pre-1995 (Camberwell), 1996 (Grimsby), 2002 (Farmington Hills), 2004 (Houston), 2006a (Laurens) and 2006b (Minerva). Recently, two more variants are reported, they are the GII.4 2008 (Apeldoorn) and GII.4 2010 (New Orleans) (Donaldson et al. 2008; Siebenga et al. 2009; Lindesmith et al. 2011). An alignment of P domain of the representative GII.4 outbreak strains is presented in Figure 12.

**Figure 12.** Sequence alignment of epidemic GII.4 NoV strains isolated 1995 to 2010. Residue 300 to 450 of the capsid P2 domain is displayed in the alignment. Amino acids involved in evolutionary site 1 is represented by blue squares, while those of site 2 is represented by a grey square. The GII.4 GrimsbyVA387 and 2004 GII.4 HoustonTCH5 strains are surrounded by black lines. Sequence alignment was performed using ClustalW2 on the EMBL-EBI database. Genebank sequence accession numbers are as follows: 1993/GII.4/Lordsdale/\(\bar{X}86557\); 1996/GII.4/GrimsbyVA387/(\(\bar{A}J004864\)); 2001/GII.4/Japan/(\(\bar{A}B294779\)); 2001/GII.4/Henry/Houston/(\(\bar{A}E310927\)); 2002/GII.4/Langen/Farmington Hills/(\(\bar{A}V485642\)); 2003/GII.4/Sakai/Asia/(\(\bar{A}B220922\)); 2004/GII.4/HoustonTCH05/(\(\bar{J}827296\)); 2006a/GII.4/Yerseke/Laurens/(\(\bar{E}F126963\)); 2006b/GII.4/DenHaag/Minerva/(\(\bar{E}F126965\)); 2008/GII.4/ Apeldoorn/(\(\bar{A}B445395\)) and 2010/GII.4/ New Orleans/(GU445325).
Today GII.4 are responsible for 70 to 80% of all NoV gastroenteritis outbreaks (Kroneman et al. 2008). In 2008, Lindesmith and coworkers proposed a model for GII.4 NoV evolution caused by structural changes of the viral capsid leading to a shift in receptor usage and ultimately immune evasion (Lindesmith et al. 2008). The emergence of a new GII.4 variant causing a new epidemic peak is usually followed by a period of stasis (Siebenga et al. 2007b; Donaldson et al. 2008). The hypothesis of herd immunity as a driving force for GII.4 evolution was later supported by Cannon and coworkers (Cannon et al. 2009). They performed a surrogate neutralization assay and measured the ability of outbreak acute- and convalescent-serum to block the binding of a panel of GII.4 NoV VLPs to HBGAs. The VLPs used represented epidemic strains and their results showed that a higher extent of HBGA-VLP blocking was achieved with each VLP corresponding convalescent outbreak sera. Reeck and coworkers have recently shown that blocking antibodies correlate to protection against NoV disease (Reeck et al. 2010).

However, the role of HBGA in epochal evolution and HBGA binding patterns for GII.4 variants have recently caused debate. In 2008, Donaldson and coworkers reported that the VLP representing the 2004-2005 GII.4 strain did not bind any known carbohydrates, and was thus suggested to have novel carbohydrate binding properties (Donaldson et al. 2008). However, this was a few years later challenged by Yang and coworkers, who reported on recombinant 2004-2005 GII.4 NoV P particles that, if yet weakly, bound to all secretor antigens (Yang et al. 2010). Just recently, by using an extended binding analysis including surface plasmon resonance, de Rougemont and coworkers concluded that the 2004-2005 GII.4 NoV VLP bound to both ABH and Lewis antigens of secretor positive individuals (de Rougemont et al. 2011).

The only crystal structure available for GII.4 NoV has until just recently been the P domain of the VA387 strain (1996 Grimsby cluster) (Cao et al. 2007). In 2011 Shanker and coworkers presented a crystal structure of the epidemic 2004 GII.4 NoV variant TCH05 (Houston) (Shanker et al. 2011). Studies of the crystal structure of the 1996 GII.4 NoV (VA387), has suggested the presence of two HBGA binding sites, referred to as site 1 and 2 (Cao et al. 2007). The location of site 1 and 2 in the 2004 GII.4 NoV (TCH05) capsid is shown in Figure 13. Co-crystallization with the A and B trisaccharide has shown that both these antigens bind with their terminal effucose to site 1. This site is highly conserved between the epidemic GII.4 variants. Site 2 is a loop constituted by aa 390 to 395 and this region is more variable and sensitive for temporal sequence variations. Shanker and coworkers have now further characterized these sites, showing that most likely all epidemic GII.4 variants bind the effucose of the ABH antigens in site 1 (Shanker et al. 2011). Thus, this binding is structurally conserved and temporal sequence variations do not affect the binding of monofucosyl ABH antigens. In contrast to the monofucose of the A, B and H epitopes that only bind to the conserved site 1, the binding of the Lea antigen involves both site 1 but also the variable site 2. Earlier studies have identified the aa residues in site 2 (392 to 395) as evolutionary “hot spots” for evolution of GII.4 NoV (Bok et al. 2009; Tan et al. 2009). Shanker and coworkers also found an unknown region of conformational flexibility in close proximity to the area of P domain dimerization and formation of the HBGA binding site. These findings may help in understanding the factors governing HBGA binding during viral capsid attachment and release in the gastrointestinal tract.
**Figure 13.** Comparison of the P domain dimers of the epidemic 2004 and 1996 GII.4 NoV variants. Orange represents the P domain dimer of the 2004 variant while the 1996 variant is displayed in cyan. The location of the N and C termini of the dimeric subunits of the 2004 P domain are indicated. Black arrows represent the HBGA binding sites of the 1996 variant. Red and black boxes indicate locations of evolutionary site 1 and site 2, respectively. The black boxes contain a conformational change in the site 2 loop between the two epidemic variants. Figure obtained from Shanker et al, with permission (Shanker et al. 2011).

The recent structural analyses of the GII.4 NoV has shed some light into evolution and the structural basis for HBGA binding of NoV. However, several questions remains and the mechanism for emergence and evolution of new NoV strains in the human population are still poorly understood.
**Norovirus vaccine**

The most suited method to test the efficacy of potential NoV vaccines would be to use human challenge studies. However, these are difficult to execute both due to ethical reasons but also due to difficulties of producing a perfectly safe NoV inoculate. Solely the nature of the NoV inoculates; stool filtrates derived from previously infected volunteers, are controversial in our modern society. Therefore, the search for a convenient animal model still continues.

One of the key issues in the development of a NoV vaccine, is the question whether a monovalent NoV vaccine would provide protection against two different genogroups. A challenge study in chimpanzees performed by Bok and coworkers reported a correlation between the presence of antibodies in sera and protection from reinfection in the chimpanzee model (Bok et al. 2011). They administrated VLP-based vaccines containing the VP1 and VP2 capsid proteins of GI (Norwalk) and GII (MD145) intramuscularly (i.m.) to the chimpanzees. The animals were then challenged with an intravenous (i.v.) injection of human stool filtrate, and subsequently stool, serum, and biopsy samples were collected. The challenge revealed that the chimpanzees were protected after homologous but not after heterologous vaccination (Bok et al. 2011). This suggests that the antibodies obtained after vaccination could not confer cross-protection against two distinct NoV genogroups. However, one confounding factor with this challenge study, is that the i.v. challenge route used differs from that of a natural infection where the viruses are spread via the fecal-oral route.

Other studies have also shown that NoV VLPs are stable and stimulate a strong immune response when given parenterally, orally, or intranasally to mice. The stability of the VLPs have been sustained even during harsh conditions such as when lyophilized or during acid conditions (pH 2.5) (Ball et al. 1998; Guerrero et al. 2001). A few phase 1 trials using GI I Norwalk VLPs have been done in humans (Tacket et al. 2003). Even trials with a plant-based oral GI I Norwalk VLP vaccine using transgenic potatoes have been performed (Tacket et al. 2000). The results from these early VLP vaccines indicated that they were safe to use, however, one major drawback was that they unfortunately only induced a weak immune response.

Recently, a novel adjuvanted NoV VLP vaccine has been tested in humans. El-Kamary and coworkers performed two phase 1 studies of a monovalent NoV GI I vaccine (El- Kamary et al. 2010). This vaccine was adjuvanted with monophosphoryl lipid A (MPL) and the mucoidherent substance chitosan, and was delivered intranasally. The results indicated that the vaccine induced virus-specific serum antibodies in most of the vaccinated individuals. Approximately one year later, the safety, immunogenicity, and efficacy of this vaccine was further investigated in a human challenge study performed by Atmar and coworkers (Atmar et al. 2011). Healthy adults were vaccinated intranasally with two doses of vaccine and then challenged with Norwalk virus. The study concludes that the vaccine provide protection against illness and infection to a homologous virus. However, this vaccine must be further tested in more sensitive populations such as young children and in elderly to confirm its effectivity. Recently, Blazevic and coworkers reported a combined vaccine consisting of NoV GII-4 VLPs and human RV recombinant VP6 (rVP6) protein that induced strong systemic cross-reactive and cross-blocking antibody responses in BALB/c mice (Blazevic et al. 2011).
To summarize, even though some progress has been made in the field of NoV vaccine during the last years, many questions remain to be solved. One of these is the duration, immunity after a natural infection is short-lived, and so far the NoV VLP vaccines have failed to evoke even the same level of serum antibody responses as observed during natural infection (Glass et al. 2009). NoVs as other RNA viruses are highly variable, and a number of different genotypes has been identified both GI and GII NoVs. A major challenge is therefore to develop a NoV vaccine that confers protection against heterologous strains, but so far this has only been seen to a very low extent. In a recent study by Lindesmith and coworkers, volunteers were immunized with GI.1 NoV VLP and were then subsequently challenged with different GI NoV strains (Lindesmith et al. 2010). The results suggest that GI.1 Norwalk virus infection induces a robust B- and T-cell response against homologous strains, but most interestingly, some infected individuals also developed immune response against heterologous GI strains. However, even if a vaccine conferring protection is developed, it is most likely that the composition of the vaccine must be constantly altered to match the epochal evolution of the globally dominating GII.4 NoV (Lindesmith et al. 2008; Lindesmith et al. 2011).
Aim of thesis

The general aim of this thesis was to investigate a potential correlation between viral evolution, host genetics and disease susceptibility.

More specifically, the aim of the individual papers were:

**Paper I.** Investigate quasispecies dynamics and identify evolutionarily conserved and thereby functionally important amino acids of the NoV capsid during chronic infection.

**Paper II.** Describe and characterize the capsid properties and susceptibility pattern of a novel GII.4 NoV strain.

**Paper III.** Describe the genetic diversity of SaV in a Central American pediatric population and investigate the role of host genetic factors and susceptibility to SaV infections.

**Paper IV.** Investigate the role of host genetic factors in susceptibility to SaV infections in a Caucasian population.
Materials

Briefly, the materials used in the enclosed papers are as follows:

**Paper I.**
The unique patient samples used in this evolutionary study consisted of fecal samples collected from an immunosuppressed heart transplant recipient who became infected with a GII.3 NoV in June 2000. This individual developed a chronic infection with vomiting and four to eight episodes of diarrhea per day lasting for several years (Nilsson et al. 2003). Fecal samples were collected from August 2000 to March 2003 (termed P1 to P17), and used for extraction of viral RNA and subsequent phylogenetic analysis of NoV capsid, Figure 14.

**Figure 14.** Flowchart describing the analysis of patient samples included in paper 1 of this thesis.

**Paper II.**
During November 2004 an outbreak of acute gastroenteritis affecting 116 individuals occurred in an elderly nursing home in El Grao de Castello’n, Spain. In order to determine the etiological agent causing the outbreak, fecal samples were collected from 26 symptomatic residents and 7 staff members. Saliva samples were also collected from 39 residents (both symptomatic patients and asymptomatic controls) and 21 health care workers (both symptomatic individuals and asymptomatic controls). This saliva was used to determine the secretor and Lewis status of the individuals included in the study.
**Paper III.**
The patients involved in this study (292 children of ≤5 years of age, with and without diarrhea) were randomly selected among 852 children that participated in a community- and hospital-based survey of sporadic acute diarrhea in León, Nicaragua from March 2005 to September 2006. Among these, 205/292 had diarrhea and 45 of these 205 individuals were admitted to hospital with severe acute gastroenteritis. The remaining 87/292 individuals were asymptomatic with no recent history of diarrhea (≤ 10 days) and were used as controls. Blood and saliva samples were also retrieved from each child to determine the ABO blood group, secretor and Lewis status. Information about symptoms such as fever (≥38°C), nausea, vomiting, loss of appetite, abdominal cramps, abdominal distension (gas) and number of loose stools during past 24 hours along with information on dehydration status and treatment plan was recorded for each child.

**Paper IV.**
During October 2008 to November 2010 a total of 64 fecal, saliva and blood samples were collected from SaV positive individuals (≥ 4 years of age) in Denmark. Fecal samples were used for SaV genotyping, while the blood and saliva samples were used to determine ABO blood group, secretor and Lewis status of the participating individuals.
Methods

Virus detection and capsid genotyping (Paper I, II, III and IV)

Due to the high sequence diversity, human Caliciviruses have conventionally been detected with PCR only targeting the more conserved polymerase region. However, with improved knowledge about NoV capsid sequences, several primer pairs have subsequently been designed to target also strain specific regions on the capsid gene. Today, the by far most used molecular detection method for NoV and SaV is conventional PCR and RT-PCR performed on viral RNA purified from stool samples of infected individuals (Atmar and Estes 2001; Logan et al. 2007; Nordgren et al. 2008; van Maarseveen et al. 2010; Shigemoto et al. 2011; Wolffs et al. 2011).

The GII.3 NoV causing the chronic infection described in Paper 1, was in an earlier study detected using the JV12 and JV13 primers targeting the polymerase region in the ORF1 (Nilsson et al. 2003), Figure 13. Later the LV6717 and LV4922 primers were used to amplify the capsid region within ORF2, which was subsequently cloned and sent for full capsid sequencing to determine the genotype. In paper I the Noro P5’ and Noro P3’ primers targeting the capsid P region (P1-1 and P2) domain was used to amplify the P region prior to cloning, Figure 15 and Table 2.

In paper II, the GII.4 NoV strain was detected as in Paper I, using the JV12 and JV13 primers. A conserved part of the S domain within the capsid gene was then amplified using the Mon381 and Mon383 primers, followed by sequencing to determine the genotype. In order to characterize aa of the capsid protein, the P2 domain was amplified and subsequently cloned using the Val fw1 and Val rv1 primers, Figure 15 and Table 2.

Figure 15. Location of the primers within the ORF1 and ORF2 used for detection and genotyping of NoV strains in Paper I and II.
Table 2. Primers for detection and genotyping of NoV

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>JV12</td>
<td>5’-ATA CCA CTA TGA TGC AGA TTA</td>
<td>Paper I and II</td>
</tr>
<tr>
<td>JV13</td>
<td>5’-TCA TCA TCA CCA TAG AAA GAG</td>
<td>Paper I and II</td>
</tr>
<tr>
<td>LV6717</td>
<td>5’-AGT ACC TTG TTC CGC TCC A</td>
<td>Paper I</td>
</tr>
<tr>
<td>LV4922</td>
<td>5’-CAC GGC CCA GCA TCC TAC A</td>
<td>Paper I</td>
</tr>
<tr>
<td>NoroP3</td>
<td>5’-TTA ATG AGA CCA TTG ATC AAA TTC</td>
<td>Paper I</td>
</tr>
<tr>
<td>NoroP5</td>
<td>5’-ATG CTT GTG CCA CCT ACT GTG GAG TCA</td>
<td>Paper I</td>
</tr>
<tr>
<td>Val fw1</td>
<td>5’-GAA CTA AAC CAT TCT CTT CTC C</td>
<td>Paper II</td>
</tr>
<tr>
<td>Val rv1</td>
<td>5’-AAG TGC TGC ACC CA CTC CTT</td>
<td>Paper II</td>
</tr>
<tr>
<td>Mon381</td>
<td>5’-CCA GAA TGT ACA ATG GTT ATG C</td>
<td>Paper II</td>
</tr>
<tr>
<td>Mon383</td>
<td>5’-CAA GAG ACT GTG AAG ACA TCA TC</td>
<td>Paper II</td>
</tr>
</tbody>
</table>

For genotyping of SaV in paper III, a nested PCR was performed using two primer pools. The "outer" primer pool consisted of two forward (the universal F13 and F14) and two reverse (R13 and R14) primers targeting the end of the RdRp and the gene encoding the capsid protein (Okada et al. 2006). The "inner" primer pool consisted of universal forward primers (F13 and F14) and genogroup specific reverse primers (G1R, G2R, G4R and G5R) (Okada et al. 2006), Figure 14 and Table 3. In the last paper (Paper IV) stool samples were screened for SaV by using real-time PCR and the primers SaV124F, SaV1F, SaV5F, and SaV1245R all targeting the polymerase region (Oka et al. 2006b; Johnsen et al. 2009). Sapovirus positive samples were further characterized by PCR and sequencing. For genotyping a nested PCR was performed amplifying part of the capsid gene (SLV5317 and SLV5749 primers) and the polypeptide capsid junction. Additionally, also another section of the polymerase gene was amplified using the JV33 and SR80 primers, Figure 16 and Table 3.

Figure 16. Location of the primers within the ORF1 used for detection and genotyping of SaV strains in Paper III and IV.
<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>F13</td>
<td>5'GAY YWG GCY CTC GCY ACC TAC</td>
<td>Paper III</td>
</tr>
<tr>
<td>R13</td>
<td>5'GGT GAN AYN CCA TTK TCC AT</td>
<td>Paper III</td>
</tr>
<tr>
<td>F14</td>
<td>5'GAA CAA GCT GTG GCA TGC TAC</td>
<td>Paper III</td>
</tr>
<tr>
<td>R14</td>
<td>5'GGT GAG MMY CCA TTC TCC AT</td>
<td>Paper III</td>
</tr>
<tr>
<td>G1R</td>
<td>5'CCC BGG TGG KAY GAC AGA AG</td>
<td>Paper III</td>
</tr>
<tr>
<td>G2R</td>
<td>5'GWG GGR TCA ACM CCW GGT GG</td>
<td>Paper III</td>
</tr>
<tr>
<td>G4R</td>
<td>5'GGG TAG CAG ATC CCA GAT AA</td>
<td>Paper III</td>
</tr>
<tr>
<td>G5R</td>
<td>5'TTG GAG GWG GTT GCT CCT GTG</td>
<td>Paper III</td>
</tr>
<tr>
<td>SaV124F</td>
<td>5'GAY CAS GCT CTC GCY ACC TAC</td>
<td>Paper IV</td>
</tr>
<tr>
<td>SaV1F</td>
<td>5'TTG GCC CTC GCC ACC TAC</td>
<td>Paper IV</td>
</tr>
<tr>
<td>SaV5F</td>
<td>5'TTT GAA CAA GCT GTG GCA TGC TAC</td>
<td>Paper IV</td>
</tr>
<tr>
<td>SaV1245R</td>
<td>5'CCC TCC ATY TCA AAC ACT A</td>
<td>Paper IV</td>
</tr>
<tr>
<td>SLV5317</td>
<td>5'CTC GCC ACCT AC RAW GCB TGG TT</td>
<td>Paper IV</td>
</tr>
<tr>
<td>SLV5749</td>
<td>5'CGG RCY TCA AAV STA CCB CCC CA</td>
<td>Paper IV</td>
</tr>
</tbody>
</table>
Lewis and ABO phenotyping of saliva and blood (Paper II, III and IV)

Antibodies towards Le\(^a\) and Le\(^b\) are routinely used to phenotype for secretor status, since non secretors cannot express Le\(^b\), only Le\(^a\) (Marionneau et al. 2001). This phenotyping is generally performed on red blood cells, even though they generally do not express H type 1 structures. However, the red blood cells have gained access to the Lewis antigens by absorbing glycosphingolipids present on plasma lipoproteins and can therefore be used for phenotyping (Marcus and Cass 1969; Ravn and Dabelsteen 2000).

Phenotyping of HBGAs can also be performed on saliva since the glycan pattern displayed in saliva are similar to that present on the epithelial cells in the gastrointestinal tract (Ravn and Dabelsteen 2000). A saliva assay has been used in a number of studies to determine the binding pattern of NoV strains (Marionneau et al. 2002; Tan and Jiang 2005; Thorven et al. 2005b; Rydell et al. 2009b).

Non secretors cannot be phenotyped for ABO in saliva since they do not express H type 1 structures. They will turn out as false positives for blood type O in the saliva assay, and must therefore be phenotyped on red blood cells. Lewis negatives do not express any of the Lewis antigens and will not bind antibodies towards Le\(^a\) and Le\(^b\) and thus turn out negative in a Lewis binding assay (Ravn and Dabelsteen 2000). However, an individual found to be Lewis negative by saliva phenotyping can be confirmed by genotyping of the FUT3 gene. Some individuals may also be positive for both the Le\(^a\) and Le\(^b\) phenotype, when genotyped these individuals are usually found to be Lewis positive weak secretors (missense mutation A385T on the FUT2 gene).

The main problem with Lewis phenotyping is usually the cross-reactivity of the antibodies used for detection (Manimala et al. 2007). We have however developed a reliable in house saliva binding assay with good specificity that are used in Paper II, III and IV. To further strengthen the results each phenotyping has been performed at least twice for each saliva sample, and in addition, all secretor phenotyping has been confirmed by genotyping of the FUT2 428 polymorphism.

Genotyping of FUT2 polymorphism (Paper II, III and IV)

Generally, the major disadvantage with genotyping is that only the known SNPs can be investigated, meaning that uncommon or yet unknown inactivating mutations may be over seen and results thus may be misinterpreted. However, the FUT2 428 mutation is the by far most common inactivating mutation found in Caucasians and thereby we chose to screen for this polymorphism in Paper II, III and IV. In addition to the 428 polymorphism, also the non sense mutation C571T and the missense mutation A385T (weak secretor) are screened for in Paper II. However, none of the investigated individuals were carriers of the mutations at nt 385 or at nt 571 in the FUT2 gene. One may speculate that these additional missense mutations also should have been investigated in Paper III, since this was a non Caucasian population. However previously, Bucardo and coworkers have while studying genetic susceptibility to NoV infection in Nicaragua, found that none of the examined individuals were homozygous or heterozygous for the 571 mutation (Bucardo et al. 2009b).
Evolutionary trace (ET) analysis (Paper I)

Today there are a wide range of computer software available for phylogenetic analysis. However, when analyzing such a complex and rapidly evolving set of quasispecies sequences as that presented in Paper I, many factors must be taken in to consideration when selecting the most appropriate method. General phylogenetic methods used to analyze evolving sequences may be found in computer programs such as PAML (Phylogenetic analysis by maximum likelihood) (Yang 1997) or BEAST (Yang et al. 1995; Drummond and Rambaut 2007). The BEAST program applying Bayesian tree reconstruction was recently successfully used by Hoffmann and coworkers to study norovirus evolution (two genotype II.4 and one II.7) in three chronically infected patients (Hoffmann et al. 2012). However, in order to study evolutionary conserved residues and their functional importance as was the purpose in Paper I, the evolutionary trace (ET) method are considered to be the single most validated approach (Wilkins et al. 2012), and was therefore our method of choice.

Statistical analyses (Paper II, III and IV)

Fisher’s exact test (with unadjusted odds ratios (OR) with 95% confidence interval) or chi-square test was used to test significant differences in distribution of secretor-positive and secretor negative individuals among symptomatic and asymptomatic/non-exposed in paper II, III and IV. These tests were also used to compare SaV-positive specimens in terms of gender, age group, clinical status (Paper IV) and histo blood group antigens in Paper III and IV. Mann-Whitney test was used to compare ELISA absorbance values between secretor-positive and secretor-negative novel GII.4 disease pattern in Paper II.

All statistical analyzes in Paper III and IV was performed in Graph pad prism 5 (Graph Pad software, Inc) for Mac OS X, while SPSS 16 for Mac was used to perform for the analyses in Paper II.
**Results and discussion**

**Paper I. Quasispecies dynamics and molecular evolution of human NoV capsid P region during chronic infection.**

NoV may cause persistent infections in immune compromised individuals, and these constitute excellent settings to investigate NoV population dynamics and evolution. In this study, we have gained access to unique patient materials collected from an immune compromised heart transplant recipient chronically infected with GII.3 NoV. We have analyzed NoV population dynamics and *in vivo* capsid gene evolution within 17 different time point isolates (collected between August 2000 and March 2003), using capsid domain sequencing and the Evolutionary trace method.

Phylogenetic analysis of the capsid sequences revealed that they had evolved in a time-dependent manner where each isolate has descended from its predecessor in time, showing a clear pattern of quasispecies distribution. A suspected “bottle neck”, or evolutionary break point, occurred between January and March 2001. Most interestingly, this evolutionary event correlated with a change of clinical symptoms where the patients vomiting subsided.

In this study we concluded that the capsid P2 domain evolved under positive selection, where aa positions of the putative carbohydrate binding pockets remained conserved. After publication of Paper 1, Hoffmann and coworkers studied evolution of GII.4 and GIL7 within chronically infected individuals and confirmed the presence of positive selection (Hoffmann et al. 2012). This feature has also been observed by Bok and coworkers in chimpanzees challenged with human GI.1 NoV (Norwalk strain) (Bok et al. 2011). Altogether these findings suggests that positive selection is a common feature within NoV quasispecies populations regardless of genotype, at least within immune suppressed individuals. One can speculate that the quasispecies population might have a different molecular makeup within an immune competent individual, due to a higher immunological pressure.

In the recent years following publication of paper I, much progress have been made in the field of NoV capsid characterization. Among other things, a crystal structure of the epidemic 2004 GIL4 NoV variant TCH05 (Houston) have recently been published by Shanker and coworkers (Shanker et al. 2011). The presence of the two HBGA binding sites originally proposed by Cao and coworkers (Cao et al. 2007), which aa positions were found conserved in our study, have recently been confirmed by other studies and are now referred to as site 1 and 2 (Yang et al. 2010; de Rougemont et al. 2011; Lindesmith et al. 2011).

The sequencing method used in Paper I is cloning followed by MegaBace, altogether a very elaborate and time consuming methodology. However, today there are other techniques commonly available, and the method of choice for such a large material would clearly be ultra deep sequencing (Su et al. 2011).
Main findings of Paper I

- The capsid P1-1 and P2 domains evolved in a time-dependent manner with a clear quasispecies distribution.

- Positive selection within capsid P2 domain during chronic infection: aa positions of the putative carbohydrate binding pocket within the P2 domain are conserved during evolution.
Paper II. The G428A nonsense mutation in FUT2 provides strong but not absolute protection against symptomatic GII.4 NoV infection.

Before publication of paper II, previously reported GII.4 NoV outbreaks were strictly secretor dependent and non secretors were completely protected against disease (Thorven et al. 2005b; Kindberg et al. 2007a; Tan et al. 2008a). However most interestingly, in this study we are the first to report of symptomatic GII.4 NoV infection of a non secretor individual (homozygous for the G428A nonsense mutation).

This rare event occurred during an outbreak of acute gastroenteritis in an elderly nursing home in El Grao de Castello’n, Spain, 2004. The outbreak was caused by a GII.4a-2004 variant virus and in total 116 individuals were infected with an attack rate of 54.2%. When performing genotyping of the FUT2 gene, we found that the G428A nonsense mutation provided strong (P<0.001) but not absolute protection against disease. One non secretor individual was symptomatically infected (termed A), and saliva from one assymptomatic non secretor (termed B) was able to bind the outbreak virus in an ELISA assay. Unfortunately, due to logistic constrains there were no saliva available from patient A at the moment of analysis. In the ELISA assay we could conclude that saliva from secretors and the non-secretor (B) as well as Lewis positive and Lewis negative individuals bound the outbreak virus strain. We were then also able to (based on Lewis genotyping and the binding data) conclude Lewis status was not a susceptibility marker for symptomatic NoV infection.

To further characterize the binding ability of the non secretors, saliva samples from patient A and B were used in a VLP binding assay. The results showed that the secretor negative saliva could not bind VLP from GI.1 (Norwalk strain), GI.3 (ChronI strain) or GII.4 (Dijon strain). The finding that saliva from patient A and B do not bind the Dijon strain is most interesting, since the Dijon strain belong to the same genotype as the outbreak strain.

The novelty of the outbreak strain was later confirmed by phylogenetic analysis showing that the outbreak strain had distinct amino acids in antigenic A (aa 296 to 298) and B (aa 393 to 395) regions of the P2 domain compared to the Dijon strain. The location of these sites within the capsid P2 domain are illustrated in Figure 17. These antigenic sites or evolutionary hot spots are variant specific epitopes observed in the P2 domain of GII.4 isolates collected between 1997 and 2006 in UK (Allen et al. 2008). Amino acid alterations in these sites are suggested to impact on the biochemical properties as well as the entire structure of the P2 domain. In paper 2, we thereby hypothesize that the aa differences seen in site A and B and additional evolutionary hot spots between the Dijon and the outbreak strain, may affect the structural and electrostatic properties of the P2 domain. In that case this would subsequently alter the binding affinities for HBGAs and perhaps explain ability of the outbreak virus to infect non secretors.

However, it is most important to remark that the novel binding pattern observed for this outbreak virus, might not solely depend on the virus but also lie on the host level. It can not be excluded that the HBGA expression is different in patient A and B, compared to other non secretors. Also it can not be ruled out that the outbreak virus binds to a novel receptor, not related to the secretor or Lewis genes.
Figure 17. P monomer of GII.4 VA378 GII.4 NoV. The S domain is illustrated in yellow while the P domain is presented in blue. The location of antigenic site A (aa 296 to 298) is displayed in red, while the antigenic site B (aa 393 to 395) is displayed in green. The capsid protein is displayed and colored using Molsoft Browser Pro.

After publication of paper II, the presence of secretor independent NoV strains were confirmed by Nordgren and coworkers. They reported on a foodborne GI.3 NoV gastroenteritis outbreak, where symptomatic disease was as likely to develop in non secretors as in secretors (Nordgren et al. 2010).

Main findings of Paper II

- The G428A nonsense mutation in FUT2 provides strong but not absolute protection against symptomatic GII.4 NoV infection.
**Paper III. Pediatric sapovirus infections in Nicaragua and host genetic susceptibility.**

While human NoV susceptibility is highly associated with secretor status and thus mutations in the *FUT2* gene, no information is yet available if host genetic factors determine susceptibility to SaV. Neither is there any information available regarding the prevalence or the genetic diversity of SaV in Central America. Therefore in this study, we describe the genetic diversity of SaV in a Central American pediatric population and investigate for the first time the role of host genetic factors and susceptibility to SaV infections.

Fecal samples were collected from 292 children (≤5 years of age) that participated in a community- and hospital based survey of sporadic acute diarrhea in León, Nicaragua from March 2005 to September 2006. Out of these fecal samples 11% were SaV positive by PCR. Phylogenetic analysis showed that the vast majority (64%) was found to be GI, followed by GII with 16% and GIV with 12%. Co-infections with GI:GII and GI:GV were found in 8% of the symptomatic children.

Blood and saliva samples were also obtained from the SaV infected children, to determine histo blood group, secretor and Lewis status. These data suggests that SaV GI and GIV can infect secretors and non-secretors, Lewis positives as well as Lewis negatives and individuals of all ABO blood groups. However, it is probably most likely that there is a genotype specific infection pattern also for SaV, as have been observed for NoV (Tan and Jiang 2005; Rydell et al. 2011b; Tan and Jiang 2011). Therefore, due to the small number of samples genotyped (12 SaV positive samples genotyped; 9 belonged to GI.1, 2 were GI.6 and 2 were GI.1) the trend observed in our study must be further analyzed using a larger sample size.

Paper II concludes that SaV diarrhea is as common as NoV diarrhea in Nicaragua (Bucardo et al. 2008). The reason for this comparably high rate of SaV infections is unknown, but probably associated with poor sanitary conditions enabling the virus to spread easily between susceptible individuals. It is also important to remark that many of the children in endemic areas are most likely malnourished, a fact that may make these individuals more susceptible for infection. This suggests that factors contributing to the severity of SaV disease in Nicaragua might not only be related to viral properties, but also to host factors and the surrounding environment.

<table>
<thead>
<tr>
<th>Main findings of Paper III</th>
</tr>
</thead>
<tbody>
<tr>
<td>• SaV infections are equally prevalent as NoV infections in the pediatric population of Nicaragua.</td>
</tr>
<tr>
<td>• No statistically significant association was found between ABO blood types, Lewis phenotypes or secretor genotype and susceptibility to SaV infection in Nicaragua.</td>
</tr>
</tbody>
</table>
Paper IV. Susceptibility to symptomatic sapovirus infection in Denmark is not associated with secretor or Lewis status.

In order to further elucidate if host genetic factors determine susceptibility to SaV, this time the host genetics in a Caucasian population was investigated (as compared to the Central American population in Paper III). During October 2008 to November 2010 a total of 64 fecal, saliva and blood samples were collected from SaV positive individuals (≥4 years of age) in Denmark.

Phylogenetic analysis was performed on 39 of the 64 SaV positive samples, and genotyping of viral strains showed that 41% (16/39) belonged to genotype GI.2, 10% was GI.1 (4/39), 2.5% was GI.5 (1/39), 8% was GII.1 (3/39), 5% was GII.4 (2/39), 18% was GIV (7/39) and 15.5% was GV (6/39).

Secretor genotyping and Lewis phenotyping revealed that in similarity to the findings in Paper III, neither secretor status nor Lewis phenotype, were associated with susceptibility to symptomatic infection with SaV. However, in contrast to the Central American population in Paper III, in this study individuals of histo-blood groups B and AB had significantly lower risk to be infected (OR 0.18, p≤0.01 and OR 0.10, p<0.05, respectively). This difference in susceptibility between the Central American and Caucasian population, remains to be elucidated and must be confirmed in a larger sample size. It is most likely due to differences in the antigenic properties observed between different SaV genotypes. In conclusion, extended analysis using full capsid sequencing is required to further elucidate the suggested association between ABO blood group and SaV susceptibility found in Paper IV.

Main findings of Paper IV

- A statistical correlation was found between susceptibility to symptomatic SaV infection and ABO histo-blood group in a Caucasian population. Most interestingly, individuals of histo-blood groups B and AB had significantly lower risk to be infected (OR 0.18, p≤0.01 and OR 0.10, p<0.05, respectively).

- Individuals of blood type B (and AB) had a significantly less risk to be infected by SaV, whereas an association to FUT2 or Lewis status to susceptibility to SaV infections was not found.
Conclusions

In the recent years, much progress has been made in the field of NoV capsid characterization. When I started my PhD studies in 2006 there were only one crystal structure available; the prototype GI.1 Norwalk strain. A year later the recombinant P protein of a GII.4 strain norovirus, the VA387, was co-crystallized with synthetic type A or B trisaccharides. Following this, last year a crystal structure of the epidemic 2004 GII.4 NoV variant TCH05 (Houston) in complex with secretor and Lewis HBGAs was published. During the period of my PhD studies, it has become evident that there is not only one but two HBG binding sites on the NoV capsid; one more conserved binding ABH antigens and the other more variable and sensitive to evolutinal pressure with affinity to Lewis antigens.

The main focus for my doctoral studies and this thesis was to elucidate the possible correlation between viral evolution, host genetics and disease susceptibility. One of the main findings presented in this thesis is the documentation of in vivo capsid gene evolution and quasispecies dynamics during chronic NoV GI.3 infection (Paper 1). In paper II, we were the first to report that the G428A nonsense mutation in the FUT2 gene provides strong but not absolute protection against symptomatic GII.4 NoV infection. In my last two papers (Paper III and IV), we were the first to investigate host genetic susceptibility factors during authentic SaV infection.

When investigating host susceptibility patterns, both the virus and host factors must be considered. Different populations have different genetic background, nutrition status, immunity and age and may also display a different susceptibility pattern. A similar feature is also observed in the viral counterpart; a factor of great importance is strain variation and genotype specific infection patterns. In Paper I, II and III only the ABO histo blood group, secretor and Lewis status has been analyzed. However, it cannot be ruled out that other unknown factors besides these investigated, may affect susceptibility to human calicivirus infection.

To summarize, the results presented in this thesis show that the success of human calicivirus infection probably is determined by a delicate interplay between virus evolution and susceptibility of the host, both genetically and immunologically.
Present and future studies

Mechanisms of persistence and evolution of GII.4 NoV
The mechanism for emergence and evolution of NoV strains, as well as protective immunity is poorly understood. In the current study we have access to a unique collection of human sera ranging from as early as 1970-ties until 2008, ideal for studying viral evolution and herd immunity to NoV.

HBGA blocking assay
In total 250 serum samples collected between 1970 to 2008 will be used in an HBGA pseudo neutralization-blocking assay to investigate the interaction of GII.4 NoV VLP representing the epidemic MD145, Houston and Grimsby strains with the H type 1 carbohydrate, Figure 18. The blocking assay will be performed essentially as described previously by Reeck and coworkers (Reeck et al. 2010). A number of these sera samples are genotyped for secretor status (FUT2 G428A) and Lewis status (FUT3 nt position 59, 202, 314, 508, 1067), as well as phenotyped to determine the ABO blood group.

Figure 18. Study design. Serum samples collected between 1970 to 2008 will be used in an HBGA blocking assay to investigate the interaction of GII.4 NoV VLP representing the epidemic MD145, Houston and Grimsby strains with H type 1 carbohydrate.

So far we have only analyzed a small fraction of our serum samples, but we have found some interesting trends regarding their ability to block binding of Houston-2002-GII.4 VLP to the H antigen. Most interestingly, we found no correlation between preexisting antibody titers and blocking antibodies towards Houston-2002-GII.4 VLP in serum collected during late 1990-ties.
We will now continue to analyze the remaining sera with Houston-2002-GII.4 VLP to confirm our preliminary results. The next step is also to extend the H-type 1 blocking assay analysis using VLPs isolated at other time points (Grimsby-1996-GII.4 VLP and MD145-1987-GII.4 VLP). So far we have performed initial blocking experiments to adjust start point serum dilutions and VLP concentrations for the Grimsby-1996-GII.4 VLP (data not shown). These preliminary results are very interesting since they indicate that some serums from the late 1970-ties have high BT50 titers (>10 000).

Altogether, we believe that this study will be most interesting and may shed some new light into evolution of GII.4 NoV well as development of host immunity.
Acknowledgements

I would like to acknowledge some very special people that has, in one way or another, stood by me and supported me during my PhD studies. A big thank you to:

My supervisor Lennart Svensson, for your unending support and always believing in me. Thank you for always taking the time to answer a question, no matter the hour of the day.

To Jorma Hinkula, for your kindness and in spite of a hectic schedule, always finding the time to give lectures in the virology course.

To Jenny Stjernberg, my dear friend and colleague, words can not express my gratitude to you! You are one of the persons I admire most, you are not only an excellent scientist but also an excellent human being. You were my light in all the tears during this winter. A special thanks also to Gunnar and Linus Stjernberg, for great friendship and keeping us company in the woods of Källstorp.

To my dear friends Louise Carlsson and Josefin Kraft, what would I have done without you? Josefin, the pink pages in this thesis are for you my friend!

To my dear friend Gabriella Andersson, for being my friend for as long as I can remember. Long distances make us meet way too seldom, but every time I see you it is like we never have been a part.

My fellow group members (past and present):

Thank you; Marie Hagbom, for your kind heart and your nice company in the office. Johan Nordgren, for your friendship and company in the lab during all these years. I hope you find your Shakira one day! Elin Kindberg, for wonderful company both in the office and on various conferences all over the world. Malin Vildevall, for the years of friendship and nice collaboration. Carolin Jönsson, for all your help and support. I miss you but I wish you the best of luck with your new job. To Claudia Istrate, for your wonderful personality! Filemon Bucardo, for an unforgettable visit to Nicaragua, it will always be in my heart. Sumit Sharma, for being the most pleasant English dictionary I ever met. Sonja Bialowas, for nice company at the Smögen conference and always having a helping hand.

To the "HIV-group": Marie Larsson, Karlhans Fru Che, Vegard and Veronica Tjomsland for nice company at various conferences and pleasant discussions over a cup of coffee or lunch. A separate and special thank you to my friend and colleague Rada Ellegård, for your positive attitude and the amazing illustrations you made for my thesis! I owe you many bars of chocolates…

To my friends and colleagues in "tjejgänget" on level 13; Amanda Nordigården, Tina Falkeborn, Camilla Halvarsson, Josefín Åhsberg, Emmy Lagerberg and Johanna Wällå. Thank you for all nice fikas and lunches. Also a BIG thank you for always inviting me to various after work activities, despite the fact that I rarely have the time to show up! ☺
To Pia Druid, “Lotta” Lenner and Annika Starkenberg, for all the nice “horse talks” in the hallway or during coffees and lunches 😊

To Mary Esping and Lotta Berglund, for invaluable support when organizing the virology course and always having a helping hand.

To our collaborators in Gothenburg Gustaf E. Rydell and Göran Larson, for nice collaborations and being my guides in the jungle of carbohydrates.

Finally, last but not least, to my family:
To my big brothers Jan-Åke and P-O Karlsson. Even though we meet way too seldom, I feel your support and I know you are always there when I need you. A special thank you for everything you have done during this winter when we lost dad and I had to put my self together and finish this thesis.

Till min älskade mamma Majvor Karlsson, tack för allt stöd, all kärlek och alla timmar (dagar) av barnvakt som du och pappa har ställt upp med. Du är den starkaste person jag känner och jag beundrar dig så mycket.

To my husband and love of my life Andreas Carlsson, I owe this thesis to you! ppm. To our son Wilhelm, for making me realize what is important in life. I love you.
References


