Fakultät für Medizin
Institut für Virologie

Genotypic and phenotypic analysis of norovirus evolution in chronically infected patients

Suliman Qadir Afridi

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines

Doctor of Philosophy (Ph.D.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Roland M. Schmid

Betreuerin: Prof. Dr. Ulrike Protzer

Prüfende der Dissertation:

1. Prof. Dr. Markus Gerhard
2. Priv.-Doz. Dr. Sabrina Schreiner

Die Dissertation wurde am 11.12.2018 bei der Fakultät für Medizin der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 09.03.2019 angenommen.

Table of Content

ZUSAMMENFASSUNG

Noroviren (NoV) sind eine der häufigsten weltweiten Ursachen für akute Gastroenteritis bei Kindern und Erwachsenen. Sie sind in der Umwelt extrem stabil und wenige Viruspartikel können eine Infektion auslösen. Aufgrund der zunehmenden Anzahl immungeschwächter Patienten werden chronische NoV Infektionen in den letzten Jahren häufiger beobachtet. Ohne verfügbare antivirale Therapie und Impfungen sind diese Infektionen sowohl wissenschaftlich als auch klinisch herausfordernd. Wir analysierten NoV Evolution in chronisch infizierten Patienten mittels „next generation sequencing“ und untersuchten Antikörperantworten mittels ELISA. Dem schlossen wir Strukturmodellierungen und molekular-dynamische Simulationen des viralen Capsids an. Ausgehend von sequentiellen Stuhlproben von Patienten mit akuten und chronischen Infektionen sequenzierten wir Capsid Gene und analysierten Selektion sowie virale Quasispezies. Auch wurde ein quantitativer ELISA zur Detektion NoV-spezifischer IgG Antikörper entwickelt. ELISA Platten wurden mit rekombinant exprimierter P-Domäne des Capsidproteins gecoatet, gebundene NoV-spezifische Antikörper wurden mit HRP-konjugiertem Protein A nachgewiesen. Der Assay wurde durch vor und nach Immunisierung mit GII.1 und GII.4 Capsidprotein entnommenen Maussera, sowie sequenziellen Sera GII.1 und GII.4 infizierter Patienten validiert. Durch GII.4 Immunisierung generierte Antikörper kreuzreagierten mit GII.1 Capsidprotein. Vorbestehende Antikörper wurden in allen vor NoV Infektion entnommenen humanen Serumproben gemessen, Antikörperkonzentrationen stiegen in der 2. Woche nach Infektion deutlich an. Auch in sequentiellen Sera chronisch infizierter Patienten stiegen Antikörperkonzentrationen gegen Capsidproteine autologer Norovirus Sequenzen über Monate und Jahre an. Dies spricht für zunehmende adaptive B-Zellantworten und passt zu folgender Ausheilung.

In einem chronisch infizierten Patienten veränderten sich NoV-spezifische Antikörperkonzentrationen anders. Nach niedrigem Signal gegenüber allen getesteten Antigenen im ersten Serum, ergab das 2. Serum ein hohes Signal gegenüber dem 1. Antigen, das mit Antigenen aus späteren Zeitpunkten deutlich abnahm. Dies deutet auf möglichen Immunescape hin. Passend dazu zeigen sequenzielle Capsid Sequenzen Mutationen in B-Zell Epitopen.

Sequenzanalyse ergab einen hohen Anteil nichtsynonymer Mutationen in 7 chronisch Infizierten (52-78 %), während sie in akut Infizierten selten waren (4.5-6.1 %). Die Variabilität viraler Quasispezies war in chronisch Infizierten erhöht, verringerte sich aber in den letzten Proben vor Ausheilung. Die

meisten Mutationen traten in der P2 Domäne des Capsidproteins auf, besonders in den Epitopen A, D und E. Ausgehend von patientenspezifischen Capsid Sequenzen wurden Strukturmodelle der Capsidproteine generiert. Veränderungen zwischen sequentiellen Proben wurden mit molekular dynamischen Simulationen analysiert. Diese ergaben veränderte Konformationen der oberflächlich gelegenen Epitope, D und E. Diese Regionen waren auch im Zeitverlauf flexibler und weniger stabil.

Die in dieser Arbeit beschriebenen genotypischen und phänotypischen Daten zeigen, dass Noroviren effizient an das humorale Immunsystem adaptieren können. Diese Befunde können für die Verbesserung der Therapie und Impfprophylaxe von NoV Infektionen hilfreich sein.

SUMMARY

Norovirus (NoV) is known to be a leading cause of acute gastroenteritis in children and adults worldwide. Extreme virus stability and the very low infectious dose facilitate extended outbreaks. As yet, there is neither antiviral therapy nor vaccine is available against NoVs. In the last years, chronic NoV infection has become a scientifically and clinically challenging condition in immunocompromised patients. In this view, we analyzed NoV evolution in chronically infected patients with next generation sequencing and determined humoral immune responses by ELISA. This was followed by structural modeling and molecular dynamics simulation of the viral capsid.

Starting from sequential stool samples of patients with acute and chronic NoV infection next generation capsid sequences were analyzed for selection and quasispecies were reconstructed. A quantitative ELISA detecting human NoV-specific IgG was developed by coating ELISA plates with recombinantly expressed P domain of NoV GII.1 capsid protein and using HRP-conjugated protein A. The assay was validated by testing mouse sera drawn before and after immunization with GII.1 and GII.4 capsid proteins as well as sequential sera from GII.g/GII.1 and GII.4 infected patients. Antibodies elicited by GII.4 immunization were cross-reactive to GII.1 capsid protein. Pre-existing NoV antibodies were detected in all serum samples before infection, but titers raised in the second week after infection. Also in sequential sera of chronically infected patients antibodies directed against the capsid proteins of the respective autologous NoV strain increased over months to years, particularly against earlier isolates. This indicated an increasing adaptive B-cell response and was in line with a subsequent cure.

In one chronically infected patient NoV-specific antibody concentrations developed differently. The first serum showed low reactivity with all sequential capsid antigens, as observed in the other patients. Interestingly the 2nd serum recognized the first antigen well but yielded steadily decreasing signals with the following capsid antigens. This indicated an immune escape of NoV varying the second and following sequential capsid proteins. The third serum produced high signals with all antigens in line with an effective humoral immune response finally leading to cure.

Sequence analysis showed that non-synonymous mutations were frequent (52-78 %) in seven chronically infected patients whereas, in acutely infected patients they were rare (4.5-6.1 %). In chronically infected patients, the variety of NoV quasispecies was found to be increased but again

decreased in late samples taken during chronic infection. Amino acid and nucleotide substitution were most frequent in the P2 domain of VP1 capsid protein, particularly in epitopes A, D and E.

Starting from patient-derived capsid sequences, structural models of the capsid proteins was generated. Molecular dynamic simulations were carried out to analyze structural changes in the sequential samples. This revealed that the epitopes D and E changed their conformation on the capsid protein surface. These regions were more flexible and less compact over time.

Genotypic and phenotypic data described in this thesis showed that NoV can adapt in particular in response to humoral immunity. The findings will be helpful to improve vaccine prophylaxis and therapy of NoV infections.

1 INTRODUCTION

1.1 Norovirus:

In October 1968 an outbreak led to the identification of NoV. Pupils of an elementary school in Norwalk, USA presented with nausea, low-grade fever, vomiting and diarrhea. The 27-nm virions were seen by electron microscope in stool samples. First the name "Norwalk virus" virus was chosen and later changed to norovirus (Kapikian et al., 1972). Norovirus is highly infectious and can cause infection even with few particles (L. Lindesmith et al., 2003). Pathogenesis is characterized by heterogeneous host virus interactions, such that all individuals are not equally susceptible to NoV infections and not all NoVs are equally pathogenic (Friesema et al., 2009; L. Lindesmith et al., 2003). NoVs have not always been associated with more severe disease in individuals with weakened immune functions, and this recent pattern is thought to be linked to genetic differences between emerging and historical human strains of NoV (Miranda de Graaf, van Beek, & Koopmans, 2016). The nucleotide sequences of the genomes of different NoV genogroups have 51–56% similarity with one another, and when comparing only ORF2 sequences rather than whole genomes the diversity between genogroups is much higher (Katayama et al., 2002; Donaldson, Lindesmith, LoBue, & Baric, 2010b). Despite differences in selection pressures between ORF1 and ORF2 and frequent recombination, the phylogeny of ORF1 has a same topology and a high genetic diversity to the phylogeny of ORF2 (Vinjé et al., 2015). Interestingly, some outbreaks are caused by strains of NoV that are genetically similar or identical to strains that were isolated 10 to 15 years earlier, raising questions about the reservoirs in which these viruses are maintained between outbreaks (Lysén et al., 2009).

1.1.1 Symptoms:

The clinical symptoms of NoV infection include watery diarrhea, vomiting, fever, headache, abdominal cramps, the presence of mucus in stool, chills and myalgia, and with incubation period of 12–48 hours. Vomiting can occur on the first day on exposure to the virus (Atmar et al., 2008; Rockx et al., 2002) and delay in gastric emptying may be a reason (Meeroff, Schreiber, Trier, & Blacklow, 1980). Diarrhea is reported to be the most frequent symptom and 18 hours after infection virus shedding in feces begins (Atmar et al., 2008). Pathology of NoV infections includes villous blunting, an increase in the number of apoptotic epithelial cells at the luminal surface of the duodenum and a

more substantial increase in the number of intraepithelial lymphocytes (Schwartz et al., 2011). In healthy individuals, NoV infection results in villous blunting and infiltration of mononuclear cells at the jejunum (Schwartz et al., 2011). It also leads to epithelial-barrier dysfunction that manifests as an increase in the number of both cytotoxic intraepithelial lymphocytes and apoptotic epithelial cells, and a decrease in protein abundance at tight junctions (Troeger et al., 2009). This dysfunction of the epithelial barrier causes diarrhea through a leak flux mechanism in which ions and water diffuse from subepithelial capillaries into the intestinal lumen (Troeger et al., 2009). Not all NoV genotypes seems to be equally pathogenic, GII.4 stains cause more severe illness than other genotypes (Friesema et al., 2009; Desai et al., 2012; Huhti et al., 2011). However, different virus shedding might be due to underlying age of infected individuals and health status rather than the virus genotype (Milbrath, Spicknall, Zelner, Moe, & Eisenberg, 2013).

1.1.2 Genome:

The genome is a 7.5 kb long RNA, protected by an icosahedral protein capsid of 38 nm in diameter (Thorne & Goodfellow, 2014). Its positive-sense RNA, directly serves as mRNA for protein translation. The genome contains three open reading frames (ORFs) that encode the proteins required for genome replication and assembly of new viral particles. The ORF1 encode for six non-structural proteins (Figure 1) including protease and RNA-dependent RNA polymerase (known as NS7) and VPg (Subba-Reddy, Goodfellow, & Kao, 2011). VPg, also known as NS5, covalently linked at the 5' end of the genome and play important role in virus replication (Subba-Reddy, Yunus, Goodfellow, & Kao, 2012). The major structural protein (VP1) and minor structural protein (VP2) are encoded by the ORF 2 and 3, respectively (Thorne & Goodfellow, 2014). The X-ray crystal structure of the prototype NoV VLPs shows that NoV capsid is composed of 180 capsid monomers forming a T=3 icosahedral capsid (Prasad et al., 1999). The capsid protein is divided into S and P domain linked together by a hinge of eight amino acids, as shown in Figure 1.1 and 1.2 (Thorne & Goodfellow, 2014). The P domain is further divided into P1 and P2 domains. The P2 domain is considered highly mutagenic and is responsible for attachment to the histo-blood group antigen. Replication of the virus occurs in the cytoplasm in close association with host-derived membrane complexes (Thorne & Goodfellow, 2014; Wobus et al., 2004). Hence VP1 is the target of the host immune system and responsible for the survival of the virus, it shows the highest mutation rate within the viral genome.

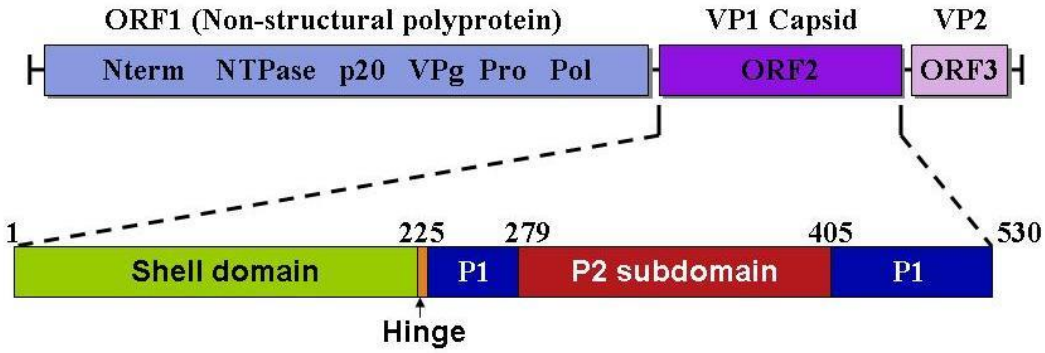

Figure 1.1: Norovirus genome. The genome consists of three open reading. ORF1 encodes a nonstructural polyprotein which is cleaved into six nonstructural proteins; ORF2 encodes VP1 capsid protein including S and P domains; and ORF3 encodes minor structural protein VP2. Adapted from Zartash et al (2015).

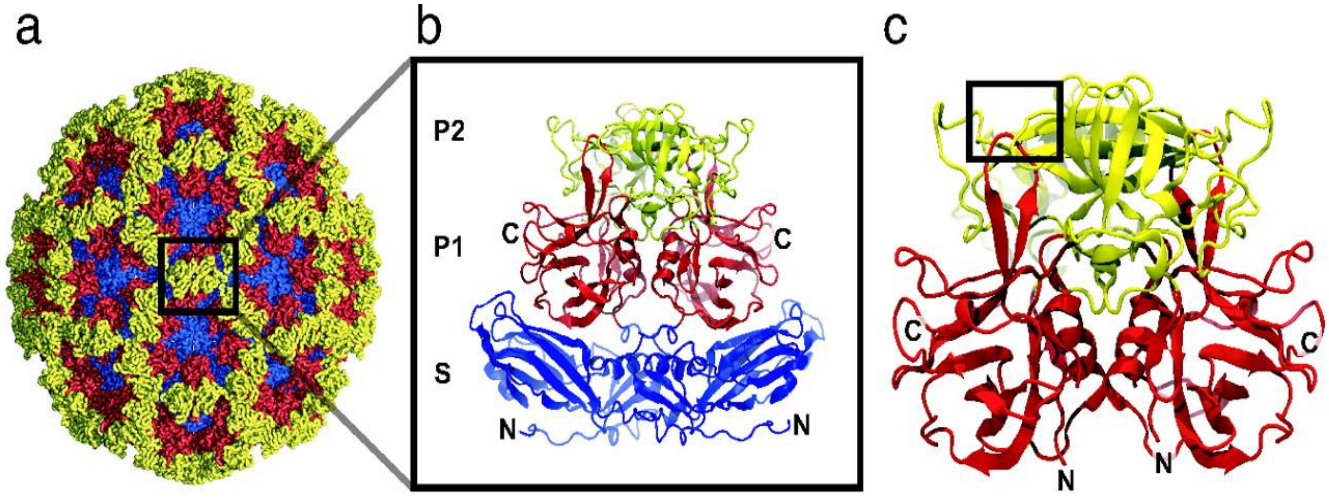

Figure 1.2: Norovirus Capsid protein. (a) The T=3 capsid structure of norovirus determined by 3.4-Å resolution formed by 90 dimers of capsid protein VP1. (b) A ribbon representation of VP1 dimer extracted from the capsid structure with S, P1 and P2 domains. (c) The structure of the P domain construct determined to 1.4-Å resolution. Adapted from Choi et al. (2008)(Choi, Hutson, Estes, & Prasad, 2008).

1.1.3 Genotypes:

NoV is divided into six genogroups, which are further subdivided into at least 30 genotypes (Vinje, 2014). Genogroups GI, GII and GIV are known to cause infections in humans (Vinje, 2014). The other genogroups be found in wide range of hosts, felines, pigs, canines, rodents, sheep's, sea lions, bats and cattle (Vinjé et al., 2015; Li et al., 2011; Wu et al., 2016). Each group contains a different number of genotypes, labeled with Arabic numbers (Figure 1.3).

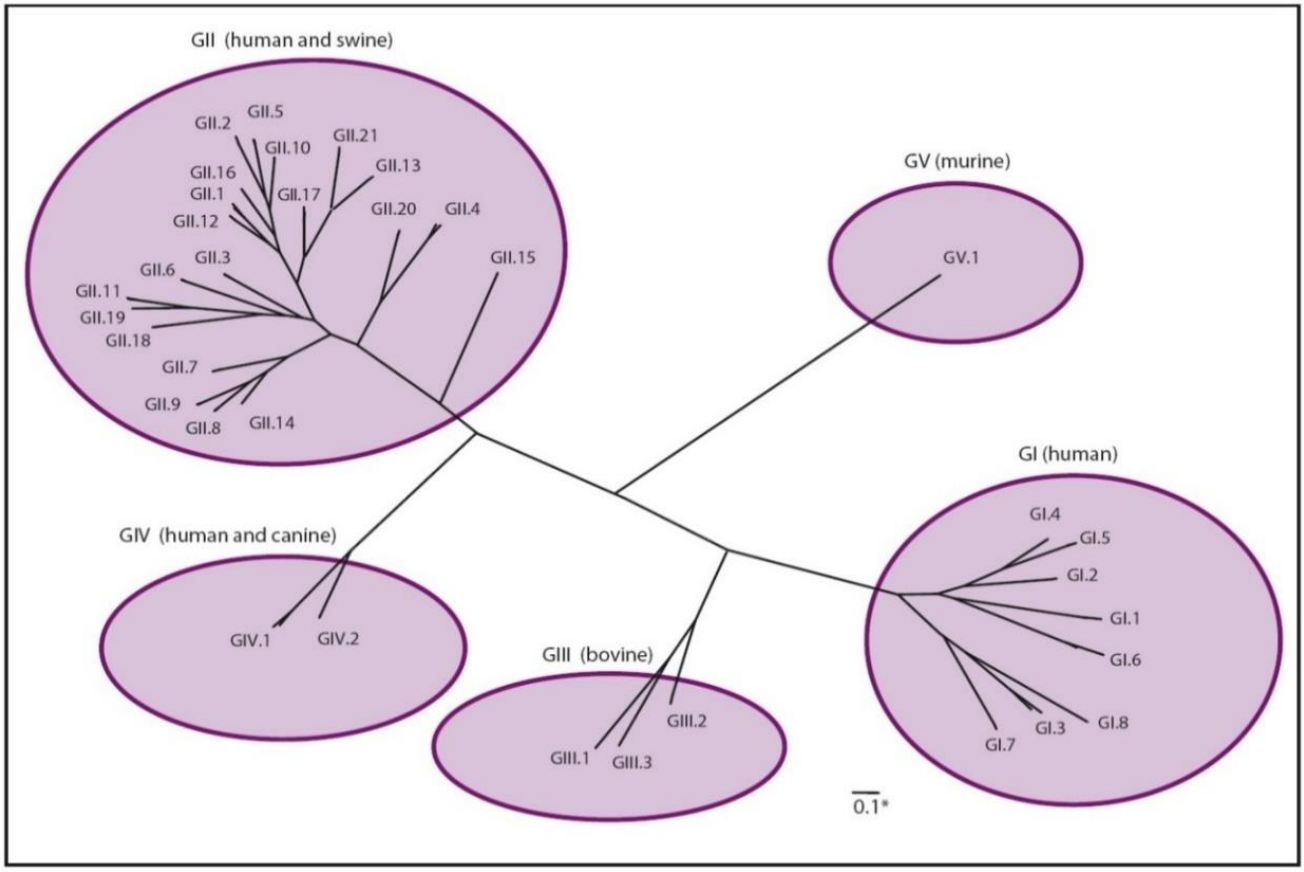

Figure 1.3: NoV genogroups. NoV is divided into five genogroups and 35 genotypes (Zheng et al., 2006).

Over the past 13 years, 70 to 80% of all reported outbreaks are cause by the GII.4 genotype variants (Hoa Tran, Trainor, Nakagomi, Cunliffe, & Nakagomi, 2013). Prevalence of genotypes differs between routes of transmission and human population (Thorne & Goodfellow, 2014). GII.4 is more often associated with transmission mediated by person to person contact than with other types of transmission, whereas, other genotypes such as GI.3, GI.6, GI.7, GII.3, GII.6 and GII.12 are more often associated with foodborne transmission (Verhoef et al., 2015). GI variants are more often associated with waterborne transmission (Lysén et al., 2009). This may be concluded that GI variants have a higher stability in water than GII variants (Miranda de Graaf et al., 2016). GII.3 strains are frequently detected in patient samples specifically in children. It evolves at a rate of 4.16 × 10 to 3 nucleotide substitutions per site per year, which is almost similar to evolution rate of GII.4 and GI strains (Denali Boon et al., 2011; Kobayashi et al., 2016).

1.1.4 Epidemiology and evolution:

NoV is known as a leading cause of acute gastroenteritis in children and adults worldwide (Hall, Glass, & Parashar, 2016; Ahmed et al., 2014). According to a recently published WHO report, NoV caused 120 million cases of diarrheal diseases, more than any other infectious agent or foodborne pathogen (Havelaar et al., 2015). GII.4 cause most of the infections in different facilities as shown in figure 1.4. Evolution of NoV is complex because of its high mutations rate. A better understanding of the genetic and antigenic evolution is of crucial importance for the development of vaccines and treatment strategies (Miranda de Graaf et al., 2016).

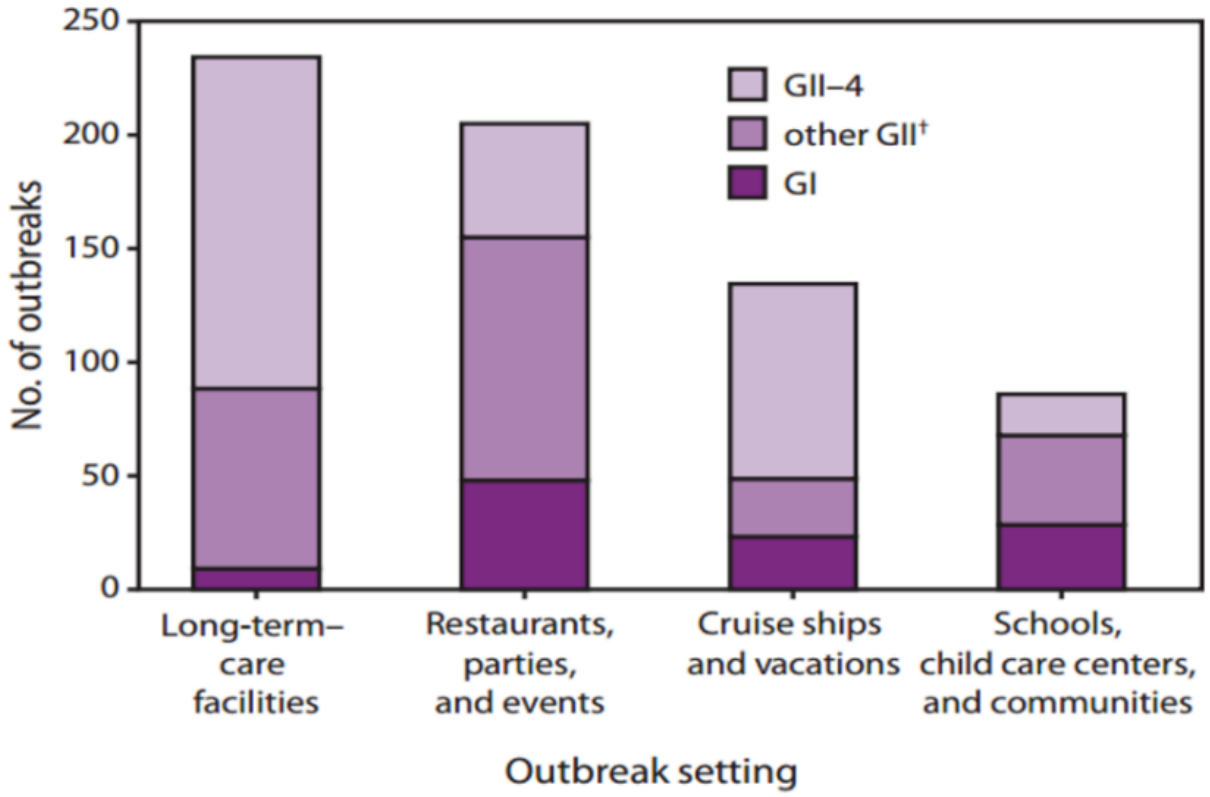

Figure 1.4: Outbreaks in United States from 1994 till 2006. The Figure illustrates the number of norovirus outbreaks laboratory-confirmed by CDC for the United States during 1994-2006 by setting and genotype. Four categories of outbreak settings are shown: 1) long-term-care facilities; 2) restaurants, parties, and events; 3) cruise ships and vacations; and 4) schools, child care centers, and communities. (Zheng, Widdowson, Glass, & Vinjé, 2010).

Studies have shown that globally circulating GII.4 variants are frequently replaced by newly emerged antigenically distinct GII.4 variants. This indicates that immunogenic pressure influences NoV evolution (Donaldson et al., 2010a). An increase in NoV outbreaks coincide with the emergence of antigenically divergent GII.4 variants (J. J. Siebenga et al., 2010). In vitro assays and bioinformatic

analyses have shown that GII.4 variants have high rate of mutations and evolution, facilitating the rise of these antigenically divergent strains (Bull, Eden, Rawlinson, & White, 2010). GII.4 isolates across the globe show some GII.4 lineages that can cause epidemics (J. J. Siebenga et al., 2009). Recently, major outbreaks of the newly emerged GII.17 variants were occurred in some parts of Asia, where GII.4 Sydney 2012 strain has been replaced by GII.17 NoV (M de Graaf et al., 2015). Although in Europe, the United States and Australia the GII.P17-GII.17 genotype has also been detected in gastroenteritis cases, but still it has not replaced GII.4 strains in these regions (M de Graaf et al., 2015). Since 1995, six antigenically variant NoV GII.4 strains caused pandemics: US 1995/96, Farmington Hills 2002, Hunter 2004, Den Haag 2006b, New Orleans 2009 and Sydney 2012 (J. S. Eden et al., 2014). Moreover, studies to describe these variances have been hampered by the lack of a cell culture system for NoV. Recently, a cell culture model has been described for human NoV (Veronica Costantini et al., 2018), and ongoing studies may provide molecular insights into the increased risk posed by emerging strains of NoV (Jones et al., 2014).

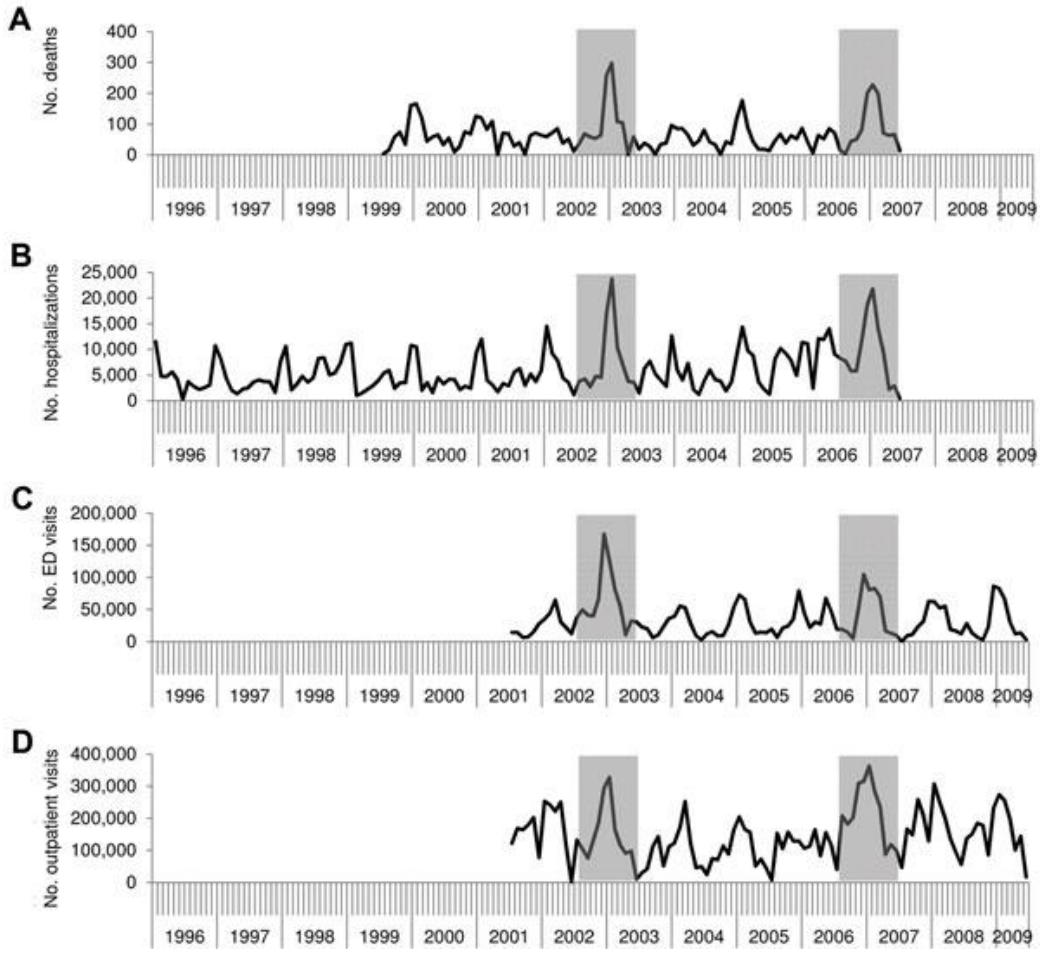

Figure 1.5: Number of norovirus-associated deaths, hospitalization in United States from 1996 till 2009. (A) norovirus-associated deaths, B) hospitalizations, (C) emergency department (ED) visits, and (D) outpatient visits across all age groups, by month and year, United States (Hall et al., 2013).

Non-GII.4 strains are subject to less adaptive pressure. Despite the same nucleotide substitution rates, the accumulation of amino acid mutations is much lower for GII.3 strains than for GII.4 strains (D. Boon et al., 2011). GII.2 strains also show limited antigenic evolution, although some genetic drift has been observed (Iritani et al., 2008; Swanstrom, Lindesmith, Donaldson, Yount, & Baric, 2014). Recombination of human NoV is common and is thought to be an important mechanism by which genetic diversity is generated (Miranda de Graaf et al., 2016). Recombination between human NoV most commonly occurs near the junction of ORF1 and ORF2 and is associated with antigenic shift and a less common site of recombination is located at the junction of ORF2 and ORF3 (John-Sebastian Eden, Tanaka, Boni, Rawlinson, & White, 2013).

1.1.5 Transmission:

NoV transmission occurs by person to person contact, either through fecal-oral route or exposure to contaminated food and water (Wikswo et al., 2014). Infected individuals shed high loads of the virus in stool (Atmar et al., 2008). High infectivity of new NoV strains can emerge and cause global epidemics (J. J. Siebenga et al., 2009). Outbreaks commonly occur on cruise ships, in the military, nursing homes, child care centers, and schools due to the extreme stability, low infectious dose and high transmissibility of the virus. Healthcare facilities and restaurants are more affected by NoV resulting in significant morbidity and mortality and substantial healthcare costs (Atmar & Estes et al., 2006). Foodborne transmission is important route for global spread (Verhoef et al., 2015) and can occur when food is contaminated by food handlers in food production (Rodríguez-Lázaro et al., 2012). For example, shellfish that are cultivated in coastal areas can be contaminated by faecal discharge (Le Guyader, Atmar, & Le Pendu et al., 2012); frozen berries are most likely to be contaminated by irrigation with sewage-contaminated water or by contact with infected personnel during harvesting and processing (Rodríguez-Lázaro et al., 2012). Outbreaks can include mixtures of virus variants, increasing the risk of viral recombination (Lysén et al., 2009). The global scale of foodborne outbreaks can be difficult to recognize, because the epidemiology of outbreaks is often tracked independently by individual countries; nonetheless, retrospective studies have shown that approximately 7% of foodborne outbreaks are part of an international event with a common source (Verhoef et al., 2011). NoV rank among the top causes of foodborne disease globally (WHO Foodborne Disease Burden Epidemiology Reference Group 2007-2015, 2015). Nosocomial transmission in hospitals is a major burden for in-patient services (Ahmed et al., 2014). Individuals are most likely to shed viral particles in considerable numbers for several weeks after the resolution of symptoms (Rockx et al., 2002). However, nosocomial outbreaks suggest that most of these outbreaks are the result of transmission from symptomatic shedders (Sukhrie et al., 2012).

1.1.6 Chronic shedders:

Immunocompromised patients, chronically infected with norovirus in a hospital setting can act as a virus reservoir and may contribute to nosocomial transmission (Sukhrie, Siebenga, Beersma, & Koopmans, 2010; Milbrath et al., 2013). As a consequence of prolonged shedding and limited immune selection pressure, these immunocompromised patients can produce numerous NoV variants (Sukhrie et al., 2010). Within host, viral variation in a chronic shedder can cause the antigenic variations that

are seen between consecutive pandemics, and some of these variants are most likely to escape herd immunity (K. Debbink et al., 2014).

1.1.7 Animal noroviruses:

There is some evidence for the transmission of NoV between different host species, however infection of humans with animal NoV has not yet been reported (Miranda de Graaf et al., 2016). Human NoVs have been detected in the stool of cattle, pigs and dogs (Mattison et al., 2007; Summa, von Bonsdorff, & Maunula, 2012). Pigs and gnotobiotic calves can become experimentally infected with human GII.4 strains (Souza, Azevedo, Jung, Cheetham, & Saif, 2008; Cheetham et al., 2006). Furthermore, canine seroprevalence to different human NoV genotypes resembles the seroprevalence in the human population and serum antibodies against bovine and canine NoVs (S. L. Caddy et al., 2015). Murine NoV is widely studied as an animal model for NoV infection, since the discovery of the first murine NoV in 2003 (Karst, Wobus, Lay, Davidson, & Virgin IV, 2003) and subsequent identification of various genetically related viral isolates from geographically diverse locations (Hsu, Riley, Wills, & Livingston, 2006; Thackray et al., 2007).

1.1.8 Norovirus infection in immunocompromised patients:

NoV infection is self-limiting in healthy individuals but in immunocompromised patients especially elderly and young children it causes severe complications (Kroneman et al., 2008; Murata et al., 2007). Prolonged infection has been reported in individuals of all ages with congenital immunodeficiencies, in patients undergoing cancer chemotherapy, individuals who are infected with HIV and transplant patients who are receiving immunosuppressive therapy (J. J. Siebenga et al., 2008; Green, 2014). Increasing number of patients hematopoietic stem cell transplants or receiving solid organ, these groups has expanded substantially in the past few decades (Miranda de Graaf et al., 2016). Younger children with inherited immune deficiencies are frequently exposed to NoV infection which can last for several months (Pierre et al., 2012). In immunocompromised individuals, chronic NoV infection arises because the individuals are unable to clear the virus following acute gastroenteritis. By contrast, to date no persistent infections have been described in healthy individuals (Miranda de Graaf et al., 2016).

1.1.9 Norovirus cell receptors:

HBGAs are considered as host receptor for NoVs (Huang et al., 2003; Hutson, Atmar, Marcus, & Estes, 2003). Human HBGAs are complex carbohydrates linked to glycolipids or glycoproteins present on mucosal epithelial cells, red blood cells or free antigens in biological fluids such as saliva or blood (L. Lindesmith et al., 2003; Marionneau et al., 2002). They are determinants of both Lewis blood group systems and the ABO blood group (Stanley, P. & Cummings, R. D 2009). In certain cell types, α (1,2)-fucosyltransferase 2 (FUT2; also known as galactoside 2-α-l-fucosyltransferase 2) adds a fucose group to precursors of HBGAs, generating H HBGAs, and subsequent reactions generate A and B HBGAs. The binding specificity of NoV VP1 to different HBGAs differs among the genotypes and genogroups (Donaldson, Lindesmith, LoBue, & Baric, 2010b), resulting in changes in the susceptibility of individuals to specific strains of NoV (Thorven et al., 2005; L. Lindesmith et al., 2003). Individuals lacking FUT2 are known as non-secretors, because A, B and H HBGAs are not present in the bodily secretions of these individuals (Le Pendu, Ruvoën-Clouet, Kindberg, & Svensson, 2006). At least 20% of Northern Europeans are non-secretors (Thorven et al., 2005). Mesoamerican ancestry children are more likely to be secretors than of European children or African ancestry (Currier et al., 2015). Non-secretors have been shown to be less exposed to infection with several GI and GII strains of NoV (Thorven et al., 2005); L. Lindesmith et al., 2003) It should be considered that although A, B and H HBGAs correspond to the A, B and O blood groups, the ABO blood group system is independent of secretor status, as FUT1 rather than FUT2 synthesizes these HBGAs for expression on the surfaces of erythrocytes (Shirato et al., 2008). Individuals with blood type B or blood type AB were also found to be less vulnerable to infection with certain strains of NoV than individuals with blood type A or blood type O (Hutson, Atmar, Graham, & Estes, 2002; Thorven et al., 2005). Similarly to human NoV, animal NoVs use carbohydrates as attachment factors. Murine NoV 1 (MNV-1) binds to terminal sialic acids, but the exact glycan structures that are bound by murine NoV differ in a strain-dependent manner (Taube et al., 2012). Bovine NoV binds to α-galactose as a terminal residue of carbohydrates (Zakhour et al., 2009). Whereas, canine NoV belonging to genogroups GIV and GVI recognize α (1,2)-fucose-containing H and A HBGAs (S. Caddy, Breiman, le Pendu, & Goodfellow, 2014). Variation in the binding of carbohydrates between animal and human NoV are likely to have a role in the observed species specificity. In addition to HBGAs, it is likely that other factors may contribute to the attachment and/or entry of NoV into host

cells. However, if such factors exist, their roles in NoV pathogenesis have yet to be characterized (Miranda de Graaf et al., 2016).

1.1.10 Laboratory diagnosis:

Real time PCR (RT-PCR) is mostly used for identification and quantitation of NoV genome since this assay has high sensitivity and specificity and provide fast result (Baert, Uyttendaele, & Debevere, 2007; Farkas et al., 2015; Fuentes et al., 2014; Neesanant et al., 2013; Yan et al., 2013).

Nevertheless, the specificity and sensitivity of RT-PCR methods vary due to high diversity in NoV genomes (Kageyama et al., 2003; Mattison et al., 2011; Schultz et al., 2011). Several commercial RT-PCR assays are available in the market (Butot et al., 2010; J Hyun, Kim, & Lee, 2014), although commercial RT-PCR assays revealed a lower detection rate of NoVs and could not detect a majority of NoV GI strains as compared with the European international assay (Butot et al., 2010). Several enzymatic immunoassays (EIAs) are commercially available for the detection of NoV GI and GII antigens in stool samples (Robilotti, Deresinski, & Pinsky, 2015). The most commonly performed EIAs are IDEIA NoV (Oxoid Ltd., Hampshire, United Kingdom) and Ridascreen Norovirus (R-Biopharm, Darmstadt, Germany). These solid-phase, sandwich-type immunoassays demonstrate a wide range of sensitivities and specificities (Robilotti et al., 2015). Sensitivities and specificities for IDEIA NoV range from 38.0 to 78.9% and 85.0 to 100.0%. Similarly, the specificities and sensitivities for Ridascreen range from 31.6 to 92.0% and 65.3 to 100.0% (Robilotti et al., 2015). Factors affecting these differences in performance may contribute to viral load present in the stool samples and viral genotypes present in the sample, as the assay antibodies show differential genotype affinities. These characteristics may also be influenced by the clinical context of collection (outbreak versus sporadic cases), patient demographics (pediatric versus adult) and the timing of collection relative to symptom onset (Robilotti et al., 2015).

Several studies which evaluate the performance of EIA for detection of NoV in outbreak investigations compared to sporadic gastroenteritis cases shows that these assays are more sensitive in the outbreak setting, particularly if multiple samples are collected (Verónica Costantini et al., 2010; Gray et al., 2007). For example, in a study by Gray et al. (Gray et al., 2007), the sensitivity of 33.3% and 44.4% was reported for the IDEIA and Ridascreen approaches when two specimens per outbreak were tested. Sensitivity increased to 80.0% for both methods if seven specimens per outbreak were tested.

1.1.11 Therapy:

Clinical intervention efforts for NoV infection are hampered by the lack of a licensed vaccine and limited success on the antiviral treatment options that are currently available (Gairard-Dory et al., 2014). Successful treatment of the disease in individuals who were chronic shedders has been observed with oral human immunoglobulin, although clearance of the virus in some patients were not observed (Florescu, Hill, McCartan, & Grant, 2008; Gairard-Dory et al., 2014). Additional studies will be required to describe whether the route of administration and/or the levels of antibodies specific to the infecting strain of NoV are important factors in human immunoglobulin treatment (Miranda de Graaf et al., 2016). In mouse model, the ability of immunoglobulin to limit infection has been also reported followed by intraperitoneal administration of immunoglobulin (Chachu et al., 2008). Another strategy to clear the virus in immunocompromised patients is by temporarily discontinuing, reducing or changing immunosuppressive drugs (Green, 2014). Antivirals, like ribavirin, nitazoxanide and interferons are most likely to inhibit NoV replication in cell culture-based replicon systems, mouse models or infected human individuals (Chang & George, 2007; Jung et al., 2012). Oral treatment with nitazoxanide also reduced the duration of symptoms in a controlled clinical trial (Chang & George, 2007). With orally administered ribavirin two chronically infected immunocompromised individuals were successfully treated, a broad-spectrum antiviral agent, although same treatment was not successful in two other patients (Woodward et al., 2015). Oral administration of natural human interferon-α to pigs hass been shown to reduce virus shedding during the treatment period compared with levels and duration of shedding in untreated pigs, however virus shedding increased after the termination of interferon-α (Jung et al., 2012). Interferon-λ orally treated to mice cured persistent NoV infections (Nice et al., 2015). Currently no clinical data are available relating to the effect of interferon treatment on the replication of NoV in humans (Miranda de Graaf et al., 2016). The development of a NoV vaccine has been affected by the lack of a cell culture system and small animal models and licensed vaccines are not yet available (Debbink, K., Lindesmith, L. C. & Baric, 2014). However, the first NoV vaccines have now completed Phase I and Phase II clinical trials (Debbink, K., Lindesmith, L. C. & Baric, 2014) 21; these vaccines are based on GI.1 VLPs or in the case of the bivalent vaccine, containing both GI.1 derived VLPs and VLPs based on the consensus sequence of several GII.4 variants (Debbink, K., Lindesmith, L. C. & Baric, 2014). In a clinical trial with healthy volunteers, intramuscular vaccination with bivalent VLPs did not significantly reduce the infection after treated with a GII.4 strain of NoV; however, the vaccination reduced the severity of vomiting and diarrhea

(Bernstein et al., 2015). As an alternative strategy to vaccination with VLPs, intranasal vaccination with a vaccine derived from only the VP1 P domain, resulted in partial protection in gnotobiotic pigs that were infected with NoV of the GII.4 genotype (Kocher et al., 2014). Vaccine development is complicated because of genetic diversity and high rate of antigenic evolution of NoV (Miranda de Graaf et al., 2016). These complications can be reduced by updating the vaccine every 2–3 years, as in the case for influenza vaccines, unless a vaccine is able to induce a broadly reactive antibody response that remains effective against new antigenic variants (Miranda de Graaf et al., 2016). Recently, bivalent vaccine containing GI.1 VLPs and GII.4 VLPs produced broadly reactive antibody responses in humans in a clinical trial, providing heterologous cross-protection to non-vaccine and vaccine variants (L. C. Lindesmith et al., 2015). A trivalent vaccine in mice combined NoV GII.4 and GI.3 genotypes with the rotavirus VP6 genotype also induced a broadly reactive immune response that provided heterologous cross-protection (Debbink, K., Lindesmith, L. C. & Baric, 2014; Tamminen, Lappalainen, Huhti, Vesikari, & Blazevic, 2013). Although the development of NoV vaccines is currently focused on the most prevalent genotypes, the recent emergence of GII. P17-GII.17 NoV in Asia should act as a warning that in future risks from NoV outbreaks might arise (Miranda de Graaf et al., 2016). Therefore, the future efficacy of vaccines against NoV may rely on the development of vaccines that are able to produce a broadly reactive immune response or on an improved understanding of NoV evolution (Miranda de Graaf et al., 2016).

1.2 Next Generation Sequencing (NGS):

The massively parallel sequencing technology known as NGS has provided large improvements over sanger sequencing. There are still limitations, especially their shorter read lengths. It makes them unsuitable for some biological problems including assembly and identification of complex genomic regions, methylation detection and gene isoform detection (Rhoads & Au, 2015).

1.2.1 PacBio:

Single-molecule real-time (SMRT) sequencing system, developed by Pacific BioSciences (PacBio), offers a new method to overcome many of these limitations (Rhoads & Au, 2015). It is also known as ‘‘PacBio sequencing” hereinafter, though ‘‘SMRT sequencing” is also used by the public. Unlike NGS, PacBio sequencing is a used for real-time sequencing and no pause is required between read steps (Schadt, Turner, & Kasarskis, 2010). These properties differentiate PacBio sequencing from

SGS, so it is classified as the third-generation sequencing (TGS). PacBio sequencing produce much longer read lengths and faster runs than SGS methods but is affected by a lower through put, higher cost per base and higher error rate (Rhoads & Au, 2015).

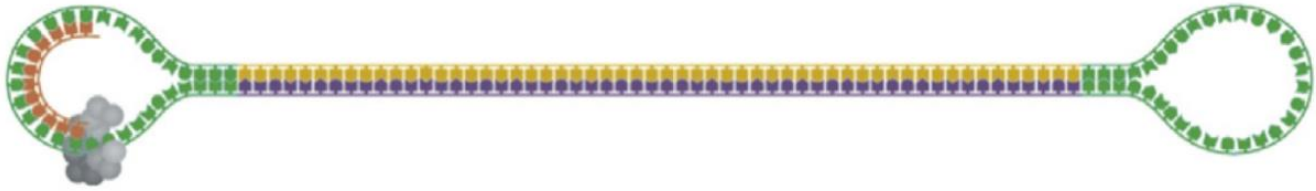

Figure 1.6: SMRTbell template. Hairpin adaptors (green) are ligated to the end of a double stranded DNA molecule (yellow and purple), forming a closed circle. The polymerase (gray) is anchored to the bottom of a ZMW and incorporates bases into the read strand (orange) (Travers, Chin, Rank, Eid, & Turner, 2010).

PacBio sequencing collects sequence information during the replication process of the target DNA molecule (Rhoads & Au, 2015). The template is known as SMRTbell, which is a closed single-stranded circular DNA that is generated by ligating hairpin adaptors to both ends of a target double-stranded DNA (dsDNA) molecule (Travers et al., 2010). When a sample of SMRTbell is added to a chip known as SMRT cell (Figure 1.6), a SMRTbell integrates into a sequencing unit called a zero-mode waveguide (ZMW), which give the smallest volume for light detection (Rhoads & Au, 2015). In each ZMW, a single polymerase is immobilized at the bottom, which binds to hairpin adaptor of the SMRTbell and start the replication (Eid et al., 2009). Four fluorescent labeled nucleotides, which produce distinct emission spectrums, are introduced to the SMRT cell. Whereas, base is held by the polymerase, a light pulse is generated that identifies the base (Eid et al., 2009). Each step of replication in all ZMWs of a SMRT cell are recorded by a ‘‘movie” of light pulses, and the pulses corresponding to each ZMW can be corelated to sequence of bases known as continuous long read (CLR)(Eid et al., 2009). The PacBio RS II platform normally produces sequencing movies of 0.5–4 h in length (Rhoads & Au, 2015). An important advantage of PacBio sequencing is the coverage of read length. While the PacBio RS sequencing with the first generation of chemistry (C1) generates mean read lengths around 1500 bp (Brown et al., 2014), the PacBio RS II system with the current C4 chemistry covers average read lengths over 10 kb (Brown et al., 2014), over half of all data reads are longer than 20 kb and highest coverage over 60 kb (Rhoads & Au, 2015). In contrast, the highest coverage of Illumina HiSeq 2500 is only paired-end 250 bp, using Rapid Run Mode system (Rhoads & Au, 2015). Because of long

read coverage of PacBio sequencing it gives advantage over Illumina sequencing, while sequencing organism with long genome like NoV.

1.3 Molecular Dynamics (MD):

A key point in the understanding of biology is to study macromolecule structure (Gelpi, Hospital, Goñi, & Orozco, 2015). Molecular interactions are basics for biological function (Tuckerman & Martyna, 2000). Biological functions are mostly affected by conformational changes in the protein. A protein molecule flexibility affects its interaction with ligand and its biological partners at different levels (Heberlé & de Azevedo, 2011). Molecular recognition rules as defined by such structural knowledge used for understanding of basic biological phenomena, transport across membranes, the building of large structures like ribosomes, like enzyme mechanisms and regulation, or how DNA read, or viral capsids and transcription is controlled. The study and prediction of protein-protein interaction networks is one of the growing fields in modern systems biology (Tuckerman & Martyna, 2000). Visual analysis of 3D structures of nucleic acid or protein obtained from the experiments, has brought enormous success in biochemistry (Gelpi et al., 2015). However, large number of structures stored at the PDB, only a partial view of 3D structure are provided for both protein and nucleic acids, and molecular dynamics can play important role in their functionality (Tuckerman & Martyna, 2000). Larger conformational changes are present in the known protein structures (Flores, 2006; Goh, Milburn, & Gerstein, 2004), and these conformational changes are a common part of an enzymes catalytic activity (Kokkinidis, Glykos, & Fadouloglou, 2012). For example, loop or domain closures may cause alteration of the active site from solvent, and change the chemical environment for substrates, or initiates the catalytic activity by bringing essential partners together (Shinoda et al., 2005; Stevens & Lipscomb, 1990). Moreover, by using molecular dynamic properties some protein characteristics can be understood. For example, transient appearance of channels in the protein structure requires diffusion of small substrates through heme-dependent enzyme molecules (Ahalawat & Murarka, 2015; Kalko, Gelpi, Fita, & Orozco, 2001). The source of the variability found may be cause of the lack of experimental data in some specific regions of protein structure, indeed, PDB study as basis for molecular flexibility has been exploited to some extent (Ahalawat & Murarka, 2015; Micheletti, 2013; Zakrzewska & Lavery, 2012).

1.3.1 History:

Molecular dynamics (MD) simulations was first developed in the late 1970s (McCammon, Gelin, & Karplus, 1977), by using simple approximations based on Newtonian physics to generate atomic motions, by reducing the computational complexity, crystallographic, a computational model of the molecular system is prepared from nuclear magnetic resonance (NMR), or homology-modeling data (Durrant & McCammon, 2011). The forces working on each of the system atoms are then estimated from an equation (Cornell et al., 1995). For instant, in order to produce the actual behavior of real molecules in motion, the energy terms explained above are parameters used to fit experimental data and quantum mechanical calculations (spectroscopy)(Cornell et al., 1995). These parameters include the identification of ideal stiffness and lengths of the springs that describe chemical bonding and atomic angles, identifying the proper van der Waals atomic radius, explaining the best partial atomic charges used for calculating electrostatic-interaction energies. Eventually, these parameters are called a 'force field' because they describe the contributions of the various atomic forces that govern molecular dynamics (Durrant & McCammon, 2011). Different force fields are commonly used in molecular dynamics simulations, including AMBER (Cornell et al., 1995) CHARMM (Brooks et al., 2009) and GROMOS (Christen et al., 2005). These principals may differ in the way but generally give same results. The positions of these atoms are moved according to Newton's laws of motion. The simulation time is highly advanced, often by only 1 or 2 quadrillionths of a second, and is repeated, typically millions of times. Because of much calculations, molecular dynamics simulations are typically performed on super computer or computer cluster using hundreds of processors in parallel (Durrant & McCammon, 2011).

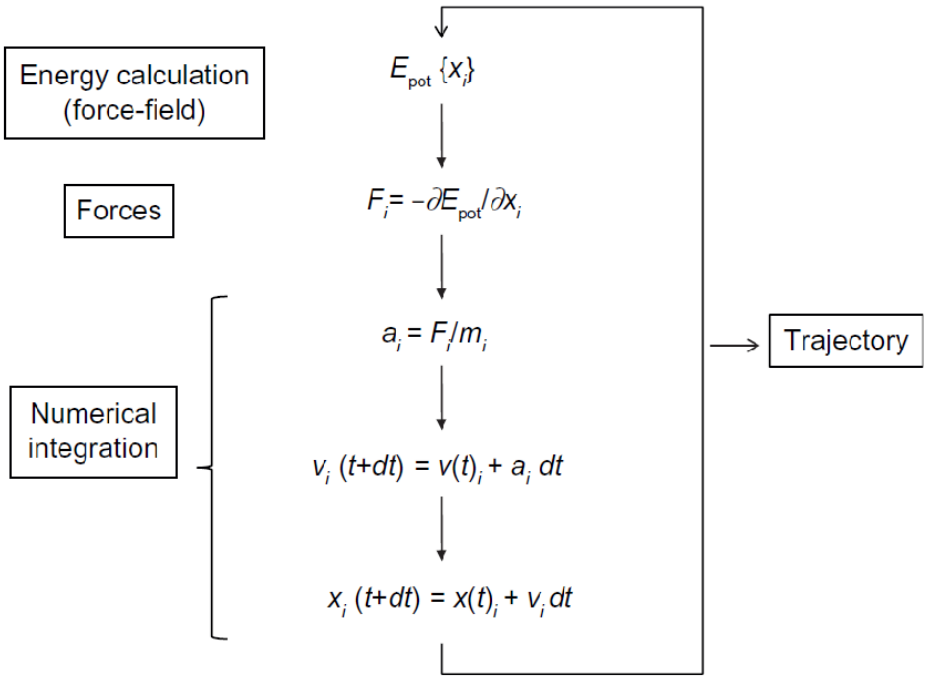

Figure 1.7: Molecular dynamics basic algorithm (Gelpi et al., 2015). E_{pot}, potential energy; t, simulation time; dt, iteration time; For each spatial coordinate of the N simulated atoms (i): x, atom coordinate; F, forces component; a, acceleration; m, atom mass; v, velocity

1.3.2 Different trajectories:

The root mean square deviation (RMSD) trajectory is used to predict the stability of the protein. A high RMSD value indicates a less stable protein structure (Sargsyan, Grauffel, & Lim, 2017). The RMSD is important in analyzing the equilibration of MD trajectories. The wild type or mutant structure is used as backbone atoms for MD simulations. The RMSD of the backbone atoms relative to the corresponding given structures for calculation (C. V. Kumar, Swetha, Anbarasu, & Ramaiah, 2014). It is used in molecular dynamics (MD) trajectories analyses, to generate the equilibration period and quality of biomolecular simulations and to cluster similar conformations (Sargsyan et al., 2017). RMSD is also used in protein structure predictions to identify the match between the modeled and NMR/X-ray protein structure as well as in docking calculations to spot the native structure among the numerous docked conformations (Kozakov, Clodfelter, Vajda, & Camacho, 2005; Mukherjee, Balius, & Rizzo, 2010; Shoichet & Kuntz, 1991). With the aim of finding whether the mutation affects the dynamic behavior of residues, RMSF which is one of MD trajectories is used as a mean describing flexibility differences among residues (A. Kumar, Rajendran, Sethumadhavan, & Purohit, 2013). Radius of Gyration (Rg) is performed in order to understand the compaction levels of the native and

mutant region in protein structure (C. V. Kumar et al., 2014). It is defined as the mass weighted root mean square distance of a collection of atoms from their common center of mass. Hence, this analysis gives us information about overall dimensions of the protein (A. Kumar et al., 2013). Hydrogen Bonds (H bonds) play a key role in the stability and molecular recognition of the protein structure (C. V. Kumar et al., 2014).

Aims of the thesis:

Due to immune suppressive therapies, an increase in chronically NoV infected patients has been observed in recent years. Besides being a clinical challenge, these chronic infections give the opportunity to study NoV evolution within an individual. I my thesis I therefore wanted to analyze how NoVs evolve on an individual level during chronic infection.

Distinct aims of my thesis are:

1. To analyze how mutations seen in pandemic NoV strains compared to mutations found in chronically infected patients. Therefore, I wanted to analyze consensus sequences reported from the years 1978 to 2014 and compare them to mutations that evolve in an infected individual over time.
2. To determine functional consequences of NoV variation, I aimed at establishing an ELISA with recombinant NoV capsid protein expressed from the NoV sequences identified in acutely and chronically infected patients. This should be used to quantify antibody responses to variant NoV capsid protein.
3. To study amino acid variation of the capsid protein, I aimed to do NGS analysis of sequential samples from chronically infected patients.
4. Finally, an aim of this thesis was to identify structural changes on NoV surface during chronic infections with computer-based molecular dynamics analyses.

2 RESULTS

2.1 Year wise alignment analysis of NoV G II.4 VP1 capsid protein:

A total of 672 NoV G II.4 VP1 capsid protein sequences from 1978 till 2014 were downloaded from the GenBank and aligned using online software clustal omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). After alignment, they were year wise distributed in groups to study different variable sites in the VP1 capsid protein, as shown in Figure 2.1. Variable sites were found to be more located in the P2 domain. Higher number of mutations were detected in the P2 domain from the years 2004, 2006 and 2010.

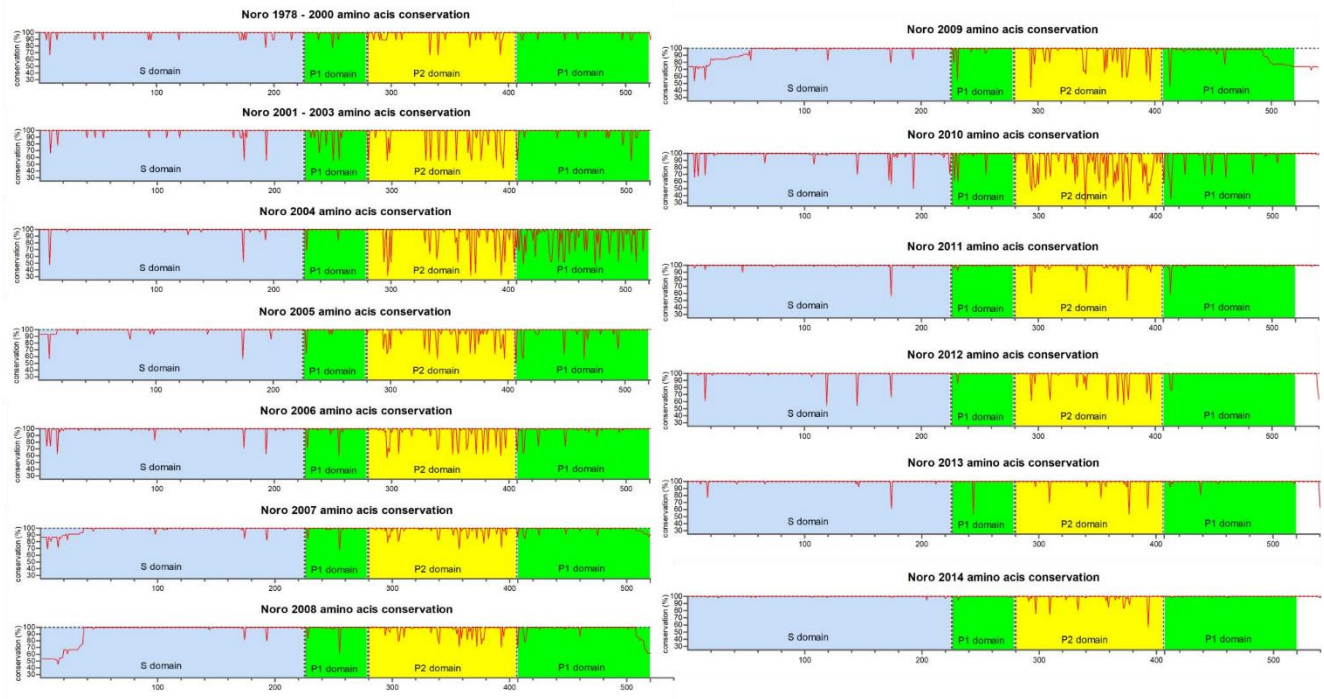

Figure 2.1: Proportion of conserved and mutated amino acids in the VP1 capsid protein from 1978, 1995, 1997, 2000- 2014. S domain (amino acids 1-225), P1 domain (amino acids 225-278 and 406-519) and P2 domain (amino acids 279-405). A high variation in the P2 domain was observed in 2004, 2006 and 2010.

2.1.1 Molecular evolution of GII.4 capsid protein:

Year wise consensus sequences were generated and aligned. Consistent with previous studies (de Rougemont et al., 2011; Parra et al., 2012), as shown in Figure 2.2, several mutations were observed in the previously described epitopes A, D and E; HBGA binding sites were conserved through the years. In the following sections, we will compare individual patient-derived sequences to this year wise analysis.

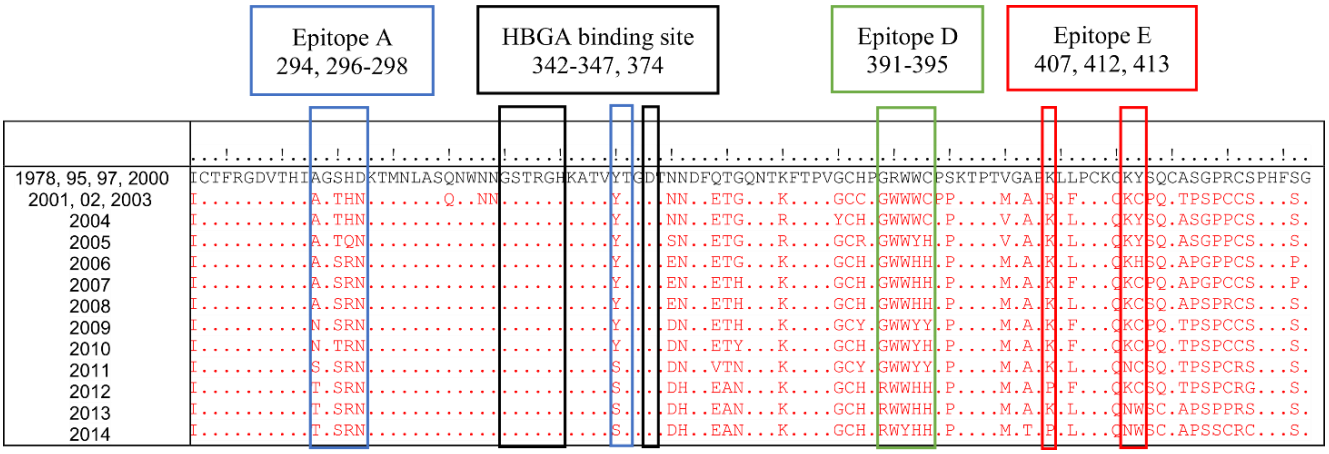

Figure 2.2: Proportion of conserved and mutated amino acids in VP1 capsid protein from 1978, 1995, 1997, 2000-2014. Numerous mutations occurs in Epitope A (blue), D (green) and E (red) in amino acid of consensus sequences of VP1 capsid protein. Histo-Blood Group Antigen binding sites (black) were conserved on consensus sequences from 1978 till 2014.

2.2 Analysis of samples drawn from patients with acute and chronic infection:

2.2.1 Expression of GII.4 P domain in *E. coli* BL21:

GII.4 P domain gene fragment was amplified using a pair of specific primers (Table 4.8). The 1007 bp PCR product (Figure 2.3A) was digested with *BamH*I and *Not*I and inserted into the expression vector pGEX-4t (Figure 2.3B). After ligation, the product was transformed into 10-beta *E. coli* competent cells. Positive clones were checked for gene insertion by RE analysis (Figure 2.3C); and the integrity of the insert was confirmed by sequencing. The recombinant plasmids were then purified and introduced into *E. coli* BL21 cells (Figure 2.3D) for protein expression. The expressed P domain was purified on the GSTrap FF column and visualized on SDS-PAGE gel by Coomassie blue staining, as shown in Figure 2.3E.

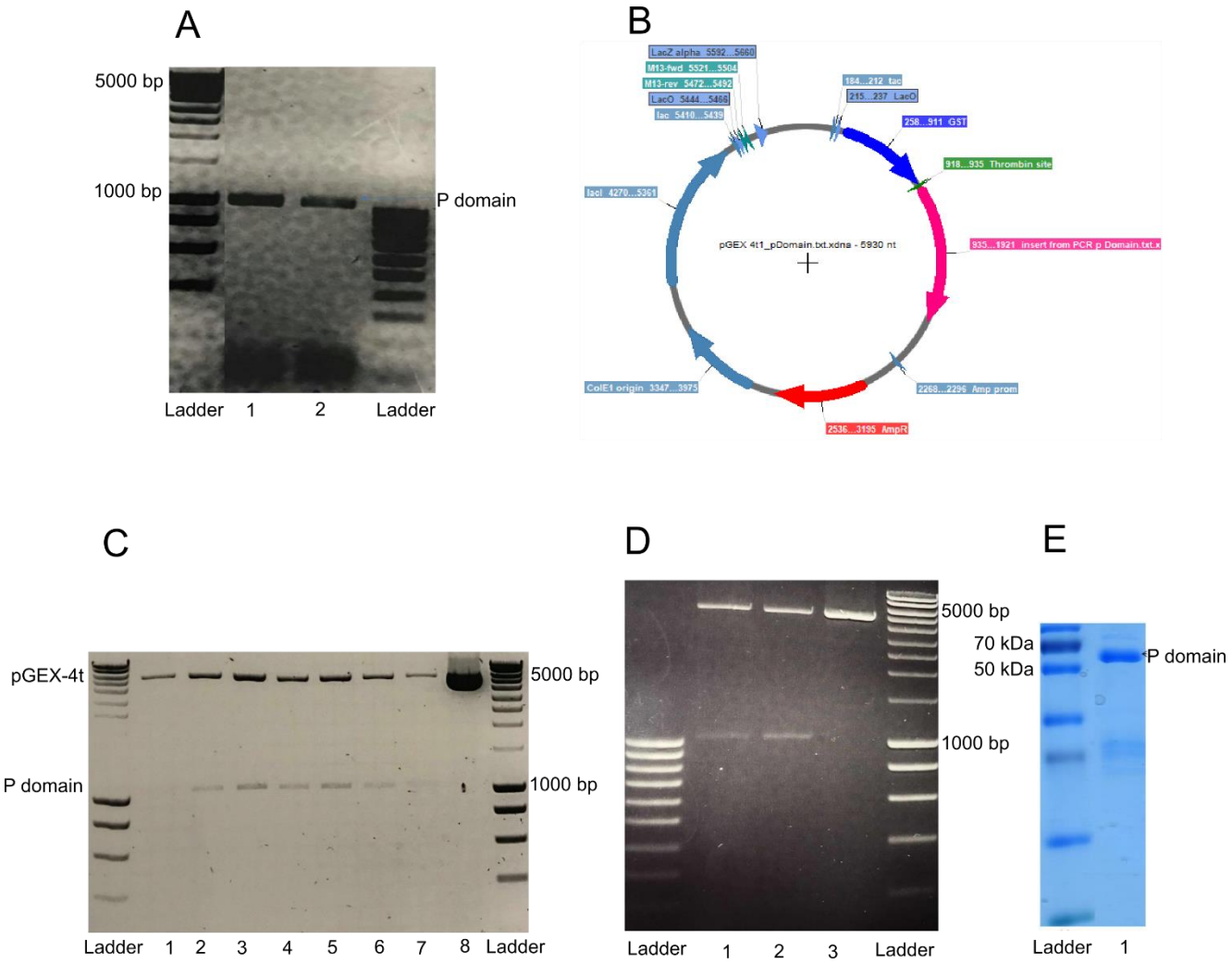

Figure 2.3: Expression of P domain capsid protein from GII.4 acutely infected patients. (A) P domain was amplified from NoV cDNA and then visualized on agarose gel (Lines 1 and 2; size: 1007 bp). (B) Schematic map of pGEX-4t-P plasmid encoding VP1 P domain: The P domain (Pink) was inserted into the *BamH*I and *Not*I restriction sites of the vector. (C) Gene insertion was confirmed by RE digestion and the results were checked on agarose gel: Recombinant plasmids (Lines 1-7); Control plasmid (Line 8). (D) Recombinant pGEX-4t-P plasmid was transformed into *E. coli* strain BL21: Samples (Lines 1 and 2); parental plasmid control (Line 3). (E) SDS PAGE of P domain expressed in *E. coli* strain BL21: The protein fused to GST tag (58 kDa) was detected on the gel (P domain 32kDa+ GST tag 28kDa).

2.2.1.1 Mouse immunization with the GII.4 P domain:

Six-week old female BALB/c mice were immunized intraperitoneally three times with 30 µg GII.4 P domain protein and 10 µg C-di-AMP (as adjuvant) with one-week interval as shown in Figure 2.4. Sera were collected one week before the first immunization and seven weeks after the last immunization, and tested for NoV-specific antibody level (Figure 2.5). We determined intra (nine replicates) and inter (six replicates) assay variation, as shown in Table 2.1.

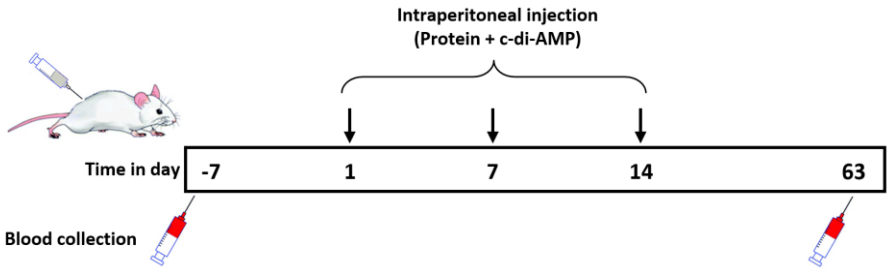

Figure 2.4: Mouse immunization. Female mice were injected three times with GII.4 P domain and C-di-AMP with one-week interval. Serum samples drawn one week before and seven weeks after immunization were tested for antibody level.

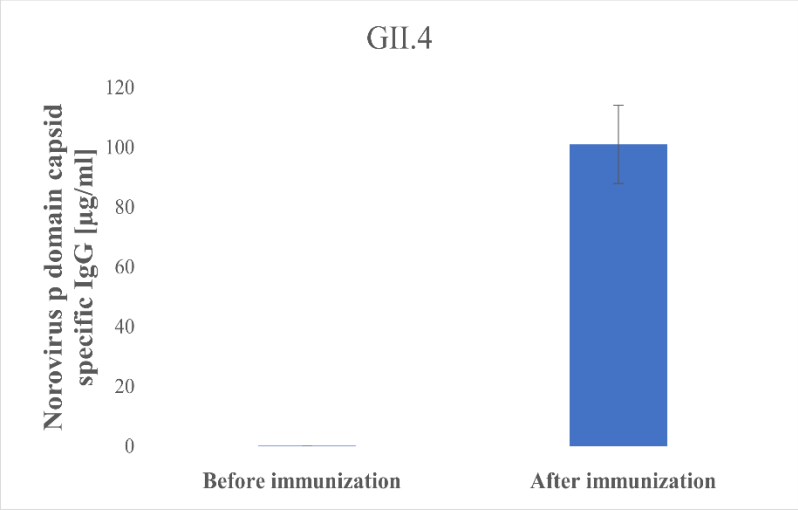

Figure 2.5: Quantitative ELISA for measurement of NoV-specific IgG level in immunized mice. ELISA plates were coated with recombinantly expressed GII.4 P domain. Serum samples from BALB/c mice injected with capsid protein of GII.4 were used to characterize ELISA. Virtually undetectable background signal (0.01 µg/ml) was detected before immunization; high signal (104 µg/ml) were measured after immunization.

2.2.2 Expression of GII.1 P domain in *E. coli* BL21:

RNA was extracted from stool samples of patients with acute infection and NoV cDNA was synthesized. The P domain gene fragment was amplified from the cDNA using specific primers and the amplified product with the size of 1036 bp was checked on agarose gel (Figure 2.6A). The PCR product was digested with *BamH*I and *Not*I and then inserted into the expression vector pGEX-4t (Figure 2.6B). After confirmation of insertion with RE digestion (Figure 2.6C) and sequencing, the recombinant plasmid was transformed into *E. coli* BL21 for protein expression (Figure 2.6D). The expressed protein was purified on the GSTrap FF column and then visualized on SDS-PAGE gel by Coomassie blue staining, as shown in Figure 2.6E.

2.2.2.1 Mouse immunization with GII.1 P domain:

Six-week-old female BALB/c mice were immunized three times with 30 µg GII.1 P domain and 10 µg C-di-AMP (as adjuvant) intraperitoneally with one-week interval (Figure 2.7). Sera were collected one week before the first immunization and seven weeks after the last immunization and used for inter (six replicates) and intra assay (nine replicates) variability (Figure 2.8).

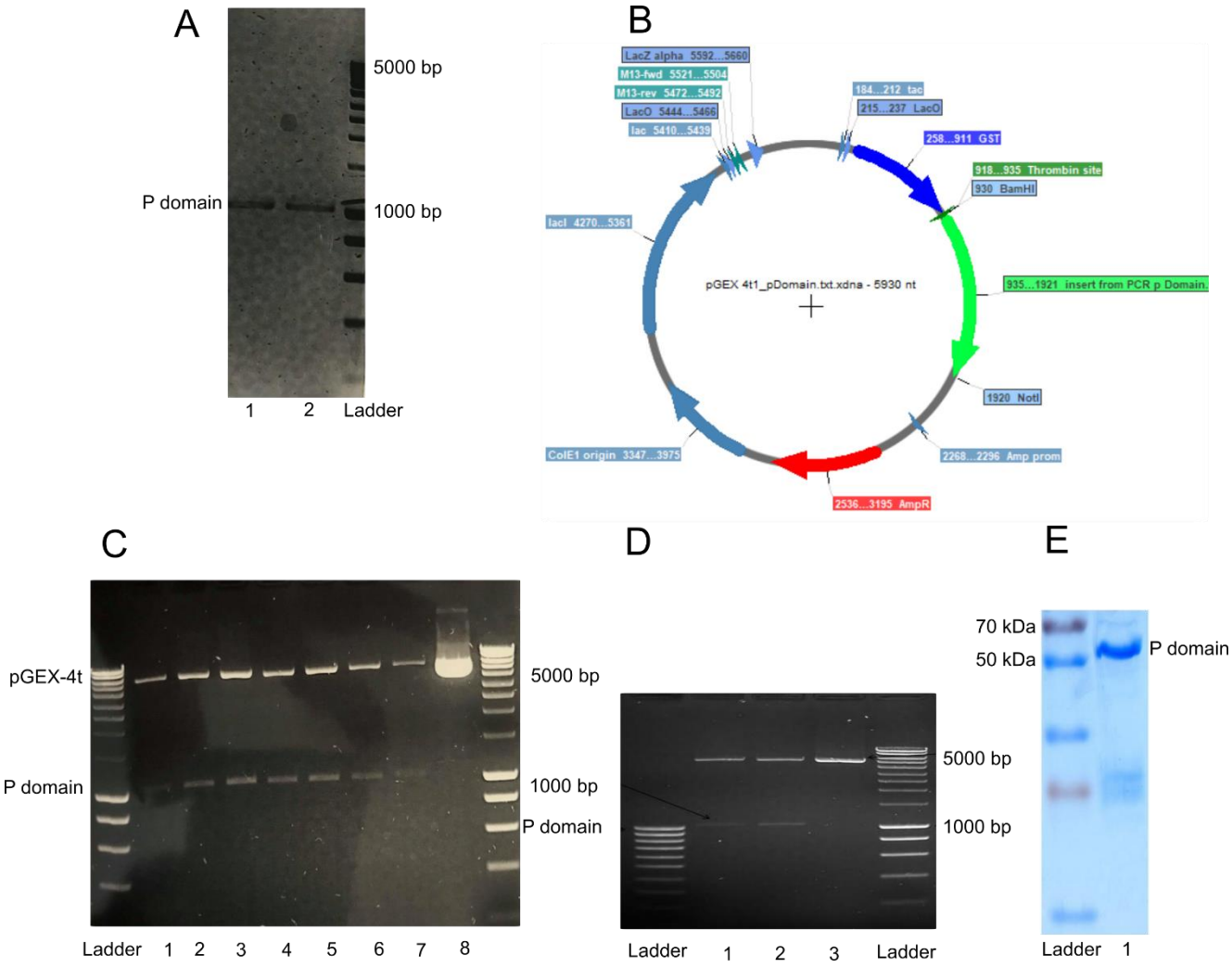

Figure 2.6: Expression of P domain capsid protein in GII.1 acutely infected patient. (A) Amplified GII.1 P domain with size of 1038 bp on agarose gel (Lines 1 and 2). (B) Schematic map of pGEX-4t-P plasmid encoding P domain: The P domain (Green) was inserted into the *BamH*I and *Not*I restriction sites of the vector. (C) Gene insertion was confirmed by RE digestion and the results were checked on agarose gel: Recombinant plasmids (Lines 1-7); Control plasmid (Line 8). (D) pGEX-4t-P recombinant plasmid in *E. coli* strain BL21: Samples (Lines 1 and 2); parental plasmid (Line 3). (E) SDS PAGE of P domain expressed in *E. coli* strain BL21: The protein fused to GST tag (58 kDa) was detected on the gel (P domain 32kDa+ GST tag 28kDa).

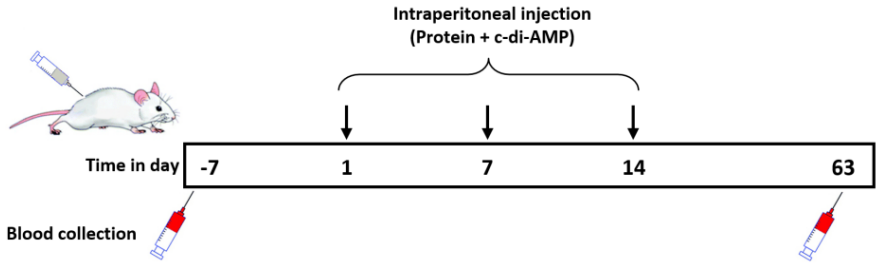

Figure 2.7: Immunization of mice. Female mice were injected three times with GII.1 P domain and C-di-AMP with one-week interval. Serum samples drawn one week before and seven weeks after immunization were tested.

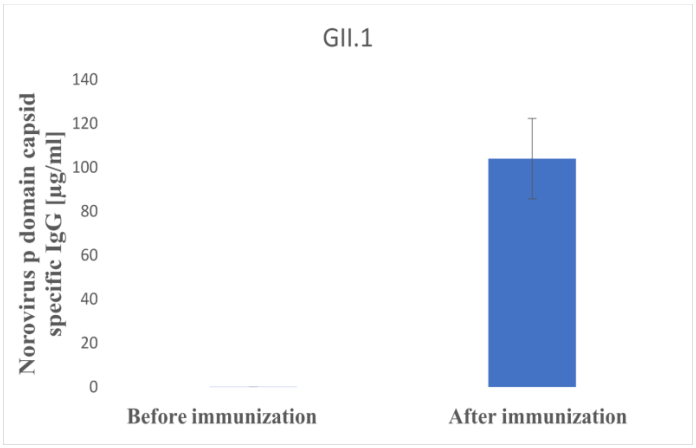

Figure 2.8: Quantitative ELISA for NoV-specific IgG. ELISA plates were coated with recombinantly expressed GII.1 P domain protein. Serum samples from BALB/c mice injected with GII.1 capsid protein were used to characterize ELISA. Virtually undetectable (0.01 µg/ml) background signal was detected before immunization; high signal (101 µg/ml) was detected after immunization

2.2.3 ELISA validation:

To validate the assay, coefficient of variation (CV) and Z value were measured for both immunized and unimmunized samples (Table 2.1). CV ≤ 20% and Z' ≥ 0.4 are considered as acceptance criterion for the assay (Iversen et al., 2004). Coefficients of variation (CV) was found to be 8.5 % for negative and 9.6 % for positive control of the GII.g/GII.1; it was 9.8 % and 8.6 % for the GII.4. Z' factor was 0.96 for the GII.g/GII.1 and 0.97 for the GII.4 (Table 2.1).

Table 2.1: Means, standard deviation, coefficient of variation and Z values calculated for GII.1 and GII.4 positive and negative controls.

	GII.1 NEG.	GII.1 POS.	GII.4 NEG.	GII.4 POS.
Mean	0.068613	1.151544	0.096709	1.1159
SD	0.005819	0.110647	0.009487	0.096839
CV	8.481466	9.608611	9.809505	8.678134
Z	0.967733		0.971431	

2.2.4 Norovirus-specific IgG level in infected patients:

The established quantitative ELISA was applied to measure serum level of NoV-specific IgG before and after infection with GII.g/GII.1 and GII.4 genotypes (Figure 2.9). Due to pre-existing NoV-specific serum antibodies, GII.10 polyclonal antibody was used as positive control for the assay. ELISA measured capsid-specific IgG in both GII.g/GII.1- (1–7) and GII.4- (8–10) infected patients. Antibody concentration varied between patients but was not significantly different between GII.1 and GII.4. Antibodies elicited by GII.4 infection were also reliably detected using GII.g/GII.1 p-domain. No increase in antibody level was observed in the first week post-infection (patient 1); a significant increase was detected in the second week (patients 3, 4 and 8) and then gradually decreased (patients 5, 6, 7, 9 and 10).

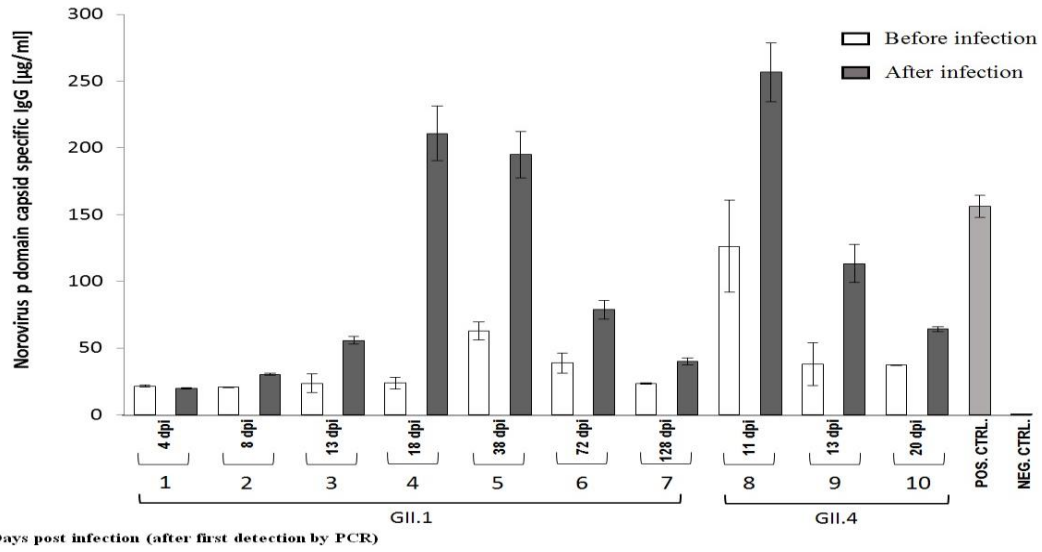

Figure 2.9: Quantitative ELISA for NoV acutely infected patients. NoV-specific IgG levels (µg/ml) were determined in 10 acutely infected patients before and after infection. GII.10 polyclonal antibodies was used as positive control. Neg. control: goat serum instead of patient serum. The error bars indicate the standard errors of the means.

2.2.5 Norovirus-specific IgG levels in sequential samples from chronically infected patients:

2.2.5.1 P domain expression:

RNA was extracted from stool samples of chronically infected patients, cDNA was synthesized and the P domain gene (1003 bp) was amplified from the cDNA using specific primers (Figure 2.10A). The PCR product was digested with *BamH*I and *Not*I and then was inserted into the expression vector pGEX-4t (Figure 2.10B). The recombinants from four patients were transformed into *E. coli* beta 10 cells followed by RE digestion and analysis on agarose gel (Figure 2.10C). The recombinant expression vectors were purified and transformed into *E. coli* BL21 cells (Figure 2.10D) for expression. The expressed proteins were purified on the GSTrap FF column; Protein concentration was measured by Bradford assay and then were confirmed on SDS-PAGE and by Western blotting (Figure 2.10E, F and G).

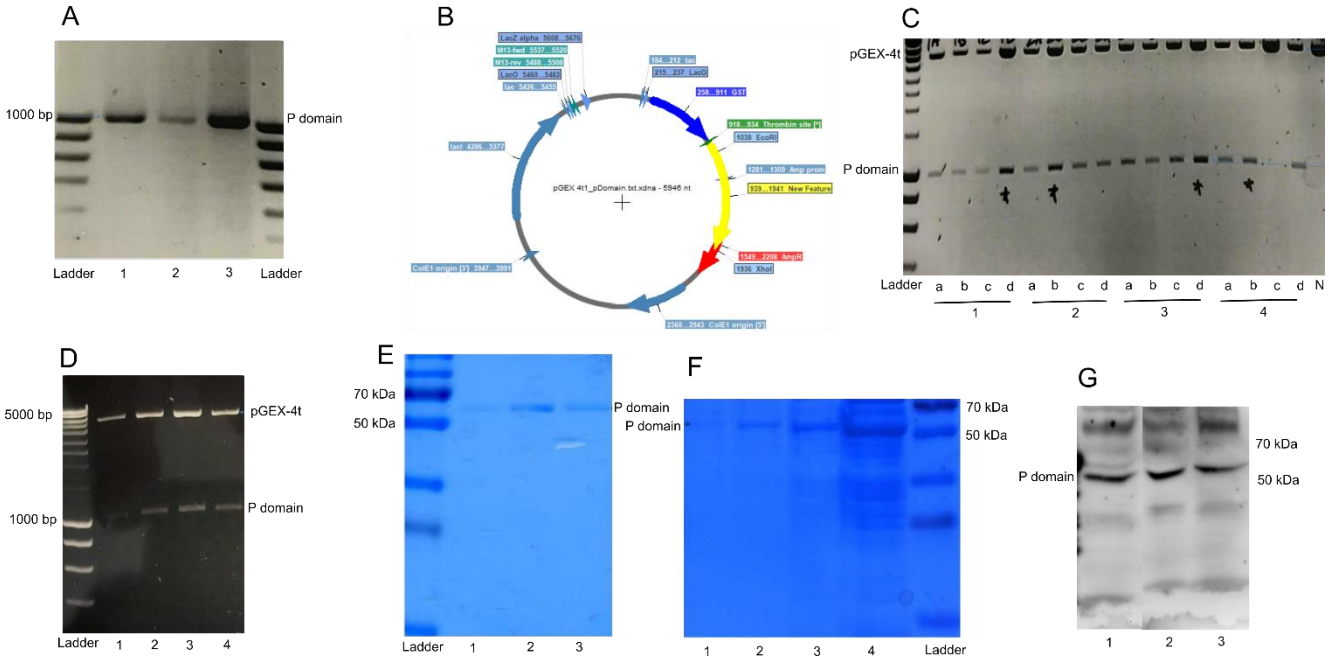

Figure 2.10: Expression of P domain capsid protein in GII.4 chronically infected patients. (A) P domain with the size of 1003 bp was visualized on agarose gel (Lines 1, 2 and 3). (B) Schematic map of pGEX-4t-P plasmid encoding VP1 P domain: The P domain (Yellow) was inserted into the BamHI and NotI restriction sites of the vector. (C) Gene insertion was confirmed by RE digestion and the results were checked on agarose gel: Recombinant plasmids (Lines 1-16); Control plasmid (Line 17). (D) pGEX-4t-P recombinant plasmid in *E. coli* strain BL21: Samples (Lines 1-4) (E) SDS PAGE of P domain expressed in *E. coli* strain BL21: The protein fused to GST tag was detected on the gel (58 kDa: P domain 32kDa+ GST tag 28kDa). (F) P domain + GST tag (1-4: 58kDa). (G) Western blotting result.

2.2.5.2 Expression of VP1 capsid gene:

Viral RNA was extracted from stool samples of chronically infected patients. cDNA was synthesized and be used to amplify the capsid gene using specific primers. Amplified fragment (1623 bp) was then checked on agarose gel (Figure 2.11A). The PCR product was digested with *EcoR*I an *Xho*I and then inserted into the pGEX-4t vector, as shown in Figure 2.11B. The recombinants from 4 patients were transformed into *E. coli* beta 10 cells followed by RE digestion and analysis on agarose gel (Figure 2.11C). Recombinant plasmids were purified and transformed into *E. coli* BL21 cells (Figure 2.11D) for expression. The expressed proteins were purified using the GSTrap FF column; Protein concentration was measured by Bradford and then were confirmed on SDS-PAGE and by Western blotting (Figure 2.11E).

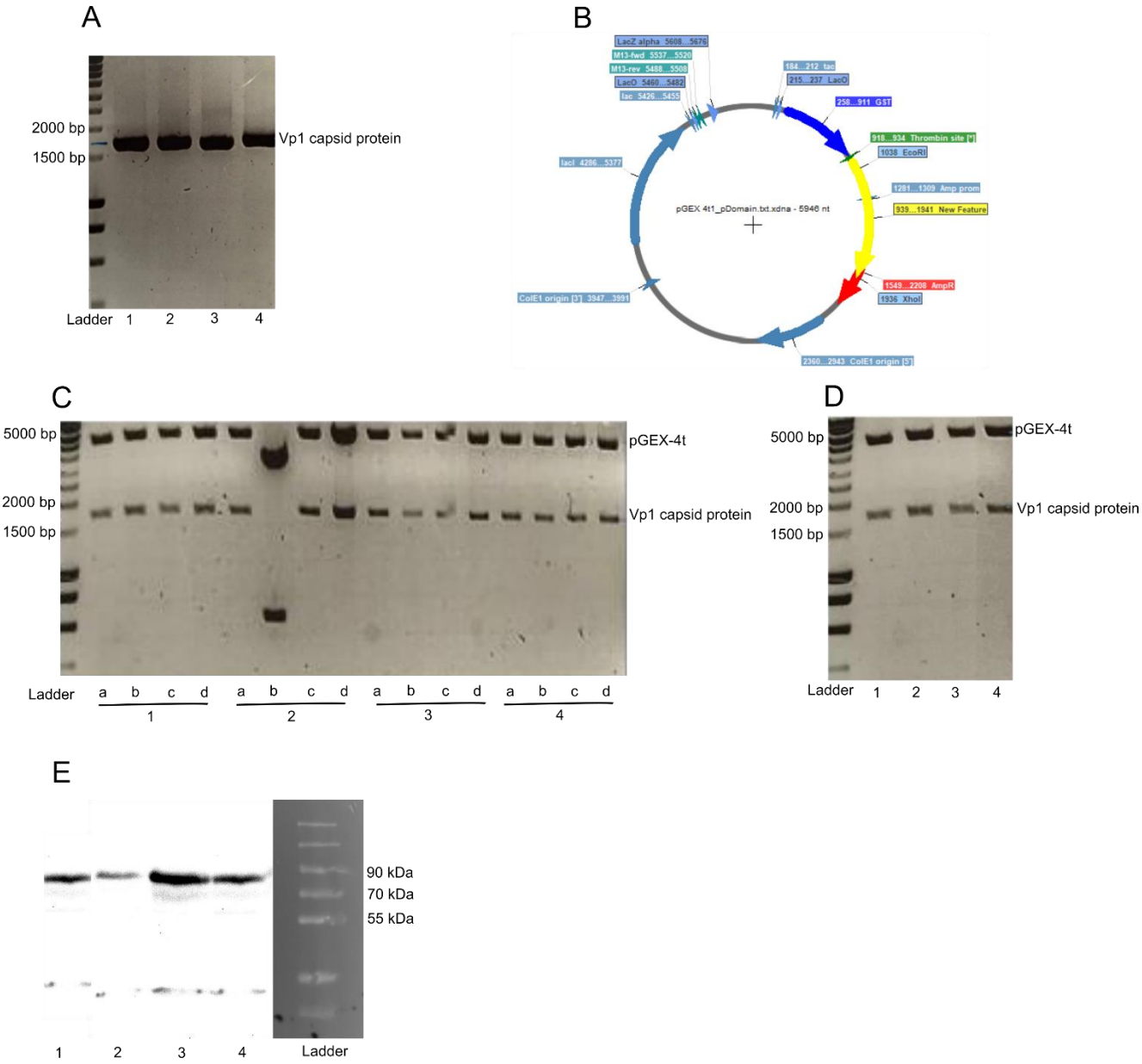

Figure 2.11: Expression of VP1 capsid protein in GII.4 chronically infected patients. (A) VP1 capsid protein of NoV with size 1674 bp visualized on agarose gel (Lines 1-4). (B) Schematic map of pGEX-4t-P plasmid encoding VP1 P domain: The P domain (Yellow) was inserted into the *EcoR*I an *Xho*I restriction sites of the vector. (C) Gene insertion was confirmed by RE digestion and the results were checked on agarose gel: Recombinant plasmids (Lines 1-16) (D) pGEX-4t-P recombinant plasmid in *E. coli* strain BL21: Samples (Lines 1-4) (E) Western blotting result (85 kDa: VP1+GST).

2.3 Quantitative ELISA for sequential samples of chronically-infected patients:

Quantitative ELISA was developed by coating the wells with the recombinantly expressed GII.4 P domain. The assay was used to quantify NoV specific IgG level in the serum samples from chronically infected patients. To check if the antigen is distributed in the wells equally, different dilutions (1:5, 1:10, 1:25) of the positive serum GII.10 were loaded into the coated wells. Antigen concentration was shown to be equal in the quadruplicate wells (Figure 2.12).

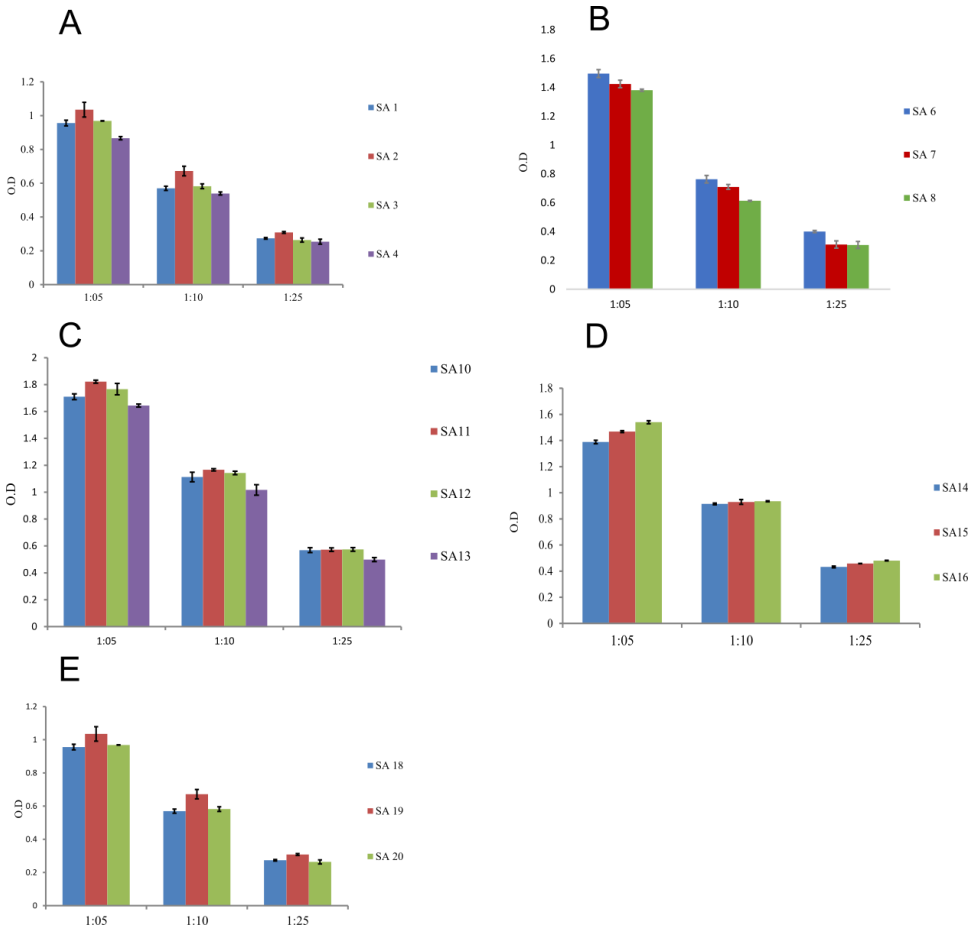

Figure 2.12: Reproducibility of ELISA plate coating for 3 capsid protein dilution. Each dilution was applied in quadruplicates. Concentration-dependent signals of anti-norovirus polyclonal antibodies. Serum dilutions 1:5, 1:10 and 1:25 were applied in duplicates. All the dilutions yielded similar ODs, indicating homogenous plate coating with recombinantly expressed VP1 P domain.

Patient 1, a bone marrow transplant: From sequential stool samples collected at days 0, 66, 121 and 273, capsid gene was amplified and expressed in *E. coli* to be coated, as antigen, into ELISA wells. Serum samples, drawn at days 133-, 148-, 256- and 430 were tested for NoV-specific antibody titer using the ELISA. First and second samples yielded low specific antibody concentration (14-34 µg/ml

for the first and 44-73 µg/ml for the second) with all the sequential antigens (Figure 2.13). In the third serum drawn at day 256, antibody concentration increased markedly to 324-, 285-, 298- and 318 µg/ml. and stayed virtually the same in the fourth serum (328-, 279-, 315- and 323 µg/ml). Apparently, the patient mounted a robust antibody response between the second and third serums.

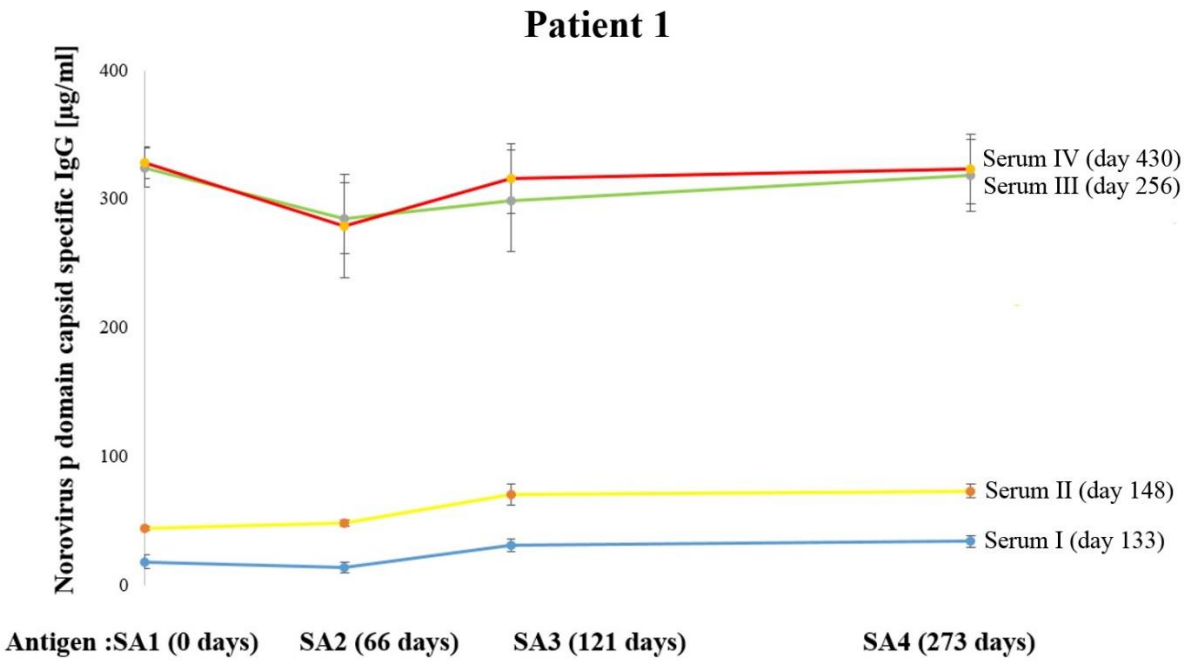

Figure 2.13: ELISA results for chronically infected patient 1. Capsid protein from days 0, 66, 121 and 273 was used as antigen for ELISA assay. Serum samples drawn at days 133, 148, 256 and 430 were tested for NoV-specific IgG level. First and second samples showed low specific antibody level with all the antigens. In contrast, high level of specific antibody was measured in the third and the fourth sera.

Patient 2, a bone marrow transplant: ELISA wells were coated with expressed capsid protein from days 0, 126 and 207. First serum sample taken at day 0 showed low specific antibody concentration (about 42 µg/ml) with all the sequential antigens. In the second serum sample, taken at day 74, antibody level was found to be gradually increased and then decreased (110-, 138- and 122 µg/ml). Whereas, in the third serum, taken at day 134, low antibody level (18-, 19- and 19 µg/ml) was detected for all the coated antigens. For the forth serum, high antibody level (233-, 253- and 237.7 µg/ml) was measured with the sequential antigens.

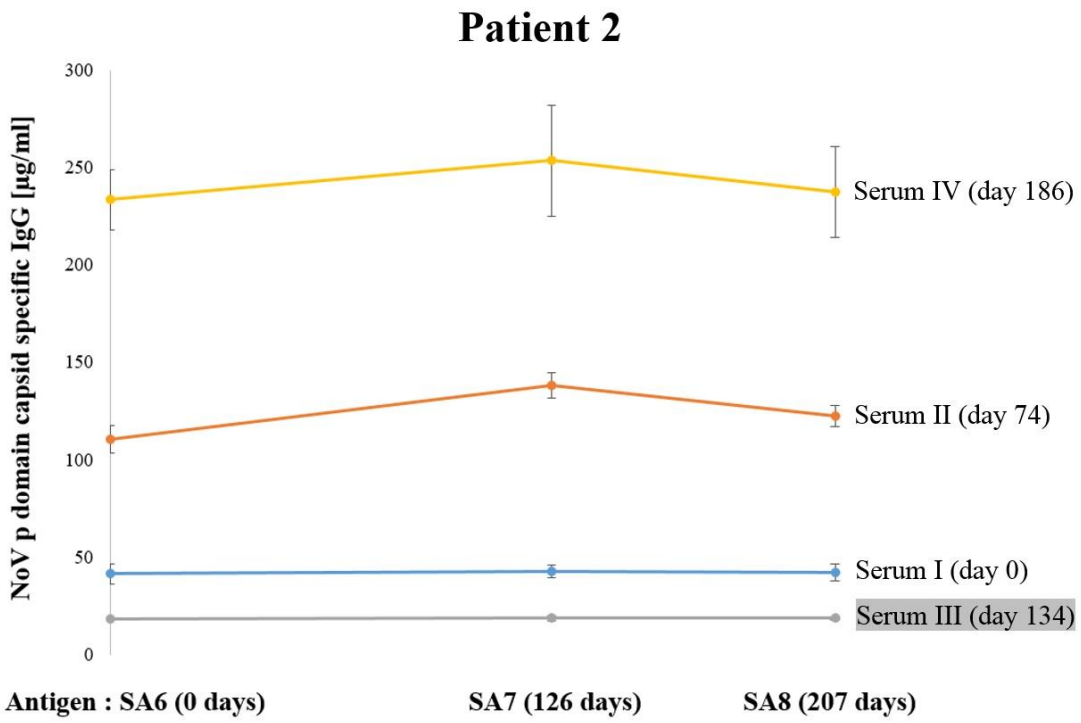

Figure 2.14: ELISA results for chronically infected patient 2. Serum samples were tested for NoV-specific IgG at day 0, 74, 134 and 186 post infection. Low antibody level was measured for the first serum sample. For the second serum, higher level of Ab was measured. Third serum showed very low concentration of specific Ab. High level of specific Ab level was detected in the fourth serum sample with all the coated antigens.

Patient 3, a bone marrow transplant: ELISA wells were coated with capsid antigens from days 0, 127, 207 and 290. Serum samples taken at days 0, 94 and 275 were then tested for NoV-specific Ab level. As shown in Figure 2.15, the first serum sample showed low antibody levels (10-, 9-, 10- and 5 µg/ml) for all the coated antigens. Interestingly, a high antibody concentration were measured in the second serum and the first antigen and then gradually decreased with the following sequential antigens (53, 26, 14 and 5 µg/ml). In contrast, high NoV-specific IgG concentration was measured in the 3rd serum with all the sequential antigens.

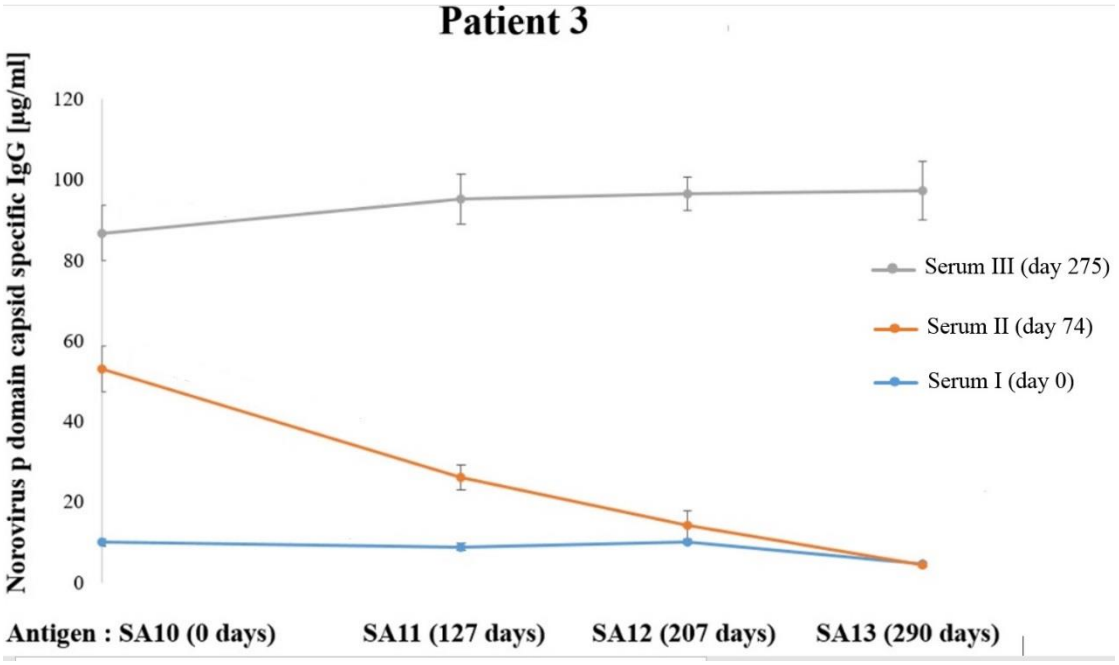

Figure 2.15: ELISA results for chronically infected patient 3. Serum samples, taken at days 0, 94 and 275, were tested for NoV-specific Ab level. First serum sample showed low level of specific IgG with coated antigens from day 0, 127, 207 and 290. For the second sample, antibody titer showed to be high in the beginning; however, it gradually reduced in interaction with the coated antigens. High level of NoV-specific IgG was measured in third serum using all the coated antigens.

Patient 4, a bone marrow transplant: Capsid antigen from days 0, 29 and 114 were coated into the ELISA wells. Serum samples taken at days 3, 95 and 120 were tested for NoV-specific Ab level. The results showed low antibody level in the first (19, 18 and 25 µg/ml) and second serum samples (19-, 20- and 26.2 µg/ml) for all the coated antigens. Higher Ab concentrations were detected in the third serum sample (Figure 2.16).

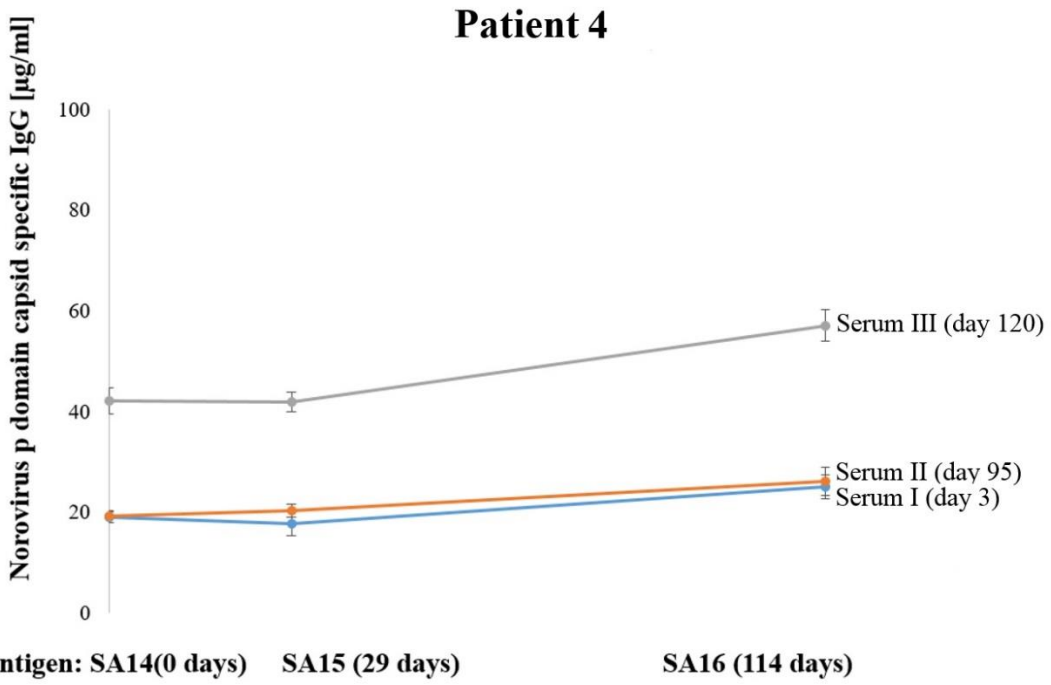

Figure 2.16: ELISA results for chronically infected patient 4. ELISA wells were coated with capsid antigen from days 0, 29, 114. Low Ab titer was detected for serum sample 1 and 2 taken at days 3 and 95, respectively. Higher Ab titers were detected of the third sample for the three coated antigens.

Patient 5, a bone marrow transplant: Antibody level was measured for serum samples taken at days 0, 22, 216 using ELISA plates coated with capsid antigens from days 0, 68 and 146. Low antibody concentrations (25, 26 and 25 µg/ml) were detected in the first serum. Compared to the first sample, an increase in Ab titer (40, 40 and 42 µg/ml) was observed in the second serum. The third serum sample showed even higher antibody level with the sequential antigens (62, 68 and 72 µg/ml).

.

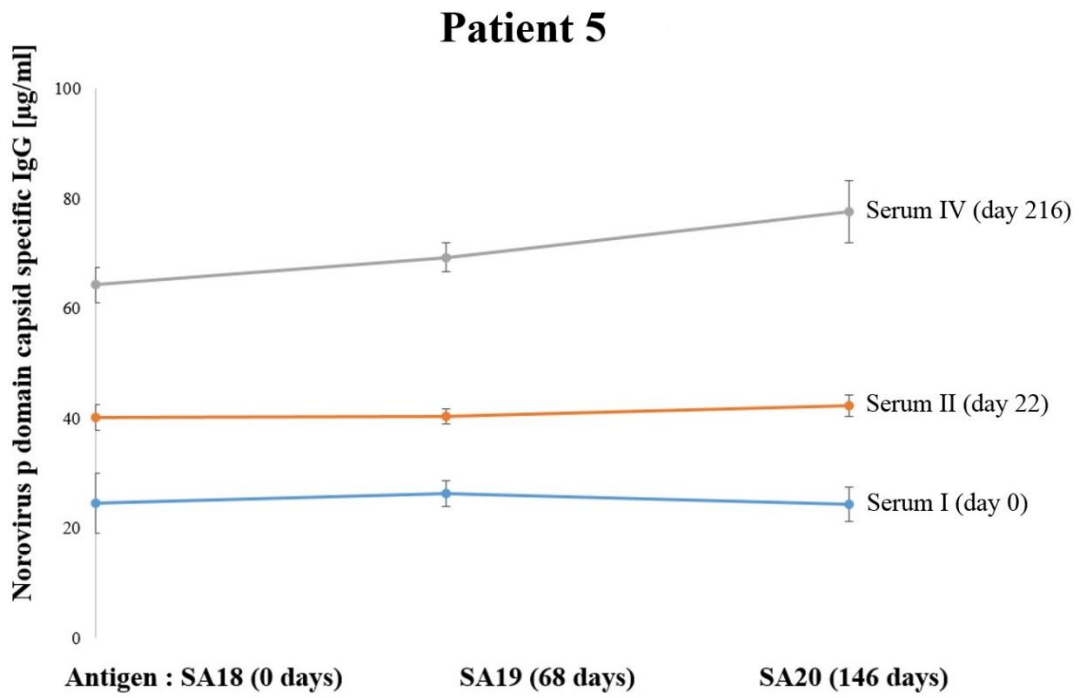

Figure 2.17: ELISA results for chronically infected patient 5. ELISA was carried out using coated capsid antigen for different time point. Serum samples were tested for NoV-specific Ab level at day 0, 22 and 216. Low Ab level was detected for the first sample; second sample showed significantly higher Ab titer when compared to the first sample; highest antibody titer was detected in the third sample.

2.4 Sanger sequencing analysis:

After ELISA, gene sequences of the P domain (in patients 1 and 2) and whole VP1 (in patients 3, 4 and 5) were checked and amino acid content were compared between the sequential samples from the chronically infected patients. Previously described epitopes A, D, E, $CD4^+$ and $CD8^+$/B cell were mapped in the proteins; mutated and conserved regions were studied, as shown in Figure 2.18. Analysis showed several mutations in the sequential sequences from the patients. However, HBGA binding sites were found to remain conserved (Figure 2.18). In patient 3, in contrast to other patients, fixed mutations at epitope A (two mutations) and E (two mutations) were detected (Figure 2.18).

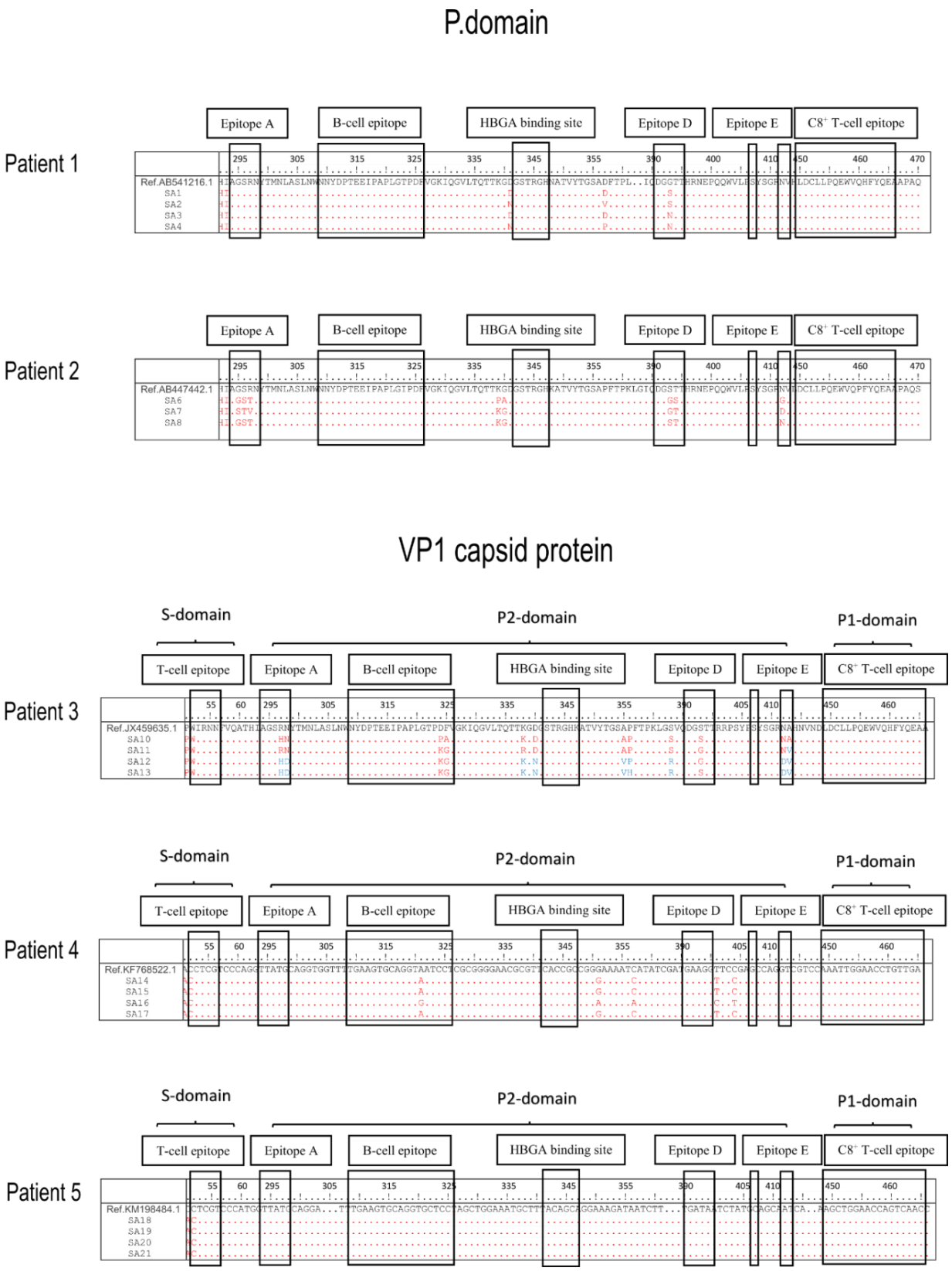

Figure 2.18: Sanger sequence analysis of capsid protein in chronically infected patients. In contrast to HBGA and CD8$^+$ T-cell epitope, some mutations were detected in the epitope A, D and E.

2.5 Next generation sequencing (NGS):

A total of 29 stool samples were collected at different time points ranging between 8 to 120 days from 10 NoV-infected patients. Out of seven patients, six were immune compromised due to bone marrow transplantation and one was HIV co-infection. The remaining three patients had acute NoV infection. Samples from the immune compromised patients were selected for a sequential analysis of the intra-host NoV population (Table 2.2).

Of these immunocompromised patients, five were infected with GII.4, while the remaining two subjects were infected with GII.3 and GII.6. The closest sequences related to these subjects were identified by executing a BLAST run against the GenBank database. The GenBank sequences GII.4/Hong (GII-4/Hong Kong/CU060025/2006/CHN; HM802539), GII.4/Miyagi (GII-4/Miyagi5/2006/JP), GII.4/Taiwan (GII-4/CGMH09/2006/TW), GII.4/Gothenburg (GII-4/P63/2012/Gothenburg/Sweden), GII.3/Vietnam (GII-30212/2009/VM), GII.6/China (China/2016/KU870455) and GII.4/Osaka (OsakaSB2-1/2014/JPA/LC133344), with highest scoring for the subjects, were considered as reference sequences for the patients 1-7, respectively.

2.5.1 Mapping of NoV sequencing reads:

The detailed information about the NGS data from the sequential samples of the chronically infected patients are described in Table 2.3. The average read length of the VP1 was 1680 bp. The average coverage ranged between 790X and 14,740X (Table 2.3).

2.5.2 Quasispecies distribution:

To understand the dynamics of NoV quasispecies over time, quasispecies with frequency > 1% were reconstructed. In the longitudinal study of the patients, the number of quasispecies was found to vary between time points. Of seven chronically infected patients, 4 patients (i.e. 1, 2, 5 and 7) showed more quasispecies at later time points than the first time point, whereas a gradual decrease in the number of quasispecies towards the end of the follow-up was observed in patients 3, 5 and 6 (Figure 2.20, Table 2.3). On the contrary, less quasispecies were found in patients with acute infection (i.e. 8, 9 and 10) as compared to chronically infected (Table 2.3).

2.5.3 Phylogenetic analysis:

To estimate the evolutionary dynamics of NoV in immunocompromised patients, phylogenetic analysis was performed using BEAST 1.8.4 on a set of viral sequences (quasispecies >1%) isolated at

varying time points. Phylogenetic trees of quasispecies from 7 chronically infected patients showed clustering of quasispecies for each sampling time; supported by strong posterior probability >0.60 (Figure 2.21). Intra-patient quasispecies distribution indicate rapid viral evolution giving rise to new pool of quasispecies over the follow up time points.

2.5.4 Genetic complexity and diversity:

To study intra-host virus evolution, VP1 consensus sequences from chronically infected patients over time were compared across varying time points. Genetic complexity estimated by Shannon entropy for the immunocompromised patients ranged between 0.015 and 0.047 (Table 2.4). The highest Shannon entropy was observed for the patients 5 and 6 (i.e. 0.044 and 0.47 respectively), whereas the lowest was estimated for the patient 7 (i.e. 0.015) (Table 2.4).

Genetic diversity was assessed by comparing the distances between early and late sequential samples from the patients (Table 2.4). Of these, three patients (2, 3, and 7) showed an increase in the genetic distances in their later samples, both at amino acid and nucleotide level. On the contrary, patients 1, 4 and 5 showed declining genetic distances in later samples. Moreover, the number of non-synonymous mutations per synonymous sites (dN) were significantly higher than the synonymous mutations per synonymous site, suggesting a positive selection (Table 2.4). In this study, we also determined positively selected sites (P >95%; P>99) in the VP1 region among the patients, where a total of 15 positively-selected sites were predicted in the patients 5 and 8 infected with G.II.3 (Table 2.5). Thirteen positively selected sites were predicted for the patients 1, 2, 3, 4, and 7) infected with G.II.4 (Table 2.5).

2.5.5 Mutational hotspot in VP1:

Mutational hotspots were identified within the VP1 capsid protein in the samples from chronically infected patients, as shown in Figure 2.19. Among the infected patients, the majority of amino acid substitutions were observed on the outer surface of the capsid protein within or neighboring the epitope A, D and E (Table 2.6 and 2.7). Except for patient 7, all the patients showed to contain a virus population with at least one amino acid change in the antigenic epitopes during the follow ups. Several amino acid changes were found within the P2 domain (Table 2.7). Three amino acid substitutions including D341N, D357V/P and P396H were observed within the P domain shared among three GII.4 patients (1, 3, 4, and 10).

In the GII.3-infected patients (5 and 8), two amino acid substitutions at position 17 and 108 were observed in the S domain. Several variable sites were detected across the VP1 capsid protein in the patient 5 (GII.3-infected) and 6 (GII.6-infected), respectively (Table 2.7). Among half of these variable sites (6 of 15 in patient 5, and 3 of 5 in patient 6) were found to occur in the P2 domain. The exact location of epitopes in GII.3 and GII.6 are not characterized.

Consensus VP1 sequences of virus populations were also assessed for mutations in the receptor binding pockets (GII.4: 343-347, 374, 442-443; GII.6: 360-363, 389, 452; and GII.3: 357-360, 386, 450-451). No amino acid changes within the receptor binding pocket sites were detected.

Table 2.2: Demographic and clinical characteristics of NoV-infected patients.

Patients	Age (years)	Gender (male/female)	NoV Genotype	Underlying condition
Chronic infection				
1	58	M	GII.4	Bone marrow transplant
2	2	M	GII.4	Bone marrow transplant
3	6	M	GII.4	Bone marrow transplant
4	60	M	GII.4	Bone marrow transplant
5	68	F	GII.3	Bone marrow transplant
6	46	M	GII.6	HIV coiinfection
7	65	M	GII.4	Bonemarrow transplant
Acute infection				
8	20	F	GII.4	-
9	85	F	GII.3	-
10	57	F	GII.2	-

Table 2.3: Sampling time points, duration, mapped reads and number of quasispecies.

Patients	Samples	Sampling timepoints (Date)	Duration in samples (Days)	No of Reads	Mapped Reads	Average Read length	Total No of Quasispecies
1	1	08.01.08	0	1965	1957	1699	15
	2	11.03.08	64	823	820	1712	12
	3	05.05.08	119	5888	5866	1686	113
	4	04.10.08	271	5507	5487	1671	125
2	6	15.01.08	0	3108	3013	1670	4
	7	19.05.08	126	2218	2150	1673	92
	8	11.12.08	332	2775	2692	1682	104
	9	16.04.09	458	7936	7834	1685	130
3	10	20.05.08	0	3231	3113	1693	16
	11	15.09.08	119	8132	7946	1683	361
	12	04.12.08	199	2384	2320	1672	137
	13	25.02.09	282	2835	2768	1681	90
4	14	03.03.11	0	2718	2619	1822	2
	15	24.03.11	22	978	940	1881	179
	16	26.05.11	85	890	862	1817	161
	17	21.07.11	141	897	863	1671	318
5	18	09.05.11	0	905	838	1957	8
	19	13.07.11	66	5841	5723	1709	58
	20	29.09.11	144	1186	1156	1945	149
	21	08.12.11	214	2304	2025	1925	32
6	281	06.07.15	0	3249	3155	1716	41
	364	28.12.15	176	3598	3493	1712	7
7	68	31.03.14	0	9323	9257	1666	16
	409	07.04.14	8	14928	14816	1651	4
	940	16.04.14	14	1205	1194	1678	1
	188	24.04.14		6711	6670	1682	12
8†	325	12.01.17	0	679	506	1654	2
9†	491	16.01.17	0	1331	1106	1676	1
10†	690	13.12.16	0	3669	3478	1811	1

†Acutely infected patients

Table 2.4: Distribution of synonymous and non-synonymous mutations in patients.

Patients	Time Interval (Days)	Genetic distance (d)		Synonymous substitution/ synonymous site (dS)	Non-synonymous substituion/ non-synonymous sited (dN)	dN/dS Ratio	Shannon Entropy (Sn)
		Nucleotide	Amino acid				
	63	0.421	0.735	0.0013	0.0028	2.15	
1	54	0.137	0.274	0.008	0.0073	0.91	0.023
	152	0.146	0.292	0.0066	0.0061	0.92	
	124	0.005	0.015	0	0.0008	0	
2	206	0.027	0.063	0.0026	0.0065	2.5	0.029
	125	0.213	0.445	0.0026	0.0098	3.76	
	118	0.042	0.079	0.0026	0.0057	2.19	
3	80	0.100	0.208	0.0053	0.0089	1.67	0.025
	83	0.060	0.134	0.0013	0.0061	4.69	
	22	0.076	0.151	0.0022	0.0042	1.90	
4	62	0.080	0.149	0.0013	0.0061	4.60	0.023
	56	0.011	0.034	0	0.0008	0	
	65	2.919	5.186	0.0031	0.0054	1.74	
5	78	0.016	0.047	0	0.0016	0	0.044
	70	0.122	0.261	0.004	0.01	2.5	
6	175	0.028	0.052	0.0026	0,0055	2.11	0.047
	7	0	0	0	0	0	
7	9	0	0	0	0	0	0.015
	8	0.077	0.232	0	0.0008	0	

Table 2.5: (A) Positively selected sites in GII.3 sequences resulting from Naïve Empirical Bayes and Bayes Empirical Bayes analysis. (B) Positively selected sites in GII.4 sequences resulting from Naïve Empirical Bayes and Bayes Empirical Bayes analysis

A)

Site	Amino acid	P value
78	N	0.985**
108	V	0.997**
272	L	0.985**
292	R	0.985**
302	D	0.985**
308	L	1.000**
311	Y	0.985**
312	Y	0.985**
404	H	0.985**
417	A	0.999**
424	M	0.998**
478	T	0.985**
515	I	0.997**
516	T	0.985**
547	I	0.985**

B)

Site	Amino acid	P value
119	V	0.662
146	V	0.736
174	S	0.959*
294	A	0.942
297	R	1.000**
339	K	0.523
340	G	0.921
341	D	0.921
357	D	0.998**
368	S	1.000**
372	E	0.996**
373	N	0.615
412	V	0.960*

*: P>95%; **: P>99%

Table 2.6: Amino acid substitutions in the GII.4 blockade epitope regions.

Patient	Genotype	Epitope A							Epitope B		Epitope C		Epitope D				Epitope E		
		294	296	297	298	368	372	373	333	382	340	376	391	393	394	395	407	412	413
1	GII.4					S	E							S					
						N	D							N					
2	GII.4		S	R	N	S	D				G	E			T	T		N	
			T	H	S	I	E				A	D			S	N		D	
3	GII.4			H	N	N	D				G			S				N	
				R	D	S	N				R			G				D	
4	GII.4	P		R															I
		S		H															M
7	GII.4																		
10†	GII.4	S				A		N	V										I
		T				E		D	M										T

†Acutely infected patients

Table 2.7: Amino acid substitutions observed across S, P1 and P2 domains of VP1 protein. Highlighted in bold are the amino acids substitutions observed at position 17 and 108 in GII.3 patients; 341, 357 and 396 in GII.4 patients.

Patient	Genotype	aa variants/mutations			
		S (1-225)	P1 (226-278)	P2 (279-405)	P1 (406-520)
1	GII.4	V119I, I179V, Y183F		**D341N, D357V/P**, V386A	V472A
2	GII.4	I75V	G255S, F257Y	G295S/N, A304T, K339R, **P357T**, N380S, R397Q, N398S	R411G, N412G/D
3	GII.4	V138I		G340R, **D341N**, K348R, Y352S, A356V, **P357H**, A365V, **H396P**	Y496F
4	GII.4	N9S, T15S, **N17S**, V146T, N164S		**N341D**, **P396H**	
5	GII.3	N78V, **V108I**	L272M	R292K, D302E, L308S, Y311H, Y312C, H404D	A417S, M424T, T478S, I515V, T516N, I547V
6	GII.6		I230V	D308N, I329V, T372A, S395G	S542T
7	GII.4	**N17S**			
8†	GII.3	**L108V**	M289T	G385E	
9†	GII.2				
10†	GII.4	T15A, P174S		I289T, S310N, **N341D**, S359A, T377V, **P396H**	I413T, A539V, L540V

† Acutely infected patients

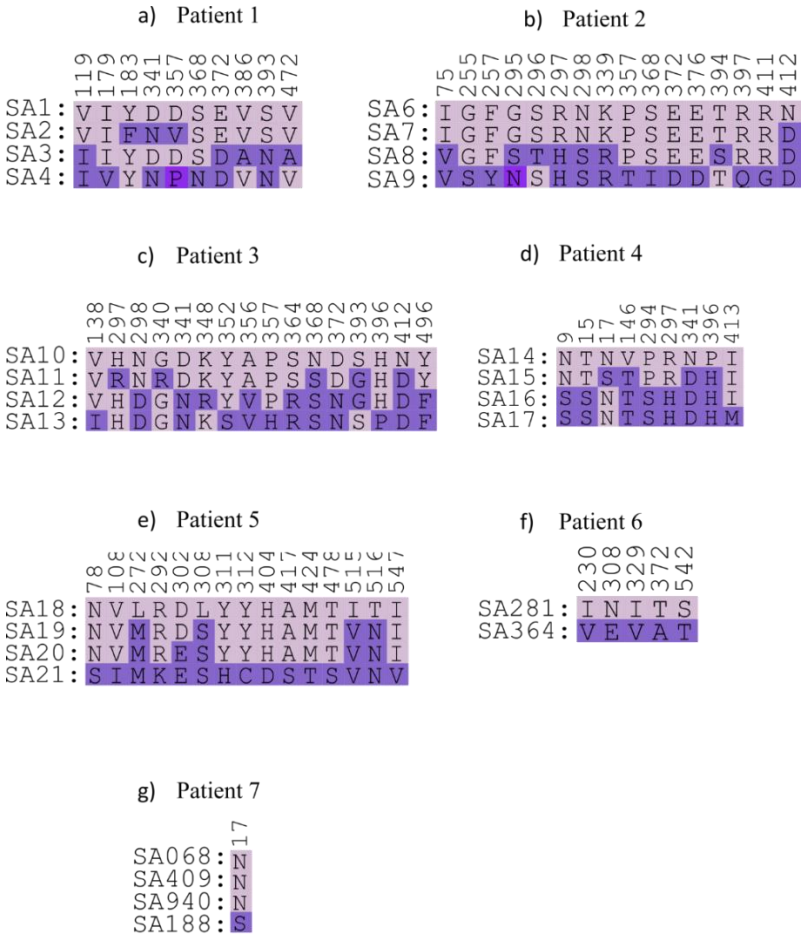

Figure 2.19: Consequential sequences of seven chronically infected patients. Mutation trajectory for seven immunocompromised patients over time. Multiply aligned VP1 consensus sequences showed variable positions for each patient (a-g). The column numbers correspond to amino acid position. The background color gets darker as the number of variations increases at position.

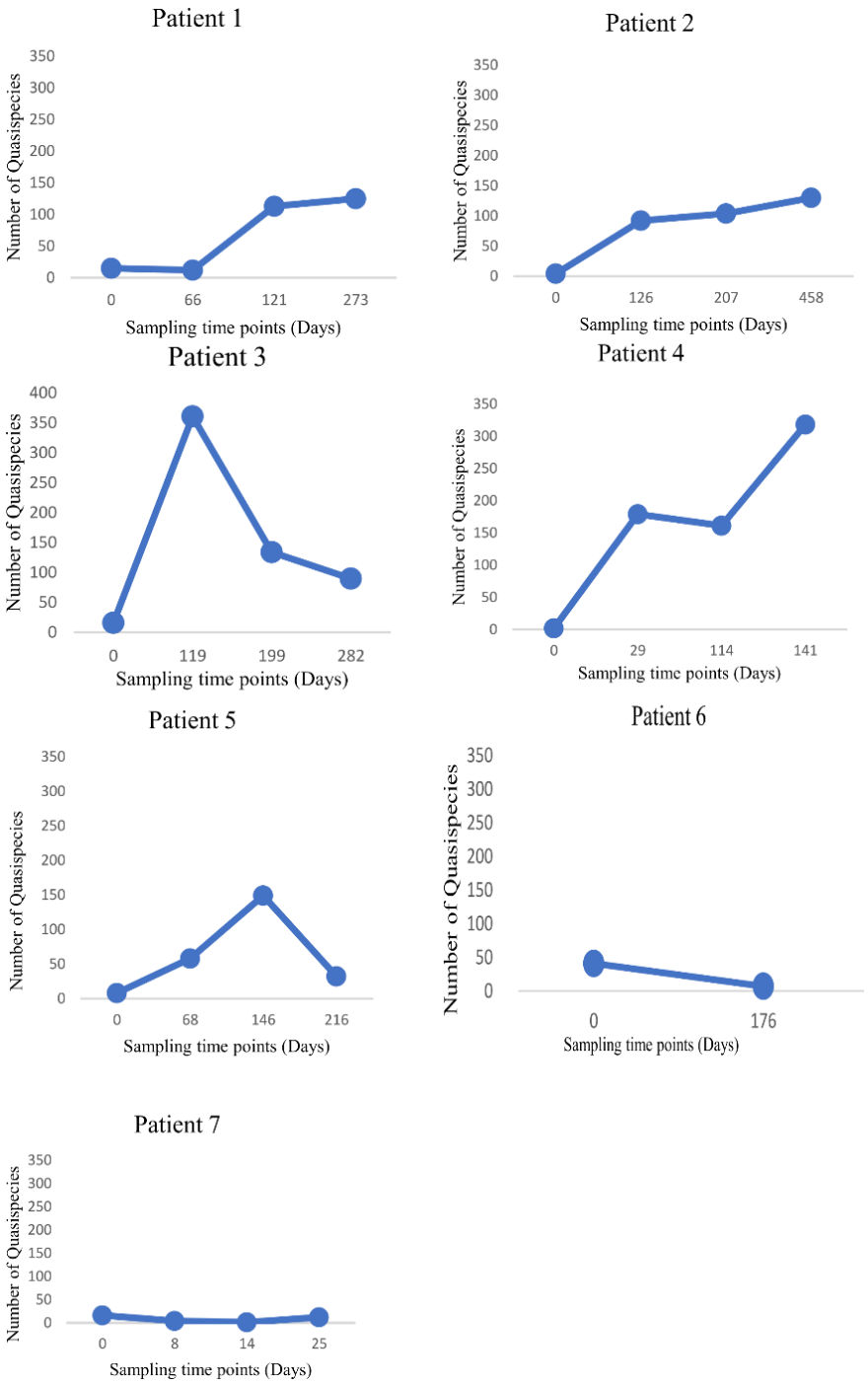

Figure 2.20: Distribution of NoV quasispecies among seven chronically infected patients across different time points. The x-axis shows the time interval between the samples and y-axis shows the number of quasispecies reconstructed.

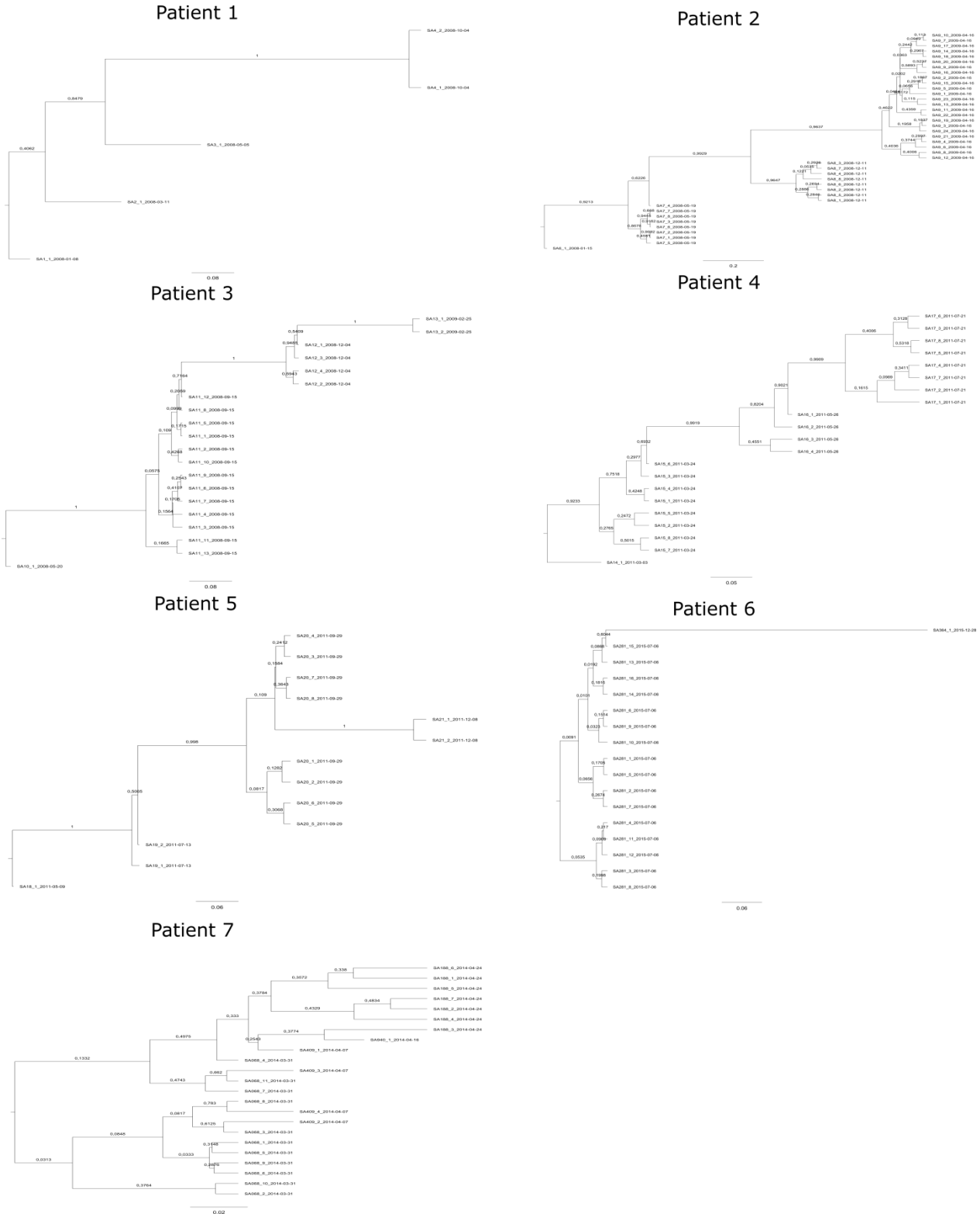

Figure 2.21: Phylogenetic analysis of VP1 sequences isolated sequentially from seven chronically infected patients. Phylogenetic trees of quasispecies >1% frequency was inferred using BEAST software. The trees showed the clustering pattern of quasispecies over different time points for each patient. The numbers above the branches indicate posterior probability of branch lengths.

2.6 Molecular Dynamics (MD) Stimulation:

2.6.1 Structure model building:

Models of VP1 structure were generated and checked with PROSA tool to confirm X-ray quality for structures (Figure 2.22). Moreover, Ramachandran plots were used for protein geometry, where > 90% favored regions, for all the models, was determined (Figure 2.23). Prediction for conserve and variable regions in the capsid protein of seven chronically infected patients showed that the S domain (pink) was conserved, whereas P1 and P2 domains contain several hypervariable regions (Figure 2.24).

2.6.2 Molecular dynamics simulations:

For the chronically infected patients, structure of the capsid protein were generated with GROMAC tool and epitopes were mapped (Figures 2.25 and 2.26). To obtain the detailed information on conformational changes in the structures, we performed molecular simulations analysis. The data revealed conformational changes in antigenic epitopes, especially in the surface-exposed epitopes in the P2 domain (Figures 2.25 and 2.26). More changes were observed in the P2E region (in the P2 domain) containing epitope E and D (Figure 2.26). In patient three, with immune escape, more conformational changes were detected in the structures from the later sequential samples, as shown in Figure 2.25.

2.6.2.1 RMSD analysis:

RMSD analysis was performed to evaluate stability of the structures. In patient 1, P2E region (containing epitope E and D) showed more fluctuation (0.8 nm) during the simulation when compared to the T-cell epitope (0.3 nm), P2C (containing HBAG) (0.3 nm), P2D (a hypervariable region next to the epitope A) (0.3 nm) and P1C (a hypervariable region in P1 domain) (0.2 nm) This revealed the instability of the capsid protein in this region, in all four sequential samples (Figure 2.27). In contrast, S domain (0.1 nm) was found to be stable for duration of simulation. Similarly, in patient 2, all the hypervariable regions were unstable for 20ns duration. In the last sample from patient 2, P2E region showed fluctuation after 15ns. In patients 3 and 4, fluctuation in all the hypervariable regions was observed except for the S domain. In patient 5 who was infected with GII.3 genotype, all the regions were stable during simulation, except for P2E region in the two initial samples. In patient 6, P2D

region showed more fluctuation than the P2E. Because of the short duration of infection, the selected regions showed less fluctuation in patient 7 than the other patients.

2.6.4 Radius of gyration (Rg) analysis:

In order to understand the compaction levels of each hypervariable sites we utilized the radius of gyration (R_g). Patient 1 showed high Rg for the P2E (containing epitope E and D), T-cell epitope and P2D regions compared to the S domain, P2C and P1C regions (Figure 2.28). However, the values were found to decrease in the last sample.

Similar pattern was observed in patients 2, 3, 4 and 7. In samples from patient 5 (GII.3-infected), the P2E, P2D, and T-cell epitope regions showed high Rg value compared to the P2C and P1C regions and the S domain. Similarly, high Rg value also detected for the P2E, P2D and T-cell epitope regions in patient 6 (GII.6-infected).

2.6.5 RMSF analysis:

We performed RMSF analysis to understand each residue contribution in flexibility of the variable site throughout the MD simulations. According to the results summarized in Figure 2.29, in all the patients except patient 7, high flexibility in P2E specifically in epitope D and E-related residues was observed. The P1C, P2C, P2D and T-cell epitope regions and the S domain tends to be less flexible. In patient seven less flexibility in all hypervariable residues was observed, probably due to short duration of infection (25 days).

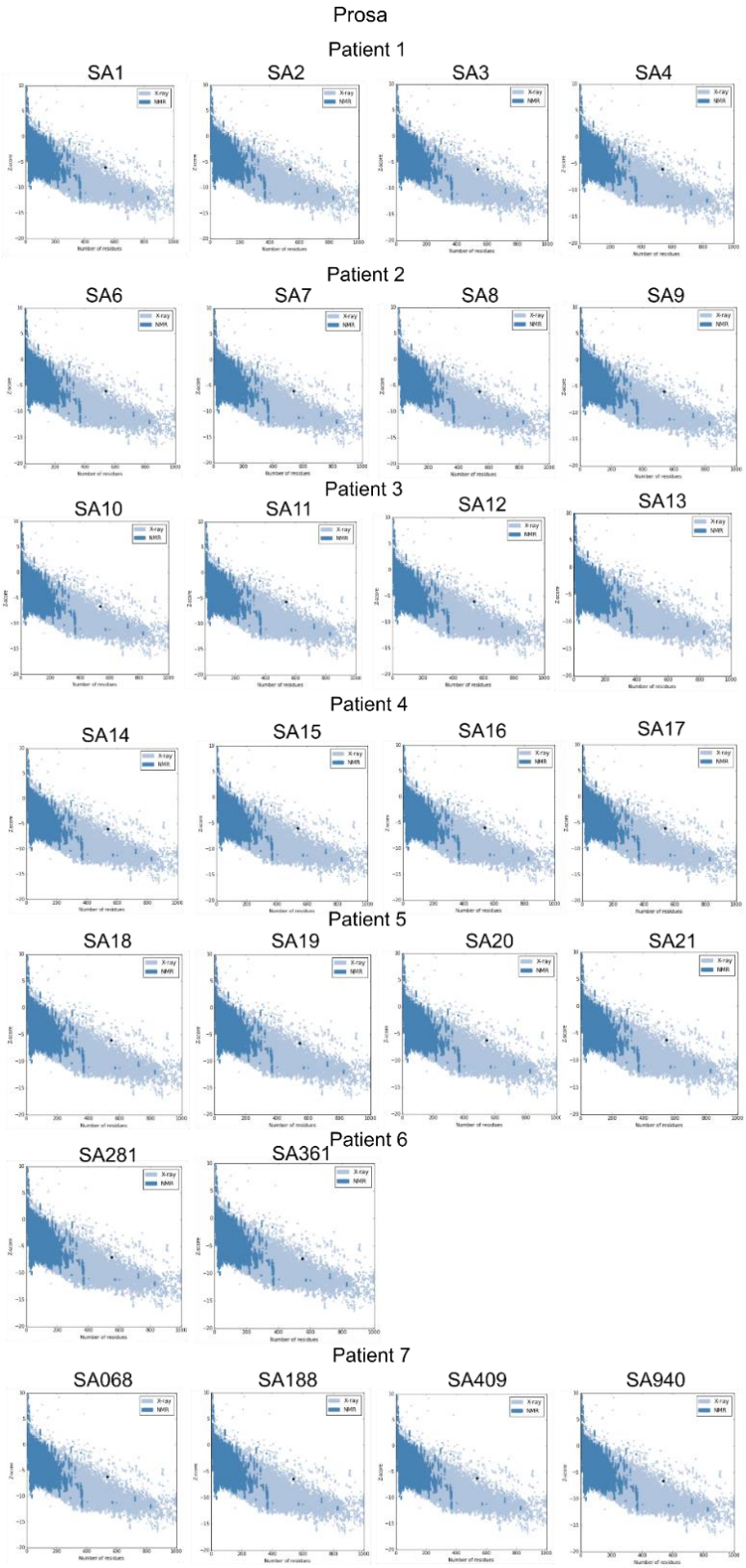

Figure 2.22: PROSA analysis for NoV VP1 capsid protein in seven chronically infected patients. (A) Black centered dots indicate all the VP1 capsid protein in seven chronically infected patients is in the range of real structure.

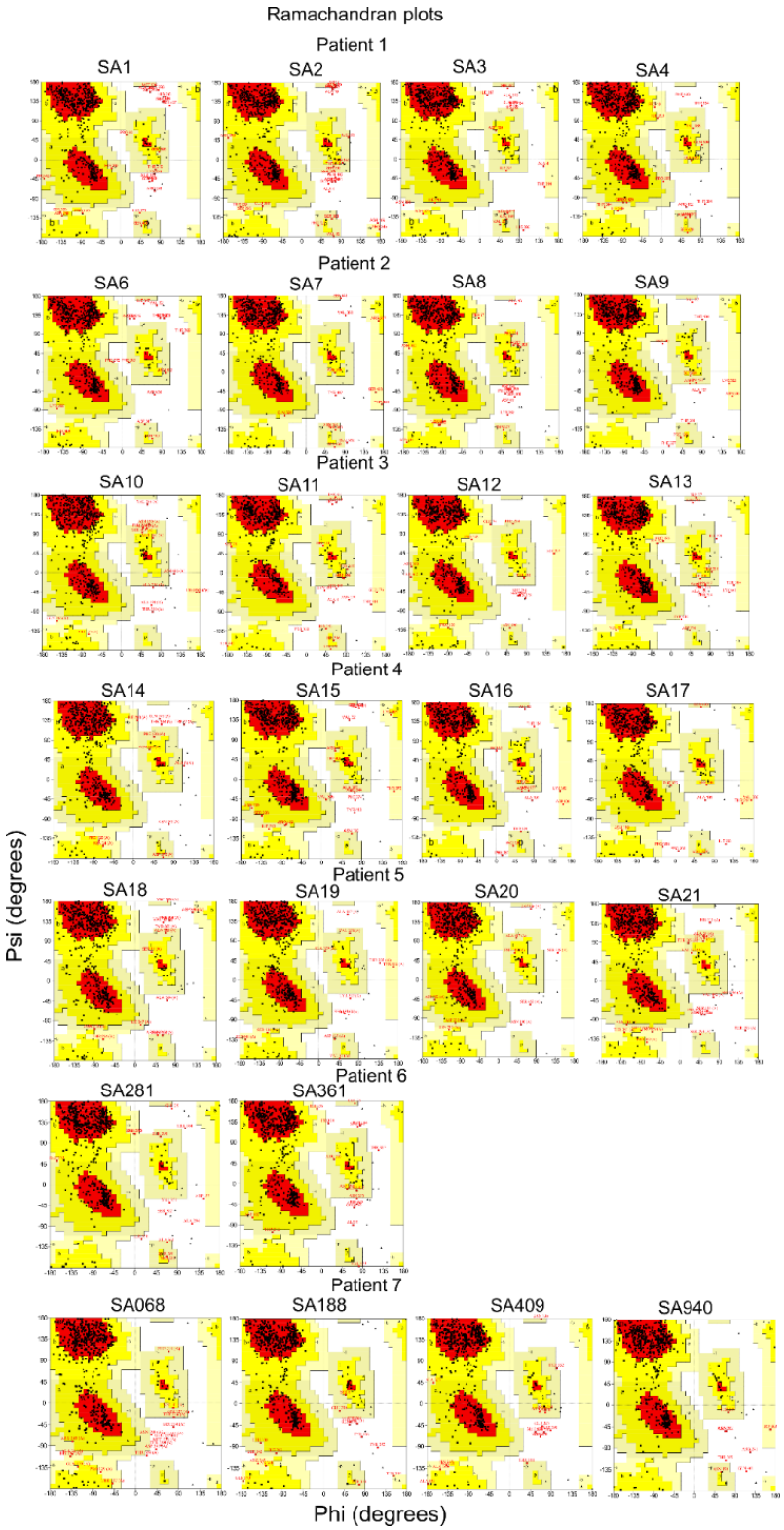

Figure 2.23: Ramachandran analysis for NoV VP1 capsid protein in seven chronically infected patients. Ramachandran plots showed the percentiles ranged from 95 to 98 indicating our capsid proteins within a range of crystal structure.

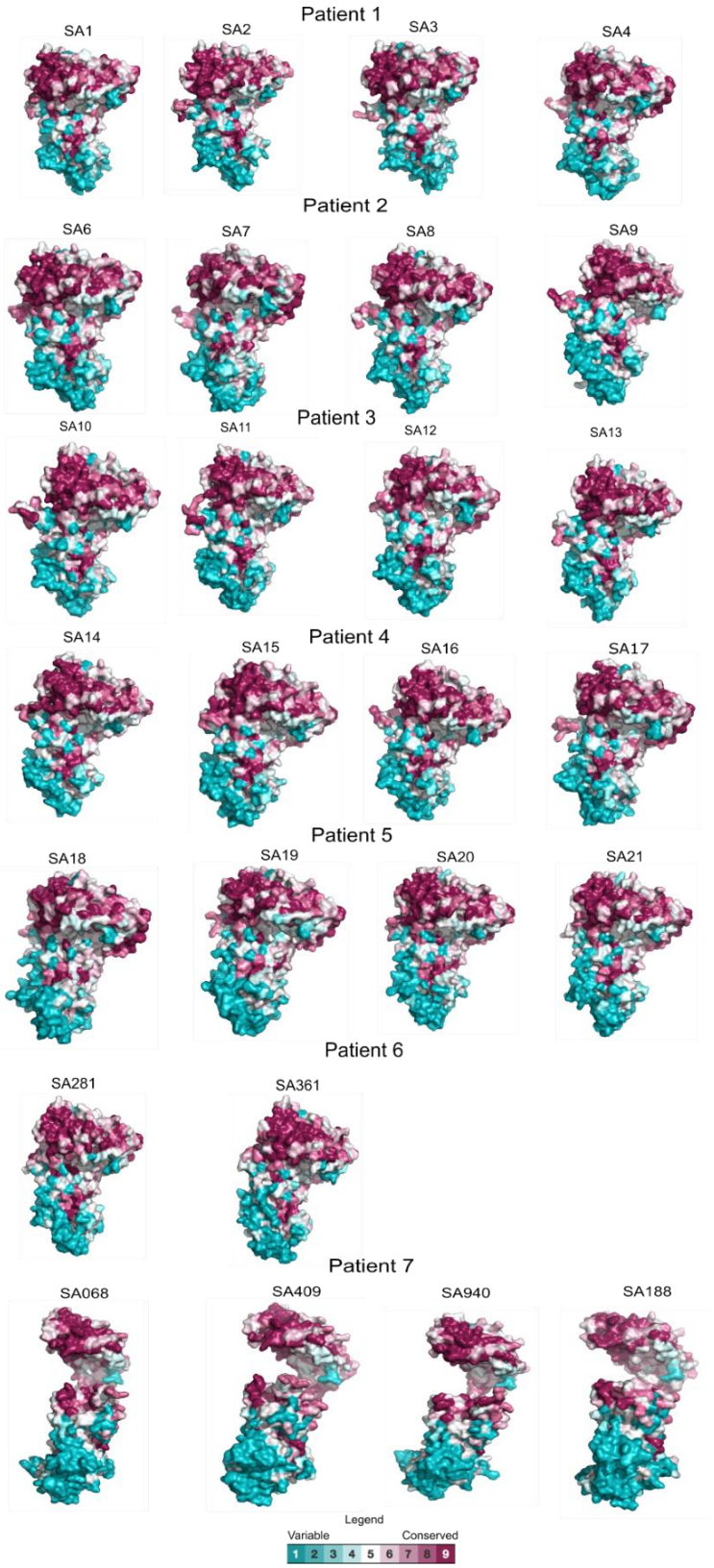

Figure 2.24: Consurf analysis for VP1 capsid protein of seven chronically infected patients. S domain (pink) showed conserved regions whereas P1 and P2 (blue) have a lot of hypervariable regions.

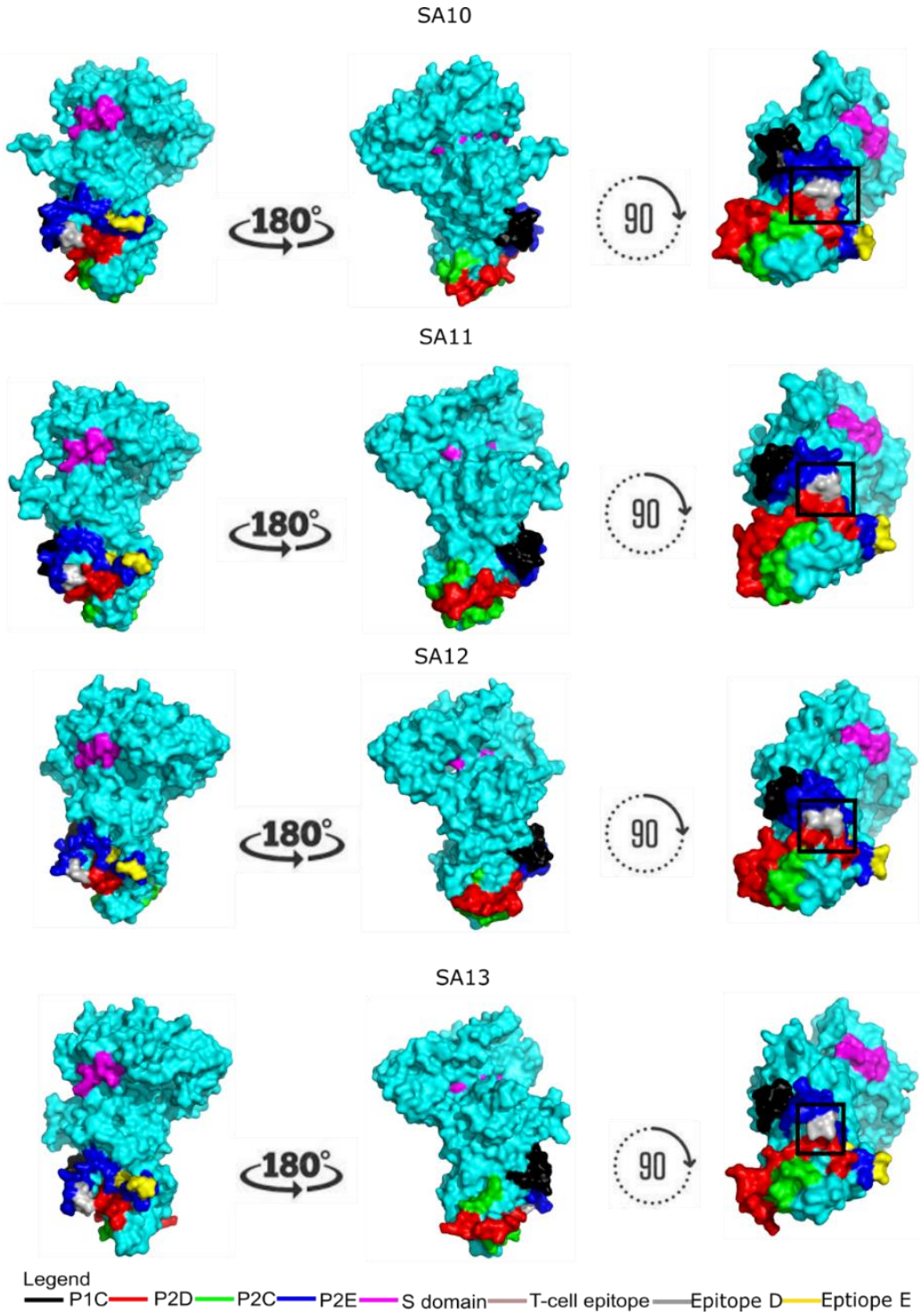

Figure 2.25: Structure of NoV VP1 capsid protein for chronically infected patient 3 generated with GROMACS tools. Different sites, (P1C: black (hypervariable region in P1 domain), P2D: red (hypervariable region neighboring epitope A in P domain), P2C: green (hypervariable region in P domain containing HBAG), P2E: blue (hypervariable region in P domain containing epitope E and D), S-domain: pink, T-cell epitope: brown, Epitope D: gray, Epitope E: yellow). In P2E more changes were observed as compared to other regions, because of the presence of epitope D and E.

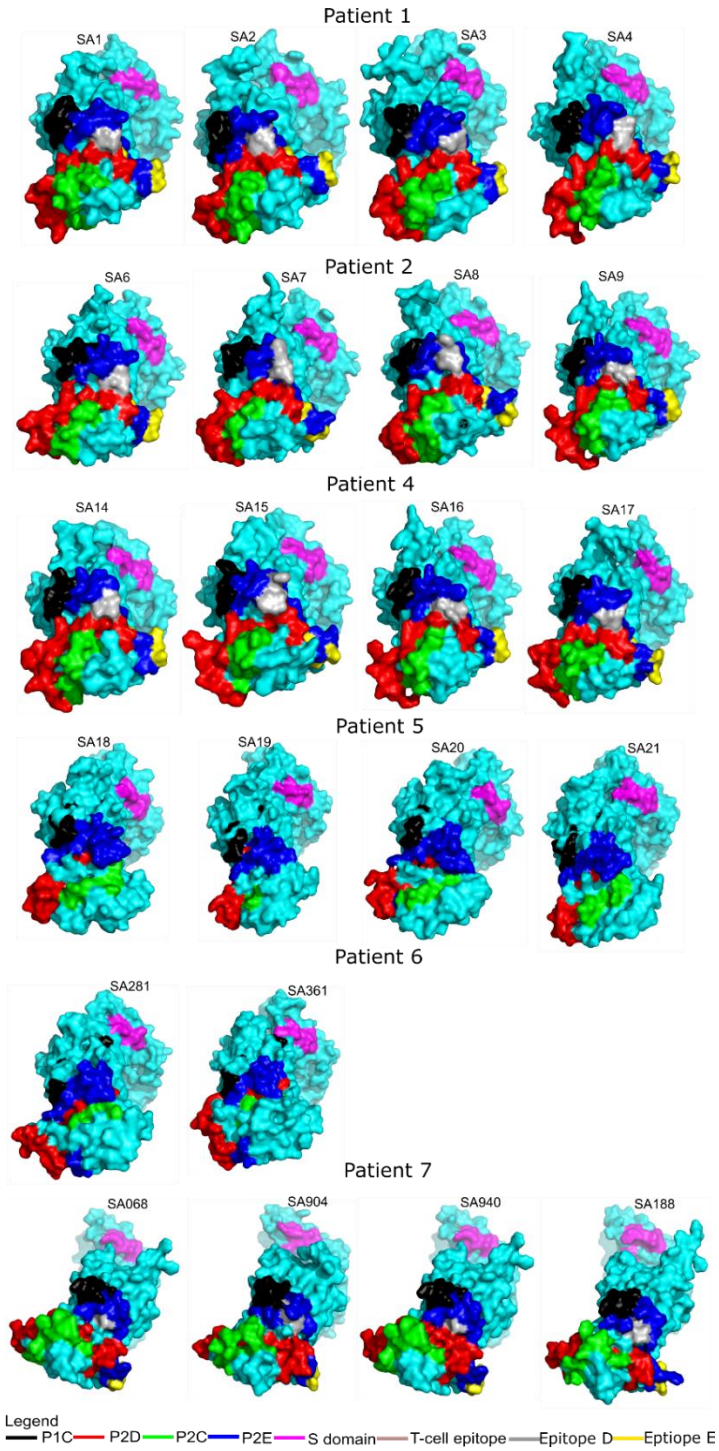

Figure 2.26: Structure of NoV VP1 capsid protein expressed from six chronically infected patients generated with GROMACS tools. Different sites, (P1C: black (hypervariable region in P1 domain), P2D: red (hypervariable region neighboring epitope A in P domain), P2C: green (hypervariable region in P domain containing HBAG), P2E: blue (hypervariable region in P domain containing epitope E and D), S-domain: pink, T-cell epitope: brown, Epitope D: gray, Epitope E: yellow). In P2E more changes were observed as compared to other regions, because of the presence of epitope D and E.

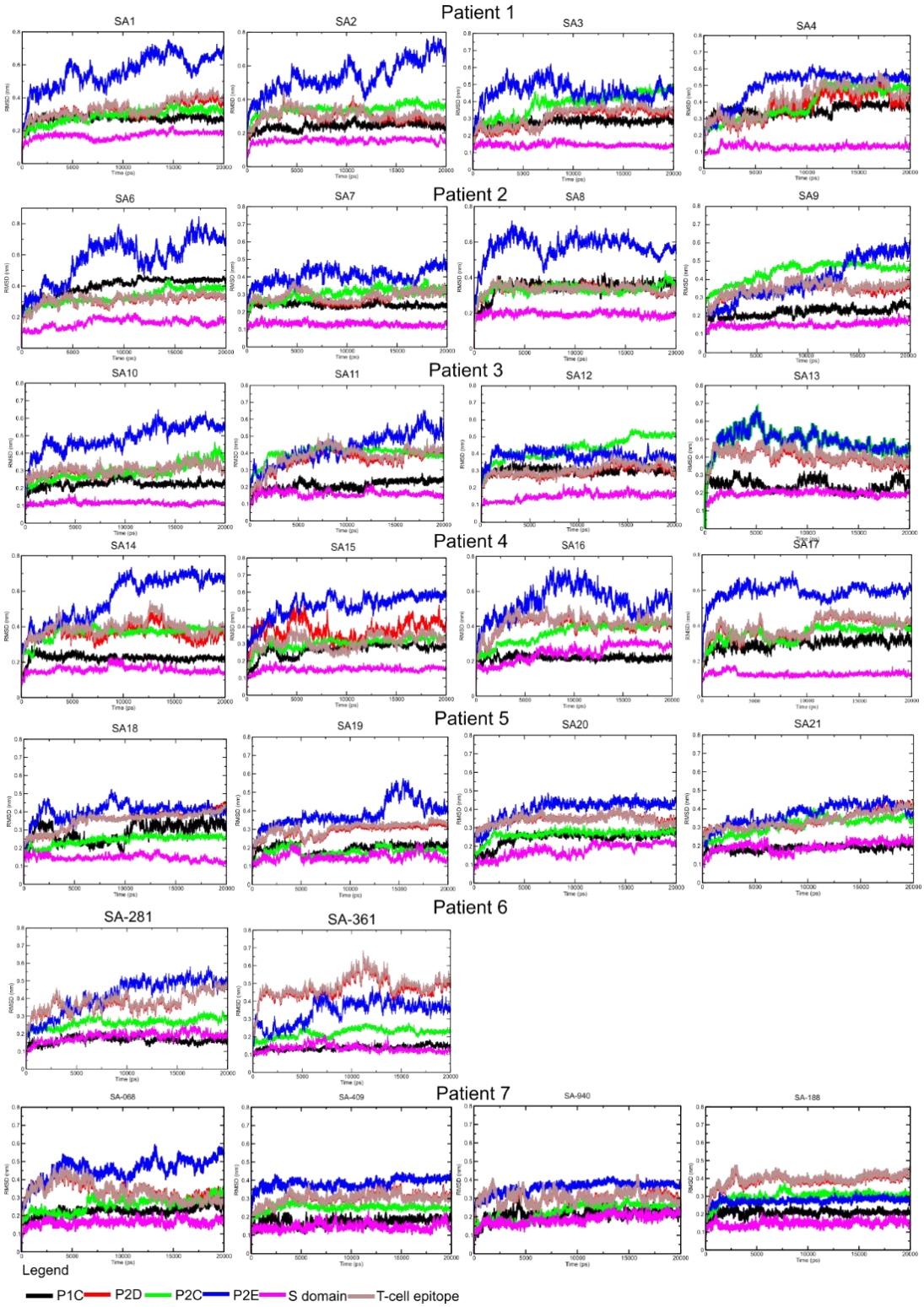

Figure 2.27: RMSD analysis for six chronically infected patients. In RMSD analysis P2E region showed a lot of fluctuation, indicating instability of NoV capsid protein in all patients followed by T-cell epitope, P2C, P2D and P1C, S domain was stable for 20ns duration.

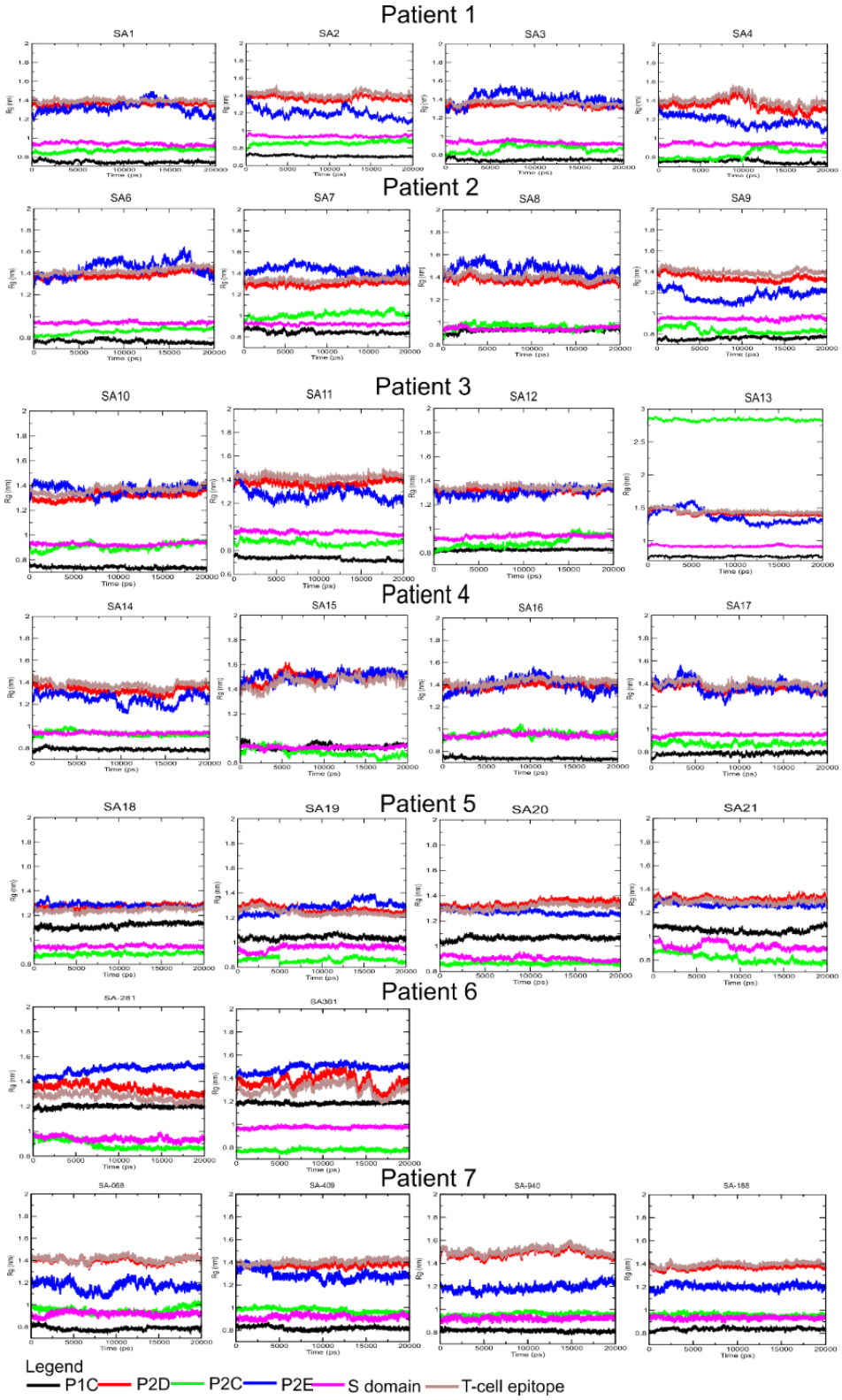

Figure 2.28: Radius of gyration analysis for seven chronically infected patients. In Rg analysis value for P2E was high followed by T-cell epitope and P2D, whereas S domain, P2C and P1C showed low Rg value throughout 20 ns duration.

Figure 2.29: RMSF analysis for seven chronically infected patients. High flexibility in P2E specifically in epitope D and E residues was observed, followed by P1C, P2C, P2D and T-cell epitope, S domain residues tends to be less flexible.

3. DISCUSSION

3.1 Molecular evolution of NoV GII.4 capsid protein:

NoVs are genetically very diverse viruses with increasing number of genotypes. However, a few genotypes such as GII.4 are more prevalent in outbreaks than others. They are not only a major cause of acute viral gastroenteritis in transplant recipients but are also responsible for prolonged chronic infection in immunocompromised patients. In this study, we characterize whole amino acid sequences of the GII.4 capsid protein available in the Genbank collected between 1974 and 2014. In the years 2004, 2006 and 2010, when several outbreaks occurred worldwide (Rowena A. Bull, Tu, McIver, Rawlinson, & White, 2006; J. S. Eden et al., 2010), higher number of mutations were observed. Variable sites were found to be located more in the P2 domain. Previous studies described three epitope sites A, D and E within the P2 domain of the GII.4 capsid protein that undergo amino acid change (Allen et al., 2009; Zakikhany et al. 2012). Our analysis revealed that amino acid changes occurred in the epitopes A (294, 296, 298), D (391-395) and E (407, 412, 413). However, HBGA binding sites (342-347, 374) were conserved through the years. It seems that continued evolution in the major blockade epitopes results in immune escape from herd immunity and the emergent of new virus variants. It is also hypothesized that antigenic changes may drive changes in HBGA affinities, altering population susceptibility patterns (Kari Debbink, Lindesmith, Donaldson, & Baric, 2012; J.-S. Eden, Tanaka, Boni, Rawlinson, & White, 2013). Interestingly, New Orleans 2009 and Sydney 2012, the most recent GII.4 pandemic variants, show both antigenic capsid mutations (antigenic drift) and intra-genotype recombination at the open reading frame 1 and 2 overlap (antigenic shift) (John-Sebastian Eden et al., 2013). Furthermore, the rise of new GII.4 variants relates to increase in the rate of infection and epidemics of gastroenteritis across the world. This phenomenon was explained by pandemic GII.4 strain, Sydney 2012 (van Beek et al., 2013). Sydney 2012 strain was first recognized in Australia (Eden et al., 2010) and after six months it was responsible for 25% of NoV-associated acute gastroenteritis outbreaks in Australia. It replaced its predecessor GII.4 strain between November 2012 and January 2013, leading to an increase in the number of outbreaks of gastroenteritis across the world (Kari Debbink et al., 2012).

3.2 Quantitative measurement of NoV-specific IgG in infected patients:

A quantitative ELISA was developed by coating the wells with recombinantly expressed P domain. Intra and inter assay variation assessment using mouse sera immunized with GII.1 and GII.4 capsid proteins confirmed the repeatability and precision of the assay. Our ELISA reliably quantitates antibodies directed against GII.4, GII.1 before and after infection. The assay detected pre-existing antibodies in all the samples. Concentrations varied, but the antibodies apparently could not protect the respective host against subsequent GII.1 and GII.4 infection. This is in line with previous reports (Rockx et al., 2002; Sakon et al., 2015; Simmons, Gambhir, Leon, & Lopman, 2013). While for highly prevalent GII.4, one can assume some degree of herd immunity (Sakon et al., 2015; Simmons et al., 2013).This is unlikely for rare GII.1 variant, as protection is only genotype-specific (Malm, Uusi-Kerttula, Vesikari, & Blazevic, 2014).Thus, virtually, all individuals exposed to the GII.1 virus within the foodborne outbreak in 2011 (Hoffmann et al., 2012) were presumably susceptible to this genotype. As described by Siebenga et al. (2010), GII.4 variants changes every 3–7 years is consistent with protective, long-term herd immunity in a substantial portion of the population. Our assay yielded similar range of capsid-specific IgG concentrations as shown in previous assays (Kavanagh et al., 2011). Atmar et al. (2015) found no correlation between antibody levels and protection of subsequent infection. On the other hand, Malm et al. (2014) reported lower GII.4 specific IgG titers in infants before GII.4 infections than their counterparts infected with other GII genotypes. We detected similar antibody concentrations in patients after GII.1 and GII.4 infection. Thus, the assay seems suitable for multiple GII genotypes as well as human and animal samples alike. Our simple and fast assay can monitor humoral immune responses in a high number of samples. It may be helpful for quickly evaluating immune responses to new vaccine candidates.

To our knowledge, we quantified IgG antibodies in sequential serum samples of chronically infected patients for the first time. ELISA plates were coated with the VP1 P2 domains from sequential samples. This allows us to monitor antibody binding to earlier, simultaneous and later capsid proteins. Antibodies in early sera bound weakly to the sequential capsid proteins. ELISA signals steadily increased in sera samples drawn at later time points, particularly with earlier antigens. This indicates an increasing adaptive B-cell response and is in line with a following cure in patients 1-4. Antibody concentrations in patient 3 developed differently. The first serum showed low signals with all the sequential capsid antigens, as observed in the other patients.

Interestingly the 2nd serum recognized the first antigen well but yielded steadily decreasing signals with the following capsid antigens. This could indicate immune escape of the second and following sequential capsid proteins. In contrast, the third serum produced high signals with all the antigens. This is in line with an effective humoral immune response leading to cure. In this patient, two fixed mutations in the epitope A and E, and one reverse mutation in the epitope D were observed. However, in other patients, no fixed mutations were detected in the epitopes, as shown in Figure 2.18. In accordance with the fixed mutations, dN/dS ratio was found to be higher in patient 3 than the other chronically infected patients.

3.3 NGS revealed genetic diversity and evolution in chronically infected patients:

We employ NGS to characterize quasispecies diversity and evolution in chronically infected patients. Sequential samples from seven immunocompromised patients over various times were analyzed. Of these patients, five were infected with GII.4, while the remaining two subjects were infected with GII.3 and GII.6. The number of quasispecies was found to be vary between time points. Low number of quasispecies was detected in the first samples of all patients. Then, it increased in all the patients with longer infections. However, a gradual decrease in number of quasispecies towards the end of the follow-up was observed in patients 3, 5 and 6. In all the patients, except patient seven who had just 25 days of infection, a peak in the number of quasispecies was observed 4-5 months post infection. Data from patient seven showed that short period of infection is not sufficient to establish high number of quasispecies.

Genetic distance was found to vary over time. Genetic distance between the quasispecies remained low; however, it was higher on amino acid level than nucleotide level, reflecting positive selection. Particularly, in patient four with documented cure, the number of quasipsecies decreased sharply in the last sample, one week before cure.

Most studies analyzing NoV intra-host genetic diversity have been restricted to identifying potential changes in the VP1 (Siebenga et al., 2008; Chan et al., 2012; Vega et al. 2014). Our data analysis revealed increasing number of nonsynonymous mutations in the capsid protein over time as reported for other viruses such as HCV and HIV (Bull et al., 2011; Ramachandran et al., 2011). Fifty-two to 78 % of the nonsynonymous mutations were observed in the P2 domain. In three immunocompetent patients with acute infections, nonsynonymous mutations were found to be rare (4.5-6.1 %) as compared to the chronically infected patients. This is in agreement with

the report of Bull et al. (2012). They observed various single nucleotide variation (SNV) with less than 2 % frequency in immunocompetent individuals with acute infection. However, in chronic infection, they detected significantly higher diversity in the sequential samples over time. Likewise, Vega et al. (2014) identified multiple SNVs within immunocompromised bone marrow transplant patients with more that 10% frequency. However, they detected no SNV throughout the viral genome in acute infection. In similar studies, Hasing et al. (2016) and Kundu et al. (2013) also showed more than 20% of nonsynonymous mutations in the immunocompromised- and bone marrow transplant patients, of these the majority were observed in the P2 domain.

In our study, accumulation of mutations across the VP1 gene including the antigenic domains (Epitopes A, C, D and E) was observed in all the immunocompromised patients. It seems that nonsynonymous mutations play a role in the evolution of these epitopes leading to antigenic changes of the capsid protein and subsequently immune escape (Nasheri, Petronella, Ronholm, Bidawid, & Corneau, 2017). In previous studies (Eden *et al.* 2014; Sabrià et al. 2018), amino acid exchanges at positions 293, 294, 356, 358, 368, 373, 376, 380, 404, 407 and 413 were suggested to be linked to the pandemic outbreaks caused by the New Orleans 2009 and Sydney 2012 variants. In the samples from chronically infected patients, we also observed amino acid changes at positions 294, 356, 368, 373, 376, 380 and 413. Our analysis showed that some mutations are genotype specific i.e. amino acid substitution at position 108 in the GII.3-infected individuals or substitution at position 357 in the GII.4-infected patients. Most fixed mutations occurred in the epitope A, D and E regions within the P2 domain. No fixed amino acid mutations were detected in patient seven, maybe due to the short duration of infection (Figure 2.19 and Table 2.6).

3.4 Molecular dynamic simulation confirmed epitope evolution in the patients:

With the rise of computing power and number of NMR / X-ray structures, molecular dynamics analyses are being done on increasingly large bimolecular systems over extensive time-scales (Sargsyan et al., 2017). The occurrence of immune escape can be predicted by observations on the effect of point mutations at the protein level. These observations can be analyzed by using advanced computational methods (C. V. Kumar et al., 2014). Experimental techniques are available to understand the dynamics of protein binding. They are connected to their spatial and

temporal resolutions, so they give only average properties instead of the motion of particles (Dror, Dirks, Grossman, Xu, & Shaw, 2012). Recently, a computational technique known as molecular dynamics simulation (MD) is applied to delineate protein motions at the atomic level (Dror et al., 2012). Popular molecular simulation suites are CHARMM, AMBER, GROMACS and NAMD; they all share some common features (Salsbury, 2010). The GROMACS package has advantages over the others as follows: I) It is simple to use as there is no usage of scripting language; II) GROMACS is designed to perform simulations on a large variety of biological structures, including proteins, chemical compounds and lipids; and III) it also has a large number of tools to do trajectory analyses (Salsbury, 2010).

In the present study, MD simulation was applied to analyze structural changes of NoV capsid protein in chronically infected patients. In all the patients, most amino acid exchanges were observed in the P2 domain, which was also predicted to be the most variable region by consurf analysis (Figure 2.24). Data analysis with the trajectories RMSD, RMSF and Rg revealed conformational changes in antigenic epitopes with the passage of time, especially in the surface-exposed epitopes in the P2 domain. More changes were observed in the P2E region (in the P2 domain) containing epitope E and D. It thought that exchange of amino acids in this region might lead to immune escape (Kari Debbink et al., 2012). This is relevant for chronically infected patients, where virus populations have time to accumulate multiple mutations. Within our chronically infected patients, one patient (patient 3) with immune escape, detected with ELISA, fixed mutations leading to structural changes in epitope regions were observed. In contrast, very few mutations were detected in the capsid proteins of patient seven, with short period of infection (25 days). As we expected, protein structures remained virtually constant in this patient (Figure 2.26).

MD analysis revealed instability in the P2 domains in all the patient's samples. Similar MD analysis has been done for Hepatitis B surface antigen by Rezaee et al. (2016). They compared the wild type G145 with the mutant 145R and showed that this mutation can influence stability and antibody binding.

The 3D structure variation of proteins plays a vital role in biological activity and interaction with other molecules (Ponsel & Bruss, 2003). RMSD and Rg values have been shown to reliably determine stability and protein compactness (Maiorov & Crippen, 1994; Tsai & Nussinov, 1997; Rezaee et al., 2016). RMSD analysis showed high fluctuation in the P2E region containing the

epitope E and D, indicating instability of the capsid protein in this region. However, the S domain was stable as expected. The patient with short time of infection showed less fluctuation as compared to longer infection (Figure 2.27). Rg analysis demonstrated more compact S domain as compared to the P2 domain. High Rg values were observed in P2E region indicating less compact P2 domain in this region in all the samples, except in the one with short infection period (25 days). Rg analysis conducted by Lobanov et al. (2008) showed the characteristics of a protein structure, number of contacts per residue, rotation angle number per residue, compactness, and radius of gyration. All these properties notionally affect the folding rate. It should be noted that the protein classes isolated according to the SCOP (A structural classification of proteins database for the investigation of sequences and structures) differ in these characteristics (Lobanov, Bogatyreva, & Galzitskaia, 2008).

There is also a correlation between the protein flexibility and Cα RMSF measures; proteins with higher RMSF values are generally known to have more flexible regions (Benson & Daggett, 2008). In our study, we observed high flexibility in the P2E region, specifically in the epitope D and E residues; S domain residues tends to be more flexible.

In conclusion, immune escape was detected in one chronically infected patient using a quantitative ELISA. NGS data analysis in the infected patients showed amino acid substitutions mostly in the P2 domain. Numerous nonsynonymous mutations (52-78 %) was observed in the capsid protein of the chronically infected patients, except for patient 7 (13.9 %) indicating positive selection. Whereas, in the acutely-infected patients, less nonsynonymous mutations (4.5-6.1 %) was detected. Structural modelling and molecular dynamic simulation revealed evolution of the antigenic epitopes in the capsid protein which could affect the binding capacity to NoV-specific antibodies, possibly leading to immune escape in chronically-infected patients. The data obtained in this thesis will help researchers to understand the genotypic and phenotypic changes that occur in antigenic epitopes during NoV infection and will help to guide vaccine development.

4.1 MATERIALS

4.1.1 Chemicals:

Table 4.1

Chemical	Supplier
Agarose	Peqlab, Erlangen Germany
Amersham ECL Prime Western Blotting Detection Reagent	Roth, Karlsruhe Germany
Ampicillin	Roth, Karlsruhe Germany
Chloramphenicol	Roth, Karlsruhe Germany
DMSO	Sigma-Aldrich, Steinheim Germany
Ethanol	Roth, Karlsruhe Germany
In-Fusion gene assembly kit	Takara, Kusatsu Japan
Isopropanol	Roth, Karlsruhe Germany
Lipofectamine 2000	Thermo Fischer Scientific, Waltham USA
Metafectene	Biontex Laboratories GmbH, Munich Germany
Methanol	Roth, Karlsruhe Germany
Milk powder	Roth, Karlsruhe Germany
NaCl	Roth, Karlsruhe Germany
Non-essential amino acids 100x	Gibco/Thermo Fisher Scientific, Waltham USA
Page Ruler Plus Prestained Protein Ladder	Thermo Fischer Scientific, Waltham USA
Penicillin/Streptomycin	Gibco/Thermo Fisher Scientific, Waltham USA
Phusion® High_Fidelity PCR master mix	Thermo Fisher Scientific, Waltham USA
Pierce RIPA Buffer	Thermo Fischer Scientific, Waltham

	USA
Polyacrylamide	Roth, Karlsruhe Germany
Protease Inhibitor (Complete)	Roche, Mannheim Germany
RotiSafe	Roth, Karlsruhe Germany
SDS	Roth, Karlsruhe Germany
SmartLadder MW-1700-10	Eurogentec, Liege Belgium
Sodium pyruvate	Gibco/Thermo Fisher Scientific, Waltham USA
Sucrose	Roth, Karlsruhe Germany
TEMED	Roth, Karlsruhe Germany
Tris HCl	Roth, Karlsruhe Germany
TrueBlue™ Peroxidase Substrate	SeraCare Life Sciences Inc., Milford USA
Trypsin	Gibco/Thermo Fisher Scientific, Waltham USA
Tween 20	Roth, Karlsruhe Germany

4.1.2 Consumables:

Table 4.2

Consumable	**Supplier**
ELISA plates	TPP, Trasadingen Switzerland
Cryo Vials	Greiner Bio One, Kremsmünster Austria
Falcon Tubes 15 ml, 50 ml	Greiner Bio One, Kremsmünster Austria
Gene Pulser® electroporation cuvettes	Biorad, Hercules USA
Immun-Bot PCDF membrane	Bio-Rad, Hercules USA
Pipette filter tips	Greiner Bio One, Kremsmünster Austria
Pipette tips	Greiner Bio One, Kremsmünster Austria

Tubes 1,5 ml, 2 ml	Greiner Bio One, Kremsmünster Austria
Whatman paper	Bio-Rad, Hercules USA

4.1.3 Commercial Kits:

Table 4.3

Kit	Supplier
GeneJet Plasmid Miniprep	Thermo Fisher Scientific, Waltham USA
NucleoSpin Blood QuickPure	Macherey-Nagel, Düren Germany
Plasmid *Plus* Midi Kit	Qiagen, Hilden Germany
QIAquick® Gel Extraction Kit	Qiagen, Hilden Germany
QIAquick® Nucleotide Removal Kit	Qiagen, Hilden Germany

4.1.4 Laboratory Equipment:

Table 4.4

Equipment	Supplier
Balance	Kern, Balingen Germany
Centrifuge 5417C	Eppendorf, Hamburg Germany
Centrifuge 5920	Eppendorf, Hamburg Germany
Chemiluminescence detector ECL Chemocam	Intas, Ahmedabad India
Freezing container	Thermo Fisher Scientific, Waltham USA
Gel chambers (SDS-PAGE)	Biometra, Göttingen Germany
Heating bath Isotemp GPD 20	Fisher Scientific, Hampton USA
Incubator	Heraeus, Hanau Germany
Nanodrop One	Thermo Fisher Scientific, Waltham USA

pH meter	Mettler Toleda, Gießen Germany
Pipette Accu jet pro	Brand, Wertheim Germany
Pipettes	Eppendorf, Hamburg Germany
Shaking incubator	Infors, Bottmingen Germany
Sterile Hood	BDK, Sonnenbühl Germany
Thermocycler Professional	Biometra, Göttingen Germany
Thermomixer C	Eppendorf, Hamburg Germany
Ultracentrifuge Optima L-90K	Beckman Coulter, Brea USA
Vortexer MS1	IKA, Staufen Germany

4.1.5 Buffers:

Table 4.5

Buffer	Components
10x SDS-electrophoresis buffer pH 8.3	30 g Tris; 24 mM 144 g Glycin; 192 mM 5 g SDS; 0,1 %
10x TBS pH 7.6	80.06 g Tris; 200 mM 24.22 g NaCl; 1,5 M
1x TBS 0.1 % (v/v) Tween20	100 ml 10x TBS 1 ml Tween20
10x Western Blot transfer buffer	30.,3 g Tris; 25 mM 144 g Glycin; 192 mM
1x Western Blot transfer buffer	100 ml 10x Western Blot transfer buffer 200 ml MeOH
50x Tris-Acetat-EDTA Electrophoresis buffer (TAE) pH 8.3	242 g Tris 57,1 ml acetic acid 0ml 0,5 M EDTA pH 8

4.1.6 Bacterial strains:

Table 4.6

Bacteria strain	Supplier
E. coli Bl21 cells	Gift from Ingo Drexler
One Shot™ TOP10 Chemically Competent *E. coli*	Thermo Fischer Scientific

4.1.7 Antibodies:

Table 4.7

Antibody	Application	Dilution	Catalog Number	Supplier
α-GII-Human serum	WB	1:5000	-	-
α-rabbit IgG (H+L)	Titration	1:5000	111-035-144	Jackson Immuno Research
GII.10 polyclonal antibody	ELISA	1:2000	-	Grant Hansman
Human-IgG	ELISA	1:40000	Ab91102	Abcam

4.1.8 Primers:

Table 4.8: GII.4 P domain capsid protein

No.	Primer	Sequence (5'→3')	Restrection enzyme
1	GII.4 -F	GGATCCTTCTTGGTGCCACCCACAGTT	BamHI
	GII.4 -R	GCGGCCGCTTATAATGCACGTCTGCGCCC	NotI

Table 4.9: GII.1 P domain capsid protein

No.	Primer	Sequence (5'→3')	Restrection

			enzyme
1	GII.4 -F	GGATCCTTCTTGGTGCCACCCACAGTT	BamHI
	GII.4 -R	GCGGCCGCTTATAATGCACGTCTGCGCCC	NotI

Table 4.10: GII.4 VP1 capsid protein

No.	**Primer**	**Sequence (5'→3')**	**Restrection enzyme**
1	GII.4 -F	GGATCC ATCGCAATCTGGCTCCCAGTTT	BamHI
	GII.4 -R	GCGGCCGC CCAGCAAAGAAAGCTCCAGCCATTA	NotI

Table 4.11: GII.4 VP1 capsid protein for NGS

No.	**Primer**	**Sequence (3'→5')**	**Barcode**
1	NgVP.1-F	TCAGACGATGCGTCATGATCGCAATCTGGCTCCCAGTT	TCAGACGATGCGTCAT
	NgVP.1-R	TCAGACGATGCGTCATCCAGCAAAGAAAGCTCCAGCCATTA	
2	NgVP.2-F	CTATACATGACTCTGCGATCGCAATCTGGCTCCCAGTT	CTATACATGACTCTGC
	NgVP.2-R	CTATACATGACTCTGCCCAGCAAAGAAAGCTCCAGCCATTA	
3	NgVP.3-F	TGTGTATCAGTACATGGATCGCAATCTGGCTCCCAGTT	TGTGTATCAGTACATG
	NgVP.3-R	TGTGTATCAGTACATGCCAGCAAAGAAAGCTCCAGCCATTA	
4	NgVP.4-F	ACACGCATGACACACTGATCGCAATCTGGCTCCCAGTT	ACACGCATGACACACT
	NgVP.4-R	ACACGCATGACACACTCCAGCAAAGAAAGCTCCAGCCATTA	
5	NgVP.5-F	CTGCGTGCTCTACGACGATCGCAATCTGGCTCCCAGTT	CTGCGTGCTCTACGAC
	NgVP.5-R	CTGCGTGCTCTACGACCCAGCAAAGAAAGCTCCAGCCATTA	
6	NgVP.6-F	TAGATGCGAGAGTAGAGATCGCAATCTGGCTCCCAGTT	TAGATGCGAGAGTAGA
	NgVP.6-R	TAGATGCGAGAGTAGACCAGCAAAGAAAGCTCCAGCCATTA	
7	NgVP.7-F	CATCACTACGCTAGATGATCGCAATCTGGCTCCCAGTT	CATCACTACGCTAGAT
	NgVP.7-R	CATCACTACGCTAGATCCAGCAAAGAAAGCTCCAGCCATTA	
8	NgVP.8-F	CGTGTCGCGCATATCTGATCGCAATCTGGCTCCCAGTT	CGTGTCGCGCATATCT
	NgVP.8-R	CGTGTCGCGCATATCTCCAGCAAAGAAAGCTCCAGCCATTA	
9	NgVP.9-F	AGAGAGAGACATGCGCGATCGCAATCTGGCTCCCAGTT	AGAGAGAGACATGCGC
	NgVP.9-R	AGAGAGAGACATGCGCCCAGCAAAGAAAGCTCCAGCCATTA	
10	NgVP.10-F	TGTCACTCATCTGAGTGATCGCAATCTGGCTCCCAGTT	TGTCACTCATCTGAGT

	NgVP.10-R	TGTCACTCATCTGAGTCCAGCAAAGAAAGCTCCAGCCATTA	

4.2 METHODS

4.2.1 General molecular biology:

4.2.1.1.1 Reverse transcription (RT):

cDNA of GII.4 P domain was generated from viral RNA, extracted from stool samples of infected patients using super script III (Invitrogen). Briefly, viral RNA was first linearized by treatment at 65 ºC for 10 min. cDNA was synthesized at 50 ºC for 45 min in a 10 µl reaction mixture including 2 µl 5X buffer, 1 µl mDDT, 1 µl dNTPs, 1 µl 3' reverse primer, 0.1 µl Superscript III and 5 µl linearized RNA.

4.2.1.1.2 Polymerase chain reaction (PCR):

PCR was carried out to amplify the target genes in 50 µl reaction mixture containing ca. 250 ng of DNA template, 2.5 µl of forward and reverse primers (table 4.8, 9 and 10) , and the Phusion High-fidelity PCR master mix (New England Biolab, Frankfurt, Germany). In a thermal cycler, following an initial denaturation at 98 ºC for 30 sec, the samples were subjected to thirty-five cycles of 10s denaturation at 98 ºC, 30s annealing at 60 ºC, 30s extension at 72 ºC, and a final extension at 72 ºC for 10 min.

4.2.1.1.3 Restriction enzyme digestion:

Restriction enzyme (RE) digestion was performed using Fast Digest Restriction Enzymes (Thermo Fisher Scientific). For this purpose, approximately one µg DNA plasmid was treated at 37 ^{0}C for 30 min with one unit RE in a reaction mixture containing 1x digestion buffer.

4.2.1.1.4 Gel electrophoresis:

PCR products and digested DNA samples were analyzed on 1% agarose gel (in TAE buffer) containing RotiSafe for DNA visualization. Electrophoresis was done at 120 V for 40-60 min and DNA fragments were visualized by UV-excitation (Bio-Rad, Germany). The size of the fragments was determined by comparison with the 1 kb Smart Ladder MW-1700-10 (Qiagen, Germany).

4.2.1.1.5 DNA purification from agarose gel:

After DNA gel electrophoresis, the respective band was cut out with a scalpel and DNA was purified with the Gene JET Gel Extraction Kit (Qiagen, Germany). One ml P1 resuspension buffer (Qiagen, Germany) was added to the gel and was placed at 50 ^{o}C for 10 min until the gel liquified. The P1 containing the amplified product was added to the column and centrifuged for 2 min. The flow through was discarded. After centrifugation 650 µl wash buffer was added. The column was centrifuged at 13000 rpm for 2 mins. The process was repeated with same conditions. Column was then added to new tube. 50 µl elution buffer was added and left for 5 min. After five min the column was centrifuged, and the DNA concentration was measured with Nanodrop (Bio-Rad, Germany).

4.2.1.1.6 Ligation:

In a total of 20 µl reaction mixture, 2 µl of 10x ligase buffer and 1 µl T4 Ligase (Qiagen, Germany) was mixed with linearized pGEX-4t vector (GE healthcare, Germany) and digested PCR product , and ligation was carried out at 4 ^{o}C overnight. The reaction was stopped at 65 ^{o}C for 10 min.

4.2.1.1.7 Transformation of *E. coli* competent cells:

One Shot™ TOP10 Chemically Competent *E. coli* (50 µl) (GE healthcare, Germany) were thawed on ice and mixed with 2,5 µl of the recombination reaction mixture. After 15 min incubation on ice, the cells were heat-shocked at 42 ^{o}C for 1 min. The cells were placed on ice for 2 min; then 500 µl of pre-warmed SOC medium was added. After 1 h incubation at 37 ^{o}C with agitation at 180 rpm, 150 µl of the culture was spread on LB agar supplemented with 100 µg/ml ampicillin. Single colonies of the transformants were selected, after overnight incubation at 37 ^{o}C, for further analysis.

4.2.1.1.8 Plasmid extraction from *E. coli*—Mini-prep:

Plasmid extraction was carried out based on the alkaline lysis methods using Thermo Fisher Scientific Gene Jet Plasmid Mini-prep Kit. Briefly, single cell colonies from overnight cultures were picked and grown at 37 ^{o}C for 16 h in 5 ml LB+antibiotics medium. The cells were harvested at 3500 rpm for 5 min and then resuspended in 250 µl of resuspending solution (25

mM Tris-HCl, 10 mM EDTA, pH 8.0, 100 µg/ml RNase I). The cells were lysed by adding 250 µl of lysis buffer (0.2 N NaOH, 1% (w/v) SDS) followed by 5 min incubation at RT. The lysate was then neutralized with 250 µl of neutralizing buffer (60% (v/v) 5M $KHCO_3$, 11.5 % (v/v) glacial acetic acid, pH 4.8). The mixture was centrifuged at high speed for 10 min at 4 ^{o}C; the supernatant containing plasmids was transferred into the spin column and after washing the pure DNA plasmids were dissolved in the elution buffer (10 mM Tris).

4.2.1.1.9 Control digestion:

Recombinant pGEX-4t plasmids containing GII.4 P domain were checked for insertion by restriction enzyme digestion. RE digestion was carried out at 37 ^{o}C for 45 min in a 25 µl reaction mixture containing 5 µl the plasmids, 1.5 µl RE and 2.5 µl 10X buffer. Digestion patterns were visualized on 1% agarose gel.

4.2.1.2 Sequencing:

Integrity and of the inserts were assured by double-stranded sequencing (GATC Biotech AG/Eurofins genomics). The sequences were evaluated by homology search against our synthetic genes using online alignment tools.

4.2.1.3 Next generation sequencing (NGS):

A total of seven chronically infected patients were processed for NGS. Patient 1, 2, 3 were isolated from Institut für Medizinische Virologie und Epidemiologie der Viruskrankheiten, Universitätsklinikum Tübingen, Tübingen, Germany, and the rest of patient samples were collected at klinikum recht der isar, Technische Universität München. cDNA of GII.4 VP1 was generated from RNA which were extracted from stool samples of acutely infected patients using super script III (Invitrogen). RNA was first linearized by placing it at 65 ºC for 10 min. A total of 10 µl mixture was generated by adding following components, 2 µl 5X buffer, 1 µl mDDT, 1 µl dNTPs, 1 µl reverse primer (table 4.8, 9 and 10), 0.1 µl superscript III and 5 µl linearized RNA. Reverse transcription was run at 50 ºC for 45 min. cDNA was then amplified with barcode-labelled primers as described in table 4.8, 9 and 10. PacBio sequencing was done at GATC, Konstanz, Germany.

4.2.1.3.1 PacBio sequencing:

One microgram of the DNA product was used to construct the PacBio DNA library using the PacBio standard 2 kb template prep protocol. Twenty-seven samples were sequenced on the PacBio RSII platform using 5 SMRT cells. C2 Polymerase was used for the sequencing reaction and 240 minutes movie windows were used for signal detection. Raw sequence data was processed and analyzed through secondary analysis protocols of PacBio's SMRT analysis software (V5.3.0). Reads for each sample were separated by specific barcodes, and circular consensus sequencing (CCS) method was used to ensure lower error rates. The CCS reads making three or greater full passes around the closed loop SMRT bell amplicon were retained (~99% accuracy). Subsequently, CCS read files for every sample were obtained in fastq format.

A consensus sequence was created for each sample using SMRT protocol, where the CCS reads are loaded into the Quiver consensus calling framework. It uses variant caller to call SNPs which automatically detects the chemistry information in the movie and makes the best consensus sequence using Quiver algorithm.

4.2.1.3.2 Quasispecies reconstruction:

CCS reads were clipped at both ends with a PHRED threshold of 30 using InDelFixer software (http://www.cbg.ethz.ch/software/InDelFixer/). The adaptors were removed, and approximate matching regions were found in the reference genome using k-mer matching. Subsequently, each read was pairwise aligned to the reference sequence using a sensitive Smith–Waterman algorithm. The reads were discarded if three or more consecutive N base calls were present and read length was less than 300 bp.

Viral quasispecies were constructed using the assembly tool QuasiRecomb 1.2 (www.cbg.ethz.ch/software/quasirecomb). QuasiRecomb implements a hidden Markov model to infer a viral quasispecies from deep-coverage NGS data using an expectation maximization (EM) algorithm for maximum *a posteriori* (MAP) parameter estimation. QuasiRecomb was employed with flags '-r 1-1653' to construct quasispecies from the entire VP1 coding region, and '-conservative' parameter was not utilized to determine the minor quasispecies and estimate their frequencies.

4.2.1.3.3 Sequence analysis:

The virus quasispecies heterogeneity was evaluated based on two parameters genetic complexity and diversity. Genetic complexity is defined in terms of probabilities of different sequences that can appear at a given time point within a population and is expressed as normalized Shannon entropy (Sn). The genetic complexity was assessed at nucleotide level using the formula:

$Sn = \sum i\varepsilon_{[A, T, C, G,]} - (p_i \ln p_i)/\ln N$, where ***pi*** represents the frequency of a particular nucleotide or deletion ***i*** at this position, and ***N*** is the total number of species within a population (Domingo E, et al 2006; Nishijima et al., 2012). Sn varied from 0 (no complexity) to 1 (maximum complexity).

The genetic diversity was evaluated by three parameters: the mean genetic distance (d), the number of synonymous substitutions per synonymous site (dS) and the number of nonsynonymous substitutions per non-synonymous site (dN). Estimation of d was conducted for all pairwise sequences using DNADIST and PROTDIST modules in PHYLIP Package v3.572 (Felsenstein J. 1993). The resulting values were corrected for alignment length and days, thus representing the genetic distance as substitutions/site/days. The values were scaled up to per 1000 days for better comparability. The calculations for dN, dS and dN/dS ratio for each pairwise sequences were computed using the Nei-Gojobori model. dN/dS >1 indicate sequences are undergoing positive selection whereas dN/dS < 1 suggest a negative selection pressure. In addition, to determine positively selected sites across VP1 region Naive Empirical Bayes (NEB) and Bayes Empirical Bayes (BEB) methods available in the CODEML module of PAML 4.7 software package were employed (Yang Z, 2007). These methods are based on a codon-substitution model where codon triplet is considered as unit of evolution. Sites having posterior probabilities (P>90%) with dN/dS >1 are likely under positive selection.

4.2.1.3.4 Phylogenetic analysis:

We estimated intra-host viral evolutionary (divergence) rate using the Bayesian Markov Chain Monte Carlo (MCMC) method in BEAST 1.8.4. The dominant quasispecies having frequency >1% were analyzed under strict molecular clock. The Hasegawa–Kishono–Yano (HKY) nucleotide substitution model with gamma-distributed rate heterogeneity among sites for each patient was employed. To estimate the NoV growth rate, an exponential-growth coalescent prior was assumed, which is intuitively appealing for viral outbreaks. Simulations were run for at least

10 million generations, and trees were sampled and stored after every 1000 generations. The first 10% (10,000,000) of each MCMC runs were discarded as burn-in. To compute and visualize the sampled trees we used the program FigTree version 1.1.1.

4.2.2 Biochemical Methods:

4.2.2.1 SDS-PAGE:

4.2.2.1.1 Sample preparation:

Separation of proteins was performed by SDS-PAGE. Protein samples were mixed with 4X Buffer (Thermo Fisher Scientific, Germany) and boiled at 95 4 °C for 5 min. Protein sample (21 µl) was mixed with 9 µl 4X buffer (Thermo Fisher Scientific, Germany). Cell debris was pelleted at 14,000 rpm, 4 °C for 10 min. In a new microtube, the supernatant was mixed with reducing SDS loading dye; heated at 95 °C for 10 min and subsequently loaded on SDS-gel.

4.2.2.1.2 Preparation of SDS-polyacrylamide gels:

SDS-polyacrylamide gels were prepared using the Mini Protein apparatus. A 12% resolving gel mixture was prepared as described in protocol. All the components were mixed gently and immediately poured into the assembled gel apparatus. The gel was overload with isopropanol to seal the gel surface, and allowed to polymerize for about 30 min at RT. Then, isopropanol overlay was removed using 3MM Whatman filter paper and 5 % stacking gel solution was added with a comb fitted and allowed to polymerize at RT. The glass plates were assembled into the electrophoresis set; both lower and upper reservoirs were filled with 1x running buffer, the combs were carefully pulled up from the gel and the gel wells were rinsed with 1x running buffer. Protein samples were loaded into the wells and SDS-PAGE was performed at 40 V until the samples reached the separation gel and then raised to 85 V.

SDS-gel preparation:

Table 4.12: Stacking and separating gel.

	Stacking gel 5 %	Separating gel 10 %
Tris 1,5 M (pH 8,8)	0,5 ml	3 ml
10 % SDS	200 µl	80 µl
40 % PAA/Bissau	240 µl	2 ml
10 % APS	15 µl	40 µl
TEMED	2 µl	7 µl
H_2O-VE	1,25 ml	3 ml
Total	2 ml	4 ml

4.2.2.2 Western blot:

After SDS-PAGE, the separated proteins were transferred on a methanol activated PVDF membrane. For this, a sheet of the PVDF membrane and four Whatman papers were soaked in transfer buffer. The transfer sandwich was then assembled from the anode to cathode as follow: 2 layers of Whatman papers, PVDF membrane (Sigma Aldrich, Darmstadt, Germany), SDS gel, 2 layers of Whatman papers. Blotting was performed at 200 mA for two hours. Afterwards, the membrane was blocked for one hour with 5 % milk powder in TBST with gentle agitation to saturate unoccupied regions on the membrane. membrane was then incubated with a primary antibody in TBST+5% milk powder at 4 °C overnight. The next day, the membrane was washed 3 times with TBST buffer, each for 10 min, before incubation with the secondary antibody (in TBST+5% milk powder) for one hour at RT. After washing with TBST buffer (3x, 10 min), ECL substrate was added to the membrane and analysis was performed with the Fusion chemiluminescence detector Fx7.

4.2.2.3 Quantitative ELISA:

P domain of HuNoV GII.1 was diluted in PBS with the final concentration of 10 µg/ml. Wells of NUNC Maxisorp ELISA plates (Thermo Fisher Scientific, Frankfurt, Germany) were coated with 100 µl (10 µg/ml) P domain capsid protein for 2 hours at room temperature (24^0C). Human

IgG was used as standard on the same plate and was diluted in PBS. The standard concentration starts with 1000 ng/ml and continues with 1:2 dilutions to 3.9 ng/ml. The wells were washed 3 times with washing buffer (PBS pH 7.2, 0.05% Tween 20). The plates were blocked with 250 µl 5% BSA (bovine serum albumin) for 2 hours at room temperature. The plates were washed 3 times with washing buffer (PBS pH 7.2, 0.05% Tween 20). 100 µl of 1:250 and 1:1000 (diluted in PBS) of the serum samples before and after NoV infection were used as primary antibody. The plates were placed for 1 hour at room temperature and were washed 5 times with washing buffer (PBS pH 7.2, 0.05% Tween 20). Hundred µl of 1:40,000 diluted in PBS protein A HRP (Thermo Fisher Scientific, Frankfurt, Germany) were added to each well and were placed at room temperature for 1 hour. The plates were washed 5 times with washing buffer (PBS pH 7.2, 0.05% Tween 20). Hundred µl of TMB dye (Sigma Aldrich, Darmstadt, Germany) were added to each well at room temperature for 5 to 10 min. Hundred µl of H_2So_4 was added to each well as stop solution. Three negative controls were taken into consideration (i) without coating (ii) without Primary antibody (iii) without secondary antibody and 1:250 polyclonal antibody was used as positive control. The signals were measured in infinite F200 (Tecan, Germany) and absorbance of 450 nm (OD_{450}) was taken into consideration.

4.2.2.3.1 Pre- and post-immunization of capsid protein in mouse:

6-week old female BALB/c mice were immunized three times with 30 µg P domain GII.1 and 10 µg C-di-AMP (as adjuvant) intraperitoneally with one-week intervals. Sera were collected one week before the first immunisation and seven weeks after the last immunisation and used, respectively, as negative and positive controls. The precision and reproducibility of the assay were tested with multiple inter and intra assays. The standard curve was generated using two-fold serial dilution of human IgG (Sigma Aldrich, Germany) ranging from 10^3- to 3.0 ng/ml.

4.2.3 Molecular Dynamics (MD):

4.2.3.1 Structure modelling, refinement and quality analysis:

Structure models of the sequences were built using I-TASSER 5.0 server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/). The X-ray crystal structure of GII.4 VP1 Protein Data Bank (PDB) accession no. “5KON” was used as a template for generating homology models. Each structural model was optimized using two energy minimization servers

Modrefiner and Chiron [https://zhanglab.ccmb.med.umich.edu/ModRefiner/]. Quality of the models was evaluated by PROSA [https://prosa.services.came.sbg.ac.at/prosa.php] and Ramachandran plots were generated using the PROCHECK server [https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/].

4.2.3.2 Preparation protein structures:

The homology model for HSPB1 was further refined using Modrefiner, which optimizes atomic details like turning side chains into favorable positions and removes steric clashes. Finally, both refined structure models using the GROMACS 4.5.5 (Van Der Spoel et al., 2005). The CHARMM27 force field is adopted for the HSPB1 protein using GROMACS software.

4.2.3.3 Solvating and energy minimization:

The protein structures were engaged in a cubic water box using the explicit TIP3 water model at a buffering distance of 1.2 nm. These complexes were subjected to minimization procedure with the Steepest Descent method for 2000 steps. All simulations were run in periodic boundary conditions with the NPT ensemble. The Berendsen coupling algorithm was used for this process, and the temperature was kept at a constant 300K with pressure at 1 bar.

4.2.3.4 Dynamics simulations:

Different epitope sites were predicted and taken into account for molecular dynamic simulation i.e. P2E region (containing epitope E and D and neighboring sites), P2C (containing HBAG binding domain and neighboring hypervariable sites), P2D (a hypervariable region next to the epitope A) and P1C (a antigenic epitope in the P1 domain).

The Particle-Mesh Ewald (PME) method is used to calculate electrostatic interactions with an interpolation of order 4 and grid spacing of 0.12 nm, and all bonds were constrained using the LINCS algorithm. Finally, the time step for all simulations was set at 2fs, and 20ns of MD simulation was performed. An analysis of MD trajectories was done with built-in GROMACS tools (Van Der Spoel et al., 2005). The program g_rms RMSD (Root mean square deviation) is used to analyze the stability of the different simulations. The program g_rmsf, root mean square fluctuation, was used to perform fluctuation analysis on C alpha atoms of each residue from the trajectory file, with the aim of finding whether the mutation affects the dynamic behavior of residues. Finally, the program g gyrate was used to analyze the radius of gyration is performed in

order to understand the levels of compaction of the native and mutant region in protein structure (Van Der Spoel et al., 2005).

4.2.4 Statistical analysis:

The data was analyzed by *t*-test and statistical significance was set at $P < 0.05$. The results were expressed as means ± standard error of mean. All the analysis was carried out using GraphPad Prism 5 and Windows Microsoft Excel 2016.

ABBREVIATIONS

aa	Amino acid
Amp	Ampicillin
bp	Base pair
CLR	Continues Long Read
CV	Coefficient of Variation
DNA	Desoxyribonucleic acid
dN	Non-synonymous mutation
dS	Synonymous mutation
ELISA	Enzyme Linked Immune Sorbent Assay
FUT2	Fucosyltransferase 2 gene
GOI	Gene of interest
HBGA	Histo-blood group antigen
HuNoV	Human norovirus
Kan	Kanamycin
kDa	Kilodalton
MCS	Multiple Cloning Site
MOI	Multiplicity of infection
MD	Molecular Dynamics
NoV	Norovirus
NGS	Next Generation Sequencing
NMR	Nuclear Magnet Resonance
ORF	Open reading frame
P domain	Protruding domain
PCR	Polymerase Chain Reaction
RE	Restriction Enzyme
RFLP	Restriction Fragment Length Pattern
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
Rg	Radius of Gyration

RNA	Ribonucleic acid
SMRT	Single Molecule Real Time Sequencing
TGS	Third Generation Sequencing
S domain	Shell domain
SGS	Single Generation Sequencing
VP1	Major capsid protein
VP2	Minor capsid protein
WT	Wild type
ZMW	Zero Mode Waveguide

5 FIGURES

5. FIGURES

6 TABLES

7 REFERENCES

Ahalawat, N., & Murarka, R. K. (2015). Conformational changes and allosteric communications in human serum albumin due to ligand binding. *Journal of Biomolecular Structure and Dynamics*, *33*(10), 2192–2204. https://doi.org/10.1080/07391102.2014.996609

Ahmed, S. M., Hall, A. A. J., Robinson, A. E., Verhoef, L., Premkumar, P., Parashar, U. U. D., … Mendelman, P. (2014). Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. *The Lancet. Infectious Diseases*, *14*(8), 725–730. https://doi.org/10.1016/S1473-3099(14)70767-4

Allen, D. J., Noad, R., Samuel, D., Gray, J. J., Roy, P., & Iturriza-Gámara, M. (2009). Characterisation of a GII-4 norovirus variant-specific surface-exposed site involved in antibody binding. *Virology Journal*. https://doi.org/10.1186/1743-422X-6-150

Atmar, R. L., & Estes, M. K. (2006). The Epidemiologic and Clinical Importance of Norovirus Infection. *Gastroenterology Clinics of North America*. https://doi.org/10.1016/j.gtc.2006.03.001

Atmar, R. L., Opekun, A. R., Gilger, M. A., Estes, M. K., Crawford, S. E., Neill, F. H., & Graham, D. Y. (2008). Norwalk virus shedding after experimental human infection. *Emerging Infectious Diseases*, *14*(10), 1553–1557. https://doi.org/10.3201/eid1410.080117

Baert, L., Uyttendaele, M., & Debevere, J. (2007). Evaluation of two viral extraction methods for the detection of human noroviruses in shellfish with conventional and real-time reverse transcriptase PCR. *Letters in Applied Microbiology*, *44*(1), 106–111. https://doi.org/10.1111/j.1472-765X.2006.02047.x

Benson, N. C., & Daggett, V. (2008). Dynameomics: large-scale assessment of native protein flexibility. *Protein Sci*. https://doi.org/ps.037473.108 [pii]\r10.1110/ps.037473.108

Bernstein, D. I., Atmar, R. L., Lyon, G. M., Treanor, J. J., Chen, W. H., Jiang, X., … Mendelman, P. M. (2015). Norovirus vaccine against experimental human GII.4 virus illness: A challenge study in healthy adults. *Journal of Infectious Diseases*, *211*(6), 870–878. https://doi.org/10.1093/infdis/jiu497

Boon, D., Mahar, J. E., Abente, E. J., Kirkwood, C. D., Purcell, R. H., Kapikian, A. Z., … Bok, K. (2011). Comparative evolution of GII.3 and GII.4 norovirus over a 31-year period.

Journal of Virology, *85*(17), 8656–8666. https://doi.org/10.1128/JVI.00472-11

Boon, D., Mahar, J. E., Abente, E. J., Kirkwood, C. D., Purcell, R. H., Kapikian, A. Z., … Bok, K. (2011). Comparative Evolution of GII.3 and GII.4 Norovirus over a 31-Year Period. *Journal of Virology*, *85*(17), 8656–8666. https://doi.org/10.1128/JVI.00472-11

Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., … Karplus, M. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, *30*(10), 1545–1614. https://doi.org/10.1002/jcc.21287

Brown, S. D., Nagaraju, S., Utturkar, S., De Tissera, S., Segovia, S., Mitchell, W., … Köpke, M. (2014). Comparison of single-molecule sequencing and hybrid approaches for finishing the genome of Clostridium autoethanogenum and analysis of CRISPR systems in industrial relevant Clostridia. *Biotechnology for Biofuels*, *7*(1). https://doi.org/10.1186/1754-6834-7-40

Bull, R. A., Eden, J. S., Rawlinson, W. D., & White, P. A. (2010). Rapid evolution of pandemic noroviruses of the GII.4 lineage. *PLoS Pathogens*, *6*(3). https://doi.org/10.1371/journal.ppat.1000831

Bull, R. A., Luciani, F., McElroy, K., Gaudieri, S., Pham, S. T., Chopra, A., … Lloyd, A. R. (2011). Sequential bottlenecks drive viral evolution in early acute hepatitis c virus infection. *PLoS Pathogens*. https://doi.org/10.1371/journal.ppat.1002243

Bull, R. A., Tu, E. T. V, McIver, C. J., Rawlinson, W. D., & White, P. A. (2006). Emergence of a new norovirus genotype II.4 variant associated with global outbreaks of gastroenteritis. *Journal of Clinical Microbiology*, *44*(2), 327–333. https://doi.org/10.1128/JCM.44.2.327-333.2006

Butot, S., Le Guyader, F. S., Krol, J., Putallaz, T., Amoroso, R., & Sánchez, G. (2010). Evaluation of various real-time RT-PCR assays for the detection and quantitation of human norovirus. *Journal of Virological Methods*, *167*(1), 90–94. https://doi.org/10.1016/j.jviromet.2010.03.018

Caddy, S., Breiman, A., le Pendu, J., & Goodfellow, I. (2014). Genogroup IV and VI canine noroviruses interact with histo-blood group antigens. *Journal of Virology*, *88*(18), 10377–10391. https://doi.org/10.1128/JVI.01008-14

Caddy, S. L., De Rougemont, A., Emmott, E., El-Attar, L., Mitchell, J. A., Hollinshead, M., … Goodfellow, I. (2015). Evidence for human norovirus infection of dogs in the United

Kingdom. *Journal of Clinical Microbiology*, *53*(6), 1873–1883. https://doi.org/10.1128/JCM.02778-14

Chachu, K. A., Strong, D. W., LoBue, A. D., Wobus, C. E., Baric, R. S., & Virgin, H. W. (2008). Antibody Is Critical for the Clearance of Murine Norovirus Infection. *Journal of Virology*, *82*(13), 6610–6617. https://doi.org/10.1128/JVI.00141-08

Chang, K.-O., & George, D. W. (2007). Interferons and Ribavirin Effectively Inhibit Norwalk Virus Replication in Replicon-Bearing Cells. *Journal of Virology*, *81*(22), 12111–12118. https://doi.org/10.1128/JVI.00560-07

Cheetham, S., Souza, M., Meulia, T., Grimes, S., Han, M. G., & Saif, L. J. (2006). Pathogenesis of a Genogroup II Human Norovirus in Gnotobiotic Pigs. *Journal of Virology*, *80*(21), 10372–10381. https://doi.org/10.1128/JVI.00809-06

Choi, J.-M., Hutson, A. M., Estes, M. K., & Prasad, B. V. V. (2008). Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. *Proceedings of the National Academy of Sciences*, *105*(27), 9175–9180. https://doi.org/10.1073/pnas.0803275105

Christen, M., Hünenberger, P. H., Bakowies, D., Baron, R., Bürgi, R., Geerke, D. P., … Van Gunsteren, W. F. (2005). The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry*. https://doi.org/10.1002/jcc.20303

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., … Kollman, P. A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society*, *117*(19), 5179–5197. https://doi.org/10.1021/ja00124a002

Costantini, V., Grenz, L. D., Fritzinger, A., Lewis, D., Biggs, C., Hale, A., & Vinjé, J. (2010). Diagnostic accuracy and analytical sensitivity of IDEIA norovirus assay for routine screening of human norovirus. *Journal of Clinical Microbiology*, *48*(8), 2770–2778. https://doi.org/10.1128/JCM.00654-10

Costantini, V., Morantz, E. K., Browne, H., Ettayebi, K., Zeng, X. L., Atmar, R. L., … Vinjé, J. (2018). Human norovirus replication in human intestinal enteroids as model to evaluate virus inactivation. *Emerging Infectious Diseases*. https://doi.org/10.3201/eid2408.180126

Currier, R. L., Payne, D. C., Staat, M. A., Selvarangan, R., Shirley, S. H., Halasa, N., … Morrow, A. L. (2015). Innate susceptibility to norovirus infections influenced by FUT2

genotype in a United States pediatric population. *Clinical Infectious Diseases*, *60*(11), 1631–1638. https://doi.org/10.1093/cid/civ165

de Graaf, M., van Beek, J., & Koopmans, M. P. G. (2016). Human norovirus transmission and evolution in a changing world. *Nature Reviews Microbiology*, *14*(7), 421–433. https://doi.org/10.1038/nrmicro.2016.48

de Graaf, M., van Beek, J., Vennema, H., Podkolzin, A., Hewitt, J., Bucardo, F., … Koopmans, M. (2015). Emergence of a novel GII.17 norovirus – End of the GII.4 era? *Eurosurveillance*, *20*(26), 21178. https://doi.org/10.2807/1560-7917.ES2015.20.26.21178

de Rougemont, A., Ruvoen-Clouet, N., Simon, B., Estienney, M., Elie-Caille, C., Aho, S., … Belliot, G. (2011). Qualitative and Quantitative Analysis of the Binding of GII.4 Norovirus Variants onto Human Blood Group Antigens. *Journal of Virology*. https://doi.org/10.1128/JVI.02077-10

Debbink, K., Lindesmith, L. C. & Baric, R. S. (2014). The state of norovirus vaccines. *Clin. Infect. Dis.*, *58*, 1746–1752.

Debbink, K., Lindesmith, L. C., Donaldson, E. F., & Baric, R. S. (2012). Norovirus Immunity and the Great Escape. *PLoS Pathogens*. https://doi.org/10.1371/journal.ppat.1002921

Debbink, K., Lindesmith, L. C., Ferris, M. T., Swanstrom, J., Beltramello, M., Corti, D., … Baric, R. S. (2014). Within-Host Evolution Results in Antigenically Distinct GII.4 Noroviruses. *Journal of Virology*, *88*(13), 7244–7255. https://doi.org/10.1128/JVI.00203-14

Desai, R., Hembree, C. D., Handel, A., Matthews, J. E., Dickey, B. W., Mcdonald, S., … Lopman, B. (2012). Severe outcomes are associated with genogroup 2 genotype 4 norovirus outbreaks: A systematic literature review. *Clinical Infectious Diseases*. https://doi.org/10.1093/cid/cis372

Donaldson, E. F., Lindesmith, L. C., LoBue, A. D., & Baric, R. S. (2010a). Viral shape-shifting: norovirus evasion of the human immune system. *Nature Reviews Microbiology*, *8*(3), 231–241. https://doi.org/10.1038/nrmicro2296

Donaldson, E. F., Lindesmith, L. C., LoBue, A. D., & Baric, R. S. (2010b). Viral shape-shifting: norovirus evasion of the human immune system. *Nature Reviews Microbiology*, *8*(3), 231–241. https://doi.org/10.1038/nrmicro2296

Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H., & Shaw, D. E. (2012). Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annual Review of*

Biophysics. https://doi.org/10.1146/annurev-biophys-042910-155245

Dunbar, N. L., Bruggink, L. D., & Marshall, J. A. (2014). Evaluation of the RIDAGENE real-time PCR assay for the detection of GI and GII norovirus. *Diagnostic Microbiology and Infectious Disease*, *79*(3), 317–321. https://doi.org/10.1016/j.diagmicrobio.2014.03.017

Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*. https://doi.org/10.1186/1741-7007-9-71

Eden, J.-S., Tanaka, M. M., Boni, M. F., Rawlinson, W. D., & White, P. A. (2013). Recombination within the pandemic norovirus GII.4 lineage. *Journal of Virology*, *87*(11), 6270–6282. https://doi.org/10.1128/JVI.03464-12

Eden, J.-S., Tanaka, M. M., Boni, M. F., Rawlinson, W. D., & White, P. A. (2013). Recombination within the Pandemic Norovirus GII.4 Lineage. *Journal of Virology*, *87*(11), 6270–6282. https://doi.org/10.1128/JVI.03464-12

Eden, J. S., Bull, R. A., Tu, E., McIver, C. J., Lyon, M. J., Marshall, J. A., … White, P. A. (2010). Norovirus GII.4 variant 2006b caused epidemics of acute gastroenteritis in Australia during 2007 and 2008. *Journal of Clinical Virology*, *49*(4), 265–271. https://doi.org/10.1016/j.jcv.2010.09.001

Eden, J. S., Hewitt, J., Lim, K. L., Boni, M. F., Merif, J., Greening, G., … White, P. A. (2014). The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012. *Virology*, *450–451*, 106–113. https://doi.org/10.1016/j.virol.2013.12.005

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., … Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, *323*(5910), 133–138. https://doi.org/10.1126/science.1162986

Farkas, T., Singh, A., Le Guyader, F. S., La Rosa, G., Saif, L., & McNeal, M. (2015). Multiplex real-time RT-PCR for the simultaneous detection and quantification of GI, GII and GIV noroviruses. *Journal of Virological Methods*, *223*, 109–114. https://doi.org/10.1016/j.jviromet.2015.07.020

Flores, S. (2006). The Database of Macromolecular Motions: new features added at the decade mark. *Nucleic Acids Research*, *34*(90001), D296–D301. https://doi.org/10.1093/nar/gkj046

Florescu, D. F., Hill, L. A., McCartan, M. A., & Grant, W. (2008). Two cases of Norwalk virus enteritis following small bowel transplantation treated with oral human serum immunoglobulin. *Pediatr Transplant*, *12*(3), 372–375. https://doi.org/10.1111/j.1399-

3046.2007.00875.x

Friesema, I. H. M., Vennema, H., Heijne, J. C. M., de Jager, C. M., Teunis, P. F. M., van der Linde, R., … van Duynhoven, Y. T. H. P. (2009). Differences in clinical presentation between norovirus genotypes in nursing homes. *Journal of Clinical Virology*, *46*(4), 341–344. https://doi.org/10.1016/j.jcv.2009.09.010

Fuentes, C., Guix, S., Pérez-Rodriguez, F. J., Fuster, N., Carol, M., Pintó, R. M., & Bosch, A. (2014). Standardized multiplex one-step qRT-PCR for hepatitis A virus, norovirus GI and GII quantification in bivalve mollusks and water. *Food Microbiology*, *40*, 55–63. https://doi.org/10.1016/j.fm.2013.12.003

Gairard-Dory, A. C., Dégot, T., Hirschi, S., Schuller, A., Leclercq, A., Renaud-Picard, B., … Kessler, R. (2014). Clinical usefulness of oral immunoglobulins in lung transplant recipients with norovirus gastroenteritis: A case series. *Transplantation Proceedings*, *46*(10), 3603–3605. https://doi.org/10.1016/j.transproceed.2014.09.095

Gelpi, J., Hospital, A., Goñi, R., & Orozco, M. (2015). Molecular dynamics simulations: advances and applications. *Advances and Applications in Bioinformatics and Chemistry*, 37. https://doi.org/10.2147/AABC.S70333

Goh, C. S., Milburn, D., & Gerstein, M. (2004). Conformational changes associated with protein-protein interactions. *Curr Opin Struct Biol*, *14*(1), 104–109. https://doi.org/10.1016/j.sbi.2004.01.005\rS0959440X04000065 [pii]

Gray, J. J., Kohli, E., Ruggeri, F. M., Vennema, H., Sánchez-Fauquier, A., Schreier, E., … Diedrich, S. M. (2007). European multicenter evaluation of commercial enzyme immunoassays for detecting norovirus antigen in fecal samples. *Clinical and Vaccine Immunology*, *14*(10), 1349–1355. https://doi.org/10.1128/CVI.00214-07

Green, K. Y. (2014). Norovirus infection in immunocompromised hosts. *Clinical Microbiology and Infection*. https://doi.org/10.1111/1469-0691.12761

Hale, A. D., Crawford, S. E., Ciarlet, M., Green, J., Gallimore, C., Brown, D. W., … Estes, M. K. (1999). Expression and self-assembly of Grimsby virus: antigenic distinction from Norwalk and Mexico viruses. *Clinical and Diagnostic Laboratory Immunology*, *6*(1), 142–145.

Hall, A. J., Glass, R. I., & Parashar, U. D. (2016). New insights into the global burden of noroviruses and opportunities for prevention. *Expert Review of Vaccines*, *15*(8), 949–951.

https://doi.org/10.1080/14760584.2016.1178069

Hall, A. J., Lopman, B. A., Payne, D. C., Patel, M. M., Gastañaduy, P. A., Vinjé, J., & Parashar, U. D. (2013). Norovirus disease in the united states. *Emerging Infectious Diseases*, *19*(8), 1198–1205. https://doi.org/10.3201/eid1908.130465

Hasing, M. E., Hazes, B., Lee, B. E., Preiksaitis, J. K., & Pang, X. L. (2016). A next generation sequencing-based method to study the intra-host genetic diversity of norovirus in patients with acute and chronic infection. *BMC Genomics*. https://doi.org/10.1186/s12864-016-2831-y

Havelaar, A. H., Kirk, M. D., Torgerson, P. R., Gibb, H. J., Hald, T., Lake, R. J., … Zeilmaker, M. (2015). World Health Organization Global Estimates and Regional Comparisons of the Burden of Foodborne Disease in 2010. *PLoS Medicine*. https://doi.org/10.1371/journal.pmed.1001923

Heberlé, G., & de Azevedo, W. F. (2011). Bio-inspired algorithms applied to molecular docking simulations. *Current Medicinal Chemistry*, *18*(9), 1339–1352. https://doi.org/10.2174/092986711795029573

Hoa Tran, T. N., Trainor, E., Nakagomi, T., Cunliffe, N. A., & Nakagomi, O. (2013). Molecular epidemiology of noroviruses associated with acute sporadic gastroenteritis in children: Global distribution of genogroups, genotypes and GII.4 variants. *Journal of Clinical Virology*. https://doi.org/10.1016/j.jcv.2012.11.011

Hoffmann, D., Hutzenthaler, M., Seebach, J., Panning, M., Umgelter, A., Menzel, H., … Metzler, D. (2012). Norovirus GII.4 and GII.7 capsid sequences undergo positive selection in chronically infected patients. *Infection, Genetics and Evolution*, *12*(2), 461–466. https://doi.org/10.1016/j.meegid.2012.01.020

Hsu, C. C., Riley, L. K., Wills, H. M., & Livingston, R. S. (2006). Persistent infection with and serologic cross-reactivity of three novel murine noroviruses. *Comparative Medicine*, *56*(4), 247–251.

Huang, P., Farkas, T., Marionneau, S., Zhong, W., Ruvoën-Clouet, N., Morrow, A. L., … Jiang, X. (2003). Noroviruses Bind to Human ABO, Lewis, and Secretor Histo–Blood Group Antigens: Identification of 4 Distinct Strain-Specific Patterns. *The Journal of Infectious Diseases*, *188*(1), 19–31. https://doi.org/10.1086/375742

Huhti, L., Szakal, E. D., Puustinen, L., Salminen, M., Huhtala, H., Valve, O., … Vesikari, T.

(2011). Norovirus GII-4 causes a more severe gastroenteritis than other noroviruses in young children. *Journal of Infectious Diseases*, *203*(10), 1442–1444. https://doi.org/10.1093/infdis/jir039

Hutson, A. M., Atmar, R. L., Graham, D. Y., & Estes, M. K. (2002). Norwalk virus infection and disease is associated with ABO histo-blood group type. *The Journal of Infectious Diseases*, *185*, 1335–1337. https://doi.org/10.1086/339883

Hutson, A. M., Atmar, R. L., Marcus, D. M., & Estes, M. K. (2003). Norwalk virus-like particle hemagglutination by binding to h histo-blood group antigens. *Journal of Virology*, *77*(1), 405–415. https://doi.org/10.1128/JVI.77.1.405-415.2003

Hyun, J., Kim, H. S., Kim, H.-S., & Lee, K. M. (2014). Evaluation of a new real-time reverse transcription polymerase chain reaction assay for detection of norovirus in fecal specimens. *Diagnostic Microbiology and Infectious Disease*, *78*(1), 40–44. https://doi.org/10.1016/j.diagmicrobio.2013.09.013

Hyun, J., Kim, H. S., & Lee, K. M. (2014). Evaluation of a new real-time reverse transcription polymerase chain reaction assay for detection of norovirus in fecal specimens. *Diagn Microbiol Infect Dis*, *78*(1), 40–44. https://doi.org/10.1016/j.diagmicrobio.2013.09.013

Iritani, N., Vennema, H., Siebenga, J. J., Siezen, R. J., Renckens, B., Seto, Y., … Koopmans, M. (2008). Genetic Analysis of the Capsid Gene of Genotype GII.2 Noroviruses. *Journal of Virology*, *82*(15), 7336–7345. https://doi.org/10.1128/JVI.02371-07

Iversen, P. W., Beck, B., Chen, Y.-F., Dere, W., Devanarayan, V., Eastwood, B. J., … Sittampalam, G. S. (2004). HTS Assay Validation. *Assay Guidance Manual*.

Jones, M. K., Watanabe, M., Zhu, S., Graves, C. L., Keyes, L. R., Grau, K. R., … Karst, S. M. (2014). Enteric bacteria promote human and mouse norovirus infection of B cells. *Science*, *346*(6210), 755–759. https://doi.org/10.1126/science.1257147

Jung, K., Wang, Q., Kim, Y., Scheuer, K., Zhang, Z., Shen, Q., … Saif, L. J. (2012). The effects of simvastatin or interferon-α on infectivity of human norovirus using a gnotobiotic pig model for the study of antivirals. *PLoS ONE*, *7*(7). https://doi.org/10.1371/journal.pone.0041619

Kageyama, T., Kojima, S., Shinohara, M., Uchida, K., Fukushi, S., Hoshino, F. B., … Katayama, K. (2003). Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *J Clin Microbiol*, *41*(4), 1548–1557.

https://doi.org/10.1128/JCM.41.4.1548

Kalko, S. G., Gelpi, J. L., Fita, I., & Orozco, M. (2001). Theoretical study of the mechanisms of substrate recognition by catalase. *J Am Chem Soc*, *123*(39), 9665–9672. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11572688

Kapikian, A. Z., Wyatt, R. G., Dolin, R., Thornhill, T. S., Kalica, A. R., & Chanock, R. M. (1972). Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. *Journal of Virology*, *10*(5), 1075–1081. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4117963

Karst, S. M., Wobus, C. E., Lay, M., Davidson, J., & Virgin IV, H. W. (2003). STAT1-dependent innate immunity to a norwalk-like virus. *Science*, *299*(5612), 1575–1578. https://doi.org/10.1126/science.1077905

Katayama, K., Shirato-Horikoshi, H., Kojima, S., Kageyama, T., Oka, T., Hoshino, F. B., … Takeda, N. (2002). Phylogenetic analysis of the complete genome of 18 norwalk-like viruses. *Virology*, *299*(2), 225–239. https://doi.org/10.1006/viro.2002.1568

Kobayashi, M., Matsushima, Y., Motoya, T., Sakon, N., Shigemoto, N., Okamoto-Nakagawa, R., … Kimura, H. (2016). Molecular evolution of the capsid gene in human norovirus genogroup II. *Scientific Reports*, *6*(July), 29400. https://doi.org/10.1038/srep29400

Kocher, J., Bui, T., Giri-Rachman, E., Wen, K., Li, G., Yang, X., … Yuan, L. (2014). Intranasal P Particle Vaccine Provided Partial Cross-Variant Protection against Human GII.4 Norovirus Diarrhea in Gnotobiotic Pigs. *Journal of Virology*, *88*(17), 9728–9743. https://doi.org/10.1128/JVI.01249-14

Kokkinidis, M., Glykos, N. M., & Fadouloglou, V. E. (2012). Protein flexibility and enzymatic catalysis. *Advances in Protein Chemistry and Structural Biology*, *87*, 181–218. https://doi.org/10.1016/B978-0-12-398312-1.00007-X

Kozakov, D., Clodfelter, K. H., Vajda, S., & Camacho, C. J. (2005). Optimal clustering for detecting near-native conformations in protein docking. *Biophysical Journal*, *89*(2), 867–875. https://doi.org/10.1529/biophysj.104.058768

Kroneman, A., Harris, J., Vennema, H., Duizer, E., Van Duynhoven, Y., Gray, J., … Koopmans, M. (2008). Data quality of 5 years of central norovirus outbreak reporting in the European Network for food-borne viruses. *Journal of Public Health*, *30*(1), 82–90.

https://doi.org/10.1093/pubmed/fdm080

Kumar, A., Rajendran, V., Sethumadhavan, R., & Purohit, R. (2013). Molecular Dynamic Simulation Reveals Damaging Impact of RAC1 F28L Mutation in the Switch I Region. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0077453

Kumar, C. V., Swetha, R. G., Anbarasu, A., & Ramaiah, S. (2014). Computational analysis reveals the association of threonine 118 methionine mutation in PMP22 resulting in CMT-1A. *Advances in Bioinformatics*, *2014*. https://doi.org/10.1155/2014/502618

Le Guyader, F. S., Atmar, R. L., & Le Pendu, J. (2012). Transmission of viruses through shellfish: When specific ligands come into play. *Current Opinion in Virology*, *2*(1), 103–110. https://doi.org/10.1016/j.coviro.2011.10.029

Le Pendu, J., Ruvoën-Clouet, N., Kindberg, E., & Svensson, L. (2006). Mendelian resistance to human norovirus infections. *Seminars in Immunology*. https://doi.org/10.1016/j.smim.2006.07.009

Li, L., Shan, T., Wang, C., Cote, C., Kolman, J., Onions, D., … Delwart, E. (2011). The Fecal Viral Flora of California Sea Lions. *Journal of Virology*, *85*(19), 9909–9917. https://doi.org/10.1128/JVI.05026-11

Lindesmith, L. C., Ferris, M. T., Mullan, C. W., Ferreira, J., Debbink, K., Swanstrom, J., … Baric, R. S. (2015). Broad Blockade Antibody Responses in Human Volunteers after Immunization with a Multivalent Norovirus VLP Candidate Vaccine: Immunological Analyses from a Phase I Clinical Trial. *PLoS Medicine*, *12*(3). https://doi.org/10.1371/journal.pmed.1001807

Lindesmith, L., Moe, C., Marionneau, S., Ruvoen, N., Jiang, X., Lindblad, L., … Baric, R. (2003). Human susceptibility and resistance to Norwalk virus infection. *Nature Medicine*, *9*(5), 548–553. https://doi.org/10.1038/nm860

Lobanov, M. I., Bogatyreva, N. S., & Galzitskaia, O. V. (2008). Radius of gyration is indicator of compactness of protein structure. *Molekuliarnaia Biologiia*. https://doi.org/10.1134/S0026893308040195

Lysén, M., Thorhagen, M., Brytting, M., Hjertqvist, M., Andersson, Y., & Hedlund, K. O. (2009). Genetic diversity among food-borne and waterborne norovirus strains causing outbreaks in Sweden. *Journal of Clinical Microbiology*, *47*(8), 2411–2418. https://doi.org/10.1128/JCM.02168-08

Maiorov, V. N., & Crippen, G. M. (1994). Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *Journal of Molecular Biology*. https://doi.org/10.1006/jmbi.1994.1017

Malm, M., Uusi-Kerttula, H., Vesikari, T., & Blazevic, V. (2014). High serum levels of norovirus genotype-specific blocking antibodies correlate with protection from infection in children. *Journal of Infectious Diseases*, *210*(11), 1755–1762. https://doi.org/10.1093/infdis/jiu361

Marionneau, S., Ruvoën, N., Le MoullacVaidye, B., Clement, M., CailleauThomas, A., RuizPalacois, G., … Le Pendu, J. (2002). Norwalk Virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals. *Gastroenterology*, *122*(7), 1967–1977. https://doi.org/10.1053/gast.2002.33661

Mattison, K., Grudeski, E., Auk, B., Brassard, J., Charest, H., Dust, K., … Booth, T. F. (2011). Analytical performance of norovirus real-time RT-PCR detection protocols in Canadian laboratories. *Journal of Clinical Virology*, *50*(2), 109–113. https://doi.org/10.1016/j.jcv.2010.10.008

Mattison, K., Shukla, A., Cook, A., Pollari, F., Friendship, R., Kelton, D., … Farber, J. M. (2007). Human noroviruses in swine and cattle. *Emerging Infectious Diseases*, *13*(8), 1184–1188. https://doi.org/10.3201/eid1308.070005

McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). Dynamics of folded proteins. *Nature*, *267*(5612), 585–590. https://doi.org/10.1038/267585a0

Meeroff, J. C., Schreiber, D. S., Trier, J. S., & Blacklow, N. R. (1980). Abnormal gastric motor function in viral gastroenteritis. *Annals of Internal Medicine*, *92*(3), 370–373.

Micheletti, C. (2013). Comparing proteins by their internal dynamics: Exploring structure-function relationships beyond static structural alignments. *Physics of Life Reviews*. https://doi.org/10.1016/j.plrev.2012.10.009

Milbrath, M. O., Spicknall, I. H., Zelner, J. L., Moe, C. L., & Eisenberg, J. N. S. (2013). Heterogeneity in norovirus shedding duration affects community risk. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268813000496

Mukherjee, S., Balius, T. E., & Rizzo, R. C. (2010). Docking validation resources: Protein family and ligand flexibility experiments. *Journal of Chemical Information and Modeling*, *50*(11), 1986–2000. https://doi.org/10.1021/ci1001982

Murata, T., Katsushima, N., Mizuta, K., Muraki, Y., Hongo, S., & Matsuzaki, Y. (2007). Prolonged Norovirus Shedding in Infants ???6 Months of Age With Gastroenteritis. *The Pediatric Infectious Disease Journal*, *26*(1), 46–49. https://doi.org/10.1097/01.inf.0000247102.04997.e0

Nasheri, N., Petronella, N., Ronholm, J., Bidawid, S., & Corneau, N. (2017). Characterization of the genomic diversity of norovirus in linked patients using a metagenomic deep sequencing approach. *Frontiers in Microbiology*. https://doi.org/10.3389/fmicb.2017.00073

Neesanant, P., Sirinarumitr, T., Chantakru, S., Boonyaprakob, U., Chuwongkomon, K., Bodhidatta, L., … Mason, C. J. (2013). Optimization of one-step real-time reverse transcription-polymerase chain reaction assays for norovirus detection and molecular epidemiology of noroviruses in thailand. *Journal of Virological Methods*, *194*(1–2), 317–325. https://doi.org/10.1016/j.jviromet.2013.08.033

Nice, T. J., Baldridge, M. T., McCune, B. T., Norman, J. M., Lazear, H. M., Artyomov, M., … Virgin, H. W. (2015). Interferon-λ cures persistent murine norovirus infection in the absence of adaptive immunity. *Science*, *347*(6219), 269–273. https://doi.org/10.1126/science.1258100

Pierre Frange Fabien Touzot Marianne Debré Sébastien Héritier Marianne Leruez-Ville Guilhem Cros Christine Rouzioux Stéphane Blanche Alain Fischer Véronique Avettand-Fenoël. (2012). Prevalence and Clinical Impact of Norovirus Fecal Shedding in Children with Inherited Immune Deficiencies. *J. Infect. Dis.*, *206*, 1269–1274. Retrieved from https://academic.oup.com/jid/article/206/8/1269/858142/Prevalence-and-Clinical-Impact-of-Norovirus-Fecal

Ponsel, D., & Bruss, V. (2003). Mapping of amino acid side chains on the surface of hepatitis B virus capsids required for envelopment and virion formation. *Journal of Virology*. https://doi.org/10.1128/JVI.77.1.416

Prasad, B. V, Hardy, M. E., Dokland, T., Bella, J., Rossmann, M. G., & Estes, M. K. (1999). X-ray crystallographic structure of the Norwalk virus capsid. *Science (New York, N.Y.)*, *286*(5438), 287–290. https://doi.org/10.1126/science.286.5438.287

Ramachandran, S., Campo, D. S., Dimitrova, Z. E., Xia, G. -l., Purdy, M. A., & Khudyakov, Y. E. (2011). Temporal Variations in the Hepatitis C Virus Intrahost Population during Chronic Infection. *Journal of Virology*. https://doi.org/10.1128/JVI.02204-10

Rezaee, R., Poorebrahim, M., Najafi, S., Sadeghi, S., Pourdast, A., Alavian, S. M., … Poortahmasebi, V. (2016). Impacts of the G145R Mutation on the Structure and Immunogenic Activity of the Hepatitis B Surface Antigen: A Computational Analysis. *Hepatitis Monthly*. https://doi.org/10.5812/hepatmon.39097

Rhoads, A., & Au, K. F. (2015). PacBio Sequencing and Its Applications. *Genomics, Proteomics and Bioinformatics*. https://doi.org/10.1016/j.gpb.2015.08.002

Robilotti, E., Deresinski, S., & Pinsky, B. A. (2015). Norovirus. *Clinical Microbiology Reviews*, *28*(1), 134–164. https://doi.org/10.1128/CMR.00075-14

Rockx, B., De Wit, M., Vennema, H., Vinjé, J., De Bruin, E., Van Duynhoven, Y., & Koopmans, M. (2002). Natural history of human calicivirus infection: a prospective cohort study. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *35*(3), 246–253. https://doi.org/10.1086/341408

Rodríguez-Lázaro, D., Cook, N., Ruggeri, F. M., Sellwood, J., Nasser, A., Nascimento, M. S. J., … van der Poel, W. H. M. (2012). Virus hazards from food, water and other contaminated environments. *FEMS Microbiology Reviews*. https://doi.org/10.1111/j.1574-6976.2011.00306.x

Sakon, N., Yamazaki, K., Nakata, K., Kanbayashi, D., Yoda, T., Mantani, M., … Komano, J. (2015). Impact of genotype-specific herd immunity on the circulatory dynamism of norovirus: A 10-year longitudinal study of viral acute gastroenteritis. *Journal of Infectious Diseases*. https://doi.org/10.1093/infdis/jiu496

Salsbury, F. R. (2010). Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. *Current Opinion in Pharmacology*. https://doi.org/10.1016/j.coph.2010.09.016

Sargsyan, K., Grauffel, C., & Lim, C. (2017). How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, *13*(4), 1518–1524. https://doi.org/10.1021/acs.jctc.7b00028

Schadt, E. E., Turner, S., & Kasarskis, A. (2010). A window into third-generation sequencing. *Human Molecular Genetics*, *19*(R2), R227–R240. https://doi.org/10.1093/hmg/ddq416

Schultz, A. C., Vega, E., Dalsgaard, A., Christensen, L. S., Nørrung, B., Hoorfar, J., & Vinjé, J. (2011). Development and evaluation of novel one-step TaqMan realtime RT-PCR assays for the detection and direct genotyping of genogroup I and II noroviruses. *Journal of Clinical*

Virology, *50*(3), 230–234. https://doi.org/10.1016/j.jcv.2010.12.001

Schwartz, S., Vergoulidou, M., Schreier, E., Loddenkemper, C., Reinwald, M., Schmidt-Hieber, M., … Schneider, T. (2011). Norovirus gastroenteritis causes severe and lethal complications after chemotherapy and hematopoietic stem cell transplantation. *Blood*, *117*(22), 5850–5856. https://doi.org/10.1182/blood-2010-12-325886

Shinoda, T., Arai, K., Shigematsu-Iida, M., Ishikura, Y., Tanaka, S., Yamada, T., … Taguchi, H. (2005). Distinct conformation-mediated functions of an active site loop in the catalytic reactions of NAD-dependent D-lactate dehydrogenase and formate dehydrogenase. *Journal of Biological Chemistry*, *280*(17), 17068–17075. https://doi.org/10.1074/jbc.M500970200

Shirato, H., Ogawa, S., Ito, H., Sato, T., Kameyama, A., Narimatsu, H., … Takeda, N. (2008). Noroviruses Distinguish between Type 1 and Type 2 Histo-Blood Group Antigens for Binding. *Journal of Virology*, *82*(21), 10756–10767. https://doi.org/10.1128/JVI.00802-08

Shoichet, B. K., & Kuntz, I. D. (1991). Protein docking and complementarity. *Journal of Molecular Biology*, *221*(1), 327–346. https://doi.org/10.1016/0022-2836(91)80222-G

Siebenga, J. J., Beersma, M. F. C., Vennema, H., van Biezen, P., Hartwig, N. J., & Koopmans, M. (2008). High Prevalence of Prolonged Norovirus Shedding and Illness among Hospitalized Patients: A Model for In Vivo Molecular Evolution. *The Journal of Infectious Diseases*. https://doi.org/10.1086/591627

Siebenga, J. J. et al. (2008). High prevalence of prolonged norovirus shedding and illness among hospitalized patients: a model for in vivo molecular evolution. *J. Infect. Dis.*, *198*, 994–1001.

Siebenga, J. J., Lemey, P., Pond, S. L. K., Rambaut, A., Vennema, H., & Koopmans, M. (2010). Phylodynamic reconstruction reveals norovirus GII.4 epidemic expansions and their molecular determinants. *PLoS Pathogens*, *6*(5), 1–13. https://doi.org/10.1371/journal.ppat.1000884

Siebenga, J. J., Vennema, H., Zheng, D., Vinjé, J., Lee, B. E., Pang, X., … Koopmans, M. (2009). Norovirus Illness Is a Global Problem: Emergence and Spread of Norovirus GII.4 Variants, 2001–2007. *The Journal of Infectious Diseases*, *200*(5), 802–812. https://doi.org/10.1086/605127

Simmons, K., Gambhir, M., Leon, J., & Lopman, B. (2013). Duration of immunity to norovirus gastroenteritis. *Emerging Infectious Diseases*. https://doi.org/10.3201/eid1908.130472

Souza, M., Azevedo, M. S. P., Jung, K., Cheetham, S., & Saif, L. J. (2008). Pathogenesis and Immune Responses in Gnotobiotic Calves after Infection with the Genogroup II.4-HS66 Strain of Human Norovirus. *Journal of Virology*, *82*(4), 1777–1786. https://doi.org/10.1128/JVI.01347-07

Stanley, P. & Cummings, R. D. (2009). Essentials of Glycobiology 2nd edn Ch. 13. In *Cold Spring Harbor Laboratory Press*.

Stevens, R. C., & Lipscomb, W. N. (1990). Allosteric control of quaternary states in E. coli aspartate transcarbamylase. *Biochemical and Biophysical Research Communications*, *171*(3), 1312–1318. https://doi.org/10.1016/0006-291X(90)90829-C

Subba-Reddy, C. V., Goodfellow, I., & Kao, C. C. (2011). VPg-Primed RNA Synthesis of Norovirus RNA-Dependent RNA Polymerases by Using a Novel Cell-Based Assay. *Journal of Virology*, *85*(24), 13027–13037. https://doi.org/10.1128/JVI.06191-11

Subba-Reddy, C. V., Yunus, M. A., Goodfellow, I. G., & Kao, C. C. (2012). Norovirus RNA Synthesis Is Modulated by an Interaction between the Viral RNA-Dependent RNA Polymerase and the Major Capsid Protein, VP1. *Journal of Virology*, *86*(18), 10138–10149. https://doi.org/10.1128/JVI.01208-12

Sukhrie, F. H. A., Siebenga, J. J., Beersma, M. F. C., & Koopmans, M. (2010). Chronic shedders as reservoir for nosocomial transmission of norovirus. *Journal of Clinical Microbiology*, *48*(11), 4303–4305. https://doi.org/10.1128/JCM.01308-10

Sukhrie, F. H. A., Teunis, P., Vennema, H., Copra, C., Thijs Beersma, M. F. C., Bogerman, J., & Koopmans, M. (2012). Nosocomial transmission of norovirus is mainly caused by symptomatic cases. *Clinical Infectious Diseases*, *54*(7), 931–937. https://doi.org/10.1093/cid/cir971

Summa, M., von Bonsdorff, C. H., & Maunula, L. (2012). Pet dogs-A transmission route for human noroviruses? *Journal of Clinical Virology*, *53*(3), 244–247. https://doi.org/10.1016/j.jcv.2011.12.014

Swanstrom, J., Lindesmith, L. C., Donaldson, E. F., Yount, B., & Baric, R. S. (2014). Characterization of Blockade Antibody Responses in GII.2.1976 Snow Mountain Virus-Infected Subjects. *Journal of Virology*, *88*(2), 829–837. https://doi.org/10.1128/JVI.02793-13

Tamminen, K., Lappalainen, S., Huhti, L., Vesikari, T., & Blazevic, V. (2013). Trivalent

Combination Vaccine Induces Broad Heterologous Immune Responses to Norovirus and Rotavirus in Mice. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0070409

Taube, S., Perry, J. W., McGreevy, E., Yetming, K., Perkins, C., Henderson, K., & Wobus, C. E. (2012). Murine noroviruses bind glycolipid and glycoprotein attachment receptors in a strain-dependent manner. *Journal of Virology*, *86*(10), 5584–5593. https://doi.org/10.1128/JVI.06854-11

Thackray, L. B., Wobus, C. E., Chachu, K. A., Liu, B., Alegre, E. R., Henderson, K. S., … Virgin, H. W. (2007). Murine Noroviruses Comprising a Single Genogroup Exhibit Biological Diversity despite Limited Sequence Divergence. *Journal of Virology*, *81*(19), 10460–10473. https://doi.org/10.1128/JVI.00783-07

Thorne, L. G., & Goodfellow, I. G. (2014). Norovirus gene expression and replication. *Journal of General Virology*. https://doi.org/10.1099/vir.0.059634-0

Thorven, M., Grahn, A., Hedlund, K.-O., Johansson, H., Wahlfrid, C., Larson, G., & Svensson, L. (2005). A homozygous nonsense mutation (428G-->A) in the human secretor (FUT2) gene provides resistance to symptomatic norovirus (GGII) infections. *Journal of Virology*, *79*(24), 15351–15355. https://doi.org/10.1128/JVI.79.24.15351-15355.2005

Travers, K. J., Chin, C. S., Rank, D. R., Eid, J. S., & Turner, S. W. (2010). A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic Acids Research*, *38*(15). https://doi.org/10.1093/nar/gkq543

Troeger, H., Loddenkemper, C., Schneider, T., Schreier, E., Epple, H. J., Zeitz, M., … Schulzke, J. D. (2009). Structural and functional changes of the duodenum in human norovirus infection. *Gut*, *58*(8), 1070–1077. https://doi.org/10.1136/gut.2008.160150

Tsai, C. J., & Nussinov, R. (1997). Hydrophobic folding units at protein-protein interfaces: Implications to protein folding and to protein-protein association. *Protein Science*, *6*(7), 1426–1437. https://doi.org/10.1002/pro.5560060707

Tuckerman, M. E., & Martyna, G. J. (2000). Understanding Modern Molecular Dynamics: Techniques and Applications. *The Journal of Physical Chemistry B*, *104*(2), 159–178. https://doi.org/10.1021/jp992433y

van Beek, J., Ambert-Balay, K., Botteldoorn, N., Eden, J. S., Fonager, J., Hewitt, J., … Koopmans, M. (2013). Indications for worldwide increased norovirus activity associated with emergence of a new variant of Genotype II.4, LATE 2012. *Eurosurveillance*.

https://doi.org/10.2807/ese.18.01.20345-en

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*. https://doi.org/10.1002/jcc.20291

Vega, E., Donaldson, E., Huynh, J., Barclay, L., Lopman, B., Baric, R., … Vinje, J. (2014). RNA Populations in Immunocompromised Patients as Reservoirs for Novel Norovirus Variants. *Journal of Virology*. https://doi.org/10.1128/JVI.02494-14

Verhoef, L., Hewitt, J., Barclay, L., Ahmed, S. M., Lake, R., Hall, A. J., … Koopmans, M. (2015). Norovirus genotype profiles associated with foodborne transmission, 1999???2012. *Emerging Infectious Diseases*, *21*(4), 592–599. https://doi.org/10.3201/eid2104.141073

Verhoef, L., Kouyos, R. D., Vennema, H., Kroneman, A., Siebenga, J., van Pelt, W., & Koopmans, M. (2011). An integrated approach to identifying international foodborne norovirus outbreaks. *Emerging Infectious Diseases*, *17*(3), 412–418. https://doi.org/10.3201/eid1703.100979

Vinje, J. (2015). Advances in laboratory methods for detection and typing of norovirus. *Journal of Clinical Microbiology*. https://doi.org/10.1128/JCM.01535-14

Vinje, J. (2014). Advances in Laboratory Methods for Detection and Typing of Norovirus. *J Clin Microbiol*. https://doi.org/10.1128/JCM.01535-14

Vinjé, J. (2015). Advances in laboratory methods for detection and typing of norovirus. *Journal of Clinical Microbiology*. https://doi.org/10.1128/JCM.01084-13

WHO Foodborne Disease Burden Epidemiology Reference Group 2007-2015. (2015). *WHO estimates of the global burden of foodborne diseases*. *Who*. https://doi.org/10.1016/j.fm.2014.07.009

Wikswo, M. (2014). Outbreaks of acute gastroenteritis transmitted by person-To-person contact-United States, 2009-2010. *American Journal of Public Health*, *104*(11), e13–e14. https://doi.org/10.2105/AJPH.2014.302321

Wobus, C. E., Karst, S. M., Thackray, L. B., Chang, K. O., Sosnovtsev, S. V., Belliot, G., … Virgin IV, H. W. (2004). Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages. *PLoS Biology*, *2*(12). https://doi.org/10.1371/journal.pbio.0020432

Woodward, J. M., Gkrania-Klotsas, E., Cordero-Ng, A. Y., Aravinthan, A., Bandoh, B. N., Liu,

H., … Kumararatne, D. (2015). The Role Of Chronic Norovirus Infection In The Enteropathy Associated With Common Variable Immunodeficiency. *The American Journal of Gastroenterology*, *110*(2), 320–327. https://doi.org/10.1038/ajg.2014.432

Wu, Z., Yang, L., Ren, X., He, G., Zhang, J., Yang, J., … Jin, Q. (2016). Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME Journal*, *10*(3), 609–620. https://doi.org/10.1038/ismej.2015.138

Yan, Y., Wang, H., Gao, L., Ji, J., Ge, Z., Zhu, X., … Chen, Z. (2013). A one-step multiplex real-time RT-PCR assay for rapid and simultaneous detection of human norovirus genogroup I, II and IV. *Journal of Virological Methods*, *189*(2), 277–282. https://doi.org/10.1016/j.jviromet.2013.02.004

Zakhour, M., Ruvoën-Clouet, N., Charpilienne, A., Langpap, B., Poncet, D., Peters, T., … Le Pendu, J. (2009). The αGal epitope of the histo-blood group antigen family is a ligand for bovine norovirus Newbury2 expected to prevent cross-species transmission. *PLoS Pathogens*, *5*(7). https://doi.org/10.1371/journal.ppat.1000504

Zakrzewska, K., & Lavery, R. (2012). Towards a molecular view of transcriptional control. *Curr Opin Struct Biol*, *22*(2), 160–167. https://doi.org/10.1016/j.sbi.2012.01.004

Zheng, D. P., Ando, T., Fankhauser, R. L., Beard, R. S., Glass, R. I., & Monroe, S. S. (2006). Norovirus classification and proposed strain nomenclature. *Virology*, *346*(2), 312–323. https://doi.org/10.1016/j.virol.2005.11.015

Zheng, D. P., Widdowson, M. A., Glass, R. I., & Vinjé, J. (2010). Molecular epidemiology of genogroup II-genotype 4 noroviruses in the United States between 1994 and 2006. *Journal of Clinical Microbiology*, *48*(1), 168–177. https://doi.org/10.1128/JCM.01622-09

Zartash et al. 2015 (http://emedicine.medscape.com/article/224225-overview).

PUBLICATIONS AND MEETINGS

Publications in peer review journals:

Suliman Qadir Afridi, Hassan Moeini, Behnam Kalili, Jochen Martin Wettengel, Oliver Quitt, Rapheala Samper Markus Gerhard, Ulrike Protzer, Dieter Hoffmann (2018). *Quantitation of norovirus-specific IgG before and after infection.* Journal of Medical Virology (in revision).

Suliman Qadir Afridi, Zainab Usman, Hassan Moeini, Sainitin Donakonda, Robert Beck, Dimitri Frishman, Percy Knole, Ulrike Protzer, Dieter Hoffmann (2018). *Longitudinal evolution of quasispecies in chronic norovirus patient*s (Manuscript in preparation).

Suliman Qadir Afridi, Hassan Moeini, Sainitin Donakonda, Zainab Usman, Jochen Martin Wettengel, Behnam Kalili, Robert Beck, Markus Gerhard, Percy Knole, Ulrike Protzer, Dieter Hoffmann (2018). *Evolution and immune escape in NoV chronically infected patient* (Manuscript in preparation).

Hassan Moeini, **Suliman Qadir Afridi**, Sainitin Donakonda, Percy Knole, Ulrike Protzer, Dieter Hoffmann (2018). *Genetic diversity and structure evolution of norovirus GII.4 capsid protein* (Manuscript in preparation).

Meetings and summer school:

One-month summer school at university of Alberta from July 1st, 2018 till August 1st, 2018. Edmonton, Canada.

Poster presentation at "11th International Virology and Infectious Diseases Congress" Vancouver, Canada.

ACKNOWLEDGEMENTS

First, I would like to thank Ulrike Protzer (supervisor) for giving me the opportunity to work as PhD in Institute of Virology, TUM. I am glad to work under the supervision of Dieter Hoffmann for letting me work on my desired project. Both of you were one of the best supervisors I ever had, and always provided support when it was needed while giving me freedom to work as independent scientist. Thank you, Markus Gerhard and Sabrina Schreiner, as members of my thesis committee, you gave great input and nice ideas during our committee meetings. Very special thanks go to Hassan Moeini for your support and help. In our weekly meetings, discussing recent results and planning future experiments. Thank you! Jochen Wettengel, Sainitin Donakonda, Stoyan Velkov, Oliver Quitt, Samuel Jeske, Zainab Asrar, Usman Saeed, Behnam kalili for helping me in my experiments. I also would like to thank all other laboratory members, including former members, naming especially Julia Hasreiter, Andreas Oswald, Konstantine, Florian Wilsch, Sebastian Altstetter, Christoph Blossey, Lisa Wolff, Antje Malo, Natalie Röder, Stefanie Prosser, Theresa Asen, Philipp Hagen, Romina Bester and Anna Kosinska, Chunkyu Ko, Martin Mück-Häusl, Daniela Stadler, Anindita Chakraborty, Martin Kächele, Julia Sacherl, Katrin Singethan, Kathrin Kappes, Felix Bohne, Thomas Michler, Lili Zhao, Win Min, Jinpeng Su, Sophia and Maarten van de Klundert. I feel like we became more than colleagues. Although the laboratory tends to be crowded, you kept the atmosphere friendly, supportive, humorous, and at the same time productive. Furthermore, I would like to thank all members of the virology diagnostics department. Thank you to Frank Thiele, Cauleen Noel, Doris Pelz and Daniela Rizzi for your support in office work. I would also like to thank my friends Abdul Moeed and Lena, Zaheer Ahmed, Fazle Azeem, Yasir Akram, Muhammad Shafiq, Zeeshan, Arsalan, Abdullah Yasin, Max Ludemann, Muhammad Yousaf Khan, Asim Farooq, Saifullah, Muhammad Anwar, Rameez Hayat, Sardar Atteq ur Rehman, Habib ur Rehman, Zahid Khan, Javeed, Matten and Farooq for your help and support.

A very important thank goes to my mother, father, my brothers and sisters. Thank you for all the love and support throughout my life, for all thoughts and laughter, what I am today is all because of you.