





Aus dem Institut für Virologie und Immunbiologie  
der Universität Würzburg  
Lehrstuhl für Virologie  
Vorstand: Prof. Dr. med. A. Rethwilm

# **Nachweis von Rekombination innerhalb des Norovirus-Capsidgens**

**Inaugural - Dissertation**

zur Erlangung der Doktorwürde der  
Medizinischen Fakultät der  
Julius-Maximilians-Universität Würzburg

vorgelegt von  
**Julia Eva Maria Münch**  
aus Münster, Nordrhein-Westfalen

Würzburg, Dezember 2006

Referent: Univ.-Prof. Dr. med. Axel Rethwilm

Koreferent: Univ.-Prof. Dr. Klaus Brehm

Dekan: Univ.-Prof. Dr. M. Frosch

Tag der mündlichen Prüfung: 02.09.2008

Die Promovendin ist Ärztin

Für meine Familie.



# INHALTSVERZEICHNIS

<b>1</b>	<b>EINLEITUNG</b>	<b>1</b>
<b>1.1</b>	<b>Das Calicivirus</b>	<b>1</b>
1.1.1	Stämme, Historie	1
1.1.2	Taxonomie: Definition der Familie „Caliciviridae“	2
1.1.3	Wirtsspektrum, Klinik einer Infektion bei verschiedenen Spezies	4
<b>1.2</b>	<b>Die Calicivirus-Gastroenteritis beim Menschen</b>	<b>4</b>
1.2.1	Stabilität, Infektiosität, Transmissionswege	4
1.2.2	Pathogenese, klinisches Bild	5
1.2.3	Immunität	6
1.2.4	Epidemiologie, Bedeutung im Verhältnis zu anderen Erregern	7
<b>1.3</b>	<b>Grundlagen</b>	<b>8</b>
1.3.1	Struktur des Genoms der CV	8
1.3.2	Genomorganisation der NV	10
<b>1.4</b>	<b>Das Capsid</b>	<b>11</b>
<b>1.5</b>	<b>Mutation und genetischer Informationsaustausch zwischen Viren</b>	<b>16</b>
<b>1.6</b>	<b>Aufgabenstellung</b>	<b>19</b>
<b>2</b>	<b>MATERIAL UND METHODEN</b>	<b>20</b>
<b>2.1</b>	<b>Material</b>	<b>20</b>
2.1.1	Stuhlmaterial	20
2.1.2	Geräte	20
2.1.3	Enzyme	21
2.1.4	Kommerzielle Testsysteme	22

2.1.5	Oligonukleotidprimer	22
2.1.6	Kompetente Bakterien und Plasmide	24
2.1.7	Medien, Puffer, Lösungen	24
2.1.7.1	Bakterienkultur	24
2.1.7.2	Plasmidaufreinigung	25
2.1.7.3	Agarose-Gelelektrophorese	26
2.1.7.4	Chemikalien	26
<b>2.2</b>	<b>Molekularbiologische Methoden, Vorgehensweise</b>	<b>27</b>
2.2.1	RNA-Extraktion aus Stuhlproben, QIAmp® DNA Stool Mini Kit	27
2.2.2	Konzentrierung der RNA mit Microcon-Tubes	28
2.2.3	Photometrische Bestimmung von RNA-Konzentration, gesamt-RNA und Reinheit der Probe	29
2.2.4	Amplifikation mittels RT-PCR (SUPERSCRIPT™ One-Step RT-PCR PLATINUM® Taq)	29
2.2.5	Amplifikation mittels nPCR (Herculase® Hotstart DNA-Polymerase)	31
2.2.6	Agarose-Gelelektrophorese zum Nachweis positiver Proben	32
2.2.7	UV-Licht-Foto	33
2.2.8	DNA-Extraktion aus Agarose Gel, QIAEX II Gel Extraction Kit	34
2.2.9	Klonierung und Transformation	35
2.2.10	Anzucht in LB-Medium	37
2.2.11	Aufreinigung mit Qiaprep Miniprep Kit	38
2.2.12	Restriktionsverdau der DNA	39
2.2.13	Sequenzierung	40
2.2.14	Primerwalking	41
<b>2.3</b>	<b>Statistische und phylogenetische Analyseverfahren, Rekombinationsnachweis und strukturelle Charakterisierung von Rekombinationsarealen</b>	<b>42</b>
2.3.1	Überblick über die Vorgehensweise zur Ermittlung von Rekombination	42
2.3.2	Statistische und phylogenetische Analyse	43
2.3.2.1	Multiples Sequenzalignment, CLUSTALX	43

2.3.2.2	Exploratory Tree Analysis (MODELTEST, LRT)	44
2.3.3	Rekombinationsnachweis	46
2.3.3.1	Sawyer's Test, GENECONV	46
2.3.3.2	Splits Tree	47
2.3.3.3	Similarity Plots	48
2.3.3.4	Bruchpunktanalyse mittels LARD	48
2.3.3.5	Phylogenetische Stammbäume und Bootstraps	50
2.3.4	Strukturelle Charakterisierung von Rekombinationsarealen	51
<b>2.4</b>	<b>GenBank Accession Numbers der neu charakterisierten Norovirus Stämme</b>	<b>51</b>
<b>3</b>	<b>ERGEBNISSE</b>	<b>52</b>
<b>3.1</b>	<b>Analyse und Erstellen der DNA-Sequenzen</b>	<b>52</b>
3.1.1	Amplifikation mittels RT-PCR und Dokumentation durch UV-Licht-Photo	52
3.1.2	Amplifikation mittels nPCR und Dokumentation durch UV-Licht-Photo	54
3.1.3	Präparierung der DNA, Klonierung, Transformation, Kultur, Aufreinigung, Sequenzierung, Primerwalking	56
<b>3.2</b>	<b>Alignment und phylogenetische Analyse der Norovirusstämme basierend auf vollständigen Capsidsequenzen</b>	<b>58</b>
3.2.1	Sequenzalignment und Screening auf Rekombinationsereignisse	60
3.2.2	Phylogenetischer Konflikt dargestellt durch SplitsTree	61
3.2.3	Zuordnung zu Parentalsträngen durch SimPlot und Exploratory Tree Analysis	62
3.2.4	Erstellen von ML-Bäumen in Umgebung der Schnittstellen	64
3.2.5	Abschätzen der statistischen Wahrscheinlichkeit für die einzelnen Knotenpunkte durch LARD	65

3.2.6	AS-Sequenzalignment mit Zuordnung der Rekombinationsschnittstellen zu den jeweiligen Capsidregionen	66
<b>3.3</b>	<b>Untersuchung der Sekundärstruktur in der Umgebung der Rekombinationsschnittstelle</b>	<b>67</b>
<b>4</b>	<b>DISKUSSION</b>	<b>69</b>
<b>5</b>	<b>ZUSAMMENFASSUNG</b>	<b>83</b>
<b>6</b>	<b>LITERATUR</b>	<b>85</b>
<b>7</b>	<b>ANHANG</b>	<b>90</b>
7.1	Abbildungsverzeichnis	90
7.2	Tabellenverzeichnis	91
7.3	Abkürzungsverzeichnis	92

# **1 EINLEITUNG**

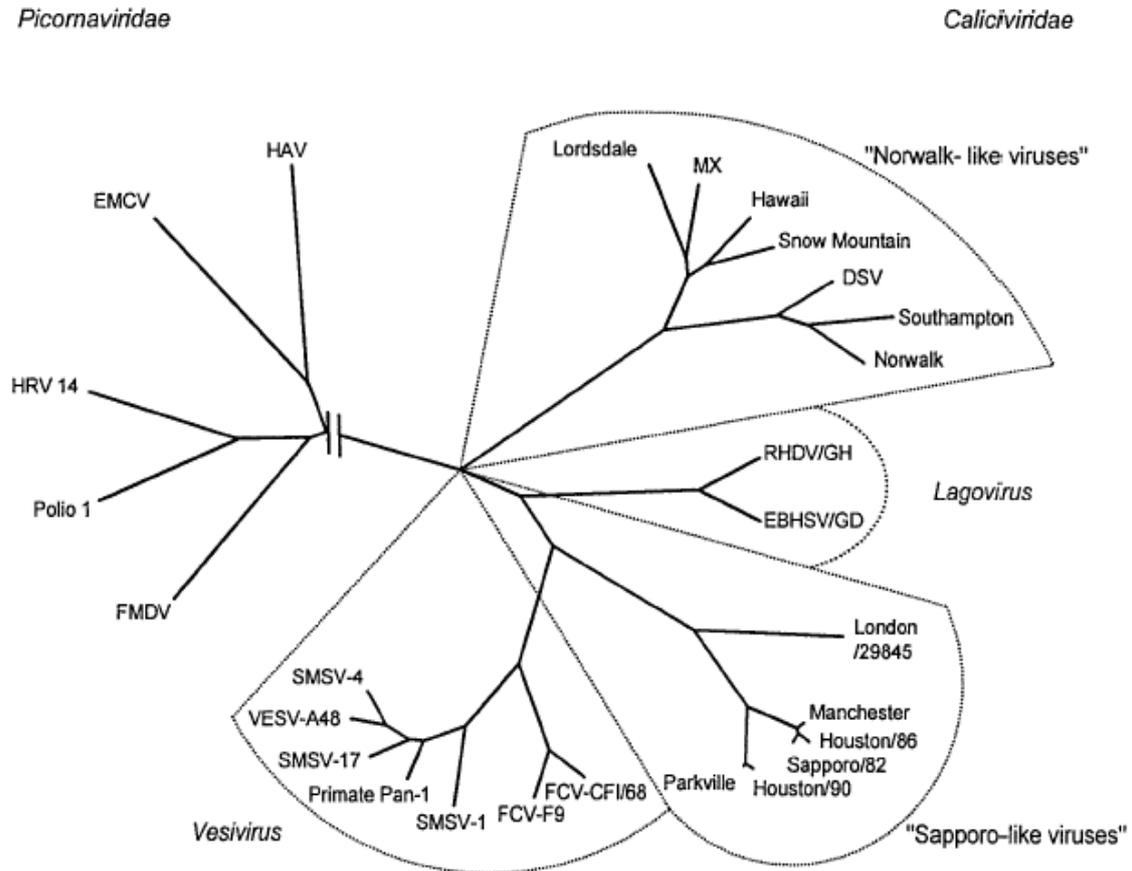
## **1.1 Das Calicivirus**

### **1.1.1 Stämme, Historie**

Caliciviren sind 27 - 38 nm große, nicht umhüllte Viren. Die genetische Information ist im Inneren des Capsids in Form eines 7.2 - 8.3 kb langen, polyadenylierten plus-Strang-RNA-Genoms gespeichert [1]. „Calici“ ist an das lateinische Wort „calix“, Kelch, angelehnt und bezieht sich auf die 32 elektronenmikroskopisch darstellbaren ikosaedrisch angeordneten Strukturen auf der Oberfläche der Viruskapsel [2].

Humane Caliciviren traten zum ersten Mal 1968 bei einem Ausbruch in Norwalk, Ohio, in Erscheinung. Die isolierten Partikel konnten 1972 mittels Immunelektronenmikroskopie (IEM) zwar dargestellt, jedoch noch nicht von Picorna- oder Parvoviren unterschieden werden und wurden daher zunächst den Picornaviren zugeordnet [3]. Im Zuge ihrer weiteren Charakterisierung zeigten sich Unterschiede in Struktur, Replikationsstrategie und physikochemischen Eigenschaften, die 1978 schließlich dazu führten, dass man sie von den Picornaviren trennte [4, 5]. Ab 1979 wurden unter dem Namen „Caliciviridae“ das humanpathogene Norwalk-Virus, der Prototyp der small round structured viruses (SRSV), und das Hepatitis-E-Virus (HEV), sowie die tierpathogenen Vertreter Vesicular-exanthema-of-swine-Virus (VESV), San-Miguel-sea-lion-Virus (SMSV) und das Feline-Calicivirus (FCV) zusammengefasst [6]. Dem VESV war erstmalig 1932 ein Ausbruch in den USA zugeschrieben worden [7] und es konnte 1968 isoliert werden, das FCV wurde 1957 in Neuseeland isoliert [8]. Den humanpathogenen Erregern wurde 1982 das Sapporo-Virus als Prototyp der „klassischen HuCV“ zugeordnet.

## 1.1.2 Taxonomie: Definition der Familie „Caliciviridae”



**Abbildung 1: Maximum likelihood Phylogramm zur Darstellung der Verwandtschaftsverhältnisse innerhalb der Caliciviridae und zu den Picornaviridae (JID 2000 Taxonomy of the Calicivirus).** Aktuelle Nomenklatur: Sapporo-like-Virus = Sapovirus; Norwalk-like-Virus = Norovirus [11],  $p < 0.05$  gilt für jede Distanz des Stammbaums, Grundlage der Analyse ist jeweils die Nukleotidsequenz des gesamten Capsidgens. Die Familie Caliciviridae (umrandet dargestellt) trennt sich phylogenetisch von der Familie Picornaviridae. Die vier Hauptäste (Clades) innerhalb der Caliciviridae bilden die Genera Norovirus, Lagovirus, Sapovirus und Vesivirus. Die drei großen Untergruppen der Picornaviridae (Enteroviren, Hepatoviren, Rhinoviren) werden durch das Hepatitis-A-Virus (HAV), Enzephalomyokarditis-Virus (EMCV), Human-Rhinovirus 14 (HRV 14) und das Foot-and-Mouth-Disease-Virus (FMDV) repräsentiert. MX = Mexicovirus, DSV = Desert-Shield-Virus, RHDV = Rabbit-haemorrhagic-disease-Virus, GH = Germany Hof, EBHSV = European-brown-hare-syndrome-Virus, FCV= Feline-Calicivirus, SMSV = San-Miguel-sea-lion-Virus, VESV = Vesicular-exanthema-of-swine-Virus

Der Beschreibung der molekular-phylogenetischen Verwandtschaftsbeziehungen kommt in Bezug auf die Caliciviren besondere Bedeutung zu, da sie sich durch das Fehlen eines Modells zur Virusanzucht den üblichen Methoden der Klassifizierung entziehen [9]. Eine Einteilung nach Serotyp oder das Herstellen einer Beziehung zwischen Genotyp und Serotyp sind somit unmöglich [10].

1979 trennte das *International Committee on Taxonomy of Viruses* (ICTV) die Caliciviren von den Picornaviren und gründete damit eine neue Virusfamilie, die Caliciviridae [6]. Folgende gemeinsame Charakterzüge innerhalb der Familie führten zu dieser Entscheidung: Vorhandensein eines einzigen großen Strukturproteins, aus dem das Capsid gebildet wird, Formation von 32 kelchförmigen Vertiefungen auf der Oberfläche der Virione, arrangiert in ikosaedrischer Symmetrie und das Fehlen einer methylierten 5'-Ende-Cap-Struktur, an dessen Stelle sich das VPg ( $10\text{-}12 \times 10^3$  kDa) befindet. Dieses ist kovalent mit der RNA verbunden und wird als essentiell für die Infektivität des Virus angesehen [4, 12]. Weitere Unterschiede zu den Picornaviridae bestehen darin, dass diese nur einen großen ORF besitzen und eine andere Reihenfolge in der Abfolge der nicht-strukturellen Proteine zeigen.

Mit der ersten Klonierung und Charakterisierung des RNA-Genoms durch Jiang et al. in den frühen 90er Jahren können RNA- beziehungsweise cDNA-Sequenzen für phylogenetische Analysen herangezogen werden. Dies führte 1998 unter anderem dazu, dass das HEV aus der Gruppe entfernt wurde [13, 14].

Vorläufig ungeachtet des breiten Wirtsspektrums, der Sequenzunterschiede zwischen den Viren der neuen Familie und der unterschiedlichen Antigenität dieser Viren, legte das ICTV 1995 nur ein Genus, das „Calicivirus“ fest. Das ICTV konkretisierte, dass es sich dabei um ein halbwegs vollständiges Genom („*reasonably complete genome*“) oder eine signifikante partielle Sequenz („*significant partial sequence*“) handeln müsse, welche mit zusätzlichen Daten über beispielweise Morphe, antigene Wirkungsweise, Wirtsspektrum und klinische Daten zu verifizieren sei [15]. Sequenzanalysen haben sich inzwischen als Hauptwerkzeug zur Differenzierung zwischen den einzelnen CV

etabliert. Der Signifikanz einer solchen Sequenz kommt bei der Zuordnung eines Stamms zu einer Untergruppe in Besonderen in Hinblick auf die Antigenität eine Bedeutung zu, die im Späteren näher diskutiert wird.

### **1.1.3 Wirtspektrum, Klinik der Infektion bei verschiedenen Spezies**

Wie aus Abbildung 1 ersichtlich, werden heute jeweils zwei tierpathogene und zwei humanpathogene Gruppen unterschieden. CV befallen dieses breite Wirtspektrum mit unterschiedlicher Klinik. Durch Vesiviren können Säugetiere (auch Meeressäuger) aber auch Amphibien, Vögel und Fische erkranken. Das FCV beispielsweise ruft bei jungen Katzen akute respiratorische Beschwerden hervor, das VESV verursacht vesikuläre Exantheme beim Schwein und Lagoviren befallen Kaninchen und Hasen. Ein besonders schweres Krankheitsbild verursacht hier das Rabbit-Haemorrhagic-Disease-Virus (RHDV), welches bei Kaninchen eine Virämie mit massiver Leberzellnekrose hervorruft, die durch Auslösen einer DIC meist tödlich verläuft. Die humanpathogenen Vertreter, die zu den Gruppen NV und SV gehören, rufen beim Menschen akute Gastroenteritiden hervor [10].

## **1.2 Die Calicivirus-Gastroenteritis beim Menschen**

### **1.2.1. Stabilität, Infektiosität, Transmissionswege**

Humane Caliciviren (HuCV) sind unter Umweltbedingungen über einen großen Zeitraum stabil. In verschiedenen Versuchen wurde das Virus unter anderem über drei Stunden bei Raumtemperatur (RT) und pH 2,7 inkubiert, bei 4 °C 18 Stunden lang 20%igem Äther ausgesetzt und 30 Minuten lang auf 60 °C erhitzt. Es konnte nachgewiesen werden, dass unter diesen Bedingungen die Infektiosität des Virus erhalten bleibt [10, 16]. Ein hygienisches Problem ergibt sich daraus, dass die zur Routinedesinfektion von Trinkwasser eingesetzten

Chlorkonzentrationen von durchschnittlich 3,75 - 6,25 mg/l die Virulenz des Virus ebenfalls nicht ausreichend mindern und es auf Grund seines geringen Durchmessers einfache Wasserfilter passieren kann [17]. HuCV sind demnach gegenüber Chlor stabiler als Enteroviren oder andere virale Durchfallerreger, beispielsweise Polio- (PV) und Rotaviren (RV) [10]. Erst ab Chlor-Konzentrationen von 10 mg/l, welche erst bei bekannter Kontamination zur Desinfektion betroffener Trinkwasseranlagen eingesetzt werden, werden HuCV inaktiviert [17].

Es konnte gezeigt werden, dass weniger als 100 Viruspartikel ausreichen, um eine akute Gastroenteritis beim Menschen hervorzurufen, welches auf einen hohen Grad an Infektiosität bedeutet [18] und so eine Problematik in Bezug auf Epidemien hervorruft. Hauptübertragungswege sind fäkalienseuchte Trinkwasser und Nahrungsmittel (Austern, Salat), kontaminierte Oberflächen und im direkten persönlichen Kontakt unter anderem das Aerosol von Erbrochenem, wenn es oral aufgenommen und geschluckt wird [10]. Austern kommt eine besondere Bedeutung zu, da sie meist roh verzehrt und häufig mit verschmutztem Wasser gedüngt werden und außerdem die Fähigkeit besitzen, das Agens zu konzentrieren [19].

Der Virusnachweis durch EM, IEM, RT-PCR, Radioimmunoassay (RIA) oder Enzymimmunoassay (EIA) gelang bisher aus Stuhl und Erbrochenem, nicht jedoch aus nasopharyngealem Sekret symptomatischer Patienten. Dies macht eine respiratorische Übertragung unwahrscheinlich, wohingegen die fäkal-orale Übertragung als häufigster Weg angesehen werden kann [10]. Die Amplifikation durch die Polymerase-Region ist die sensitivste und daher heute am häufigsten angewendete Nachweismethode. Die Nachweisgrenze liegt hier zwischen 10 und 100 Virusgenomen [18].

### **1.2.2 Pathogenese, klinisches Bild**

HuCV werden oral aufgenommen und passieren aufgrund ihrer Säurefestigkeit ohne Virulenzverlust den Magen. Es gibt Hinweise dafür, dass der primäre

Replikationsort im oberen Intestinaltrakt liegt. In bioptisch gewonnenem Material aus dem proximalen Dünndarm, vor allem Jejunum, wird bei histologisch intakter Mukosa, elektronenmikroskopisch eine Verkürzung, Verbreiterung und Abstumpfung der Mikrovilli nachgewiesen. Es kommt zu einer Infiltration mit Mononukleären Zellen und zytoplasmatischen Vakuolisierungen. Diese Veränderungen werden auch dann beobachtet, wenn die infizierten Probanden asymptomatisch sind. Erbrechen und Übelkeit erkrankter Patienten, welche trotz histologisch unverändertem Magen auftreten kann, werden am ehesten durch Magenentleerungsstörungen verursacht [10].

Die akute HuCV-Gastroenteritis verläuft beim ansonsten gesunden Patienten so mild, dass es nicht zu einer stationären Einweisung kommt. Ein bis zwei Tage nach der oralen Aufnahme werden die Patienten zunächst durch Bauchkrämpfe, Übelkeit und allgemeines Krankheitsgefühl auffällig, nach einigen Stunden kommen Diarrhöen oder Erbrechen (selten beides), Kopfschmerzen und Myalgien hinzu. Der Patient verweigert die Nahrungsaufnahme. Der Krankheitsverlauf bei Kindern wird häufiger durch Erbrechen bestimmt und von höherem Fieber begleitet als beim Erwachsenen. Insgesamt hält die Symptomatik etwa zwei Tage an. Die Ausscheidung der Krankheitserreger ist eine Woche nach Krankheitsbeginn beendet. Eine Langzeitausscheidung beim immunkompetenten Patienten kann bis zu acht Monate andauern [10]. Im März 2003 wurde der Fall eines mit Azathioprin und Cyclosporin immunsupprimierten Patienten publiziert, der dauerhaft über mehrere Jahre CV-RNA ausgeschieden hat und über diesen Zeitraum auch persistierend an Diarrhoen litt. Bis zum Ende des Beobachtungszeitraums konnte eine Virusausscheidung mittels PCR und Elektronenmikroskopie nachgewiesen werden [20].

### **1.2.3 Immunität**

Die Immunität gegenüber den HuCV ist noch weitgehend ungeklärt. Verschiedene Resistenzmuster konnten bis heute beobachtet werden. Eine

Kurzzeitimmunität schützt den Patienten sechs bis vierzehn Wochen vor einer Neuinfektion mit dem selben Serotyp [21]. Die Langzeitimmunität betreffend hat eine Studie gezeigt, dass die Hälfte der getesteten Probanden sowohl bei der Erst- als auch bei einer Zweitinfektion im Abstand von 27 bis 42 Monaten gastrointestinale Symptome entwickeln, die andere Hälfte wird weder bei der Erst- noch bei der Zweitinfektion klinisch auffällig. Gerade nicht erkrankte Patienten weisen vor der Zweitinfektion einen sehr geringen oder gar keinen Antikörpertiter im Serum auf [22]. Eine Erklärungsmöglichkeit ist, dass bei unterschiedlichen Probanden verschiedene genetisch determinierte Virusrezeptoren im Intestinaltrakt existieren, durch die das Virus nur schlecht oder gar nicht in die Zelle übertreten kann [23].

#### **1.2.4 Epidemiologie, Bedeutung im Verhältnis zu anderen Erregern**

Seit 1995 damit begonnen wurde, die RT-PCR mit einer Mischung unterschiedlicher Primer zur Routinediagnostik von HuCV zu nutzen, wurde deutlich, dass sie in ihrer Rolle als weltweiter Erreger akuter Gastroenteritiden (AGE) in ihrer Häufigkeit stark unterschätzt worden waren. Frankhauser et al. identifizierte sie 1998 sogar in 94 % einer Serie von 90 Ausbrüchen nicht-bakterieller AGE in den USA [24]. Ähnlich hohe Prävalenzen fand man in Studien in den Niederlanden [25], in Großbritannien [26], Japan [27] und Australien [18, 28]. Prävalenzstudien mit Hilfe von Antikörperbestimmungen in der Bevölkerung zeigten, dass Ak gegen NV in den USA von der frühen Kindheit bis ins Erwachsenenalter langsam erworben werden, so dass im Alter von 50 Jahren circa 50 % der Individuen seropositiv sind [10]. In Staaten wie Bangladesh sind dagegen schon im Alter von vier Jahren und in Ga-Rankuwa, Südafrika, schon mit zwei Jahren nahezu 100 % der Kinder Ak-positiv [10]. Zusammenfassend handelt es sich beim HuCV um ein Virus, das auf jedem Kontinent vorkommt und eine ausgesprochen hohe Anzahl an Subtypen aufweist. Die genaue weltweite Verteilung ist noch nicht vollständig geklärt. Die hohe Frequenz, mit der neue Subtypen entdeckt werden und das Fehlen eines

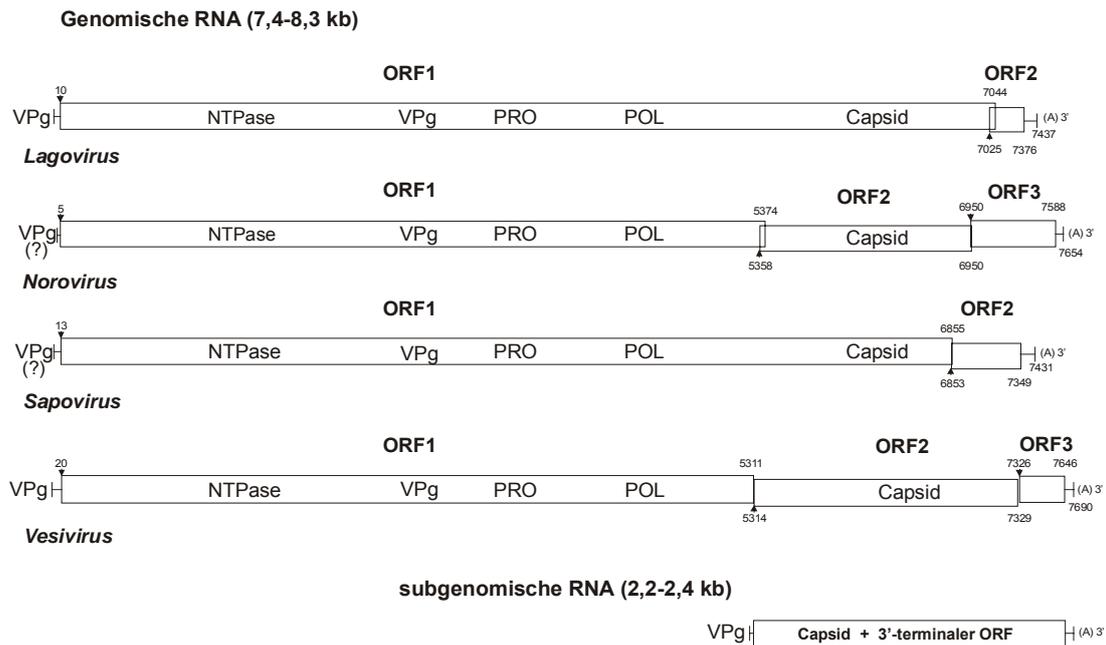
Anzuchtmodells, erschweren die Arbeit verschiedener Forschungsgruppen, die sich mit der Entwicklung eines geeigneten Impfstoffs beschäftigen.

## **1.3 Grundlagen**

### **1.3.1 Struktur des Genoms der CV**

Hinsichtlich von Genomorganisation und Sequenz sind jeweils Lagovirus (LV) und SV sowie Vesivirus (VV) und NV eng miteinander verwandt. Die Abfolge der genetischen Informationen für strukturelle und nichtstrukturelle Proteine ist bei allen vier Genera weitgehend identisch [9]. An das 5'-Ende der tierpathogenen CV ist ein Protein, das VPg, angehängt, für dessen Vorhandensein es bei den humanpathogenen CV jedoch noch keine eindeutigen Beweise gibt. Darauf folgen in 5'-3'-Richtung jeweils die Informationen von Nucleosidtriphosphatase (NTPase), VPg, Proteinase (Pro), RNA-abhängige-RNA-Polymerase (Pol), Capsid und einem basischen Protein, dessen Funktion noch unklar ist. Folglich geht bei allen CV der Capsidinformation die Information für ein langes nichtstrukturelles Protein voraus. Dieses wird als Ganzes abgelesen. Die nichtstrukturellen Virusproteine entstehen durch Spaltung dieses Vorläuferproteins mit Hilfe einer einzigen Cystein-Protease (Pro) [29].

Am unteren Bildrand der Abbildung 2 (s. u.) befindet sich die subgenomische RNA. Hierbei handelt es sich um ein etwa 2,3 kb langes RNA-Molekül, das die genetische Information für das Capsid und ORF3 enthält. Es wird davon ausgegangen, dass es auch bei den humanpathogenen Genera (SV und NV) vorkommt, wenngleich sein Vorhandensein erst für tierpathogene Vertreter bewiesen werden konnte [29]. Bei diesen befindet es sich neben dem Genom im Virusinneren und dient vermutlich als weitere Matritze für die Capsidsynthese. Auch die subgenomische RNA weist am 5'-Ende ein VPg auf und ist am 3'-Ende polyadenyliert [29].



**Abbildung 2: Schematische Darstellung von Genomorganisation und Benutzung der Open-Reading-Frames (ORF) und subgenomischer RNA anhand exemplarischer Caliciviren der einzelnen Genera [9].** Lagovirus: RaCV/LV/RHDV/GH/1988/GE (GenBank accession no. M67473), Norovirus: HuCV/NV//Norwalk/1968/US (GenBank accession no. M87661), Sapovirus: HuCV/SV/Manchester/1993/UK (GenBank accssecion no. X86560), Vesivirus: FeCV/VV/FCV/F9/1958/US (GenBank accession no. M86379). Pfeilspitzen markieren die Stelle, an der sich das erste AUG des ersten ORF befindet. NTPase = Nucleosidtriphosphatase, VPg = Virusprotein, PRO = Protease, POL = Polymerase. Die schraffierte Fläche im Vesivirus-Genom markiert die Leadersequenz des Capsid-Vorläuferproteins.

Im Gegensatz zu SV und LV, die nur zwei ORF besitzen, zeigen NV und VV drei ORF [29]. Innerhalb dieser zwei möglichen Organisationsformen verschmilzt das Capsidgen entweder mit ORF1 zu einem einzigen großen Polyprotein-Gen (LV und SV) oder das Capsidgen ist in einem eigenen ORF kodiert, der in einigen Basenpaaren mit ORF1 überlappt (VV und NV) (siehe Abbildung 2).

### 1.3.2 Genomorganisation der NV

Humane NV werden in zwei Genogruppen unterteilt (GGI und GGII) [29]. Eine dritte Gruppe (GGIII) wird durch das tierpathogene Jena-Bovine-Virus (JBV) präsentiert [30]. Unterschiede zwischen den zuerst genannten bestehen darin, dass die Gesamtlänge des Genoms der GGII-Viren durchschnittlich 200 nt kürzer ist und dass diese Viren in der 5'-Region bemerkenswerte Sequenzunterschiede zu GGI aufweisen. Sie reflektieren möglicherweise Differenzen in Bezug auf Sekundärstruktur und Regulatorsignale [29].

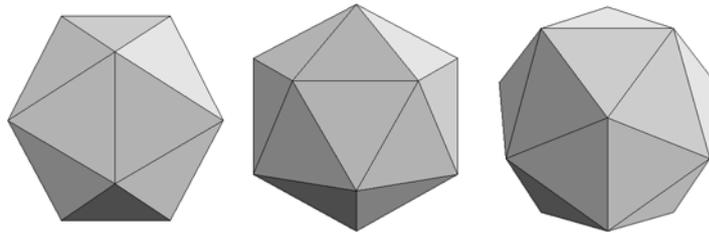
Zwischen ORF1 und ORF2 befindet sich bei beiden Genogruppen ein Frameshift, so dass das 5'-Ende des Capsids das 3'-Ende der Polymerase-Region überlagert. Bei Viren der GGI überschneiden sich diese Regionen mit 17 nt, bei denen der GGII mit 20 nt [29].

In ORF1 beider Genogruppen beginnen 4 nt upstream vor der Grenze zu ORF2 repetitive Sequenzen von 24-26 nt Länge. Es wird vermutet, dass diese Sequenzen Regulatorfunktion besitzen und die subgenomische RNA begrenzen. Eventuell spielen sie auch eine Rolle bei der Replikation und Transkription des Virus oder bilden Signale zur Verpackung der RNA ins Capsid [29]. Der zweite Frameshift, zwischen ORF2 und ORF3, hat zur Folge, dass der dritte ORF wieder im gleichen Leserahmen wie ORF1 liegt. Bei beiden Genogruppen beträgt dieser letzte Frameshift nur eine Base [29].

Zu GGI gehören beispielsweise, mit einem Genom von ca. 7,7 kb, das Norwalk-Virus (NorV), Southhampton-Virus (SouV), Desert-Shield-Virus (DSV), Queens-ArmsVirus (QAV), Winchester-Virus (WinV), Musgrove-Virus (MusV), Sindlesham-Virus (SinV), Hesse-Virus (HesV) und Chiba-Virus (ChiV). Zu GGII werden unter anderem das Lordsdale-Virus (LorV), Hawaii-Virus (HV), Mexico-Virus (MX), Melksham-Virus (MeIV), Hillingdon-Virus (HilV), Leeds-Virus (LeeV), Snow-Mountain-Virus (SMV), Toronto-Virus (TorV) und Bristol-Virus (BriV), Seacroft-Virus (SeaV), Amsterdam-Virus (AmsV), Carlton-Virus (CarV), Mt-Keira-Virus (MtKV) und das Camberwell-Virus (CamV) gezählt [10, 30].

## 1.4 Das Capsid

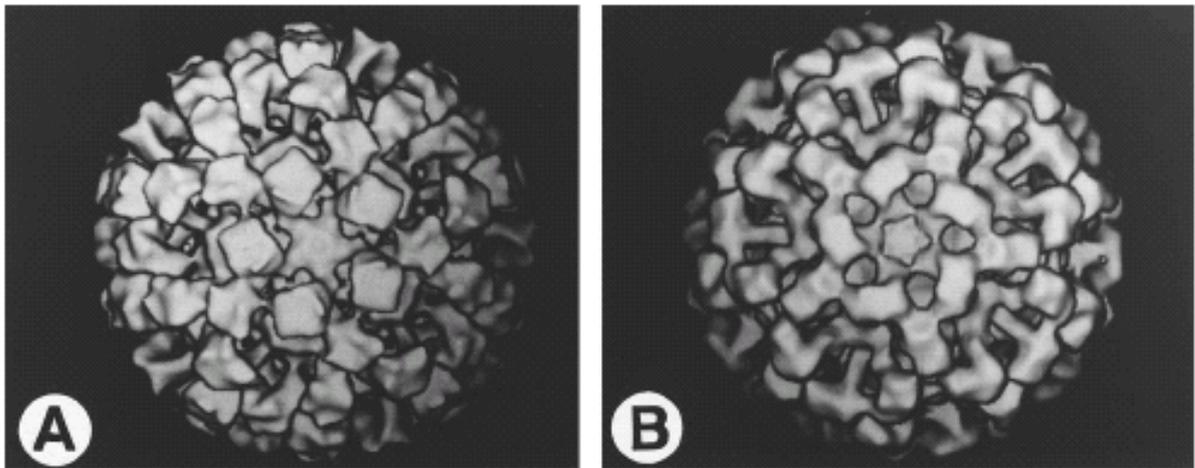
Alle Viren mit kubischer Symmetrie sind Ikosaeder, das heißt, durch zwanzig gleichseitige Dreiecksflächen begrenzte Körper. Dreht man einen Ikosaeder, so fällt der Blick abwechselnd auf die Achsen der zwei-, drei- und fünffachen Symmetrie [31].



**Abbildung 3: Ikosaeder aus unterschiedlichen Perspektiven**

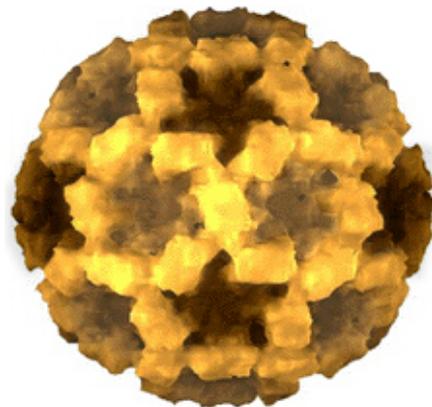
Auch das NV-Capsid hat ein ikosaedrisches Grundgerüst mit einer molekularen Masse von circa 58 kDa. Es ist dabei aus 180 Einzelkopien, also 90 Dimeren eines einzigen Virusproteins zusammengesetzt, welches somit alle funktionellen Aufgaben wie strukturelle Integrität, Immunogenität und Infektivität erfüllt [32]. Die strukturelle Zusammensetzung des Capsids aus vielen identischen Kopien eines einzigen Proteins kommt bei pflanzenpathogenen Viren relativ häufig vor. Unter den tier- und humanpathogenen Erregern gibt es dagegen außer dem CV nur noch ein einziges Virus, das Nodavirus, das eine ähnliche Strukturidee aufweist [32].

Neunzig Dimere bilden gemeinsam einen Mantel, aus dem neunzig bogenförmige Kapsomere herausragen. Auch wenn keine RNA verpackt wird, ergibt sich durch Kultivieren im Baculovirus diese typische Struktur, die rekombinanten „*virus-like particles*“ (rVLPs), die unter anderem in der Forschung am Capsid Verwendung finden. Auch kleinere rVLPs aus beispielsweise nur sechzig Proteinen können durch selbständige Zusammenlagerung entstehen.



**Abbildung 4: Elektronenmikroskopische Darstellung des Capsids [9].** Die Kelchartigen Strukturen auf der Oberfläche sind charakteristisch für Caliciviren. A: Bildliche Rekonstruktion von Pan-1 (primate Calicivirus) durch Kryo-Elektronenmikroskopie (KEM) B: Bildliche Rekonstruktion von rekombinanten Norwalk „virus-like particles“ durch KEM.

Die neunzig bogenförmigen Kapsomere befinden sich an allen zweifachen Symmetrieachsen des  $T = 3$  ikosaedrischen Gittergerüsts und sind so angeordnet, dass sich die Kavitäten an allen drei- und fünffach gefalteten ikosaedrischen Achsen befinden. So entstehen insgesamt 32 Kavitäten, die in der Aufsicht wie Kelche aussehen [10].



**Abbildung 5: Geometrie des Virus.** Die fünfeckigen Kelche sind zur Verdeutlichung der Geometrie verdunkelt, und die sechseckigen Kelche aufgehellt dargestellt.

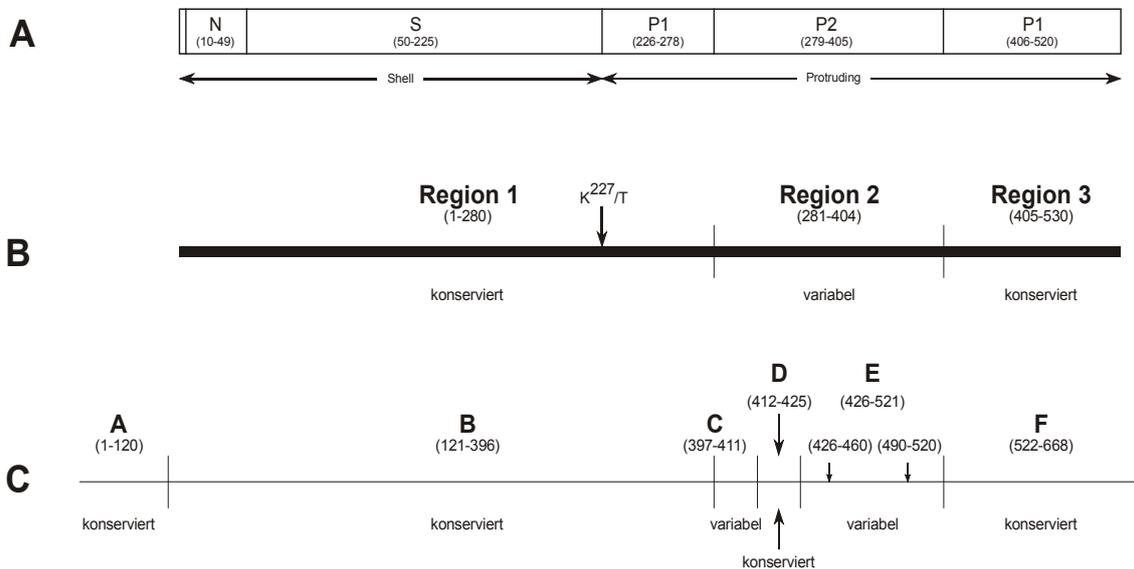
Zwölf der Kelche befinden sich in den fünffachen Symmetrieachsen, mit einer kleinen Erhebung mittig auf dem Grund des Kelchs (Abbildung 7a: 5). Die anderen zwanzig Kelche liegen in den dreifachen Symmetrieachsen und sind zentral flach (Abbildung 7a: 3). Ihre Anordnung läßt sich mit einem Fußball vergleichen: Um ein Fünfeck (schwarz, fünffache Symmetrieachse) herum liegen jeweils fünf Sechsecke (weiß, dreifache Symmetrieachse). Die Kapsomere spiegeln die Ecken der Fünf- und Sechsecke wieder.

Wird der Radius von innen nach außen verfolgt, so befindet sich zwischen 105 und 145 Å eine kontinuierliche Massendichte, der Mantel um die RNA. Von dort aus ragen die Bögen bis 190 Å heraus und sind an der höchsten Stelle zu einer rechteckigen Plattform (50 x 70 Å) abgeflacht (Abb. 7a und 7b). Die dazwischen liegenden Kavitäten sind 40 Å tief und betragen 90 Å im Durchmesser. Bemerkenswert ist, dass diese Abfolge von außen nach innen auch genau der Abfolge der einzelnen Domänen von N- nach C-terminal innerhalb des Capsidproteins entspricht.

Elektronenmikroskopisch können zwei Teile differenziert werden: Die S-Domäne, die als klassische achtfach gefaltete  $\beta$ -Faltblatt-Struktur dem Mantel (Shell) um die RNA entspricht, und die P-Domäne, die die bogenförmigen Ausläufer (Protruding Region) bildet. Vor der S-Domäne befindet sich N-terminal die sogenannte N-Domäne, die das Kapsid von innen ausfüllt. Bei S und N handelt es sich um konservierte Regionen, N gilt sogar als hochkonserviert innerhalb aller verschiedenen Caliciviren [10].

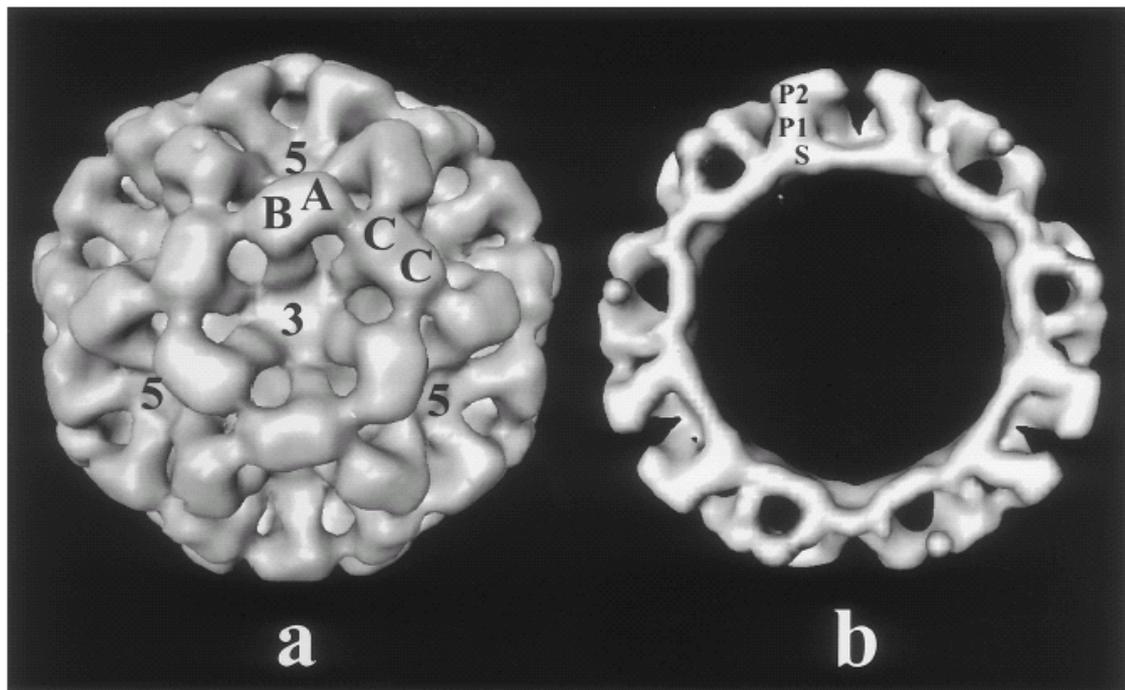
Aus Kryo-EM-Rekonstruktionen ist bekannt, dass die C-terminale P-Domäne zwei unterschiedliche Subdomänen enthält, wobei sich die erste, P1, wiederholt (P1 - P2 - P1) (Abbildung 6a) [32]. Die P2-Domäne ist die Region des Capsidproteins, die die höchste Sequenzvariabilität unter den NV aufweist. Aus ihr entsteht der Teil des Bogens, der am weitesten herausragt und eine Plattform bildet. P1 dagegen formt die Seitenpfeiler des Bogens. Die Exposition der P2-Region auf der Plattform des Bogens steht in Einklang mit der Annahme, dass es sich bei ihr um die Antikörper-Bindungsstelle handelt. Der Region E innerhalb von P2 scheint diesbezüglich auch eine große Bedeutung

zu kommen, da sich gezeigt hat, dass sich beim Feline CV die Neutralisationsspezifität ändert, wenn diese Region alteriert [10].



**Abbildung 6a, 6b und 6c: Das Capsidprotein als Strang mit Darstellung der variablen und konservierten Regionen.** 6a: Primärstruktur des Capsidproteins. (N = N-Terminus, S = Shell-Domäne, P1 und P2 = Protruding-Regions). 6b: Das Capsidprotein grob eingeteilt nach variablen und konservierten Regionen. Hierbei korrespondieren die Abbildungen 6a und 6b im jeweiligen Größenverhältnis und den Längen der Abschnitte. K-227/T markiert die Trypsin-Spaltstelle zwischen Lysin und Threonin. 6c: Detaillierte Darstellung der variablen und konservierten Regionen anhand eines Beispiels aus der Gruppe der Vesiviren. In der Region E sind mit zwei Pfeilen die Stellen gekennzeichnet, anhand denen bei FCV der Wechsel der Neutralisationsspezifität gezeigt werden konnte.

Auf Basis der Lokalisation und unter Berücksichtigung der ikosaedrischen Symmetrieachsen werden die 90 Kapsomere in zwei Subtypen unterteilt. Typ 1 besteht aus einer A- und einer B-Untereinheit und Typ 2 aus zwei C-Untereinheiten. Die A/B-Dimere umgeben die Achsen der fünffachen Symmetrie, während die C/C-Dimere zwischen den pentametrischen Ringen der A/B-Dimere lokalisiert sind. So gibt es bei 60 A/B Dimeren nur 30 C/C-Dimere [32].



**Abbildung 7a und 7b.** 7a: Elektronenmikroskopische Darstellung des Norovirus Capsids. Mit „A“, „B“ und „C“ werden unterschiedliche Formen der Quartärstruktur der Protruding-Region bezeichnet. „3“ und „5“ markieren die Lokalisationen drei- und fünfacher Symmetrie. 7b: Zentraler Querschnitt eines rVLPs. „S“ steht für shell, und „P“ für protruding Region [32].

Trotz identischer Primärstruktur können unterschiedlich konfigurierte Proteine entstehen, da S- und P-Domäne bei Aminosäure 225 durch eine anpassungsfähige Verbindung verknüpft sind. Diese ermöglicht Änderungen in der relativen Orientierung der beiden Domänen zueinander, so dass sich A, B und C in ihrer Konformation unterscheiden und so in das ikosaedrische Gitter passen. Dazu kommt, dass in Bezug auf das N-terminale Ende bei B die ersten neun und bei A und C die ersten 28 Aminosäuren funktionell verändert sind. In der Quartärstruktur des Capsids interagiert dieser längere funktionelle Arm von B mit der benachbarten S-Domäne der C Untereinheit. Zwischen A und B gibt es dagegen diverse Wasserstoffbrückenbindungen, die zwischen C und C nicht existieren [32].

Trotz dieser weitreichenden Erkenntnisse ist noch nicht bekannt, wie es zu genau dieser Faltung der Polypeptidkette kommt, worin das Signal zu Konformationsänderung zwischen A, B und C besteht, wo genau die Stamm-

Spezifität, Antigenität und Rezeptor-Bindungsstelle lokalisiert sind und wie die selbständige Zusammenlagerung (*self-assembly*) des Virus ermöglicht wird [32]. Dennoch wird an dieser Stelle bereits klar, warum bereits kleinste Veränderung in der Basenabfolge des Capsidgens weitreichende Folgen in Bezug auf Konformation und Antigenität haben können.

## 1.5 Mutationen und genetischer Informationsaustausch zwischen Viren

Wie alle Lebewesen unterliegen auch Viren den Gesetzen der Evolution. Mutationen verändern das Genom und natürliche Selektion entscheidet, welche der entstandenen Neumutationen am besten lebensfähig ist und sich somit im Laufe der Zeit am stärksten vermehren wird [33].

Mutationen können verschiedene Ursachen haben. Nicht nur durch chemisch-physikalische Einwirkung (UV-Licht, Nitrit, Röntgenstrahlen) kommt es zu Defekten mit fehlerhafter Reparatur, auch bei jeder regelrechten Replikation treten mit gewisser Wahrscheinlichkeit verändernde Substitutionsereignisse auf. Da die RNA-abhängige RNA-Polymerase keine Korrekturlesefähigkeit besitzt (fehlende 3'-5'-Exonuclease-Aktivität) [34] ist diese Wahrscheinlichkeit für RNA-Viren verhältnismäßig hoch. Im Vergleich: RNA-Viren: 1/10.000, DNA-Viren: 1/100.000 bis 1/100.000.000, zelluläre DNA: 1/10.000.000.000 [33]. Zusätzlich haben verschiedene Viren durch Rekombination, Reassortierung und Komplementation die Möglichkeit, genetisches Material untereinander auszutauschen [33]. Komplementation ist die Wiederherstellung eines zuvor durch Mutation inaktiv gewordenen Genprodukts durch Austausch von genetischen Informationen. Das fehlende oder defekte Gen kann durch ein Helfervirus oder das Wirtsgenom bereitgestellt werden [33]. Diese Methode wird von DNA-Viren angewendet. Reassortierung ist eine Methode, die nur von segmentierten Viren, beispielsweise von Influenza-A-Viren angewendet wird. Bei der Doppelinfection einer Zelle werden die einzelnen Segmente des Genoms getrennt vervielfältigt und anschließend Teile beider Genome in das neue Virus eingebaut [33]. Rekombination wird sowohl durch segmentierte als

auch unsegmentierte Viren durchgeführt. Hierbei wird eine Donorsequenz in ein kontinuierliches Akzeptor-Molekül eingefügt [35]. Es gibt verschiedene Unterformen: Bei der homologen Rekombination wird eine bestimmte Sequenz im Akzeptormolekül durch eine homologe Donorsequenz ersetzt. Da diese Form nur bei homologen Parentalsträngen vorkommt, bleibt die Struktur der Sequenz unverändert [35]. Zu Strukturveränderungen kommt es dagegen bei aberranter homologer Rekombination, da hier Sequenzen zwischen verwandten Viren ohne striktes Alignment ausgetauscht werden und bei nichthomologer Rekombination, dem genetischen Austausch zwischen unverwandten Sequenzen [35]. Die meisten Viren rekombinieren nach dem *copy-choice* Modell. Nach diesem Modell entstehen Hybrid-RNA-Stränge, wenn die RNA abhängige RNA-Polymerase während der Replikation von einem RNA-Strang auf einen anderen übergeht. Homologe Rekombination entsteht, wenn die Replikase genau dort weiterliest, wo sie den anderen Strang verlassen hat. Ist dies nicht der Fall, kommt es zu nichthomologer oder aberranter homologer Rekombination [35]. Das *copy-choice* Modell wurde ursprünglich für Polioviren aufgestellt [36].

Viren sind in ihrer Entwicklung stark vom Wirt abhängig. Da das Immunsystem des Wirts einen hohen Selektionsdruck ausübt, haben Virusspezies mit hoher Diversität einen Selektionsvorteil. Diversität kann bedingt sein durch eine hohe Vermehrungsrate mit einem schnellen *turnover* (bis zu 90 % pro Tag), Rekombination, Koplementation und vor allem bei RNA-Viren durch die Neigung zu fehlerhafter Replikation. Im Zuge dieser Anpassung kann sich auch das Wirtsspektrum verbreitern. Virusspezies mit besonders hoher Diversität innerhalb der Mitglieder, zum Beispiel HI-Viren, werden auch Quasispezies genannt [37].

Durch den von Jiang et al. publizierten Verdacht der Entdeckung einer natürlich vorkommenden Rekombinanten innerhalb der humanpathogenen Caliciviren erschienen verschiedene Aspekte in neuem Licht. Sowohl die Verwandtschaftsverhältnisse und die enorme Diversität, als auch mögliche Escape-Mechanismen bei mangelnder Entwicklung von dauerhafter Immunität gegen diese Viren müssen neu betrachtet werden [38]. Die Möglichkeit dieser

Viren, aufgrund ihrer besonderen Capsidstruktur durch Konformationsänderung auf kleine Alterationen sensibel reagieren zu können, werden besonders im Hinblick auf ihre Antigenität und die Identifikation und Zuordnung detektierter Stämme anhand der Polymerase-Region neu diskutiert.

Diese Arbeit zeigt das Vorkommen neuer Norovirusstämme innerhalb von Deutschland und ihre verwandschaftliche Beziehung untereinander, untermauert das natürliche Auftreten von Rekombination innerhalb von Noroviren durch Nachweis zweier rekombinanter Stämme und diskutiert verschiedener Methoden zur taxonomischen Zuordnung von Einzelstämmen.

## 1.6 AUFGABENSTELLUNG

Die hohe genetische Diversität humaner Caliciviren und deren Klassifikation im Hinblick auf Mutations- und Rekombinationsraten innerhalb verschiedener Genomabschnitte ist noch nicht vollständig geklärt. Die gewöhnlich verwendete Methode zur Zuordnung zu einem CV-Stamm ist zur Zeit die Sequenzanalyse der RNA-Polymerase-Region. Vor dem Hintergrund der Beschreibung eines natürlich vorkommenden rekombinanten Virusstamms durch Jiang et al. ist zu klären, ob Rekombination humaner Caliciviren frequent natürlich vorkommt oder ein isoliertes Ereignis darstellt. Bei gehäuftem natürlichem Vorkommen in immunologisch bedeutsamen Regionen des Virusgenoms ist fraglich, ob die Untersuchung der Polymerase-Region für eine eindeutige Virusbestimmung ausreicht, oder ob zur Klassifikation nah verwandter Stämme die Verwendung der Capsidregion erwägt werden muss [25].

Die Ziele dieser Arbeit werden wie folgt zusammengefaßt:

- 1) Phylogenetische Analyse humaner Caliciviren anhand der Sequenzanalyse der Capsidregion. Das Untersuchungsgut ist Virus-RNA aus Stuhlproben an Norovirus-Gastroenteritis erkrankter Kinder im Alter von bis zu vier Jahren aus fünf deutschen Städten (Freiburg, Hamburg, Bochum, Erlangen und Dresden). Die RNA der Capsidregion wurde isoliert, amplifiziert, kloniert, sequenziert und phylogenetisch analysiert.
- 2) Identifikation von rekombinanten Virusstämmen, den zugehörigen Parentalviren und der exakten Rekombinationsschnittstelle.
- 3) Untersuchung der Rekombinanten in Hinblick auf typische Eigenschaften von Viren, bei denen Rekombinationsereignisse gehäuft nachgewiesen werden können.
- 4) Diskussion der Tragweite von Rekombinationsereignissen innerhalb verschiedener Genomabschnitten im Hinblick auf die Klassifikation und immunologische Bedeutung.

## **2 MATERIAL UND METHODEN**

### **2.1 Material**

#### **2.1.1 Stuhlmaterial**

Die verwendete Stuhlproben wurden durch zwanzig Arztpraxen in den Städten Hamburg, Bochum, Freiburg, Erlangen und Dresden gesammelt und stammten von bis zu vier Jahre alten Kindern, die sich mit akuter Gastroenteritis vorgestellt hatten. Der Zeitraum der Probengewinnung erstreckte sich über 14 Monate. Bis zur Verarbeitung wurden die Proben bei  $-20\text{ }^{\circ}\text{C}$  aufbewahrt [39]. Im Vorfeld dieser Studie waren alle Proben anhand der Polymerase-Region auf das Vorhandensein von CV-RNA untersucht worden [40]. Die insgesamt 119 Proben, die ein positives Testergebnis bezüglich NV lieferten, wurden als Material zur Charakterisierung der kompletten Capsidsequenz gewählt.

#### **2.1.2 Geräte**

Analysewaage	Sartorius
Agarose-Gelelektrophoresesystem	Pharmacia Biotech, Hoefer
Blockthermostat	Eppendorf
BAS cassette 2040	Fujifilm
Brutschrank	Heraeus
DNA-Sequenzierer	Applied Biosystems
Geldokumentationssystem	Biorad
Heizblock	Eppendorf
Imaging Plate 2040	Fujifilm
Kunststoffküvetten	Hartenstein

Liquid Scintillation counter	Wallac
PCR-Cycler	Perkin Elmer, Eppendorf
Photometer	Eppendorf
Polyacrylamid-Gelelektrophoresesystem	Hoefer
Schüttelinkubator	New Brunswick Scientific
Spannungsquelle	Pharmacia Biotech
SpeedVac	Heraeus, Eppendorf
Tischzentrifuge	Biofuge fresco
UV-Küvetten	Eppendorf
Vortexer	Scientific Industries
Wasserbad	Memmert
Zentrifuge	Heraeus, Eppendorf

### 2.1.3 Enzyme

<i>Taq</i> DNA Polymerase in Storage Buffer B	Promega
SuperScript™ One-Step RT-PCR with Platinum® <i>Taq</i>	Invitrogen
SuperScript™ One-Step RT-PCR for Long Templates	Invitrogen
RNasin® (40 U/μl)	Promega
Promega Rnasin® Ribonuclease Inhibitor	Promega
Herculase® Hotstart DNA Polymerase	Stratagene

## 2.1.4 Kommerzielle Testsysteme

ABI PRISM® BigDye® Terminator v1.1 Cycle Sequencing Kit	Applied Biosystems
Microcon® Concentration	Millipore
QIAEX II Gel Extraction Kit	Qiagen
QIAmp® DNA Stool Mini Kit	Qiagen
QIAprep® Miniprep Kit	Qiagen
Zero Blunt® TOPO® PCR Cloning Kit for Sequencing	Invitrogen

## 2.1.5 Oligonukleotidprimer

RT-PCR	Sequenz 5'-3'
Calman 1 (CalORF1-1)	5'-GCA CAC TGT GTT ACA CTT CC-3'
Calman 6 (CalORF3-2)	5'-GCTGAA GAA CCT AGT CTC G-3'
Calman 8 (CalORF3-6)	5'-TCT TTA TCA TGT TGG AAA GAA C-3'
Calman 10 (CalORF3-4)	5'-AAG AAG TGG CAG TAG GAG C-3'
Calman 12 (CalORF3-8)	5'-TTG ATG GCT CCA GCT CCA GCA TTG-3'
Primerwalking	Sequenz 5'-3'
capseq-1	5'-GAG AGC CGT TAT CGT CTG-3'
capseq-2	5'-ATT GCG TTG CGCTCA CTG-3'
capseq-3	5'-GA TCG GCC TTC CAA CAA GT-3'
capseq-4	5'-CTT CCT CGC TCA CTG ACT CG-3'
capseq-5	5'-CCA GAC CAA GAG CCA ATG-3'
capseq-6	5'-CCA GAC AAG AGC CAA TGT TC-3'
capseq-7	5'-CTG AGT AGT TTG GGG AGC ACC-3'
capseq-8	5'-GCA ACC GGC AGG AAC TTA GG-3'
capseq-9	5'-ATG GAT TT TAC GTG CCC AG-3'
capseq-10	5'-AGT CA CGT CTG GTG TCA-3'
capseq-11	5'-CGG CCC AGC ATT CTA CAG-3'
capseq-12	5'-CTG GAG TG CTT CTT GGT AG-3'
capseq-13	5'-GGT ATC CAC AGA ATC AGG-3'
capseq-14	5'-ATG ACC AGT CTA CCT GAG-3'

capseq-15	5'-CCA GAC AAG AGC CAA TGT TC-3'
capseq-16	5'-CTG GCT TTA ATG GAA ATC -3'
capseq-17	5'-GAC TGG GGT GAA TTT CGT G-3'
capseq-18	5'-GTG TGA GCC ACT GTA CAT AG-3'
capseq-19	5'-GAG AGC CGG TAT CGT CTG TT-3'
capseq-20	5'-CTG GAG CTG CTT CTT GGT AG-3'
capseq-21	5'-CTG ATC TCA GCC ACC GCG AA-3'
capseq-22	5'-GGC ATA GTG GAC CTG AAG AA-3'
capseq-23	5'-CCA CCA CGC TAG GAG ATA GA-3'
capseq-24	5'-AGT TGG ACC CAG AGA AAT TG-3'
capseq-25	5'-GGC AAG TGT GTA GAA CTG GT-3'
capseq-26	5'-CCG CCC TGA CAA AAC TGA AG-3'
capseq-27	5'-GGC ATA GTG GAC CTG AAG AA-3'
capseq-28	5'-TCC TGC GGA GAA CTG TGA CC-3'
capseq-29	5'-TAG AAT GCT GGG CCG TGG AGT-3'
capseq-30	5'-GTT GTT GGT GCC GCT ATT-3'
capseq-31	5'-CGA CAA GAG ACT GTG AAG AC-3'
capseq-32	5'-CCA CCA GGT TAC TAT GTT CC-3'
capseq-33	5'-TTG GTA TCT GCG CTC TGC TG-3'
capseq-34	5'-CCA CCA GGT TAC TAT GTT CC-3'
capseq-35	5'-GGT GAG CAT GCC TTG GAT CT-3'
capseq-36	5'-ACA TGG CGA TAG CTA GAC TG-3'
capseq-37	5'-TCA TGA CTG CCT GCA ATG TG-3'
capseq-38	5'-AAG AGA CAG GAT GAG GAT-3'
capseq-39	5'-ATG TTG ACA GCS GAC ACT GG-3'
capseq-40	5'-GTT GTT GGT GCC GCT ATT GC-3'
capseq-41	5'-CGT GAG GAC TCG ACA AGA GA-3'
capseq-42	5'-TGA AGT GCA GGT AAT CCT CG-3'
capseq-43	5'-GGT TAT TAC CGT CCT GGA TG-3'
capseq-44	5'-TCC TCG CTC ACT GAC TCG CT-3'
capseq-45	5'-TGT AAC TTC AGA GAG CGC AC-3'
capseq-46	5'-AAT TGC GTT GCG TC ACT GC-3'
capseq-47	5'-GAG AGA GCC GTT ATC GTC TG-3'

capseq-48	5'-CCA GAC AA AGC CAA TGT TC-3'
capseq-49	5'-TGT GCC GGT CTC CGT TAT CG-3'
capseq-50	5'-TCG AAT GAC GCC AAC CCA TC-3'
capseq-52	5'-TCT GAT GTC CTT GGC TCT GG-3'
TOPO-Cloning-Reaktion	Sequenz 5'-3'
M13 Forward	5'-GTA AAA CGA CGG CCA G-3'
M13 Reverse	5'-CAG GAA ACA GCT ATG AC-3'
T3	5'-ATT AAC CCT CAC TAA AGG GA-3'
T4	5'-TAA TAC GAC TCA CTA TAG GG-3'
TopoBlunt-Seq 1	5'-CGC CAA GCT CAG AAT TAA CC-3'
TopoBlunt-Seq 2	5'-CGA CGG CCA GTG AAT TGT AA-3'

## 2.1.6 Kompetente Bakterien und Plasmide

<i>E. coli</i> TOP 10 Zellen	Invitrogen
<i>E. coli</i> DH5 $\alpha$ -T1	Gibco BRL
<i>E. coli</i> JM 109	Invitrogen
PCR <sup>®</sup> 4Blunt-TOPO <sup>®</sup>	Invitrogen

## 2.1.7 Medien, Puffer, Lösungen

### 2.1.7.1 Bakterienkultur

LB Medium (flüssig)	2% LB Broth base 0,1% NaCl
LB Agar	1% Trypton 0,5% Hefeextrakt 10 mM NaCl 2% Agar

Ampicillin-Stocklösung	100 mg Ampicillin/ml Lösung
1-fach LB-Medium mit Ampicillin	20 ml 5-fach LB-Medium Ad 1l A. bidest bei 55 °C Ampicillin dazu 1ml Ampicillin- Stock/l Medium
LB-Ampicillin Agarplatten	55°C: 1ml Ampicillin-Stocklösung dazu
LB-Platten für JM 109 E.coli	10 ml IPTG (2g+10 ml H <sub>2</sub> O) + Filtern oder jede fertige Platte mit 100µl IPTG ausplattieren
LB-X-Gal-Platten	+ 4 ml X-Gal
SOC-Medium	2% Trypton 0,5% Hefeextrakt 10 mM NaCl 2,5 mM KCl 10 mM MgCl <sub>2</sub> 10 mM MgSO <sub>4</sub> 20 mM Glucose
Glycerol-Stocks	70% ige Glycerol-Lösung

### 2.1.7.2 Plasmid-Aufreinigung

TE – Puffer	10mM Tris/HCl pH 8.0; 0.5 mM EDTA + 1/10 RNase A ( 2mg/ml )
Natrium-Acetat	
Isopropanol	
70 % Ethanol	

### 2.1.7.3 Agarose-Gelelektrophorese

6 x DNA-Ladepuffer	30 % Glycerol 0,25 % Bromphenolblau 0,25 % Xylencyanol
TAE-Puffer	40 mM Tris-Base 1 mM EDTA 20 mM Essigsäure pH 8,0
Agarose-Gel	1 %, mit Ethidiumbromid

### 2.1.7.4 Chemikalien

Reagenz	Firma
Agar	ICN
Agarose	Cambrex
$\beta$ -Mercaptoethanol	Roth
Bromphenolblau	Sigma
CaCl <sub>2</sub>	Calbiochem
Dimethylformamid	Merck
dNTP's (je 10 mM)	Promega
EDTA	AppliChem
Essigsäure	Merck
Ethidiumbromid	Roth
GeneRuler <sup>®</sup> 1, 3 und 10 kb DNA Ladder	MBI Fermentas
Glucose	Merck
Glycerol	Roth
Glycin	Roth
HCl	Merck
Hefeextrakt	Roth
Isopropanol	Merck

Isopropyl-3-D-Thiogalaktosid (IPTG)	Roth
Kaliumacetat	Merck
KCl	Fluka Chemika
LB Broth base	Invitrogen
Magermilchpulver	Lasana
Methanol	Merck
MgCl <sub>2</sub>	Merck
MgSO <sub>4</sub>	Merck
MnCl <sub>2</sub>	Merck
NaCl	Merck
NaOH	Merck
SDS	Roth

## 2.2 Molekularbiologische Methoden, Vorgehensweise

### 2.2.1 RNA-Extraktion aus Stuhlproben mit QIAmp® DNA Stool Mini Kit

Zu Beginn der Extraktion von RNA aus Stuhlproben wurden 180-200 mg gefrorener Stuhl in 1,4 ml ASL-Puffer aufgetaut. Bei ausreichender Stuhlmenge wurde jeder Extraktionsschritt gleichzeitig mit einem Duplikat durchgeführt, um für spätere Konzentrierungsschritte eine ausreichende Lösungsmenge zur Verfügung zu haben. ASL-Puffer lysiert Lipopolysaccharide und zerstört so Zellwandbestandteile im Stuhlmaterial. Um mechanische Schäden an der RNA weitgehend zu vermeiden, wurde die Probe vorsichtig geschwenkt und mit einem Glasspatel umgerührt bis sich der Stuhl vollständig gelöst hatte. Die entstandene Suspension inkubierte zur Zelllyse für 5 min. im auf 70 °C vorgewärmten Heizblock und wurde nach erneutem Schwenken 1 Minute zentrifugiert. Alle Zentrifugationsschritte der Extraktion erfolgten bei 20.000 xg (~14.000 rpm). 1,2 ml des Überstands wurden in ein neues Eppendorf-Röhrchen (EpR) mit einer halben Tablette InhibitEX™ gefüllt, das Pellet wurde verworfen. InhibitEX™ bindet Nukleinsäure schädigende Substanzen und PCR-

Inhibitoren. Die Tablette wurde durch Schwenken gelöst und die Suspension inkubiert bei Raumtemperatur (RT) 1 min. lang. Um die Flüssigkeit besonders gründlich von Verunreinigungen zu befreien, wurde diese zweimal je 3 min. zentrifugiert und das Pellet nach jedem Schritt verworfen. Je 200 µl des Überstands wurden in neue EpR mit 15 µl Proteinase K überführt und mit 200 µl AL-Puffer so lange geschwenkt, bis sich eine homogene Flüssigkeit bildete. Die Proben wurden kurz zentrifugiert und inkubierten im Anschluß 10 min. bei 70 °C. Nach Zusatz von 200 µl 96 %igem Ethanol wurde das gesamte Lysat eines EpR auf eine QIAmp Elutionssäule pipettiert und eine Minute lang zentrifugiert. Die Elutionssäule bindet Nukleinsäuren (NS) mit Hilfe einer Silikatgel-Membran. Nacheinander erfolgten zwei Waschschrte mit je 500 µl Puffer AW1 und AW2. Nach Verwendung von Puffer AW1 wurde eine Minute und nach Puffer AW2 drei Minuten lang zentrifugiert und das Filtrat nach jedem Zentrifugationsschritt verworfen. Im letzten Schritt empfiehlt der Hersteller, die NS mit Hilfe von 200 µl salzarmem AE-Puffer aus dem Silikatgel der Säule zu lösen. Um an dieser Stelle bereits eine Konzentrierung vorzunehmen, wurden statt 200 µl nur 50 µl AE-Puffer verwendet. Nach Inkubation von einer Minute wurde die Säule in ein frisches EpR gestellt und eine Minute lang zentrifugiert. Alle Elutionen von gereinigten Nukleinsäuren aus ursprünglich der selben Stuhlprobe wurden zu circa 150-200 µl vereinigt, zum Schutz der RNA vor RNAsen mit RNAsin (Promega, 40 U/µl) versetzt und unmittelbar für die Amplifikation verwendet. Extraktion, Preparation des Reaktionsansatzes und Amplifikation wurden in separaten Räumen durchgeführt.

### **2.2.2 Konzentrierung der RNA mit Microcon-Tubes**

Mindestens 400 µl Eluat wurden auf eine Microcon-Säule pipettiert und 12 min. bei 500 x g zentrifugiert. Die Säule wurde invertiert und 3 min. bei 1000 x g zentrifugiert. Dadurch wurde die Gesamtmenge des Eluats auf wenige µl reduziert. Erst im Anschluß wurde RNAsin hinzugefügt.

### 2.2.3 Photometrische Bestimmung von RNA-Konzentration, Gesamt-RNA und Reinheit der Probe

Spektrophotometrische Messungen ermöglichen unter anderem die Bestimmung der Nukleinsäurekonzentration in einer Lösung, da Nukleinsäuren bei einer bestimmten Wellenlänge monochromatisches Licht proportional zu ihrer Konzentration absorbieren. DNA und RNA besitzen gleichermaßen bei 260 nm ein Absorptionsmaximum. Die Proben wurden abhängig vom Gesamtvolumen 1:10 bis 1:40 mit A. bidest auf 200 µl verdünnt. Die verdünnte Probe wurde in eine Quarzglasküvette überführt und bei OD<sub>260</sub> photometrisch gegen den Leerwert (A. bidest) gemessen. Unter Berücksichtigung der Verdünnung wurde die Konzentration der NS anhand des Lambert-Beerschen Gesetzes ermittelt:

$$c (\mu\text{g/ml}) = OD_{260} \times F \times V$$

c = Konzentration, V = Verdünnungsfaktor, F = Multiplikations-Koeffizient (ssRNA: 40 µg/ml). F sagt aus, dass eine Absorption von 1 bei 260 nm Wellenlänge einer ssRNA-Konzentration von 40 µg/ml entspricht. Die Gesamtmenge an Nukleinsäuren errechnet sich nach:

$$NS\text{-gesamt } (\mu\text{g}) = c \times \text{Volumen (ml)}$$

Eine zusätzliche Messung bei 280 nm läßt eine Aussage über Verunreinigung mit Proteinen oder Phenol (Maximum bei 320 nm) zu. Bei sauberer Aufreinigung sollte der Quotient OD<sub>260</sub>/OD<sub>280</sub> für DNA bei 1,8 und für RNA bei 2,0 liegen [41].

### 2.2.4 Amplifikation mittels RT-PCR (SUPERSCRIPT™ One-Step RT-PCR with PLATINUM® Taq)

Die Polymerase-Kettenreaktion (PCR) wurde 1984 von Mullis entwickelt und dient der schnellen Amplifikation von DNA-Fragmenten mittels hitzestabiler Taq-Polymerase. Zwei Oligonukleotidprimer (sense- und antisense, 20-30 nt lang) legen den zu vervielfältigenden Sequenzbereich fest. Aufgrund der

Spezifität der Primer kann die PCR auch als Nachweismethode genutzt werden. Der Standard-PCR-Zyklus im Thermocycler setzt sich aus den Phasen Denaturierung (ca. 94°C), Annealing (45-64°C) und Elongation (ca. 72°C) zusammen. Die Anzahl der Zyklen ist variabel. Die Reverse-Transkriptase-PCR (RT-PCR) ist eine Sonderform der PCR zur Amplifikation von RNA. Im ersten Elongationsschritt (hier 30 min. bei 50°C) schreibt die Reverse-Transkriptase-Polymerase RNA in komplementäre DNA (cDNA) um. In den folgenden Schritten wird die cDNA durch die Taq-Polymerase weiter vervielfältigt.

Der hier verwendete RT/Platinum *Taq* Mix enthält SuperScript II H-Reverse Transcriptase (SSIIRT) und PLATINUM® *Taq* DNA Polymerase (TDNAP), welche bei Raumtemperatur durch Antikörper (Ak) blockiert ist. Dadurch bedingt beginnt die Elongationsphase mit einem „hot start“ und es erhöhen sich Sensitivität, Spezifität und Effektivität der PCR. Im Anschluß an die RT-Elongation wird der Ansatz auf 94°C erhitzt. Bei dieser Temperatur lösen sich die Ak von der *Taq* DNA Polymerase, die RNA/cDNA-Hybride schmelzen und die SSIIRT wird inaktiviert. Der Versuchsablauf erfolgte laut Herstelleranweisung (s.u.). Bei der Festlegung der Elongationszeit wurde nach Angaben des Herstellers 1kb pro Minute als Schreibgeschwindigkeit der *Taq* DNA Polymerase angenommen.

#### Komponenten

RT/Platinum <i>Taq</i> Mix	SuperScript II H-Reverse Transcriptase <i>Taq</i> DNA Polymerase
2X Reaction Mix	Puffersystem, Mg <sup>2+</sup> , Stabilisatoren, dNTPs (200 µM)
MgSO <sub>4</sub>	5 mM und 50 mM

<u>Ansatz für eine Reaktion</u>	<u>Volumen</u>	<u>Konzentration</u>
2X Reaction Mix	25 µl	
RNA Template	7 µl	10 pg-100 pg
Primer sense (10 pmol/µl)	5 µl	0,2-0,4 µM
Primer antisense(10 pmol/µl)	5 µl	0,2-0,4 µM
RT/ <i>Taq</i> -Mix	2 µl	
A.bidest (RNase frei)	6 µl	
→ RT-PCR-Ansatz	50 µl	

<u>Zyklus</u>	<u>Temp.</u>	<u>Zeit</u>		<u>Temp.</u>	<u>Zeit</u>		<u>Temp.</u>	<u>Zeit</u>
1 x	50 °C	30 Min.	→	94 °C	2 Min.			
40 x	94 °C	30 Sek.	→	55 °C	30 Sek.	→	68 °C	3 Min. 10 Sek.
1 x	72 °C	7 Min.	→	4 °C	∞			

### 2.2.5 Amplifikation mittels nPCR (Herculase® Hotstart DNA Polymerase)

Um Sensitivität und Spezifität des Nachweises durch Amplifikation zu erhöhen, wurde eine nested PCR (nPCR) angeschlossen. Dazu wurden jeweils zwei neue Primerpaare gewählt, deren Sequenz innerhalb der in der RT-PCR amplifizierten Sequenzen lagen. Um die Wahrscheinlichkeit des Falschablesen bei nPCR zu reduzieren, wurde Herculase® Hotstart DNA polymerase (Stratagene) verwendet. Das Polymerasegemisch (s.u.) ist unterhalb der Denaturierungstemperatur mit Ak blockiert („hot start“) und enthält eine 3'-5'-Exonuclease-Aktivität mit Korrekturlesefähigkeit. Die *Pfu* DNA Polymerase generiert ein PCR-Produkt mit stumpfen Enden. Gemäß der Herstellerempfehlung wurde eine Elongationstemperatur von 72°C bei einer Elongationszeit von 1min. pro kb Template und eine Denaturierungstemperatur von 92°C gewählt. Die Schmelztemperatur ( $T_m$ ) der gewählten Primer lag unter 60°C.

#### Komponenten

Enzyme	Herculase® Hotstart DNA polymerase (5 U/μl) <i>Pfu</i> DNA polymerase PCR-Verstärkungs-Faktor <i>Taq 2000™</i> DNA polymerase
10 x Herculase® reaction buffer	
DMSO	
dNTP mix	25 mM je dNTP Konzentration für < 10kb: 200μM pro dNTP

<u>Ansatz für eine Reaktion</u>	<u>Volumen</u>
Herculase Puffer	5 µl
dNTP mix	0,4 µl
DNA-Template	100 - 250 ng
Primer (sense)	100 ng
Primer (antisense)	100 ng
Herculase Hotstart-Polymerase	0,5 µl
A.bidest (RNase frei)	ad 50 µl
→ nPCR-Ansatz	50µl

<u>Zyklus</u>	<u>Temp.</u>	<u>Zeit</u>		<u>Temp.</u>	<u>Zeit</u>		<u>Temp.</u>	<u>Zeit</u>
1 x	95 °C	2 Min.						
10 x	95 °C	30 Sek.	→	P-Tm -5 °C	30 Sek.	→	72 °C	3 Min. 10 Sek.
20 x	95 °C	30 Sek.	→	P-Tm -5 °C	30 Sek.	→	72 °C	3 Min. 10 Sek. + 10 Sek./Zyklus
1 x	4 °C	∞						

### 2.2.6 Agarosegel-Elektrophorese zum Nachweis positiver Proben

Die Agarosegel-Elektrophorese ist eine effektive Methode, um DNA-Fragmente von 0,5-25 kb Länge voneinander zu trennen und anhand ihrer Länge zu identifizieren. Mittels Verdünnungsreihe einer definierten Menge NV-RNA war im Vorfeld unserer Untersuchungen die untere Nachweisgrenze von RT-PCR und anschließender Analyse durch Gelelektrophorese für die von uns erwarteten 2,4 kb bis 2,8 kb großen Fragmente ermittelt worden. Bei RT-PCR mit angeschlossener nPCR lag diese Grenze bei 84 ng CV-RNA in einem jeweils 7 µl RT-PCR-Ansatz. Für ein 1% Agarosegel wurden 1 g Agarose in 100 ml TAE-Puffer durch Erhitzen in der Mikrowelle gelöst und 3 µl Ethidiumbromid (1mg/ml) hinzugefügt. Nach Abkühlen auf 50 °C wurde das Gel in eine Form mit eingesetztem Kamm gegossen und nach Erstarren in eine Elektrophoresekammer mit TAE-Puffer überführt. In die durch die Kammzinken

entstandenen Taschen wurden Größenmarker und Proben pipettiert und eine Spannung von 120 Volt angelegt. Durch die Phosphatgruppen der DNA migrieren die Stränge im elektrischen Feld zur Anode hin und trennen sich auf, da die Gelmatrix DNA-Fragmente proportional zu ihrer Größe bremst. Die DNA war nach ca. 1,5 h ausreichend weit migriert und wurde unter UV-Licht betrachtet und fotografiert. Der Fluoreszenzfarbstoff Ethidiumbromid lagert sich zwischen den Basen der DNA ein und leuchtet bei Anregung durch UV-Licht auf. Die Länge der DNA-Fragmente wurde mit Hilfe des Größenmarkers abgeschätzt.

#### Marker

1µl Gene Ruler™ 1kb DNA Ladder  
+ 1µl 6x Loading Dye Lösung  
+ 4 µl destilliertes Wasser

#### Probe

10 µl Probe  
+ 2µl 6x Loading Dye

### **2.2.7 UV-Licht-Foto**

Die Fotografie ethidiumbromidhaltiger Agarosegele erfolgte unter UV-Licht der Wellenlänge 256 nm bis 312 nm. Bei dieser Wellenlänge fluoresziert Ethidiumbromid, das sich durch Interkalation in der DNA angereichert hat. Die Gele wurden mit einer CCD-Schwarz-Weiß-Videokamera aufgenommen, ausgedruckt und abgespeichert. Proben, die optisch keine unerwünschten Banden im UV-Licht-Foto der Gel-Elektrophorese zeigten, wurden direkt in die Klonierung eingesetzt. Waren neben diesen PCR-Produkten Fragmente unerwünschter Länge zu erkennen, wurde die Proben erneut durch Gelelektrophorese aufgetrennt, unter kurzzeitiger UV-Licht-Kontrolle mit einem sterilen Skalpell ausgeschnitten und mittels QIAEX II Gel Extraction Kit von Gel-Resten befreit. Im Anschluß wurden auch diese Proben kloniert und transformiert.

### 2.2.8 DNA-Extraktion aus Agarosegel mit QIAEX II Gel Extraction Kit

Die Proben wurden auf einem 1%-Agarosegel elektrophoretisch aufgetrennt. Der Reinigungsvorgang beruht auf Elution des Agarosegels in Puffer QX1 und selektiver Adsorption von NS an QIAEX II Silica-Gel Partikeln in Gegenwart von chaotropischen Salzen. Auf diese Weise wird die DNA schonend, da ohne Einsatz von Phenol-Extraktion oder Ethanol-Präzipitation von Salzen, Agarose, Polyacrylamiden, Farbstoffen, Proteinen und Nukleotiden gereinigt. Die DNA kann laut Hersteller direkt in Restriktionsverdau, Ligation oder PCR weiterverarbeitet werden. Durch die geringe Entstehung von Scherkräften bei nur 30 Sekunden langer Zentrifugation mit 10.000 xg wurden lange DNA-Fragmente geschont. Die DNA-Bande wurde mit einem sterilen Skalpell aus dem Agarosegel geschnitten und in ein 1,5ml EpR überführt. Für 100 µg Gel wurden 300 µl Puffer QX1 und 10 µl QIAEX II hinzugefügt. Während der zehninütigen Inkubation bei 50°C wurde der Inhalt in zweiminütigen Abständen gemischt. Bei diesem Vorgang muss die Farbe des pH-Indikators im Puffer QX1 gelb bleiben. Im Anschluß wurde die Probe 30 sec. lang (10.000 xg) zentrifugiert. Der Überstand wurde verworfen und das Pellet mit 500 µl Puffer QX1 gewaschen und resuspendiert. Der Ansatz wurde erneut 30 sec. lang zentrifugiert und der Überstand verworfen. Es folgen zwei Waschschrte mit je 500 µl ethanolhaltigem Puffer PE. Nach Resuspendierung wurde der Ansatz jeweils 30 sec zentrifugiert und der Überstand verworfen. Anschließend trocknete das Pellet 10-15 min. an der Luft, bis es weiß wurde. Das Lösen des Pellets erfolgte durch Zusatz von 20 µl H<sub>2</sub>O, Resuspendierung und Inkubation der Elution für 5 min. bei RT. Nach erneuter Zentrifugation von 30 sec. wurde der Überstand mit der DNA in ein sauberes EpR überführt und bei -20 °C gelagert.

## 2.2.9 Klonierung und Transformation

Die Klonierung ist einerseits eine Alternative zur PCR um große DNA-Abschnitte zu vervielfältigen. Andererseits stellt sie sicher, dass im Folgenden während einer Versuchsreihe nur mit einem einzigen Klon einer Sequenz weitergearbeitet wird. Es kann davon ausgegangen werden, dass eine Bakterienzelle nur ein einziges Hybrid-Molekül aufnimmt [42]. Der Einbau von fremder DNA (Insert) in einen Vektor heisst ‚Ligation‘. Das Produkt aus Vektor und Plasmid-DNA wird ‚rekombinante DNA‘ genannt. Wird dieses Hybrid-Molekül in eine Empfängerzelle eingeschleust, wird von ‚Transformation‘ oder ‚Transfektion‘ gesprochen. Die amplifizierte DNA wurde in den aktivierten Plasmidvektor pCR® 4Blunt-TOPO® eingeschleust und kompetente E. coli mit der rekombinanten DNA transformiert. Das im Vektor enthaltene Ampicillin-Resistenzgen ermöglichte das Wachstum auf ampicillinhaltigen LB-Agar-Platten, während durch das Vorhandensein des Ampicillins im Agar verhindert wurde, dass sich andere Bakterien ungehindert ausbreiten konnten. Die Expression der LacZ $\alpha$ -*ccdB*-Genfusion verhinderte wiederum das Wachstum von Kolonien, die ein Plasmid ohne Insert enthielten (s.u.). Der hier verwendete Plasmidvektor pCR® 4Blunt-TOPO® besitzt ein Kanamycin- und ein Ampicillin-Resistenzgen. pUC ori ist die Startstelle für die Replikation des Vektors im Bakterium. Die direkte Selektion der rekombinanten Zellen wird durch Expression des Gens *ccdB* ermöglicht, dessen Produkt für E.coli toxisch ist. Dieses ist innerhalb des Vektors mit dem C-Terminus des LacZ $\alpha$ -Fragments fusioniert. Ligation mit dem blunt-end-PCR-Produkt verhindert die Expression der LacZ $\alpha$ -*ccdB*-Genfusion und erlaubt so nur das Wachstum von positiven Rekombinanten. Ein blue-white-Screening ist also bei Verwendung dieses Vektors nicht erforderlich.

### Bestandteile:

SOC-Medium

One Shot ®-Zellen (E. coli)	TOP 10
	DH5-alpha-T1
	JM 109

Salt Solution	1,2M NaCl	
	0,06 M MgCl <sub>2</sub>	
PCR® 4Blunt-TOPO®		
10 ng/μl Plasmid-DNA in:	50 % Glycerol	pH 7,4, 25 °C
	50 mM Tris-HCl	
	1mM EDTA	
	2 mM DTT	
	0,1 % Triton X-100	
	100 μg/ml BSA	
	30 μM Bromophenolblau	
dNTP-Mix	12,5 mM pro Base	pH 8,0
Primer-Konzentration	0,1 μg/μl TE-Puffer	pH 8,0
Primer	M 13 forward	
	M 13 reverse	
	T3 Primer	
	T7 Primer	
Control Template		
Control PCR Primers		
pUC18+19 Kontroll-DNA	10 ng/μl + 10 pg/μl	

Der Plasmidvektor pCR® 4Blunt-TOPO® ist herstellerseits linearisiert; an dem 3'-Ende ist die Topoisomerase I kovalent gebunden, die dort das Zucker-Phosphat-Rückgrat hinter dem Motiv 5'-CCCTT innerhalb der LacZα-ccdB-Genfusion spaltet. An dieser Stelle wurde im Folgenden das Insert eingefügt. Die freiwerdende Spaltungsenergie wurde in einer kovalenten Bindung zwischen dem 3'-Phosphat des gespaltenen Strang und einem Tyrosyl-Rest (TYR-274) der Topoisomerase I konserviert und stand für die anschließende Ligation zur Verfügung. Der Ansatz wurde in einem EpR zusammengefügt und 5 min. bei RT vermengt. Anschließend wurde das Reaktionsgemisch in einen Isolationsbehälter mit Eis gestellt.

Zur Vorbereitung für die Transformation von chemisch kompetenten E. coli Bakterien mit Plasmiden wurden die Zellen auf Eis aufgetaut. Ein 42 °C-Wasserbad wurde vorbereitet und das SOC-Medium auf RT angewärmt. Die Temperatur der selektiven LB-Ampicillin-Platten (100 µg/ml) betrug 37 °C.

TOPO® Cloning Reaction für chemisch kompetenten E.-coli:

PCR-Produkt	0,5- 4 µl
Salzlösung	1 µl
Steriles Wasser	ad 5 µl
TOPO Vector	1 µl
Endvolumen	6 µl

Es wurden 2 µl der Klonierungsreaktion in den Behälter mit E. coli überführt und das Gemisch für 5-30 min. zum Inkubieren in einen mit Eis gefüllten Isolationsbehälter gestellt. Durch einen 30 sekündigen 42 °C-Hitzenschock im Wasserbad wurde die Aufnahme der DNA in die Zellen induziert, welche anschließend wieder zurück auf das Eis gestellt wurden. 250 µl SOC-Medium wurden hinzugefügt, die Röhrchen verschlossen und bei 37 °C mit 200 rpm in horizontaler Lage geschüttelt. Nach einer Stunde wurde der Ansatz auf den angewärmten Agarplatten verteilt. Um die Wahrscheinlichkeit für ein gleichmäßiges Wachstum beziehungsweise eine günstige Koloniedichte zu erhöhen, wurden je 50 µl des Ansatzes auf eine und der Rest auf eine andere Platte verteilt. Die Kolonien wuchsen bei 37 °C über Nacht im Brutschrank. Nach der Inkubation wurden die Platten in den 4 °C-Kühlschrank gestellt.

**2.2.10 Anzucht in LB-Medium**

Aus jeder Petrischale wurden fünf Einzelkolonien gewählt, mit denen mittels Pipettenspitze jeweils ein Kulturröhrchen mit 5 ml LB-Medium angeimpft wurde. Die Röhrchen inkubierten unter Schütteln bei 154 rpm im 37 °C-Brutraum. Die

Petrischalen wurden für weitere Versuche im 4 °C-Kühlschrank gelagert. War über Nacht keine deutliche Trübung eingetreten, galt die Anzucht als fehlgeschlagen. Um im Anschluß an die Inkubation das weitere Wachstum in den positiven Kulturröhrchen einzuschränken, wurden diese im 4 °C-Raum aufbewahrt.

### **2.2.11 Aufreinigung mittels QIAprep Miniprep Kit**

Die Aufreinigung mittels QIAprep Miniprep Kit ist ein Verfahren zur Präparation von Plasmid-DNA für den Restriktionsverdau. Die aufgereinigte DNA ist für Sequenzanalysen geeignet. Prinzip ist die alkalische Lyse der Bakterienzellwand und die Bindung von DNA an an Silizium im salzigen Milieu. Die DNA wird an eine Silicatgel-Membran gebunden, gewaschen und anschließend in salzarmer Lösung wieder freigesetzt. Sie ist anschließend frei von Endonukleasen, Salzen, Ethanol, SDS und denaturierten Proteinen und Zelldebrid und kann im Anschluß an Restriktionsverdau und Gelelektrophorese mit UV-Licht-Photo direkt sequenziert werden. Zur Vorbereitung der Aufreinigung wurden 2 ml LB-Medium der Übernachtskultur in ein 2 ml EpR überführt und bei 13.000 x g 15 sec. zentrifugiert. Der Überstand wurde in ein Auffanggefäß gegossen und das EpR auf Zellstoff ausgeklopft. Dem Pellet wurde erneut 1 ml des Kulturmediums hinzugefügt und der Vorgang wiederholt. Die im Kulturröhrchen verbleibenden 0,75 ml LB-Medium wurden zu gleichen Teilen mit 70 % Glycerol versetzt und als Rückstellprobe im Kühlschrank bei – 80 °C aufbewahrt. Das bakterielle Pellet wurde in 250 µl Puffer P1 mit RNase A gründlich resuspendiert. 250 µl Puffer P2 wurden hinzugefügt und das EpR vier bis sechs mal invertiert. Puffer P2 enthält NaOH und SDS, welches Phospholipid- und Proteinkomponenten der Zellmembran löst. Das alkalische Milieu denaturiert chromosomale DNA und Proteine. Da auch Plasmid-DNA angegriffen wird, wurde dieser Vorgang nach Herstellerempfehlung auf weniger als 5 min. beschränkt. 350 µl Puffer N3 wurden hinzugefügt, das EpR erneut vier bis sechsmal invertiert und 10 min. zentrifugiert. Jeder Zentrifugationsschritt

dieser Methode erfolgt bei 10.000 xg (~ 13.000 rpm). Durch Puffer N3 wird das Lysat neutralisiert und auf salzreiche Bindungskonditionen eingestellt. Größere DNA-Stränge, SDS, denaturierte Proteine und Zelldebrid präzipitieren. Der Überstand wurde nach der Zentrifugation auf eine Säule (QIAprep Spin Column) überführt. QIAprep Spin Säulen enthalten eine Silicatgel-Membran, die selektiv DNA in salzreichem Puffer bindet und bei salzarmem Puffer wieder freisetzt. Die Säule wurde 30 sec. zentrifugiert und der Durchfluß verworfen. Zur Entfernung von Endonukleasen wurde die Säule mit 500 µl PB-Puffer gewaschen. Die Zentrifugation wurde wiederholt und der Durchfluß erneut verworfen. Um Salze und Ethanol auszuwaschen wurden 750 µl PE-Puffer auf die Membran aufgetragen und nacheinander je 30 sec. und 1 min. lang zentrifugiert. Dann wurde sie in ein steriles 1,5 ml EpR gestellt und die Mitte der Membran wurde mit 50 µl RNase/DNase-freiem Wasser (pH 7,0-8,5) bzw. Puffer EB (10mM Tris-Cl, pH 8,5) befeuchtet. Die Säule inkubierte daraufhin 5 min. bei RT und wurde anschließend 1 min. lang zentrifugiert. Um eine Hydrolyse zu vermeiden wurde der DNA-haltige Durchfluss bei -20°C aufbewahrt.

### **2.2.12 Restriktionsverdau der DNA**

Restriktionsendonukleasen sind Phosphodiesterasen bakteriellen Ursprungs, die DNA an spezifischen Palindrom-Sequenzen spalten. Das hier verwendete Restriktionsenzym EcoRI spaltet DNA spezifisch an der Sequenz GAATTC. Die Enden der Bruchstücke sind daher genau definiert. Aus Puffer, H<sub>2</sub>O und Restriktionsenzym wurde ein Mastermix erstellt. Pro Ansatz wurden 7 µl Mastermix mit 3 µl DNA vermischt. Der Reaktionsansatz inkubierte für eine Stunde bei 37 °C im Heizblock. Anschließend wurde mittels Agarosegel-Elektrophorese die Länge des Inserts ermittelt und anhand dessen abgeschätzt, ob es sich um das gewünschte PCR-Produkt handeln kann.

### 2.2.13 Sequenzierung

Der Sequenzierung liegt das Prinzip der 1975 von Sanger entwickelten Didesoxykettenabbruchmethode zugrunde. Einem PCR-ähnlichem Ansatz wurden zu gleichen Teilen die vier fluoreszenzmarkierten Basen als Didesoxynucleotide (ddNTPs) hinzugefügt. Ohne die Hydroxygruppe ist eine DNA-Verlängerung durch eine DNA-Polymerase in der PCR nicht möglich und es kommt zum Kettenabbruch an der Position der Sequenz, an der die jeweilige Base als ddNTP nach dem Zufallsprinzip eingebaut worden war. Im Gegensatz zur einer PCR wird bei der Sequenzierung nur ein Primer verwendet und die DNA nur linear amplifiziert. Die Sequenz des Primers ist durch die Verwendung des spezifischen Restriktionsenzym genau festgelegt. Im Anschluss an die Amplifizierung wurde der Ansatz mittels Gelelektrophorese aufgetrennt. Die dabei eingebauten Didesoxynucleotide wurden im Laser-Fluorometer angeregt und entsendeten infolge Fluoreszenz-Markierung Licht verschiedener Wellenlänge. Diese Veränderung ermöglicht auch die Automatisierung der DNA-Sequenzierung. Unter Verwendung des ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) wurden die Proben im ABI PRISM 310 Sequenzierer nach Herstellerprotokoll analysiert. Zur Amplifikation wurde die AmpliTaq DNA Polymerase FS I verwendet, welche pro Sequenzierungsvorgang ca. 600 Basen bearbeitet kann. Die Qualität der Sequenz wurde durch Betrachten ihrer Fluoreszenzmaxima mittels ABI 310 Software überprüft. Durch Primerwalking wurde der gesamte Strang in beide Richtungen vollständig sequenziert. Hierzu wurden aus den neuen Sequenzdaten Oligonukleotidprimer abgeleitet, deren Sequenzen im Materialteil einzeln aufgeführt sind. Zur Bestätigung der erhaltenen Daten wurden verschiedene cDNA-Klone sowie ungeklonte nPCR-Produkte jeweils in beiden Richtungen sequenziert und die Ergebnisse verglichen.

### Reaktionsansatz für eine Sequenzierung:

Premix (Big Dye)	2 µl
Primer (10 pmol/µl)	1 µl
DNA-Template	1,25 µl
H <sub>2</sub> O	5,75 µl
Reaktionsansatz	10 µl

Durch Resuspendieren wurde das Harz in den Säulen gelöst, der obere Deckel der Säule gelockert und der untere Deckel verworfen. Die Säule wurde in einem offenen 2 ml EpR 1 min. bei 2000 xg zentrifugiert. Der Durchfluß wurde verworfen und die in der Säule verbleibende Flüssigkeit auf Zellstoff ausgeklopft. Dann wurde die Säule in ein neues 1,5 ml EpR gestellt und ihr Deckel verworfen. 10 µl des Sequenzierungsansatzes wurden in die Mitte der Säule pipettiert. Nach erneuter Zentrifugation (1 min. bei 2000 xg) befand sich die gereinigte Probe im sterilen EpR und wurde 1-2 h in der Speed-Vac vakuumgetrocknet und anschließend in dieser Form bei -20 °C aufbewahrt oder direkt weiterverwendet. Zur Weiterverarbeitung wurde das Sediment in 15 µl Template Supression Reagent (TSR) resuspendiert und die DNA 2 min. bei 90 °C denaturiert. Der gesamte Ansatz wurde in ein Sequenzier-Tube überführt und mit einem Gummideckel versehen. Bis zur Sequenzierung wurde das Tube bei 4 °C aufbewahrt.

Die DNA im Ansatz wurde in 30 Zyklen 10 sec. auf 96 °C erhitzt, dann 5 sec. auf 55°C abgekühlt und anschließend 4 min. bei 60 °C amplifiziert. Im Anschluß wurden die Ansätze gemischt, auf ein Sequenziergel aufgetragen und elektrophoretisch aufgetrennt. Das Ergebnis wurde mittels Laser-Fluorometer ausgewertet.

### **2.2.14 Primerwalking**

Primerwalking ist ein Verfahren zur Sequenzierung langer DNA-Stränge, da die zur Sequenzierung verwendete DNA-Polymerase nur eine Anzahl von etwa 600 Nukleotiden sicher fehlerfrei liest und synthetisiert. Es wird eine Sequenz im

Endbereich des korrekt sequenzierten Strangs gewählt und anhand dieser ein neuer 20 nt Oligonukleotidprimer generiert. Um die Wahrscheinlichkeit für eine Korrekte Sequenzanalyse zu erhöhen, wurden alle Stränge in beide Richtungen sequenziert und jeweils miteinander verglichen.

## **2.3 Statistische und phylogenetische Analyseverfahren, Rekombinationsnachweis und strukturelle Charakterisierung von Rekombinationsarealen**

### **2.3.1 Überblick über die Vorgehensweise zur Ermittlung von Rekombination**

Unsere Vorgehensweise zur Aufdeckung von Rekombinationsereignissen lässt sich wie folgt gliedern. Zunächst untersuchten wir alle alignen Sequenzen im Hinblick auf Rekombinationsereignisse, ohne dabei die Rekombinations-schnittstellen exakt zu bestimmen. Um der Aussage der einzelnen Screeningmethode zusätzliches Gewicht zu verleihen, wurde der gesamte Datensatz parallel mit unterschiedlicher Methoden untersucht. Hierzu verwendeten wir Exploratory Tree Analysis, Similarity Plots und Sawyer's Test. Erstere sind häufig zum Nachweis von Rekombinationsereignisse bei RNA-Viren verwendete Methoden [43-45]. Für Sawyer's Test konnten eine hohe Potenz und Zuverlässigkeit in der Aufdeckung von Rekombinationsereignissen durch verschiedenen Computer-Simulations Studien nachgewiesen werden [46, 47].

Dieser Schritt zur Groborientierung war notwendig, da die im Folgenden verwendete Methode LARD gut zur genauen Lokalisation der Schnittstelle geeignet ist, jedoch fehleranfällig wird, wenn sich Rekombinationsereignisse in der Mitte der Sequenz befinden, so dass im Vorfeld Kenntnisse über die Lokalisation des Bruchpunktes gesammelt werden müssen [48]. LARD bestimmt die exakte Lokalisation der Rekombinationsschnittstellen unter Verwendung von *Likelihood Ratio* Analysen und einer Monte-Carlo Simulation

[44]. Zum Schluss wurden die Capsidregionen beidseits der jeweiligen Schnittstelle durch phylogenetische Analysen überprüft. Ein ähnliches Vorgehen wurden bereits durch andere Forschungsgruppen erfolgreich durchgeführt [43, 45]. Zuletzt wurde die RNA-Sekundärstruktur der Rekombinationsregion ermittelt. Die Hinweise auf Rekombination wurden ausschließlich als stichhaltig angesehen, wenn alle erwähnten Methoden ein kongruentes Ergebnis lieferten.

## **2.3.2 Statistische und phylogenetische Analyse**

### **2.3.2.1 Multiples Sequenzalignment, CLUSTALX**

Zu Beginn einer phylogenetischen Analyse werden in einer Menge an DNA-Sequenzen durch Erstellen eines multiplen Sequenzalignments orthologe Positionen einander zugeordnet. Es muss zunächst das Alignment gefunden werden, das die Gesamtkosten für alle paarweisen Alignments minimiert. Eine einzige Substitution innerhalb einer Sequenz kann dabei mehrfach zu den Gesamtkosten des Alignments beitragen, während *Gaps* in der Kakulation zu geringeren Kosten als Punktmutationen führen [49].

Der Verwandtschaftsgrad zwischen Sequenzen kann mit Hilfe eines multiplen Alignments rekonstruiert und mittels distanz- oder charakterbasierenden Methoden klassifiziert werden. Distanzmethoden, zum Beispiel die Neighbor-joining Methode, sind algorithmische Methoden. Der Baum entsteht aus einer Distanzmatrix durch progressives Clustering und erfordern einen geringeren Rechenaufwand als charakterbasierende Methoden. Diese benötigen einen hohen Rechenaufwand, da jede einzelne Sequenzposition (*character*) bewertet und das Ergebnis statistisch gesichert wird. Hier wurden ML-Methoden verwendet, da sie eine direkte Aussage über die Wahrscheinlichkeit des Baumes erlauben [35].

Das multiple Sequenzalignment wurde mittels CLUSTALX unter Benutzung der Standardeinstellungen durchgeführt [50]. CLUSTALX berechnet nach

paarweiser Alignment der einzelnen Sequenzen eine Distanzmatrix, welche die Basis zur Erstellung des Gesamtalignments bildet. Gemäß der errechneten Distanzen werden die Sequenzen zu einem sogenannten ‚Guide Tree‘ zusammengefügt [51].

### **2.3.2.2 Exploratory Tree Analysis (MODELTEST, LRT)**

Bei der explorativen Baumanalyse wurde ein 300 Basen großes Gleitfenster in 150-Basen-Schritten entlang des Capsidgen-Alignments bewegt. Für jede einzelne Region wurden dabei ML-Bäume konstruiert und der wahrscheinlichste Baum ausgewählt. Voraussetzung ist die Verwendung eines passenden Substitutionsmodells, welches die Likelihood-Funktion maximalisiert. Dieses wurde durch MODELTEST in Zusammenarbeit mit PAUP (phylogenetic analysis using parsimony) anhand des vorliegenden Datensatzes gewählt. PAUP berechnet und speichert für jedes Modell die Likelihood-Werte, MODELTEST wendet im Anschluß auf diese Werte einen Likelihood-Ratio Test (LRT) (Likelihood-Quotienten Test) oder Akaike’s Information Criterion (AIC) an, bis sich keine signifikante Verbesserung mehr ergibt [52]. Durch AIC kann der zu erwartende Klassifikationsfehler unter Annahme der Normalverteilung geschätzt werden. Das LRT vergleicht zwei Modelle direkt miteinander. Dabei ist ein Modell M1 (z. B. JC-Modell) ein Spezialfall des anderen Modells M2 (z.B. K2P), das heißt, es entspricht einander, wenn für das Modell mit der größeren Anzahl an Freiheitsgraden einzelne Parameter festgelegt werden. Also ist M2 in der Regel das ‚bessere‘ Modell, da es an den tatsächlichen Datensatz besser angepaßt werden kann. Der LRT prüft nun, ob M2 signifikant besser als M1 ist, also ob es vertretbar ist, durch Einführung weiterer geschätzter Parameter eine größere Ungenauigkeit bezüglich der Ergebnisse in Kauf zu nehmen. Das Ergebnis dieser Kalkulationen ist das optimale Substitutionsmodell nach LRT oder AIC, mit Hilfe dessen PAUP die Analyse durchführt. Folgende Parameter können für die Analyse genutzt und festgelegt werden: Frequenz der einzelnen Basen (A, C, G und T), Anzahl der zu verwendenden Substitutionstypen,

typspezifische Substitutionsrate, positionsspezifische Substitutionsrate nach Gammaverteilung mit acht Rate-Kategorien, Form-Parameter der Gammafunktion, Korrektur für invariante Positionen. Diese Parameter werden anhand des Datensatzes abgeschätzt. Alle Bäume, die von PAUP berechnet werden, sind prinzipiell ungewurzelt, können aber durch Einsatz einer Außengruppe nachträglich gewurzelt werden.

Wenn ein Strang in unterschiedlichen Capsidregionen verschiedene phylogenetische Positionen einnimmt, das heißt, wenn er sich nach Weiterrücken des Gleitfensters mit einem anderen Strang oder einer anderen Gruppe innerhalb des Baumes zusammenlagerte, wurde dies als Hinweis auf Rekombination gewertet. Diese Lokalisationen des Sequenzalignments wurde durch eine Similarity Plot Analyse mittels Simplot weiter untersucht, wie bereits bei anderen Forschungsgruppen beschrieben worden ist [53].

Die verschiedenen in Frage kommen Substitutionsmodelle unterscheiden sich deutlich in ihrer Komplexität und damit auch in der bei Verwendung zu erbringenden Rechenleistung. Die Bandbreite reicht vom wahrscheinlich einfachsten, dem Jukes-Cantor Modell (JC) über erweiterte Modelle wie dem Kimura 2 Parameter Modell (K2P), Felsenstein 1981 Modell (F81) und Hasegawa, Kishino und Yano 1985 Modell (HKY85) bis zum General Reversible Modell (REV). Während Ersteres voraussetzt, dass sich die Wahrscheinlichkeit für eine Substitution über die Zeit konstant verhält, jede Form der Substitution mit gleicher Wahrscheinlichkeit auftritt und alle vier Basen mit gleicher Häufigkeit vorkommen, erlaubt letzteres eine Höhere Rate an Transversionen versus Transitionen, Unausgewogenheit in der Basenzusammensetzung und ferner, dass für alle sechs Möglichkeiten der Substitution eine andere Wahrscheinlichkeit gilt. Dies kommt dem tatsächliche Sachverhalt relativ nahe. Jedoch muß jeder dieser flexiblen Parameter anhand des zu bearbeitenden Datensatzes abgeschätzt werden, um entsprechend in die Kalkulation mit einbezogen werden zu können. Da dies nur annäherungsweise geschehen kann, kann jeder zusätzliche Parameter in der Gesamtheit (zu großen sampling errors) zu einem Verlust der statistischen Potenz führen [49]. Es wird also anhand der vorliegenden Sequenzen das

Evolutionsmodell gewählt, das mit der Mindestanzahl an variablen Parametern auskommt und trotzdem eine ausreichend genaue Einschätzung liefert. Die relative Passgenauigkeit der Modelle wird abgeschätzt indem für jedes Modell die Wahrscheinlichkeit, den vorliegenden Datensatz zu erhalten, ermittelt wird. Nicht berücksichtigt wird die Inhomogenität der Substitutionsraten bezogen auf die verschiedenen Orte der Sequenz. Mit Hilfe dieser Methoden wird eine Likelihood-Matrix bezüglich der Substitutionsereignisse erstellt aus der ein phylogenetischer Baum generiert werden kann [49]. Die zur Verfügung stehenden Methoden lassen sich in distanzbasierte und charakterbasierte Verfahren unterteilen [54]. Den hier verwendeten Programmen liegt ein charakterbasiertes Verfahren, die Maximum-Likelihood-Methode (ML-Methode) zugrunde. Diese wird benutzt, um unbekannte Parameter zu schätzen, von denen eine bekannte Wahrscheinlichkeitsfunktion für einen stochastischen Prozeß abhängt. Das bedeutet, anhand einer festen Stichprobe, der Sequenzdaten, werden unbekannte Parameter so geschätzt, dass der Wert der Likelihood-Funktion sein Maximum erreicht. Dabei steht „Likelihood“ nicht für die Wahrscheinlichkeit den richtigen Baum zu erhalten, sondern dafür, dass der erhaltene Baum die gesammelten Daten hervorgebracht hat [49]. Somit kann mit diesem Modell die Wahrscheinlichkeit  $L$  für den Erhalt der beobachteten Daten berechnet werden. Die Vorteile der Maximum-Likelihood Methode gegenüber distanzbasierten Verfahren (z. B. Maximum Parsimony) liegen in der kleinen Varianz der Ergebnisse, die selbst bei kleinen Sequenzen relativ zuverlässig sind, darin, dass das Verfahren statistisch wohl begründet ist und dass alle Sequenzinformationen genutzt werden [54]. Der Nachteil liegt in der hohen erforderlichen Rechenleistung.

### **2.3.3 Rekombinationsnachweis**

#### **2.3.3.1 GENECONV, Sawyer's Test**

Das erstellte Gesamtalignment wurden mittels GENECONV bearbeitet. Das Programm verwendet Sawyer's Test [55].

Sawyer's Test ist eine statistische Methode zur Erkennung des Austauschs von genetischem Material zwischen NS-Strängen und somit zum Nachweis von Rekombinationsereignissen geeignet. Das Programm untersucht ein gegebenes Alignment von zwei NS-Strängen auf das Vorhandensein langer Segmente identischer Nukleotidabfolge, die einen Übertritt von genetischem Material andeuten können [55]. Der Test nutzt dazu das Ungleichgewicht der Verteilung identischer Segmente innerhalb der alignierten Sequenzen, *imbalance of distribution*, und überprüft gleichzeitig das Vorhandensein von heterogenen Substitutionsraten entlang des Genoms. Dabei ist eine erfolgreiche Durchführung unabhängig von der Vorabwahl potentiell zu einer monophyletischen Gruppe gehörenden Stränge [56]. Sawyer's Test toleriert innerhalb der Alignierung keine *Gaps*, so dass diese vor der Anwendung entfernt werden mußten [56].

GENECONV listet in Frage kommende Ereignisse ihrem P-Werten nach abgestuft und in einer Kalkulationstabelle auf [57]. Die P-Werte aller globalen Fragmente wurden durch Multiple Comparison Tests bezüglich aller möglichen Sequenzpaare korrigiert. Gemäß der Standardeinstellungen wurden 10.000 Permutationen durchgeführt. Als statistisch signifikant bezüglich der Rekombinationsereignisse wurde ein globaler Permutations P-Wert von  $<0,05$  festgelegt.

### **2.3.3.2 Splits Tree**

Die von uns verwendete Version 2.4 des Programms SplitsTree (Wetzel und Hudson) nutzt die Split-Decomposition Methode von Bandelt und Dress um phylogenetische Bäume zu konstruieren [58]. Diese Methode berücksichtigt, dass phylogenetische Daten in Evolutionsanalysen unterschiedliche, zum Teil widersprüchliche Signale enthalten, die zur Konstruktion baumartiger Netzwerke, anstelle sich dichotom verzweigender Bäumen genutzt werden können [58]. Bei der phylogenetischen Baumkonstruktion wurde die Kimura 3-ST-Methode angewendet. Bildete sich innerhalb des anhand unserer Daten

konstruierten Baumes ein solches Netzwerk, bestand der Verdacht auf ein Rekombinationsereignis. Die Knotenpunkte des Baumes wurden mittels Bootstrap-Analyse (1000 Permutationen) untersucht und die Lokalisationen der netzartigen Verzweigungen wurden durch Exploratory Tree Analysis und mit SimPlot erneut bearbeitet.

### **2.3.3.3 Similarity Plots**

SimPlot berechnet paarweise für jede einzelne Capsidregion die prozentuale Differenz zwischen der fraglichen Mosaiksequenz und einer Gruppe von Referenzsträngen und stellt diese graphisch dar. Auf der X-Achse werden die Similarity Scores in Prozent abgetragen, auf der Y-Achse deren exakte Position in Basenpaaren. Die einzelnen Regionen entstanden, indem ein 200 Basen großes Gleitfenster in 20-Basen-Schritten entlang des Nukleotidalalignments bewegt wurde. Anhand der resultierenden Ähnlichkeitsprofile (*similarity profiles*) konnte abgelesen werden, welche Stämme alternierend eine starke Sequenzähnlichkeit mit dem fraglichen Rekombinationsstamm zeigen. Bewegte sich die Sequenzidentität zwischen 90 und 95% wurden die betroffenen Stränge als mögliche Parentalstämme angesehen. Aus den Daten kann zusätzlich für jede Position ein phylogenetischer (neighbor-joining) Baum erstellt werden [58]. Rekombinationsverdächtige Stämme wechseln dabei innerhalb der Topologie des Baumes die Position [58].

### **2.3.3.4 Bruchpunktanalyse mittels LARD**

Nach Identifikation der wahrscheinlichsten Parentalsequenzen schätzten wir mittels ML-Methode und Likelihood Ratio Test die exakten Rekombinationsschnittstellen ab. Dazu wurde LARD (*Likelihood Analysis of Recombination in DNA*) Version 1.0 verwendet [43-45].

LARD errechnet für das Sequenzalignment einer möglichen Rekombinanten und deren potentiellen Parentalstämmen den optimalen Bruchpunkt einer Rekombination [59]. Dabei besagt die Arbeitshypothese H1, dass auf jeder Seite des Rekombinationsbruchpunktes zwei unterschiedliche Bäume existieren, nämlich die des jeweiligen Parentalstamms [48]. Folglich zerteilt das Programm die Strämme am potentiellen Bruchpunkt und konstruiert für die Sequenzen jeder Seite einen separaten ML-Baum. Dieser Vorgang wird für jede mögliche Region, in der es in den vorangegangenen Untersuchungen Hinweise auf Rekombination gab, durchgeführt. Die ML-Werte der einzelnen Bruchpunkte  $b$  werden aufgelistet. Der Höchste Wert wird erwartet, wenn die Sequenz am tatsächlichen Bruchpunkt unterbrochen wurde, da dann dem rekonstruierte Baum am ehesten die tatsächliche Phylogenie zu Grunde liegt [59].

Um die statistische Signifikanz des Ergebnisses zu überprüfen, führt LARD einen Likelihood Ratio Test und eine Monte-Carlo (MC) Simulation durch [35]. Dazu wird der Likelihood-Wert des fraglichen Bruchpunkts mit dem Likelihood-Wert des ungebrochenen Sequenzalignment verglichen. Die MC-Simulation ist ein stochastischer Simulationsalgorithmus, der der Generierung synthetischer Daten zur statistischen Auswertung dient [60]. Der Ansatz simuliert hier Sequenzen basierend auf dem bekannten Datensatz unter Ausschluß der Möglichkeit der Rekombination, also unter Annahme der Nullhypothese H0, jedoch unter Berücksichtigung der lokalisationspezifischen „*rate heterogeneity*“ der vorliegenden Daten [59]. Diesbezüglich ermöglicht das Programm sowohl das Einstellen von unterschiedlichen Substitutionsraten für jede Codonposition als auch die Berücksichtigung von unterschiedlichen Substitutionsraten für einzelne Sites [61]. Der erhaltene Datensatz wird mit den gleichen Verfahren analysiert wie die Originaldaten [44]. Dann wird für jedes in Frage kommende  $b$  die Differenz  $\Delta L$  der Wahrscheinlichkeiten  $L$  beider Hypothesen H0 und H1 errechnet ( $\Delta L = L(H1) - L(H0)$ ). Der Wert  $b$ , der zum größten Wert von  $\Delta L$  führt, ist die wahrscheinlichste Rekombinationsschnittstelle. Um die statistische Signifikanz der Likelihood Ratio zu überprüfen, wurde nun durch X<sup>2</sup>-Test und MC-Simulation in Seq-Gen [62] ermittelt, ob  $\Delta L$  signifikant größer als Null ist [48, 49, 63, 64].

### 2.3.3.5 Phylogenetische Stammbäume und Bootstraps

Im Anschluss an die Identifikation der Rekombinationsschnittstellen wurden unter Verwendung der entsprechenden Evolutionsmodelle für jede rekombinante Sequenz separate ML-Bäume konstruiert. Es wurde ein passendes, anhand der Daten ausgewähltes Substitutionsmodell verwendet und die bei der Konstruktion entstehenden phylogenetischen Konflikte mittels Bootstrapanalyse durch 1000 Permutationen überprüft. Die Gefahr falsch positiver Resultate durch Heterogenität der Substitutionsraten an verschiedenen Lokalisationen wurde durch Abschätzen der *among-site* Variation anhand des vorhandenen Datensatzes minimiert. Unter diesen Umständen wurde ein phylogenetischer Konflikt als statistisch signifikant angesehen, wenn der entsprechende Bootstrap-Wert größer als 75 % von 1000 Permutationen an bedeutenden Knotenpunkten war. Diese Vorgehensweise ist angelehnt an Chare et al. [43].

Im folgenden Schritt wurden der Stichprobenfehler der zuvor detektierten phylogenetischen Konflikte mittels Pseudoreplikation durch Bootstrap-Analyse abgeschätzt (Felsenstein, 1985). Dieses Verfahren gewinnt aus den vorhandenen Daten einer Matrix unabhängige Stichproben (Pseudomatrizen), indem aus dem originalen Datensatz zufällig Positionen (Matrixspalten) gezogen und wieder zurückgelegt werden. Durch dieses Verfahren können manche Positionen mehrmals in der Pseudostichprobe vorhanden sein, andere kommen dagegen gar nicht vor. Dieser Vorgang wird so oft wiederholt, bis die Pseudomatrix die gleiche Größe wie die Originalmatrix aufweist. Der Prozess der Generierung von Pseudostichproben wurde hier 1000 mal wiederholt, aus jedem Datensatz wurde ein separater Stammbaum erstellt und aus allen Bäumen ein Konsensusbaum. Die Relevanz der einzelnen Knotenpunkte sowie die der Grundzüge des gesamten erstellten Baumes kann anhand des Bootstrap-Wertes, der die Häufigkeit des Merkmals unter den generierten Bäumen in Prozent angibt, abgeschätzt werden. Je größer der Wert, desto stärker werden Genauigkeit und Aussagekraft unterstützt [54]. Wurde ein phylogenetischer Konflikt durch einen Bootstrapwert von mehr als 75 % der 1000 Replikationen belegt, nahmen wir diesen als tatsächlichen phylogenetisch

relevanten Knotenpunkt an. Dieses Vorgehen wurde bereits durch andere Gruppen beschrieben [43]. Um die Wahrscheinlichkeit falsch positiver Resultate durch *rate-heterogeneity* zu minimieren, wurde die Substitutionsrate der einzelnen Sites anhand der Sequenzdaten abgeschätzt und in die Berechnungen bezüglich der phylogenetischen Inkongruenzanalyse einbezogen.

#### **2.3.4 Strukturelle Charakterisierung von Rekombinationsarealen**

Studien haben gezeigt, dass der Sequenzbereich auf beiden Seiten eines Bruchpunkts besondere Charakteristika bezüglich der Sequenzkomposition aufweist. Wir haben die Rekombinations-Hotspots im Hinblick darauf und auf den Prozentsatz an übereinstimmenden Nukleotiden in beiden Parentalsequenzen untersucht. Zusätzlich wurde die wahrscheinliche RNA-Sekundärstruktur der Rekombinationsregion in einem Bereich von > 50 nt, wie durch anderen Gruppen durchgeführt, bestimmt [65-68].

Die voraussichtliche RNA Sekundärstruktur der Rekombinationsregion wurde mittels Mfold Web Server für Nucleic Acid Folding and Hybridisation Prediction ermittelt [69]. Mfold basiert auf dem Prinzip der Energieminderung, mit dem Ziel, die thermodynamisch optimale Sekundärstruktur einer gegebenen Nukleinsäuresequenz zu berechnen. Dazu wird ein sogenannter *Energy dot plot* erstellt [69]. In der Version Mfold 3.1 erfolgen alle Berechnungen für eine Temperatur von 37°C [70].

#### **2.4 GenBank Accession Numbers der neu charakterisierten Norovirus Stämme**

Die vollständigen Sequenzen der untersuchten Stämme sind unter den GenBank Accession Numbers AY532111 bis AY532135 zugänglich. Die einzelnen Nummern sind im Ergebnisteil aufgeführt.

### **3 ERGEBNISSE**

Aus unseren klinischen Proben konnten fünfundzwanzig bisher nicht beschriebene NV-Capsidsequenzen ermittelt werden. Die Proben stammten von Stuhlproben an GE erkrankter Kinder aus fünf deutschen Städten, Freiburg, Erlangen, Dresden, Hamburg und Bochum, und waren im Zeitraum von 14 Monaten, März 1997 bis April 1998, gesammelt worden.

#### **3.1 Erstellen und Analyse der DNA-Sequenzen**

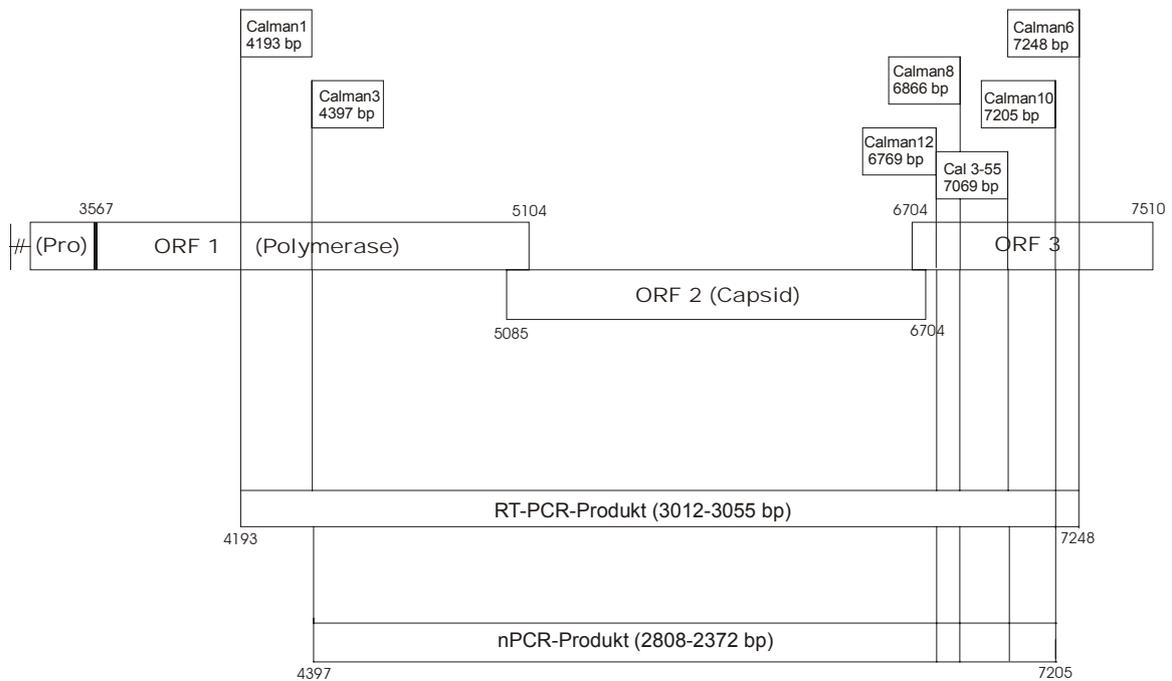
##### **3.1.1 Amplifikation mittels RT-PCR und Dokumentation durch UV-Licht-Photo**

Die aus Stuhlproben gewonnenen NS wurden mittels RT-PCR amplifiziert. Folgende Primer wurden in Kombination verwendet: Calman-1 (forward) und Calman-6 (reverse), sowie Calman-1 (forward) und Calman-10 (reverse), siehe Methodenteil. Das erwartete RT-PCR-Produkt war etwa 3 kb lang (Tabelle 1, Abbildung 8) und umfasste jeweils den gesamten ORF2 (1,6 kb), den größten Teil des ORF3 (0,8 kb) und 0,8 kb des ca. 5 kb langen ORF1. Im von uns amplifizierten Fragment waren somit die Informationen der gesamten Capsid-Region, des ORF3 und mehr als die Hälfte der ca. 1,5 kb langen Polymerase-Region von ORF1 enthalten.

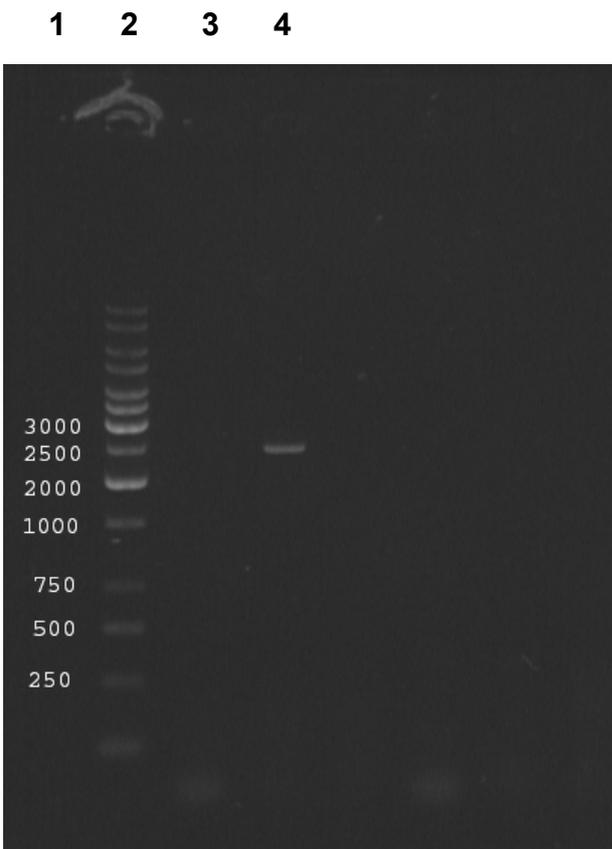
Im Anschluss an die RT-PCR wurde mittels Gelelektrophorese, UV-Licht-Photo und Fragmentlängenvergleich zu einem Standard ermittelt, welche Ansätze das erwartete Fragment enthalten konnten. Dieser Schritt wurde als grobes Screening gewertet und diente darüber hinaus dazu auszuschließen, dass unerwünschte Fragmente, die spätere Reaktion stören könnten, in optisch nachweisbarer Menge mit amplifiziert worden waren.

Methode	antisense→	Calman 6	Calman 10	Calman 8	Calman 12
	sense ↓				
RT-PCR	Calman 1	3,055	3,012	2,673	2,576
nPCR	Calman 3	2,851	2,808	2,469	2,372

**Tabelle 1: Länge der amplifizierten Fragmente in Kilobasen.**



**Abbildung 8: Schematische Darstellung des NV-Genoms.** Die vertikalen Linien begrenzen das Amplikon an der Stelle, an der die jeweiligen Primer binden. Unterhalb des Namens des Primers ist die genaue Position in bp (= Basenpaaren) angegeben und im unteren Bildteil sind die erwarteten PCR-Produkte mit Längenangabe in bp abgebildet. Die Länge der RT-PCR-Produkte liegt zwischen 3012 nt und 3055 nt, und die des n-PCR-Produkts zwischen 2808 nt und 2372 nt.

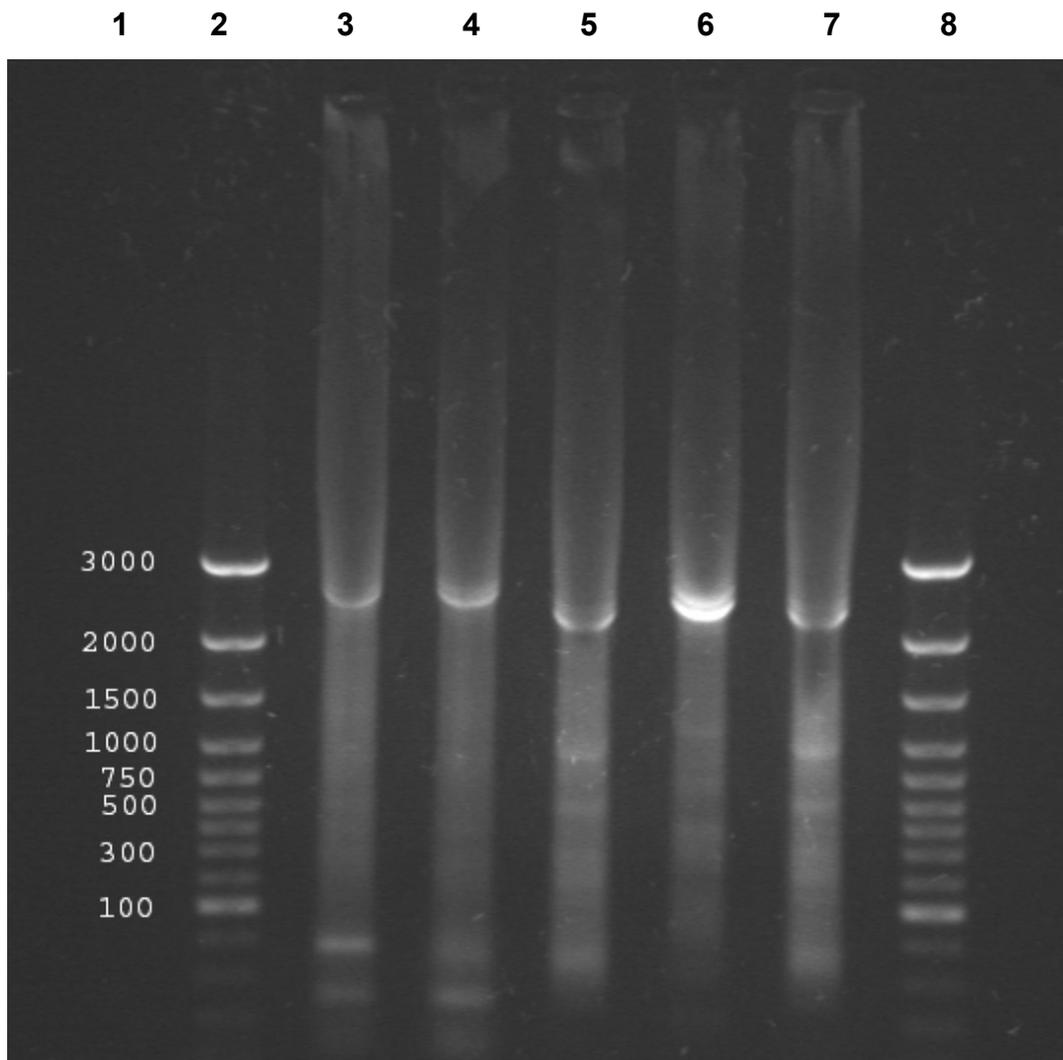


- 1: Länge in Basen
- 2: Längenstandard
- 3: Negativ-Kontrolle
- 4: Probe (Bochum 188)

**Abbildung 9: UV-Licht-Photo der RT-PCR Proben in der Gel-Elektrophorese.** Die Fragmentlängen werden im Vergleich zu einem Standard (links im Bild) abgeschätzt. Hat das amplifizierte Fragment die erwartete Länge, kann es sich um die gewünschte Sequenz handeln.

### 3.1.2 Amplifikation mittels nPCR und Dokumentation durch UV-Licht-Photo

Um die Fragmentspezifität zu erhöhen und die amplifizierte NS-Menge zu vergrößern, wurde jede Probe mittels nPCR mit NV-spezifischen Primern, die sich innerhalb der zuvor amplifizierten Sequenz anlagerten, erneut bearbeitet. Als antegrader Primer wurde in allen Ansätzen Calman-3 und als reverser Primer jeweils Calman-8 und Calman-12 verwendet. Alle Produkte der RT-PCR Primer-Kombination Calman-1 und -6 wurden zusätzlich mit Calman-3 (forward) und Calman-10 (reverse) amplifiziert.

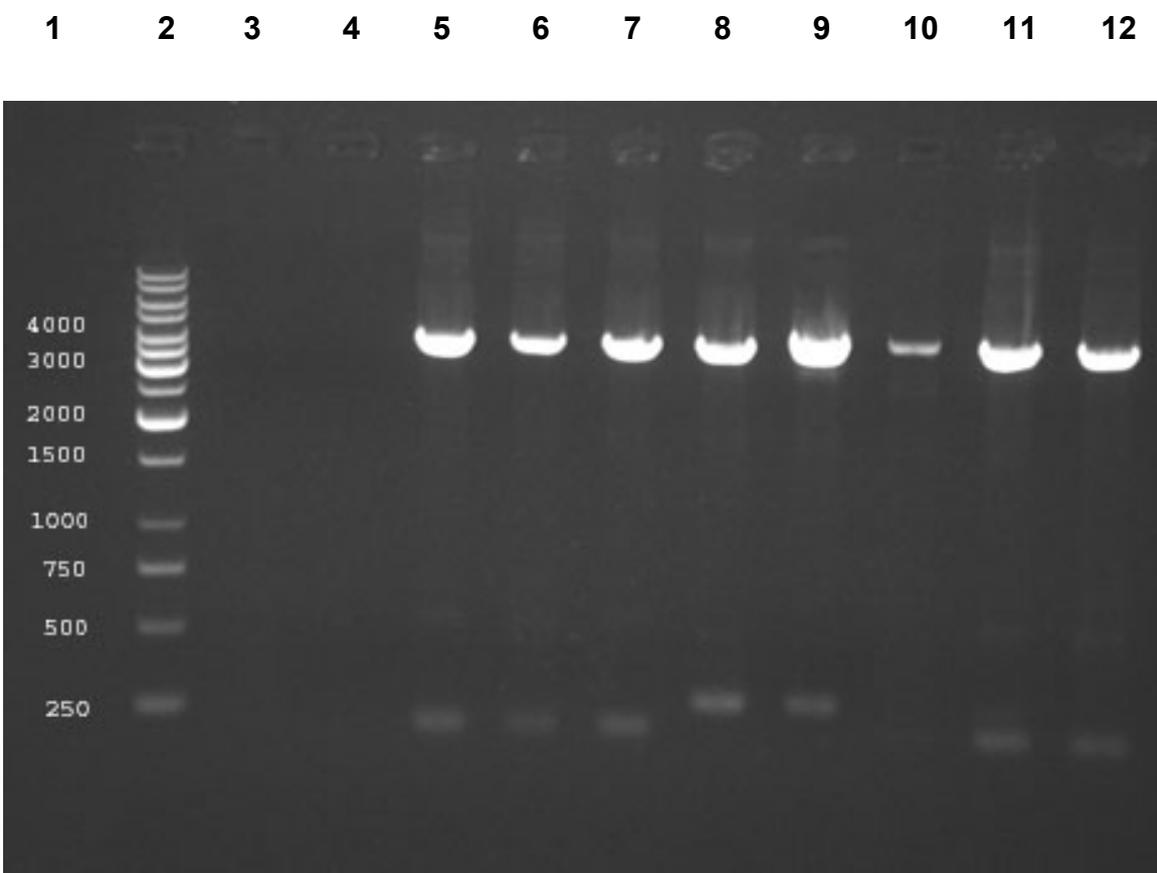


**Abbildung 10: UV-Licht-Photo der Gel-Elektrophorese der nPCR. Probe: Dresden174 amplifiziert mit unterschiedlichen Primerpaaren.** Der Längenstandard befinden sich jeweils rechts und links im Bild (2 und 8). 1 = Länge in bp, 2 und 8 = Längenstandards, Kombination der Primerpaare: 3 = C1-3/C3-4, 4 = C1-3/C3-55, 5 = C1-3/C3-8, 6 = C1-3/C3-6, 7 = C1-3/C3-8.

Die im UV-Licht-Photo zu erwartenden Fragmente lagen zwischen 2,37 kb und 2,82 kb und umfassten so die gesamte Capsidsequenz und Teile von ORF1 und ORF3, siehe Tabelle E1.

### 3.1.3 Präparierung der DNA, Klonierung, Transformierung, Kultur, Aufreinigung, Sequenzierung, Primerwalking

Das im Vektor enthaltene Ampicillin-Resistenzgen ermöglicht den entsprechenden E. coli das Wachstum auf LB-Agar-Ampicillin-Platten, während die Expression der LacZ $\alpha$ -ccdB-Genfusion das Wachstum von Kolonien, die ein Plasmid ohne Insert enthalten verhindert. Es wuchsen darum nur E. coli Kolonien, die ein Plasmid mit gewünschtem Insert aufgenommen hatten.



**Abbildung 11: UV-Licht-Photo der Gel-Elektrophorese nach Aufreinigung.** 1 = Länge in bp, 2 = Längenstandard, 3 = Negativ-Kontrolle, 4 = Erlangen/06 0675, 6 - 12 = Erlangen/05 118

Im Anschluss an Klonierung und Transformation wurde LB-Medium angeimpft und inkubiert als Übernachtskultur. Kulturröhrchen mit deutlicher, homogener Trübung wurden weiterverarbeitet. Die Plasmid-DNA wurde mittels QIAprep Miniprep Kit (Qiagen, Hilden) für die Sequenzierung aufgereinigt und im

Restriktionsverdau durch die Restriktionsendonuklease EcoR1 bei dem Motiv GAATTC gespalten. Im UV-Lichtfoto nach Gelelektrophorese wurden die Fragmentlängen abgeschätzt (Abb. 11).

Die Proben wurden mittels ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) sequenziert. Es konnten pro Sequenzierungsvorgang ca. 600 Basen bearbeitet werden. Die Qualität der Sequenz wurde durch Betrachten ihrer Fluoreszenzmaxima mittels ABI 310 Software überprüft. Durch Primerwalking wurde das gesamte Amplikon in beide Richtungen sequenziert. Die Sequenzen der Primer sind im Methodenteil gesondert aufgeführt. Die Daten wurden durch Sequenzieren von cDNA-Klonen sowie unklonierten nPCR-Produkten verifiziert. Es wurden keine Konsensussequenzen erstellt. Die von uns erstellten Sequenzen sind unter den in Tabelle 2 aufgeführten GenBank accession numbers einzusehen.

<b>Stamm</b>	<b>Herkunft</b>	<b>Genet. Cluster</b>	<b>GenBank accession no.</b>
Freiburg057/1997/GE	Deutschland	II/4	AY532111
Freiburg204/1997/GE	Deutschland	II/4	AY532112
Freiburg253/1998/GE	Deutschland	II/4	AY532113
Erlangen195/1997/GE	Deutschland	II/4	AY532114
Dresden153/1997/GE	Deutschland	II/4	AY532115
Dresden245/1997/GE	Deutschland	II/4	AY532116
Dresden267/1997/GE	Deutschland	II/4	AY532117
Dresden319/1997/GE	Deutschland	II/4	AY532118
Bochum024/1998/GE	Deutschland	II/4	AY532119
Bochum026/1997/GE	Deutschland	II/4	AY532120
Bochum031/1997/GE	Deutschland	II/4	AY532121
Bochum108/1997/GE	Deutschland	II/4	AY532122
Bochum136/1998/GE	Deutschland	II/4	AY532123
Bochum220/1997/GE	Deutschland	II/4	AY532124
Bochum224/1998/GE	Deutschland	II/4	AY532125
Bochum272/1998/GE	Deutschland	II/4	AY532126
Bochum339/1997/GE	Deutschland	II/4	AY532127
Hamburg048/1997/GE	Deutschland	II/4	AY532128

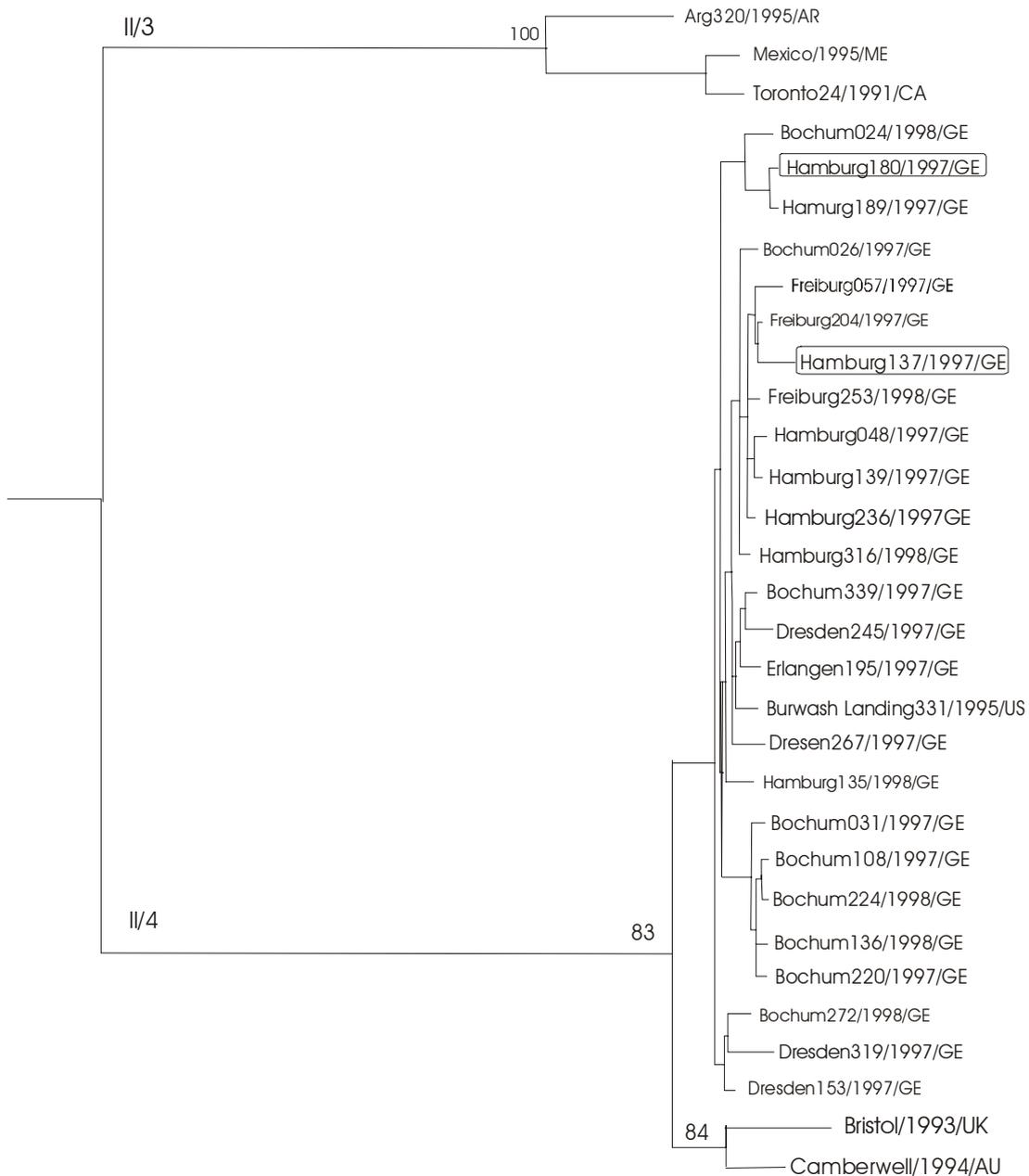
Hamburg135/1998/GE	Deutschland	II/4	AY532129
Hamburg137/1997/GE	Deutschland	II/4	AY532130
Hamburg139/1997/GE	Deutschland	II/4	AY532131
Hamburg180/1997/GE	Deutschland	II/4	AY532132
Hamburg189/1997/GE	Deutschland	II/4	AY532133
Hamburg236/1997/GE	Deutschland	II/4	AY532134
Hamburg316/1998/GE	Deutschland	II/4	AY532135

**Tabelle 2: GeneBank accession numbers der von uns amplifizierten und sequenzierten Stämme.** Alle Stämme waren in deutschen Städten gesammelt worden und clusterten in GG II/4 (s.u.).

### **3.2 Alignment und Phylogenetische Analyse der Norovirusstämme basierend auf vollständigen Capsidsequenzen.**

Bei der Untersuchung der NV-Capsidsequenzen im Hinblick auf Rekombination mittels Sawyer's Test, Exploratory Tree Analysis und Similarity Plots wurden zwei wahrscheinlich rekombinante Stämme und deren Parentalstämme identifiziert. Sie stammten jeweils von unterschiedlichen Patienten. Im Folgenden werden die Ergebnisse der Identifikationsschritte beschrieben. Sawyer's Test (Tabelle 3) zeigte deutliche Hinweise auf Genkonversionsereignisse, die im Einklang mit dem Clusterwechsel einzelner Stämme bei der *Exploratory Tree Analysis* entlang des Sequenzalignments standen.

Durch Similarity Plots und Bruchpunkt Analyse wurden die Rekombinanten identifiziert, die entsprechenden Parentalstämmen zugeordnet und die Rekombinationsschnittstelle genau lokalisiert (Abbildung 14, Tabelle 4). Durch Konstruktion von ML-Bäumen beidseits jeder einzelnen Rekombinationsschnittstelle (Abbildung 15) wurden die Ergebnisse bestätigt. Hohe Bootstrap-Werte von mehr als 80 % bei 1000 Permutationen an den entsprechenden Schnittstellen verifizierten diese und stützten die Hinweise auf phylogenetischen Konflikte innerhalb der jeweiligen ML-Bäume. Wir konstruierten für beide von uns gefundenen Rekombinanten und deren Parentalstämme insgesamt fünf ML-Bäume.



**Abbildung 12 : Evolutionäre Verwandtschaftsverhältnisse innerhalb der NV-GG II/4 basierend auf der Nukleotidabfolge der gesamten Capsidsequenz.** Der ML-Baum wurde unter Benutzung des General Time Reversible Substitutionsmodells, einer Gamma-Verteilung für amog-site Variation (mit acht Raten) sowie Korrektur für nicht variante Genorte und der Basenzusammensetzung der einzelnen Sequenzen erstellt. Die einzelnen Faktoren wurden anhand des von uns gewonnenen Datensatzes abgeschätzt. An den einzelnen Knotenpunkten sind die Bootstrapwerte als prozentualer Anteil von 1000 Permutationen eingetragen. Um eine verbesserte Übersicht zu erhalten, ist der Baum mittig gewurzelt und jeder einzelne Ast skaliert gezeichnet. Die rekombinanten Stämme Hamburg180/1997/GE und Hamburg137/1997/GE sind durch Umrandung optisch hervorgehoben.

Der oben gezeigte ML-Baum liefert einen Überblick über die verwandtschaftlichen Verhältnisse aller von uns isolierten und sequenzierten neuen NV-Stämme mit Angabe der Bootstrap-Werte als Prozentanteil von 1000 Permutationen. Die rekombinanten Stämme Hamburg180/1997/GE und Hamburg137/1997/GE sind durch Kästchen mit Pfeil optisch hervorgehoben. Alle von uns gefundenen Stämme clusterten gemeinsam mit den von uns in späteren Schritten als Außengruppe zum Wurzeln verwendeten Stämme Bristol/1993/UK und Camberwell/1994/AU innerhalb der NV GG II/4. Die Abbildung stellt weitere Stämme der Genogruppe II dar, deren Capsidsequenzen aus der GenBank Datenbank stammten. Bezogen auf die phylogenetische Darstellung von 100 bereits bekannten in der GenBank verfügbaren Sequenzen, handelte es sich um das Cluster 4 der Genogruppe II bei insgesamt fünfzehn unterschiedlichen genetischen Clustern [71]. In Cluster II/4 befanden sich 37 Sequenzen. Hohe Bootstrap-Werte unterstützten die Architektur der relevanten Knotenpunkte.

### **3.2.1 Sequenzalignment und Screening auf Rekombinationsereignisse**

Mit den Capsidsequenzen wurde durch CLUSTALX ein multiples Sequenzalignment erstellt und durch Bearbeitung mit GENECONV die wahrscheinlichsten Partner für eine stattgehabte Genkonversion, beziehungsweise Rekombination festgelegt. GENECONV untersucht mittels Sawyer's Test ein Sequenzalignment auf lange Abschnitte identischer Nukleotidabfolge und ermittelt so Regionen, die zwischen verschiedenen Stämmen ausgetauscht worden sein können. Die Ergebnisse von Sawyer's Test lieferten deutliche Hinweise auf Genkonversionsereignisse und sind in Tabelle 3 dargestellt. In dieser wird die Sequenz des möglichen rekombinanten Stamms gezeigt, die jeweils zu einem Abschnitt der Capsidregion eines anderen Stamms identisch ist.

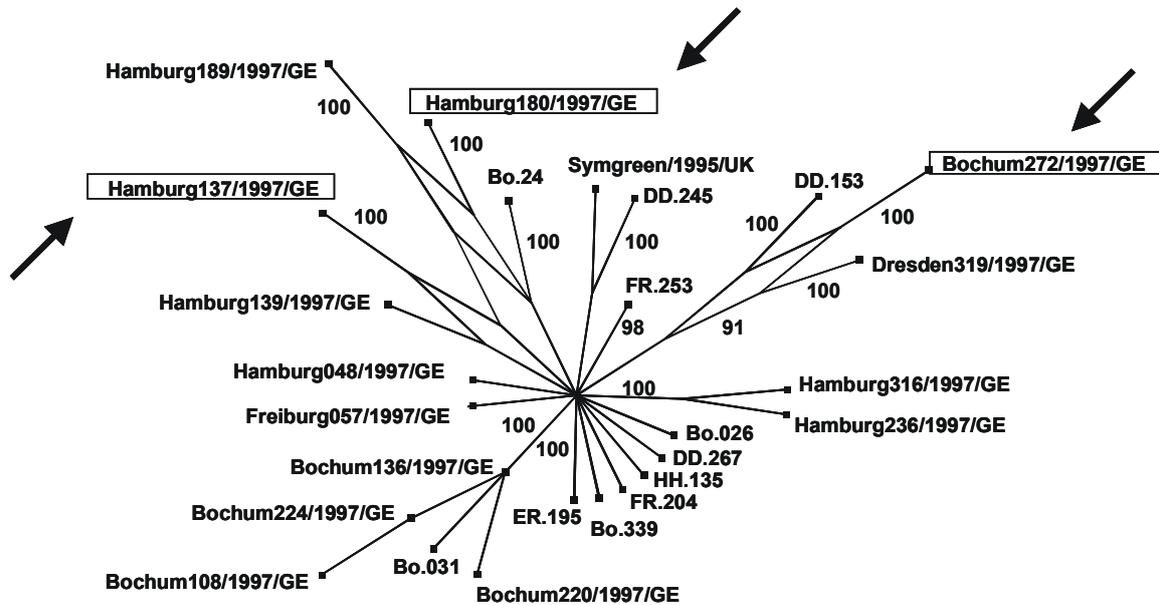
Rekombinante Stämme	Parentalstränge	Capsidregion (nt)	Länge (nt)	P-Wert
Hamburg180/1997/GE	Bochum024/1998/GE	580–761	182	0.0007
Hamburg137/1997/GE	Hamburg139/1997/GE	1166–1696	531	0.0100

**Tabelle 3: Sawyer's Test.** Capsidregionen, in denen sich die Nukleotidabfolge der angegebenen Stämme entsprechen und die Länge dieser identischen Sequenzen. Der aufgeführte P-Wert ist jeweils der globale P-Wert für den gesamten im Interesse liegenden Bereich des Alignments.

### 3.2.2 Phylogenetischer Konflikt dargestellt durch SplitsTree

SplitsTree konstruiert anhand der Sequenzdaten einen phylogenetischen Baum, der an verschiedenen Stellen netzwerkartige Verzweigungen zeigte. Jede dieser Verzweigungen ist ein Hinweis auf widersprüchliche Informationen in Bezug auf die zugrunde liegenden Verwandtschaftsverhältnisse. Die Phylogenie des Baumes zeigt somit, dass keine eindeutige Beziehung im Sinne einer dichotomen Verzweigung zwischen den entsprechenden Stämmen hergestellt werden konnte.

Dieses Phänomen trat zwischen den Stämmen Hamburg180/1997/GE, Hamburg189/1997/GE und Bochum024/1998/GE, zwischen Hamburg137/1997/GE, Hamburg189/1997/GE und Hamburg139/1997/GE sowie Bochum272/1987/GE, Dresden319/1997/GE und Dresden153/1997/GE auf. Da die im Folgenden verwendeten Methoden nur in den zwei erstgenannten Stämmen mit ausreichender Wahrscheinlichkeit Rekombinationsereignisse nachweisen konnten, wurde der Stamm Bochum272/1987/GE nicht als rekombinanter Stamm in die Endergebnisse aufgenommen.



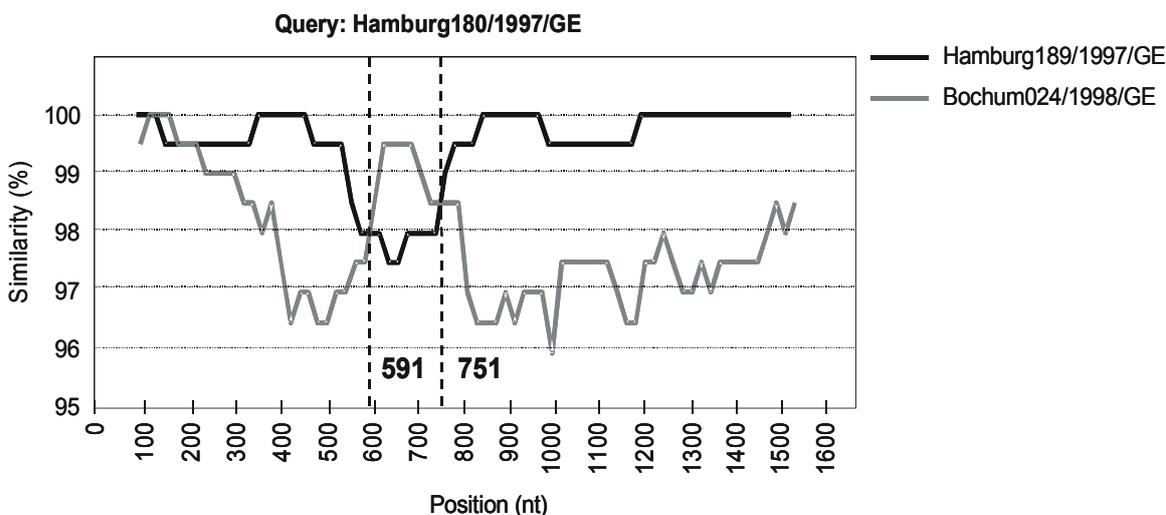
**Abbildung 13:** Durch SplitsTree anhand der Sequenzen erstellter phylogenetischen Baum. Folgende Abkürzungen stehen für die Sequenzen: Bo.024 = Bochum024/1998/GE, Bo.026 = Bochum026/1997/GE, Bo.031 = Bochum031/1997/GE, Bo.339 = Bochum339/1997/GE, ER.195 = Erlangen195/1997/GE, FR.204 = Freiburg204/1997/GE, FR.253 = Freiburg253/1998/GE, HH.135 = Hamburg135/1998/GE, DD.153 = Dresden153/1997/GE, DD.245 = Dresden245/1997/GE, DD.267 = Dresden267/1997/GE.

### 3.2.3 Zuordnung zu Parentalsträngen durch SimPlot und Exploratory Tree Analysis

Mittels SimPlot und Exploratory Tree Analysis untersuchten wir die Sequenzen, die in Abbildung 13 netzartigen Verzweigungen zeigten. Bei der Exploratory Tree Analysis wird ein 300 bp großes Gleitfenster in 150 bp-Schritten entlang des Sequenzalignment bewegt. *Branchswapping* deutet dabei auf Rekombinationsereignisse hin. SimPlot errechnet die *similarity scores* zwischen den Sequenzen zweier Stämme und stellt diese graphisch dar. Dabei wird die Sequenz eines möglichen Hybridstammes gleichzeitig mit beiden Sequenzen der zu untersuchenden Parentalstämme verglichen und deren verwandtschaftliches Verhältnis anhand dieser *similarity scores* in Form zweier Graphen abgebildet. Der eine Graph zeigt das Ergebnis des Vergleichs zwischen Parentalstamm 1 (Abb. 14: Hamburg189/1997/GE) und dem

Hybridstrang, der zweite die *similarity scores* zwischen Parentalstamm 2 (Abb. 14: Bochum024/1998/GE) und dem Hybridstrang (Abb. 14: Hamburg180/1997/GE). Die Programme identifizierten alle in Frage kommenden Rekombinanten sowie deren Parentalstämme und legten jeweils die wahrscheinlichste Schnittstelle für das Rekombinationsereignis fest (Tabelle 4). Für jede Schnittstelle wurde die jeweilige *Likelihood Ratio* errechnet.

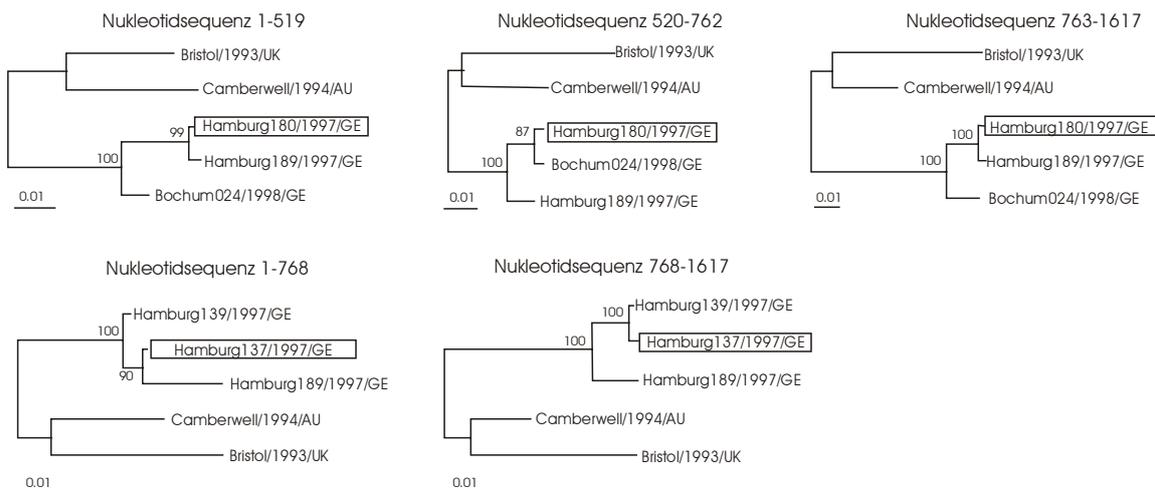
Für Mosaiksequenzen gilt, dass diese für einen bestimmten Abschnitt hohe *similarity scores* zu dem einen Parentalstamm zeigen, und nach passieren der Rekombinationsschnittstelle, zu dem anderen Stamm. Dieses Phänomen konnten wir für die Stämme Hamburg180/1997/GE (Abbildung 14) und Hamburg137/1997/GE (nicht dargestellt) zeigen. Die für ein Rekombinationsereignis in Frage kommenden Lokalisationen wurden hinsichtlich ihrer Wahrscheinlichkeiten mittels  $\chi^2$ -Test untersucht und bei  $p < 0,05$  als signifikant angenommen.



**Abbildung 14: Similarity Plot des rekombinanten Stamms Hamburg180/1997/GE mit den Parentalstämmen Hamburg189/1997/GE und Bochum024/1998/GE.** Die Analyse zeigt auf der vertikalen Achse das prozentuale Maß an Nukleotidentität in der Capsidsequenz, die in nt-Positionen auf der horizontalen Achse aufgetragen wurde. Das 200 nt Gleitfenster, wurde in 20-nt-Schritten entlang der Sequenz fortbewegt. Die abgeschätzte nt-Position der Rekombinationsschnittstellen (519 und 751) ist durch gestrichelte Linien markiert. An diesen Punkten sinken die similarity scores der einen Kurve ab, während die der anderen gleichzeitig steil ansteigen.

### 3.2.4 Erstellen von ML-Bäumen in Umgebung der Schnittstellen

Jeweils vor und nach den angenommenen Rekombinationsschnittstellen wurden ML-Bäume konstruiert. Es wurde beobachtet, dass die Mosaiksequenzen von Baum zu Baum die Zugehörigkeit zu einem bestimmten Cluster wechselten.



**Abbildung 15: ML-Bäume der rekombinanten Stämme Hamburg180/1997/GE und Hamburg137/1997/GE.** Die Parentalstämmen und die NV-Stämme Bristol/1993/UK und Camberwell/1994/AU sind als Außengruppen dargestellt. Die ML-Bäume wurden unter Verwendung des General Time Reversible Modells der Nukleotidsubstitution erstellt. Dabei wurde eine Gammaverteilung mit among site-variation und 8 Raten sowie Korrektur für nicht variable Sequenzpositionen und die Basenzusammensetzung der Stämme berücksichtigt, beziehungsweise zur Konstruktion verwendet. Die Daten wurden anhand des vorliegenden Sequenzdatensatzes abgeschätzt. Alle Astlängen sind skaliert dargestellt und an allen relevanten Knotenpunkten sind die Bootstrapwerte im Prozentanteil von 1000 mathematischen Replikationen aufgeführt

Wie aus Abbildung 15 ersichtlich clustert der rekombinante Stamm Hamburg180/1997/GE vor der Nukleotidposition 519 mit Hamburg180/1997/GE und in 3'-Richtung davon mit Bochum024/1998/GE. In 3'-Richtung von Nukleotidposition 763, dem zweiten Bruchpunkt, wechselt der Stamm in seine ursprüngliche Position zurück. Ähnlich clustert der rekombinante Stamm

Hamburg137/1997/GE vor der Nukleotidposition 768 mit dem Parentalstrang Hamburg189/1997/GE und in 3'-Richtung mit Hamburg139/1997/GE. Der Positionswechsel der Stämme beidseits der vermutlichen Bruchpunkte ist ein definitiver Hinweis auf ein Rekombinationsereignis.

Die in Abb. 15 gezeigten widersprüchlichen Phylogenien beidseits der P1-1 Domäne wurden durch hohe Bootstrap-Werte von >85% der 1000 durchgeführten Permutationen unterstützt. Ähnliches konnte für den Stamm Hamburg137/1997/GE gezeigt werden, der widersprüchliche phylogenetische Ergebnisse jeweils bezüglich der Stämme Hamburg139/1997/GE und Hamburg189/1997/GE zeigte.

### 3.2.5 Abschätzen der statistischen Wahrscheinlichkeit für die einzelnen Knotenpunkte durch LARD

Um die Ergebnisse weiter zu verifizieren, führten wir die Untersuchung mit einem zusätzlichen Programm, LARD, durch. Das Programm erstellte ML-Bäume und führte zur statistischen Einschätzbarkeit der einzelnen Knotenpunkte eine Bootstrap-Analyse durch. LARD überprüfte sowohl die Lokalisationen der Schnittstellen als auch deren statistische Signifikanz und kam unabhängig von den vorangegangenen Untersuchungen zu dem selben Ergebnis.

Rekombinanter Stamm	Parentalstämme	Capsidregionen (nt)	Breakpoint (nt)	LR	P-Wert
Hamburg180/1997/GE	Hamburg189/1997/GE Bochum024/1998/GE	1 - 519, 763 - 1617 520 - 762	519 762	8.355 5.858	<0.001 <0.010
Hamburg137/1997/GE	Hamburg189/1997/GE Hamburg139/1997/GE	1 - 768 769 - 1617	768	5.607	<0.010

**Tabelle 4: Breakpoint Analyse der Rekombinanten und ihrer wahrscheinlichen Parentalstämme.** LR = likelihood ratio; p-Wert wurde mittels  $X^2$  test mit 3 df bestimmt.

### 3.2.6 AS-Sequenzalignment mit Zuordnung der Rekombinations-schnittstellen zu den jeweiligen Capsidregionen

Der rekombinante Stamm Hamburg180/1997/GE zeigte bezüglich der Nukleotidzusammensetzung hohe Ähnlichkeit mit dem Stamm Hamburg189/1997/GE in den Regionen, die die Information für die Shell-, P2- und P1-2-Domäne enthalten (Nukleotide 1-519 und 763-1617, Abbildung 16). In der Region, die für die P1-1 Region kodiert, ist der NV-Strang Hamburg180/1997/GE mit dem Strang Bochum024/1998/GE nahezu identisch. Aus Abbildung 16 ist ersichtlich, dass sich an der Nukleotidposition 768 sowohl eine Rekombinationsschnittstelle als auch die Kopplungsstelle zwischen den Domänen P1-1 und P2 befindet. Die Position dieser Schnittstelle wird durch Likelihood Ratio Analysen und phylogenetische Analysen mittels ML-Baum-Konstruktion auch hier signifikant unterstützt (Tabelle 4).

	<b>SHELL</b>		
HH137	MKMASNDANPSDGSTANLVPEVNNEVMALEPVVGA AIAAPVAGQQNI IDPWIRNNF		
HH180	MKMASNDANPSDGSTANLVPEVNNEVMALEPVVGA AIAAPVAGQQNVIDPWIRNNF		
	<b>SHELL</b>	<b>100</b>	
HH137	VQAPGGFTVSPRNAPGEILWSAPLGPDLNPYLSHLARMYNGYAGGFVQVILAGNA		
HH180	VQAPGGFTVSPRNAPGEILWSAPLGPDLNPYLSHLARMYNGYAGGFVQVILAGNA		
	<b>SHELL</b>		
HH137	STAGKIIFAAVPPNFPTEGLSPSQVTMFPHIIVDVRQLEPVLIPDPVRNNFYHYNQSN		
HH180	STAGKIIFAAVPPNFPTEGLSPSQVTMFPHIIVDVRQLEPVLIPDPVRNNFYHYNQSN		
	<b>SHELL</b>	<b>200</b>	<b>P1-1</b>
HH137	STIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFIFLVPPTVESRI KPFTVPILTV		
HH180	STIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFIFLVPPTVESRTKPFTVPILTV		
	<b>P1-1</b>		<b>P2</b>
HH137	EEMSNSRFPIPLEKLYTGRSSAFV VQPQNGRCTTG VLLGTTQLSAVNICTFRGDVTHIA		
HH180	EEMSNSRFPIPLEKLYTGRSSAFV VQPQNGRCTTG VLLGTTQLSAVNICTFRGDVTHIA		
	<b>P2</b>	<b>300</b>	
HH137	GSHDYTMNLASQNMNNYDPTEEIPAPLGTPDFVVGKIQGMLTQTTREDGSTRAHKATV		
HH180	GSHDYTMNLASQNMNNYDPTEEIPAPLGTPDFVVGKIQGMLTQTTREDGSTRAHKATV		
	<b>P2</b>	<b>400</b>	
HH137	STGSAHFTPKLGSVQYTTDTNND FQTGQNTKFTPVGVIQDGNNHQNEPQQWVLP		
HH180	STGSVHFTPKLGSVQYTTDTNND FQTGQNTKFTPVGVIQDGNNHQNEPQQWVLP		

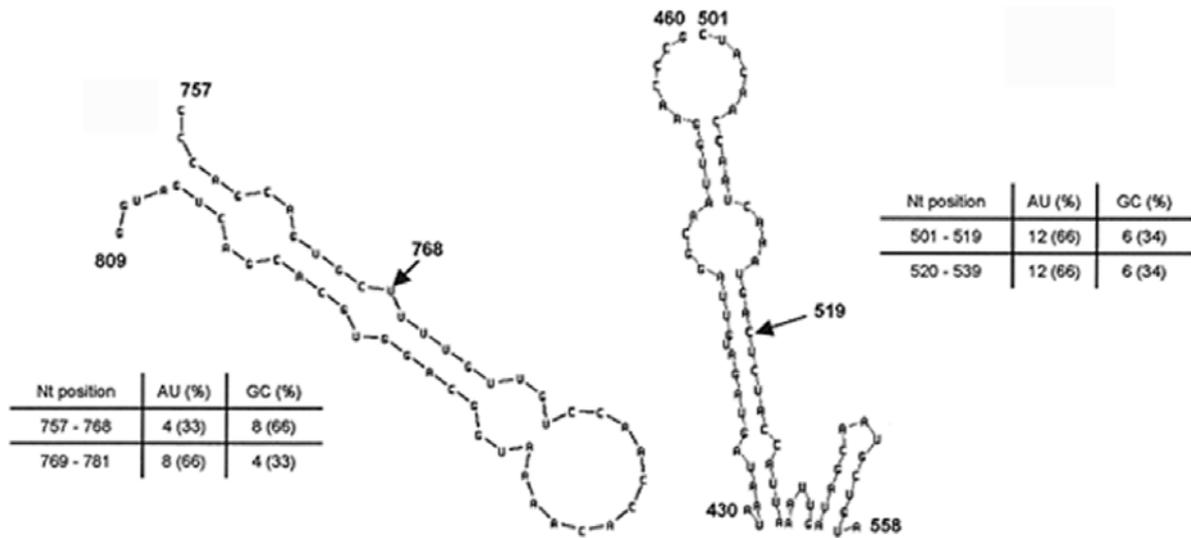
	P2	P1-2	
HH137	GRTGHNVHLAPAVAPTFPGEQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFYQEA		
HH180	GRTGHNVHLAPAVAPTFPGEQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFYQEA		
	P1-2		500
HH137	APAQSDVALLRFVNPDTGRVLFECKLHKSGYVTV AHTGPHDLVIPNGYFRFDSWVN		
HH180	APAQSDVALLRFVNPDTGRVLFECKLHKSGYVTV AHTGPHDLVIPNGYFRFDSWVN		
	P1-2		545
HH137	QFYTLAPMRNGAGRRRAL		
HH180	QFYTLAPMRNGAGRRRAL		

**Abbildung 16: Aminosäuren-Alignment der vollständigen Capsidsequenzen.** Die Lokalisationen der Rekombinationsschnittstellen sind unterstrichen, fett gedruckt und grau hinterlegt. Die jeweils erste Aminosäure einer Domäne ist fett gedruckt hervorgehoben, Aminosäuren, die sich innerhalb der beiden Stränge nicht entsprechen, sind unterstrichen. HH137 = Hamburg137/1997/GE; HH180 = Hamburg180/1997/GE.

### 3.3 Untersuchung der Sekundärstruktur in der Umgebung der Rekombinationsschnittstelle

Die voraussichtliche RNA Sekundärstruktur der Rekombinationsregion wurde mittels Mfold Web Server für *Nucleic Acid Folding and Hybridisation Prediction* ermittelt (Abbildung 17) [69]. Mfold basiert auf dem Prinzip der Energieminderung, mit dem Ziel die thermodynamisch optimale Sekundärstruktur einer gegebenen Nukleinsäuresequenz zu berechnen. Dazu wird ein sogenannter ‚Energy dot plot‘ erstellt [69]. In der Version Mfold 3.1 erfolgen alle Berechnungen für eine Temperatur von 37 °C [70].

Die Nukleotidzusammensetzung und Struktur der dargestellten Abschnitte entsprach typischen Regionen, die sich in der Umgebung von Rekombinationsschnittstellen befinden (siehe Diskussion).



**Abbildung 17: Charakterisierung der Zusammensetzung der Sequenz und der RNA-Sekundärstruktur der wahrscheinlich rekombinanten Region.** Links im Bild ist der Stamm Hamburg137/1997/GE (Nukleotid 757 - 809) und rechts der Stamm Hamburg180/1997/GE (Nukleotid 430 – 460 und 501 – 558) dargestellt. Die Rekombinationsschnittstellen sind jeweils durch einen Pfeil hervorgehoben und die Nukleotidzusammensetzung tabellarisch dargestellt.

## 4 DISKUSSION

Diese Studie weist das Auftreten von Rekombinationsereignissen innerhalb der Capsidregion bei 2 von 25 zuvor noch nicht beschriebenen Norovirus-Sequenzen nach. Die genetische Information für diesen Nachweis wurde aus Stuhlproben isoliert, die in Deutschland zwischen 1997 und 1998 gesammelt worden sind. Die Proben stammten von Patienten im Alter von bis zu vier Jahren und wurden im Vorfeld unserer Studie auf das Vorhandensein von HuCV-RNA überprüft. In allen Proben konnte diese anhand der Polymerase-Region nachgewiesen werden. Unter den von uns amplifizierten und sequenzierten Stämmen befanden sich zwei rekombinante Viren, deren Rekombinationsschnittstellen innerhalb der Capsidregion des viralen Genoms lagen. Nach unserem Informationsstand wurden Rekombinationsereignisse in dieser Region bis zu diesem Zeitpunkt noch nicht beschrieben, konnten jedoch zu einem späteren Zeitpunkt durch andere Methoden bestätigt werden [72]. Unsere Ergebnisse weisen also im Einklang mit späteren Studien darauf hin, dass homologe Rekombination innerhalb des NV-Capsidgens kein isoliertes Ereignis darstellt.

Molekulare Evolutionsforschung untersucht, wie aus der sequenziellen Abfolge von vier NS-Grundbausteinen der gesamte Bauplan für einen vollständigen Organismus entsteht und wie die Weitergabe der genetischen Information erfolgt [42]. Da dieser Prozeß durch das Auftreten von Substitutionen und Rekombination nicht fehlerfrei abläuft, zeigen die heutigen Organismen in ihrem Genom Spuren ihrer molekulargenetischen Geschichte. Die genetische Information ist somit ein Dokument der Evolutionsgeschichte jedes einzelnen Organismus [42]. Durch analytische Verarbeitung dieser Informationen können Aussagen über diese Entwicklungsgeschichte getroffen werden.

Phylogenetische Bäume repräsentieren Hypothesen zu Verwandtschaftsverhältnissen verschiedener Taxa. Als hypothetischer Vorgänger ist dabei die Wurzel des Baumes zu sehen. Da die NS-Sequenz dieses Vorgängers in der molekularen Evolutionsforschung häufig nicht bekannt ist, sind die meisten Bäume wurzellos. Also repräsentiert der aufgestellte Baum die

wahrscheinlichsten Verwandtschaftsverhältnisse aller Taxa untereinander, lässt aber generell nur Vermutungen zu deren Entwicklung zu [54]. In diesem Rahmen sind NS-Sequenzen als Grundlage für phylogenetische Analysemethoden ausgesprochen wertvoll, da sie eine digitalisierbare exakte Darstellung des Aufbaus und der Arbeitsanweisung aller funktionellen Bestandteile des Organismus enthalten. Die Informationen bezüglich ihrer Evolutionsgeschichte sind jedoch aus folgenden Gründen nur unter Vorbehalt anwendbar [49]: Für jeden einzelnen Genlokus kommen vier verschiedene Zustände in Frage, nämlich A, T/U, G und C. Werden homologe Genorte zweier verwandter NS-Sequenzen verglichen, kann streng genommen nur gefolgert werden, ob sich diese Loci entsprechen oder nicht. Der maximale Unterschied, der beobachtet werden kann, ist „1“ [49]. Daraus folgt, dass sich alle Informationen bezüglich eines einzelnen Genortes auf den jetzigen Zustand und gegebenenfalls die Tatsache beschränken, dass sich etwas verändert hat, nicht wie oft und aus welchem Urzustand heraus. Es gibt hier also keine Information bezüglich Substitutionsrate und Ursprungssequenz. Das bedeutet, sobald an einem Ort mehr als ein Substitutionsereignis stattgefunden hat, geht die Verbindung zu vorherigen Substitutionen verloren [49]. Anhand dieses einen Genlokus können dann keine Rückschlüsse auf die Entwicklung der Sequenz geschlossen werden.

Zur phylogenetischen Analyse werden darum größere Sequenzabschnitte betrachtet. Evolutionsmodelle erleichtern dabei die Rekonstruktion vergangener Ereignisse und lassen Vermutungen über zukünftige Veränderungen zu. Anhand der durch NS-Sequenzen gegebenen Informationen wird versucht, Gesetzmäßigkeiten zu erkennen, mit deren Hilfe Evolutionsmodelle entworfen werden. Durch diese soll die Entwicklung von Organismen annäherungsweise rekonstruiert werden [49]. Sie sind somit ein wichtiges Instrument der Evolutionsforschung und dienen auch wesentlich der Detektion von Rekombinationsereignissen.

Werden aus einem Sequenzdatensatz, der Mosaiksequenzen enthält, phylogenetische Bäume erstellt, führt dies zu scheinbar widersprüchlichen Analyseergebnissen. Der phylogenetische Konflikt ist mathematisch darstellbar

und bildet den Angriffspunkt der zur Rekombinationsanalyse verwendbaren Methoden. Beispiele für graphisch dargestellte phylogenetische Konflikte unserer Studie sind netzwerkartigen Strukturen innerhalb der durch SplitsTree erstellten Bäumen und *branchswapping* zwischen ML-Bäumen, die unterschiedliche Abschnitte rekombinanter Sequenzen darstellen. ML-Bäume, die einen hohen Rechenaufwand benötigen, gelten als Goldstandard in der Rekombinationsanalyse.

RNA-Viren können durch Reassortierung und Rekombination genetische Informationen austauschen. Reassortierung kann nur durch segmentierte Viren, wie z. B. das Influenza A Virus durchgeführt werden [35]. Rekombination kann dagegen sowohl durch segmentierte Viren als auch durch Viren mit anderer Genomorganisation, wie das Calicivirus, benutzt werden, tritt jedoch praktisch nur bei Viren mit Plus-Strang RNA auf. Allgemein wird bei der Rekombination eine Donor-Nukleotidsequenz in ein kontinuierliches Akzeptormolekül eingefügt [35]. Verschiedene Rekombinationsformen wurden bei RNA-Viren nachgewiesen: homologe, aberrant homologe und nichthomologe Rekombination [35]. Bei homologer Rekombination wird die Donor-Sequenz so in das Akzeptormolekül eingepasst, dass die Struktur des Genoms unverändert bleibt. Im Falle aberranter homologer Rekombination tauschen ähnliche Viren Sequenzen ohne striktes Alignment aus. Die genetische Abfolge innerhalb des Stranges verändert sich. Nichthomologe Rekombination ist schließlich der genetische Austausch zwischen praktisch unverwandten Viren. Da weder durch uns noch in anderen Studien Strukturveränderungen innerhalb des Genoms rekombinanter CV nachgewiesen werden konnte [38, 71], gehen wir davon aus, dass der Austausch genetischer Information zwischen NV mittels homologer Rekombination stattfindet, oder dass mögliche nichthomologe Rekombinationsereignisse zu nicht replikationsfähigen Nachkommenviren führen, die der Detektion entgehen.

Die meisten Plus-Strang RNA-Viren benutzen zum genetischen Informationsaustausch ein *copy-choice* Modell. Dabei findet das eigentliche

Rekombinationsereignis in Folge eines sogenannten *template switch* der viralen Replikase statt. Dazu löst sich die RNA-abhängige RNA-Polymerase während des Replikationsvorganges von dem Matrizenstrang, lagert sich im Falle homologer Rekombination an die orthologe Sequenzposition eines anderen Strangs an und nutzt dessen genetische Information zur Synthetisierung des neuen Strangs [35]. Es resultiert eine Mosaiksequenz, die auf beiden Seite des sogenannten Rekombinationsbreakpoints die genetischen Informationen unterschiedlicher parentaler Genome enthält. Wir gehen bezüglich unserer Ergebnisse von homologer Rekombination nach dem *copy-choice* Modell aus. Das Genom eines Virus besteht im Wesentlichen aus proteinkodierenden Sequenzen, so auch der Capsid-Bereich, den wir untersucht haben. Auf Proteinen kann sowohl ein positiver als auch ein negativer Selektionsdruck lasten. Durch ihre Funktionalität ist nicht an allen Genorten jede denkbare Veränderung möglich. Theoretisch kann eine einzige Substitution die Tertiärstruktur eines Enzyms so verändern, dass die Virusreplikation in einem Ausmaß gestört wird, dass eine Vermehrung unmöglich oder dessen Eliminierung durch den Wirtsorganismus erleichtert wird. Andererseits kann eine durch Substitution oder Rekombination entstandene Veränderung an bestimmten Genorten durch Änderung der Antigenität des Virus einen Selektionsvorteil darstellen. Möglich wäre dies an einem Ort an dem das Immunsystem des Wirtsorganismus angreift, wie dies zum Beispiel bezüglich der von uns untersuchten P-Region vermutet wird. Diese Annahme wurde bereits 2003 bei Nilsson et al. in Anbetracht einer Kummulation von Veränderungen in der P2-Region geäußert [20].

Somit wird Rekombination viraler RNA durch viele verschiedene immunologische und intrazelluläre Faktoren beeinflusst, welche das Auftreten funktionsfähiger rekombinanter Viren fördern oder hemmen können [35]. Es ist verständlich, dass verschiedene Genorte bezüglich ihrer Substitutionsraten divergieren oder sich eine Rückmutation oder Rekombination begünstigt durchsetzt. Dieser Umstand äußert sich in der *rate-hererogeneity* verschiedener Orte eines Stranges.

Grundvoraussetzung für ein Rekombinationsereignis ist jedoch die Koinfektion eines Wirts durch zwei unterschiedliche Virusstämme, die eine einzige Zelle gleichzeitig infizieren. Das Zustandekommen eines solchen Ereignisses unterliegt diversen äußeren Einflüssen. So können Koinfektionen zum einen dadurch verhindert werden, dass sich die natürliche Verbreitung unterschiedlicher viraler Stämme zeitlich und/oder räumlich nicht überlappt. Zum anderen ist es möglich, dass das Immunsystem des Wirts beide Erreger so schnell eliminiert, dass die Wahrscheinlichkeit für die Koinfektion einer einzelnen Zelle sinkt oder die Infektion mit dem einen Erreger die Wahrscheinlichkeit für eine Zweitinfektion (z. B. durch Interferenz) erniedrigt.

Der am stärksten limitierende Faktor für einen *template switch* bei homologer Rekombination ist ein zu großer Sequenzunterschied zwischen den potentiellen Rekombinationspartnern [35]. Bezogen auf die zugänglichen GenBank Daten des NV befanden sich alle von uns analysierten Stämme im Cluster 4 der Genogruppe II [71]. Jedes der von uns nachgewiesenen Rekombinationsereignisse spielte sich zwischen NV dieser Genogruppe ab und alle Parentalstämme stammten aus dem gemeinsamen Cluster II/4. Die jeweiligen Parentalstämme waren somit untereinander relativ eng verwandt. Auch in der parallel verlaufenden Studie unter Verwendung der allgemein zugänglichen GenBank Daten konnten keine Rekombinationsereignisse nachgewiesen werden, die zwischen NV aus unterschiedlichen Genogruppen stattgefunden hatten [71].

Koinfektionen, die durch Viren aus unterschiedlichen Genogruppen verursacht werden, konnten für andere GE-Erreger bereits nachgewiesen werden [73, 74] und sind für das NV ebenfalls denkbar. Dies ist also wahrscheinlich nicht die Ursache dafür, dass Rekombinationen zwischen zu weit entfernten Viren nicht stattfindet. Als Grund für die Beschränkung auf die selbe Genogruppe ist vorstellbar, dass der *teplate switch* der viralen Replikase erleichtert wird, wenn sich die Stämme ähnlich sind. Spielte die Ähnlichkeit der Sequenzen für den Rekombinationsvorgang eine untergeordnete Rolle, würde homologe Rekombination in weitaus geringerem Maße auftreten als es tatsächlich der Fall ist. Allein die Tatsache, dass orthologe Sequenzpositionen offenbar erkannt

werden, betont die Bedeutung des genetischen Ähnlichkeitsgrads der Sequenzen. Andererseits führt aberrante homologe Rekombination durch Veränderung der genetischen Abfolge häufiger zu Fehlfunktionen der Strukturgene, so dass es denkbar ist, dass das Virus durch eine solche Funktionsstörung die Replikationsfähigkeit verliert und darum nicht natürlich in Erscheinung tritt.

Wie bereits erwähnt, konnten wir innerhalb des von uns untersuchten Sequenzdatensatzes zwei rekombinante Stämme nachweisen und in einer parallel durchgeführten Studie, die alle öffentlich zugänglichen Stämme der GenBank Datenbanken untersuchte, wurden acht rekombinante Stämme detektiert [71]. Damit ergaben sich Virusrekombinanten bei fast 10 % der untersuchten Fälle. In Anbetracht der Tatsache, dass Rekombination zwischen Viren ein äußerst seltenes Ereignis darstellt, ist dies eine große Anzahl. Verglichen jedoch mit Viren, die extrem zur Rekombination neigen, wie z. B. dem HI-Virus, ist die Anzahl dennoch gering. HI-Viren können einen bis sieben genetische Austausche pro Replikationszyklus durchführen [75].

Die Nomenklatur unserer Virusstämme in Anlehnung an eine durch Green et al. beschriebene, ebenfalls auf der vollständigen CV-Capsidsequenz basierenden Einteilung vorgenommen [10]. Danach ist beispielsweise Hamburg180/1997/II/GE der 180. Stamm, der in Hamburg 1997 als Probe gesichert wurde und gehört laut phylogenetischem Analyseergebnis der Genogruppe II an. Die Bezeichnung der beiden Rekombinanten Stämme, die in unserer Studie nachgewiesen werden konnten, lautet dem zufolge: Hamburg180/1997/GE (HH180) und Hamburg137/1997/GE (HH137).

Der Stamm HH180 zeigte zwei, der Stamm HH137 nur eine Rekombinationschnittstelle. Daraus lässt sich schließen, dass im Fall des Stammes HH180 entweder in einem Replikationsvorgang zweimal in Folge ein *template switch* der viralen Replikase stattgefunden hat, oder, vielleicht wahrscheinlicher, es kam während dieser oder einer späteren Koinfektion zur Zweitrekombination zwischen einem bereits einfachrekombinanten Stamm und einem Parentalstamm.

Durch Rohayem et al. wurden genetische Rekombinationen in vier verschiedenen Clustern der GGII nachgewiesen, jedoch nur eine einzige Rekombinante innerhalb von GGI. Gleichzeitig ist GGII häufiger in Epidemien vertreten als GGI, welches eventuell ein Grund für diese Beobachtung und gleichzeitig für die Tatsache ist, dass uns gelingen konnte, innerhalb einer Gruppe von nur 25 Capsidsequenzen der NV GGII zwei rekombinante Viren nachweisen zu können. Wir verwendeten dabei verschiedene analytische Methoden, die dieses Ergebnis jeweils mit hoher Wahrscheinlichkeit bestätigten. Insgesamt nahm der Anteil rekombinanter Viren an der Gesamtanzahl in beiden Studien etwa 10 % ein.

Um diese offensichtliche Neigung des NV zur Rekombination genauer erklären zu können, haben wir die Stämme hinsichtlich Struktur und typischer Merkmale rekombinanter Sequenzen untersucht. Es zeigte sich, dass das NV-Genom viele dieser besonderen Charakteristika aufweist und sich somit zur Rekombination auch durch seine strukturelle Beschaffenheit anbietet (siehe unten).

Ob ein rekombinantes Virus natürlich in Erscheinung tritt, hängt jedoch neben der Entstehung noch von weiteren limitierenden Faktoren ab. Da verschiedene virale Stämme die Replikation in einer Zelle gegenseitig behindern, muß zur effektiven Vermehrung funktionsfähiger rekombinanter Viren eine reinigende Selektion die Replikation solcher Stämme hemmen, die gar nicht oder nur bedingt funktionsfähig sind.

Alle bereits erwähnten äußeren Faktoren wirken Rekombinationsereignissen direkt oder indirekt entgegen. Ist es dem neuen Virus jedoch gelungen, die Zelle in ausreichender Anzahl zu verlassen, kommt die ‚escape‘-Wirkung gegenüber dem Immunsystem durch Rekombinationsereignisse innerhalb immunologisch bedeutsamer Sequenzregionen zum tragen und wirkt sich begünstigend auf die Ausbreitung des entstandenen rekombinanten Virus aus. Das Virus kann zunächst nicht durch spezifische Antikörper eliminiert werden und erhält somit einen Selektionsvorteil den eigenen Parentalsequenzen gegenüber, da diese vor dem Rekombinationsereignis mit gewisser Wahrscheinlichkeit bereits mit dem Immunsystem des Wirts in Kontakt getreten sind. Daraus resultiert

vermutlich eine vermehrte Replikation der Mosaiksequenzen, so dass das natürliche Auftreten von rekombinanten Stämmen mit Bruchpunkten in immunologisch bedeutsamen Regionen begünstigt wird. Dieser Zusammenhang konnte auch bei einem durch Nilsson et al. beschriebenen immunsupprimierten Patienten beobachtet werden. Während das Virus beim immunkompetenten Patienten nach einer gewissen Zeit eliminiert wird, war offensichtlich in diesem Falle das Immunsystem des Wirts nicht in der Lage, mit ausreichender Geschwindigkeit auf die Entstehung neuer Stämme zu reagieren, so dass sich diese immer wieder vermehren konnten. Es herrschte ein Gleichgewicht zwischen der Immunabwehr des Wirts und der Fähigkeit des Virus, durch ‚escape‘-Mechanismen diesem zu entkommen. So wurden immer wieder neue Stämme beobachtet, die Veränderungen in der P2-Region zeigten. Offensichtlich scheint eine Veränderung in diesem Bereich einen Selektionsvorteil darzustellen, so dass sich die entsprechenden Viren bevorzugt vermehren können [20].

In Studien zu natürlich in der Bevölkerung zirkulierenden NV-Strängen wurde festgestellt, dass typische epidemiologische Charakteristika des NV einen positiven Einfluß auf das Auftreten von Rekombinationsereignissen haben. NV kommen z. B. ausgesprochen häufig vor [74, 76] und verbreiten sich rapide innerhalb der Bevölkerung. Dies kommt sowohl durch die hohe Viruslast der infektiösen Faeces [77] als auch die starke Virulenz des Erregers zustande und steigert die Wahrscheinlichkeit für Koinfektionen besonders mit solchen Viren, die häufig Epidemien verursachen, wie die der GGII. Durch Iritani et al. wurde die Existenz natürlich zirkulierender Arg320-like Stränge nachgewiesen [76]. Dies kann als Beweis gewertet werden, dass rekombinante Stränge natürlich vorkommen, hier am Beispiel der Bevölkerung von Osaka City. Rekombination von NV scheint also unter natürlichen Bedingungen nicht nur möglich zu sein sondern auch tatsächlich stattzufinden, so dass es wahrscheinlich ist, daß NV alle genannten Voraussetzungen zur Rekombination erfüllen. Diese Beobachtungen stehen in Einklang mit unseren Ergebnissen.

Bezüglich der einzelnen von uns verwendeten Methoden konnte durch verschiedene Arbeitsgruppen eine grosse Leistungsfähigkeit und Zuverlässigkeit bei der Erkennung von Rekombinationsereignissen zwischen RNA-Viren nachgewiesen werden [43-45].

Wegen der bereits erwähnten Mängel bei isolierter Verwendung [46, 78, 79], wurde durch uns ein kombinierter Ansatz benutzt. Somit wurden der selbe Datensatz zur Erfassung und korrekten Identifizierung von Rekombinanten und den dazugehörigen Parentalsträngen, sowie zur exakten Lokalisierung der entsprechenden Rekombinationsschnittstellen hinsichtlich statistischer, phylogenetischer und struktureller Gesichtspunkte mit verschiedenen Methoden untersucht. Trotz der Verwendung zuverlässiger Methoden kam es in einem Fall zu einem nicht Übereinstimmen der durch die einzelnen Analysen gelieferten Ergebnisse. Dem Stamm Bochum272/1997/GE wurde durch Splits Tree die Parentalsequenzen Dresden153/1997/GE und Dresden319/1997/GE zugeordnet. Dies wurde in diesem Ansatz durch hohe Bootstrap-Werte unterstützt, konnte jedoch durch die anderen Methoden nicht mit Sicherheit bestätigt werden und wurde darum nicht als Rekombinationsereignis akzeptiert. Diese Vorgehensweise, das heißt, den Ausschlusses von Rekombinationsereignissen, die nicht durch alle Methoden bestätigt werden konnten, wurde bereits durch andere Arbeitsgruppen diskutiert, die ebenso vorgegangen sind [43-45].

Übereinstimmende Ergebnisse lieferten hingegen alle Methoden bezüglich der Mosaiksequenzen Hamburg137/1997/GE (HH137) und Hamburg180/1997/GE (HH180). Beide Stämme besaßen jeweils einen Bruchpunkt innerhalb der Region P1-1 und HH180 zeigte eine zusätzliche Schnittstelle innerhalb der S-Domäne. Wie bereits erwähnt, unterstützt diese Beobachtung die Annahme einer immunologisch angetriebenen Evolution des NV, die durch Rekombination innerhalb des Capsidgens begünstigt zu werden scheint. Wie bereits erwähnt, wurde eine ähnliche Hypothese 2003 durch Nilsson et al. vertreten, welcher eine Anhäufung von Mutationen innerhalb der P2-Domäne eines NV-Stranges beobachteten. Auch hier wurde eine Begünstigung dieser Mutationen durch immunologischen Druck vermutet [20].

Verschiedene Arbeitsgruppen beschrieben in der Vergangenheit bereits die Auswirkungen, die das Auftreten von Rekombinationsereignissen innerhalb immunologisch bedeutsamer viraler Regionen haben kann, und diskutierten dies bezüglich anderer RNA-Viren [35, 36, 43, 80, 81]. Die Vorteile genetischen Austauschs scheinen entweder in der Ansammlung vorteilhafter, oder der Eliminierung schädlicher Gene oder Sequenzen zu liegen [35]. Wir vertreten unseren Ergebnissen zufolge die Ansicht, dass die hohe Variabilität des NV-Capsidgens, zu einem wesentlichen Anteil durch Rekombinationsereignisse bedingt ist und Rekombination darum möglicherweise eine Schlüsselrolle in der virulenten Wirkung dieser Erreger spielt. Durch diese Fähigkeit scheint es NV in besonderem Maße möglich zu sein, der Immunantwort des Wirts entkommen zu können, welches möglicherweise zum Teil die Problematik erklärt, geeignete Impfstoffe und eine dauerhafte Immunität gegen das NV zu entwickeln (siehe Einleitung).

Die verschiedenen Regionen des CV-Capsids unterscheiden sich in ihrer Sequenzvariabilität. Prasad et al. konnten zeigen, daß die S-Domäne des CV-Capsid Gens die geringste Variabilität aufweist. Die P1-Domäne verhält sich über die Zeit mäßig variabel und die P2-Domäne, die für den am weitesten herausragenden Teil des Capsids kodiert, ist hochvariabel [82]. Da dessen Struktur für das Immunsystem am leichtesten zugänglich ist, wird seit langem vermutet, dass sich dort die Antikörperbindungsstelle des CV-Capsids befindet, welches bis jetzt noch nicht bewiesen werden konnte. Trifft dies zu, ist die P2-Domäne die eigentliche Determinante der Stammidentität gegenüber dem Immunsystem.

Bezüglich dieser Domäne konnte des weiteren gezeigt werden, dass ein monoklonaler Antikörper, der einen Abschnitt der P2-Domäne zwischen den Positionen 300 und 384 erkennt, das Binden des NV an die Wirtszelle verhindern kann. Diese Beobachtung impliziert eine Beteiligung dieser Region an der Zellbindungsaktivität des Virus [82] und betont dessen Bedeutung für die Virulenz des Erregers. Möglich ist auch, dass die P1-Region eine wesentliche Rolle in der Antigenität des Virus spielt, da diese für die Bögen des Capsids

kodiert. Es ist denkbar, dass Sequenzunterschiede in diesem Bereich zu Konformationsänderungen führen und auch auf diese Weise die immunologische Identität des Virus so verändert wird, dass ein Selektionsvorteil entsteht. Beide von uns isolierten rekombinanten Stämme wiesen Rekombinationsschnittstellen innerhalb der P1-Region auf.

Wie bereits durch andere Gruppen festgestellt wurde, ist das Auftreten von homologer Rekombination, das auf dem *copy-choice*- und *template-switching*-Mechanismus basiert, stark von der spezifischen Nukleotidzusammensetzung und RNA-Sekundärstruktur der Rekombinationsregion abhängig. Nagy et al. stellten fest, dass vor allem solche viralen Sequenzen zur Rekombination neigen, die nahezu gleiche Anteile an G/C und A/U oder mehr als 60% G/C enthalten, wenn sich diese Region vor einer A/U-reichen Region ( $\geq 60\%$  A/U) befindet. Diese Erkenntnisse basierten auf Modellen zu homologer Rekombination des Brome Mosaic Bromovirus (BMV) [65, 83], des Tomato Bushy Stunt Tombusvirus [84, 85] und des Turnip Crinkle Carmovirus [68, 86]. Zusätzlich wurde festgestellt, dass A/U- und G/C-reiche Abschnitte ähnlicher Länge, wenn sich diese auf beiden Seiten der Schnittstelle befinden, das Auftreten von Rekombinationsereignissen fördern. Sequenzen die diese Eigenschaften zeigen, werden ‚*homologous recombination activators*‘ genannt [87]. Nagy et al. vermuteten, dass Sequenzähnlichkeit allein die homologe Rekombination in diesem Fall des Brome Mosaic Bromovirus nicht ausreichend fördert [88], so dass zusätzlich bestimmte strukturelle Charakteristika zur Rekombination benötigt werden könnten. Insbesondere unterstützt angeblich das Vorhandensein von lokalen Konformationsfluktuationen und/oder einer stabilen Haarnadelstruktur in einem A/U-reichen Sequenzabschnitt die Anlagerung der RNA-abhängigen RNA-Polymerase durch Erleichterung der lokalen Basenpaarung zwischen Donor- und Akzeptor-RNA-Strang [68].

In unserer Studie wurde die unmittelbare Umgebung (50 nt) beidseits der identifizierten Bruchpunkte hinsichtlich Struktur und Zusammensetzung untersucht. Wir stellten fest, dass HH137 strangaufwärts von der Rekombinationsschnittstelle einen hohen G/C-Gehalt aufwies, während die

Sequenz strangabwärts in Richtung des 3'-Ende prozentual mehr A/U enthielt. HH180 zeigte dagegen sowohl am 3'- als auch am 5'-Ende der Rekombinationsregion A/U-reiche Sequenzen. In beiden Fällen waren die A/U und G/C-reichen Sequenzen von ähnlicher Länge. Darüber hinaus konnten durch Analyse der RNA-Sekundärstruktur am 3'-Ende der Rekombinationsregion stabile Haarnadelstrukturen nachgewiesen werden. Folglich zeigten die rekombinanten Stränge wesentliche Charakteristika von *'homologous recombination activators'*.

Dem Begriff der Stammidentität aus immunologischer Sicht kommt eine besondere Bedeutung zu, die Einfluß auf die Terminologie des Virus nehmen kann. Bis zum jetzigen Zeitpunkt bildet die Grundlage zur Festlegung der Identität eines einzelnen Stamms die genetische Information der Polymeraseregion des NV. Wir vermuten, dass das NV mittels seiner Struktur und besonderen Eigenschaften in der Lage ist, durch Rekombination von immunologisch bedeutsamen Regionen die Antigenität des Einzelstammes zu verändern, so dass das Virus hochvariabel erscheint [10]. Zum Routinenachweis von NV-RNA ist es zweckmäßig, die Polymeraseregion zu verwenden. Die unbeschädigte Isolierung und Amplifikation des gesamten NV-Capsids ist sehr aufwendig und darum zur Routinediagnostik nicht geeignet. Die hohe Variabilität des NV-Genoms im Capsidbereich wirft jedoch im Zusammenhang mit der Taxonomie des NV Fragen auf. Gerade weil das Capsidgen unseren Ergebnissen zu folge gut zum Nachweis von Rekombinationsereignissen zwischen zirkulierenden NV-Strängen geeignet ist, entstehen in Bezug auf die Taxonomie neue Gesichtspunkte. Vor dem Hintergrund des voranschreitenden Wissens über NV ist deren phylogenetische Klassifikation allein auf der Basis partieller NS- oder AS-Capsidsequenzen oder anderer Regionen des Virus diskussionswürdig.

Rekombinationsereignisse zwischen NV kommen jedoch nicht ausschließlich innerhalb des Capsidgens vor. Mit dem rekombinanten Virus Arg320/1995/AR konnten Jiang et al. bereits 1999 einen Stamm nachweisen [38], der bezüglich der Polymeraseregion (ORF 1) und des Capsidgens (ORF 3) Unterschiede in Sequenzidentitäten und Phylogenien zeigte. Als Parentalstämme wurden

Lorsdale/1993/UK und Mexico/1989/MX identifiziert [76]. Der Rekombinationsbruchpunkt von Arg320/1995/AR befand sich am Übergang zwischen ORF1 und ORF2. Durch Similarity Plots und phylogenetische Inkongruenz wurde diese Annahme unterstützt, trotzdem konnte die genaue Rekombinationschnittstelle bisher nicht identifiziert werden. Auch hier zeigten die RNA-Struktur der Rekombinationsregion sowohl eine Haarnadelstruktur als auch die für *homologous-recombination activators* typische Nukleotidzusammensetzung. Dies ist ein weiterer Hinweis darauf, dass Rekombination innerhalb des NV-Genoms nicht nur möglich sondern auch wahrscheinlich ist. Darüber hinaus wird anhand dieses Beispiels deutlich, dass die Sequenz der Polymerase-Region auch in diesem Fall zu Klassifizierung eines Stammes nicht ausreicht, da diese vom Stamm Lorsdale/1993/UK abstammt, während die Antigenität durch die Capsid-Region bestimmt wird und damit der des Stamms Mexico/1989/MX entspricht.

Zusammenfassend zeigen unsere Ergebnisse, dass Rekombination zwischen Noroviren der Genogruppe II natürlich vorkommt und vermutlich relativ häufig ist. Von 25 durch uns isolierten, sequenzierten und phylogenetisch analysierten NV-Capsidsequenzen zeigten zwei Stämme eindeutige Rekombinationsmuster. Die verschiedenen von uns zum Rekombinationsnachweis verwendeten Methoden (Sawyer's Test, Similarity Plots, Splits-Tree Analysis, Likelihood-Ratio Analysis und Strukturanalysen) wiesen in beiden Fällen kongruent darauf hin, dass ein Rekombinationsereignis innerhalb des Capsidgens stattgefunden haben muß und bestätigten dieses Ergebnis jeweils mit hohen Wahrscheinlichkeitswerten. Ein weiterer NV-Stamm, der keine eindeutige Bestätigung durch die Methoden erhielt, wurde nicht als rekombinanter Stamm in die Ergebnisse aufgenommen. Über Ausmaß und Rate von Rekombinationsereignissen zwischen NV insgesamt kann allerdings anhand dieser geringen Anzahl von Proben keine absolute Aussage getroffen werden. Hinter der relativen Häufigkeit mit der in dieser Studie und durch Rohayem et al. rekombinante Sequenzen innerhalb der NV GGII nachgewiesen werden konnten, vermuten wir ursächlich den Umstand, dass NV-Epidemien in den

meisten Fällen durch Viren dieser Genogruppe verursacht werden. Bereits mehrfach konnte nachgewiesen werden, dass Stämme der NV GGII und in besonderem die des Clusters II/4 ca. 80 % aller NV-Epidemien verursachen [24, 89]. Mit steigender Präsenz eines Stamms erhöht sich auch die Wahrscheinlichkeit für Ko- bzw. Superinfektionen, die durch diesen verursacht werden. Dies ist, wie bereits erwähnt, die wichtigste Grundvoraussetzung für ein Rekombinationsereignis. Diese Annahme wird durch Beobachtungen von Twiddy et al. unterstützt, die nachwiesen, dass eine erhöhte Dichte in geographischer und zeitlicher Stichprobennahme die Wahrscheinlichkeit einer Detektion von rekombinanten Stämmen erhöht [45]. Unklar ist zum jetzigen Zeitpunkt noch, ob die Häufigkeit von Rekombinationsereignissen in dieser Gruppe die zunehmende Virulenz der Erreger verursacht oder durch sie begründet wird. Möglicherweise trifft beides zu. Ebenfalls kann aktuell noch nicht klar beurteilt werden, ob die in verschiedenen Ländern beobachtete höhere Frequenz der NV GGII-Viren von biologischen Faktoren wie Tropismus, Replikation und Pathogenität abhängt.

Zu Beginn der Diskussion wurde bereits ein Einblick in die Prinzipien, Gesetzmäßigkeiten und Herausforderungen der phylogenetischen Analyse gegeben. In Anbetracht unserer Ergebnisse und in Anlehnung an Empfehlungen anderer Forschungsgruppen [35] schlagen wir vor, dass zur Erstellung von NV-Stammbäumen nur solche Analyseprogramme verwendet werden, die Rekombinationsereignisse erkennen und darstellen können. Andernfalls besteht die Gefahr, dass der durch ein Programm erstellte Baum nicht die wahrscheinlichsten Verwandtschaftsverhältnisse des Datensatzes widerspiegelt.

## 5 ZUSAMMENFASSUNG

Diese Arbeit zeigt das Vorkommen von fünfundzwanzig neuen Norovirusstämmen in Deutschland und stellt ihre phylogenetischen Verwandtschaftsverhältnisse dar. Bei zweien dieser Stämme handelte es sich um natürlich entstandene rekombinante Viren, deren Rekombinationsbruchpunkte innerhalb der Capsidregion lagen. Die Schnittstellen wiesen in ihrer unmittelbaren Umgebung charakteristische Eigenschaften rekombinanter Viren auf. Die Virulenz und die Antigenität des Erregers, sowie die Methoden der taxonomischen Zuordnung wurden in Bezug auf die Tragweite dieser Ergebnisse diskutiert.

Die genetische Information für den Virusnachweis wurde aus 119 NV-positiven Stuhlproben von bis zu vier Jahre alten Kindern isoliert und sind in den Städten Hamburg, Bochum, Freiburg, Erlangen und Dresden zwischen 1997 und 1998 gesammelt worden. Durch Amplifikation und Sequenzierung wurde die komplette Capsidsequenz von fünfundzwanzig zuvor noch nicht beschriebenen NV-Stämmen ermittelt und mit Maximum-Likelihood Analysen ihre verwandtschaftliche Beziehung als phylogenetischer Baum dargestellt. Durch verschiedene, zum Rekombinationsnachweis geeignete Methoden (Exploratory Tree Analysis, Similarity Plots, Splits Tree und Sawyer's Test) wurden der Datensatz auf das Vorkommen von Rekombinationsereignissen untersucht. Splits Tree lieferte zunächst Hinweise auf insgesamt drei rekombinante Stämme: Hamburg180/1997/GE (HH180), Hamburg137/1997/GE (HH137) und Bochum272/1987/GE (BO272). Durch die im Anschluss verwendeten Methoden (s.o.) wurden jedoch nur die zwei Erstgenannten als Mosaiksequenzen bestätigt, so dass der Stamm BO272 nicht in die Endergebnisse aufgenommen wurde.

Sim Plot und Exploratory Tree Analysis ordneten den rekombinanten Stämmen die jeweiligen Parentalstämme zu. Dem zufolge entstanden beide Rekombinanten aus Viren der NV-Genogruppe II/4, HH180 aus Hamburg189/1997/GE und Bochum024/1998/GE und HH137 aus

Hamburg189/1997/GE und Hamburg139/1997/GE. Durch das Programm LARD wurden die genauen Lokalisationen der jeweiligen Rekombinationsbruchpunkte ermittelt. Diese befanden sich beim Stamm HH180 an den Sequenzpositionen 519 nt und 762 nt und beim Stamm HH137 an der Position 768 nt. Vor und nach diesen Bruchpunkten wurden Maximum-Likelihood-Bäume aus dem rekombinanten Stamm, den beiden Parentalstämmen und einer Außengruppe erstellt. Mit hohen Bootstrapwerten für die einzelnen Verzweigungen wechselten beide rekombinanten Stämme nach jedem Bruchpunkt ihre phylogenetische Zugehörigkeit zu einem anderen Parentalvirus innerhalb des konstruierten Baumes.

Um beurteilen zu können, ob Noroviren generell dazu neigen, genetische Information auszutauschen, wurde die Struktur der Rekombinationsregion analysiert und mit der Struktur von Viren, bei denen Rekombinationsereignisse gehäuft nachgewiesen werden konnten, verglichen. Beide Stämme zeigten typische Eigenschaften von sogenannten ‚*homologous recombination activators*‘. In Bezug auf die Genomorganisation befanden sich bei jeder Rekombinanten eine der Schnittstellen im Bereich der Protruding-Region der Capsidsequenz, in der sich vermutlich die Antikörper-Bindungsstelle des Virus befindet.

Diese Studie zeigt also im Einklang mit späteren Ergebnissen, dass homologe Rekombination innerhalb des NV-Capsidgens kein isoliertes Ereignis darstellt. Im Hinblick auf immunologische Bedeutung und Klassifikation wurde die Tragweite von Rekombinationsereignissen innerhalb verschiedener Genomabschnitten diskutiert und alternative Lösungen vorgeschlagen. Es ist fraglich, ob bei natürlichem Vorkommen von Rekombination in wahrscheinlich immunologisch bedeutsamen Regionen des Virusgenoms die Sequenzierung der Polymerase-Region zur Klassifikation ausreicht, oder ob die Verwendung der Capsidregion sinnvoller wäre.

“*A bifurcating tree is a hypothesis about how taxa are related, not a truism.*”  
(Posada: Recombination in Evolutionary Genomics).

## 6 LITERATUR

1. Cubitt WD et al. Virus Taxonomy, Classification and Nomenclature of Viruses; Sixth Report of the ICTV. *Arch Virol* 1995; 10: 359-363
2. Appleton H. Small round viruses: Classification and role in food-borne infections. *Ciba Found Symp* 1987; 128: 108-125
3. Kapikian AZ et al. Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. *J Virol* 1972; 10: 1075-1081
4. Burroughs JN, Brown F. Physico-chemical evidence for re-classification of the caliciviruses. *J Gen Virol* 1974; 22: 281-286
5. Ehresmann DW, Schaffer FL. RNA synthesized in calicivirus-infected cells is atypical of picornaviruses. *J Virol* 1977; 22: 572-576
6. Mathews RE. Classification and nomenclature of viruses. *Intervirology* 1979; 12: 129-296
7. Traum J. Vesicular exanthema of swine. *J Am Vet Med Assoc* 1936; 88: 316-327
8. Fastier LB. A new feline virus isolated in tissue culture. *Am J Vet Res.* 1957; 18: 382-389
9. Green KY et al. Taxonomy of the Calicivirus. *The Journal of infectious Diseases* 2000; 181: 322-330
10. Green KY, Chanock RM, Kapikian AZ. Human Caliciviruses. *Fields Virology* 2001; Chapter 27: 841-874
11. Mayo MA. Virus Taxonomy. *Arch Virol* 2002; 147: 1071-1076
12. Black DN et al. The structure and replication of calicivirus RNA. *Nature* 1978; 274: 614-615
13. Berke T, Matson DO. Reclassification of the Caliciviruses into distinct genera and exclusion of hepatitis E virus from the family on the basis of comparative phylogenetic analysis. *Arch Virol* 2000; 145: 1421-1436
14. Jiang X, Wang K, Estes MK. Sequence and genome organisation of Norwalk virus. *Virology* 1993; 195: 51-61
15. Maniloff J. Identification and classification of viruses that have not been propagated. *Arch Virol* 1995; 140: 1515-1520
16. Dick CP, Johnson RP, Yamashiro S. Sites of persistence of feline calicivirus. *Res Vet Sci* 1989; 47: 367-373
17. Monroe SS, Ando T, Glass RI. Introduction: Human Calicivirus- An Emerging Pathogen Whose Time Has Come. *The Journal of infectious Diseases* 2000 181: 249-251
18. Glass RI et al. The Epidemiology of Enteric Caliciviruses from Humans: A Reassessment using New Diagnostics. *The Journal of Infectious Diseases* 2000; 181: 254-261
19. Berg DE et al. Multi-State Outbreaks of Acute Gastroenteritis Traced on Fecal-Contaminated Oyster Harvested in Louisiana. *The Journal of Infectious Diseases* 2000; 181: 381-386
20. Nilsson M, Hedlund K, Thorhagen M, Larson K, Johansen A, Ekspong A, Svensson L. Evolution of human calicivirus RNA in vivo: accumulation of

- mutations in the protruding P2 domain of the capsid leads to structural changes and possibly a new phenotype. *J Virol* 2003; 77: 13117-13124
21. Wyatt RG, Dolin R, Blacklow NR. Comparison of three agents of acute infectious nonbacterial gastroenteritis by cross-challenge in volunteers. *The Journal of Infectious Diseases* 1974; 129: 709-714
  22. Parrino TA, Schreiber DS, Trier JS. Clinical immunity in acute gastroenteritis caused by Norwalk agent. *N Engl J Med* 1977; 297: 86-89
  23. Baron RC et al. Serological response among teenagers after natural exposure to Norwalk Virus. *The Journal of Infectious Diseases* 1984; 150: 531-534
  24. Frankhauser RL et al. Molecular epidemiology of Norwalk-like-viruses in outbreaks of gastroenteritis in the United States. *The Journal of Infectious Diseases* 1998; 178: 1571-1578
  25. Vinjé J et al. Molecular detection and epidemiology of small round structured viruses in outbreaks of gastroenteritis in The Netherlands. *The Journal of Infectious Diseases* 1996; 174: 610-615
  26. Norcott JP et al. Genomic diversity of small round structured viruses in the United Kingdom. *J Med Virol* 1994; 44: 280-286
  27. Diseases, N.I. Viral gastroenteritis of children, Japan, 1993-98. *Infectious Agents Surveillance Report, Ministry of Health and Welfare, Japan* 1998; 19 no.11
  28. Wright PJ et al. Small round structured (Norwalk-like) viruses and classical human calici viruses in southeastern Australia, 1980-96. *J Med Virol* 1998; 55: 312-320
  29. Clarke IN, Lambden PR. Organization and Expression of Calicivirus Genes. *The Journal of Infectious Diseases* 2000; 181: 309-316
  30. Schuffenecker I et al. Genetic Classification of "Sapporo-like viruses". *Arch Virol* 2001; 146: 2115-2132
  31. Kayser FH. et al. *Medizinische Mikrobiologie Thieme* 1998: 385
  32. Prasad BVV, Hardy ME, Estes MK. Structural Studies of Recombinant Norwalk Capsids. *The Journal of Infectious Diseases* 2000. 181: 317-321
  33. Schwytzer M. Allgemeine Virologie; Genetik und Evolution der Viren. [www.vetvir.unizh.ch/Lehre/pdf\\_files/virusevol03.pdf](http://www.vetvir.unizh.ch/Lehre/pdf_files/virusevol03.pdf)
  34. Hunt M, Viral Genetics. [www.med.sc.edu:85/pptvir2002/GENETICS-2002.ppt](http://www.med.sc.edu:85/pptvir2002/GENETICS-2002.ppt), 2002; University of South Carolina School of Medicine
  35. Worobey M, Holmes EC. Evolutionary aspects of recombination in RNA viruses. *Journal of General Virology* 1999; 80: 2535-2543
  36. Worobey M, Rambaut A, Holmes EC. Widespread intraserotype recombination in natural populations of dengue virus. *Proc Natl Acad Sci USA* 1999; 96: 7352-7357
  37. Kräusslich H-G. Subvirale Agentien und Virusevolution. [www.hyg.uni-heidelberg.de/molvir/index.html](http://www.hyg.uni-heidelberg.de/molvir/index.html) 2003
  38. Jiang X et al. Characterisation of a novel human calicivirus that may be a naturally occurring recombinant. *Arch Virol* 1999; 144: 2377-2387
  39. Poppe M, Ehlken B, Rohwedder A, Lugauer S, Frank HD, Stehr K, Rieger CH, Petersen G, Lorkowski G, Karmaus W, Werchau H, Henker J, and the RoMoD-study group. Epidemiologie und Klinik von Rotavirus-

- Gastroenteritiden bei hospitalisierten Säuglingen und Kleinkindern in Deutschland. *Monatschr. Kinderheilkunde* 2002; 150 491-496
40. Rohayem J, Berger S, Juretzek T, Herchenroder O, Mogel M, Poppe J, Henker J, Rethwilm A, A simple and rapid single-step multiplex RT-PCR to detect Norovirus, Astrovirus and Adenovirus in clinical stool samples. *J Virol* 2004; 118: 49-59
  41. Thumfart, JO. Studien zur Genomorganisation und Polyproteinprozessierung bei Caliciviren von Kaninchen und Katzen. *Fakultät für Chemie und Pharmazie der EKU Tübingen, Dissertation* 2001; 1-154
  42. Haeseler v. A, Liebers D. *Molekulare Evolution*. Vol. 1. 2003, Frankfurt am Main
  43. Chare ER, Gould EA, Holmes EC, Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *J. Gen. Virol.* 2003; 84: 2691-2730
  44. Holmes EC, Worobey M, Rambaut A. Phylogenetic Evidence for Recombination in Dengue Virus. *Biology and Evolution* 1999; 16: 405-409
  45. Twiddy SS, Holmes EC. The extent of homologous recombination in members of the genus Flavivirus. *J. Gen. Virol.* 2003; 84: 429-440
  46. Posada D, Crandall KA, Holmes EC. Rekombination in Evolutionary Genomics. *Annu. Rev. Genet.* 2002; 36: 75-97
  47. Wiuf C, Christensen T, Hein J. A simulation study of the reliability of recombination detection methods. *Mol. Biol. Evol.* 2001; 18: 1929-1939
  48. Husmeier D, Wright F. Probabilistic Divergence Measures for Detecting Interspecies Recombination. *Bioinformatics* 2001; 1: 1-8
  49. Page RDM, Holmes EC. *Molecular Evolution: A Phylogenetic Approach*. Blackwell Science 1998
  50. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG, The CLUSTAL\_X windows interface: flexible strategies for multiple alignment aided by quality analysis tools. *Nucleic Acids Res.* 1997; 25: 4876-4882
  51. Gibson TJ, Higgins DG, Thompson JD, Jeanmougin F, ClustalX: Programme d'alignement multiple global avec interface graphique. *Plate-Forme Bio-Informatique de Strasbourg*.
  52. Klusmann-Kolb A, Metzler D, Printzen Ch, *Phylogenetische Analysen biologischer Sequenzdaten* 2005
  53. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999; 73: 152-160
  54. Kramer A et al. *Phylogenetische Bäume, Evolution und Mathematik, Seminararbeit zur Vorlesung Mathematik/Informatik B, 2.Semester. Molekulare Biotechnologie, Universität Heidelberg. www.iwr.uni-heidelberg.de/~Markus.Kirkilionis/ProjekteB/Phylogenie.pdf.*
  55. Sawyer S. GENECONV: a computer package for the statistical detection of gene conversion. *distributed by the author; Dept. Math. Wash. Univ. St. Louis, available at: http://www.math.wustl.edu/sawyer, 1999*

56. Sawyer S. Statistical Test for Detecting Gene Conversion. *Mol. Biol. Evol.* 1989; 6: 526-538
57. Sawyer S. *GENECONV Molecular Biology Computer Program*. 2004
58. Fan X et al. Liver Transplantation With Hepatitis C Virus-Infected Graft: Interaction between Donor and Recipient Viral Strains. *Hepatology* 2003; 38: 25-33
59. Worobey M, Holmes EC, Homologous Recombination in GB Virus C/Hepatitis C Virus. *Mol. Biol. Evol.* 2001; 18: 254-261
60. Pantle U *Monte-Carlo-Simulation*. 2003
61. Rambaut A: *LARD - Likelihood Analysis of Recombination in DNA*
62. Rambaut A, Grassly NC. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput. Appl. Biosci.* 1997; 13: 235-238
63. Graur D, Li W-H. Fundamentals of molecular evolution. *Sinauer Associates* 2000
64. Nei M, Kumar S. Molecular Evolution and phylogenetics. *Oxford University Press* 2000
65. Nagy PD, Bujarski JJ. Efficient system of homologous RNA recombination in brome mosaic virus: sequence and structure requirements and accuracy of crossovers. *J Virol* 1995; 69: 131-140
66. Nagy PD, Bujarski JJ, Engineering of homologous recombination hotspots with AU-rich sequences in brome mosaic virus. *J Virol* 1997; 71: 3799-3810
67. Nagy PD, Pogany S, Simon AE. RNA elements required for RNA recombination function as replication enhancers in vitro and in vivo in a plus-strand RNA virus. *Embo J.* 1999 18: 5653-5665
68. Nagy PD, Zhang C, Simon AE. Dissecting RNA recombination in vitro: role of RNA sequences and the viral replicase. *Embo J.* 1998; 17: 2392-2403
69. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003; 31: 3406-3415
70. Gräf S, *Skript zum Praktikum: Angewandte Bioinformatik II* 2003
71. Rohayem J, Münch J, Rethwilm A. Evidence of Recombination in the Norovirus Capsid Gene. *J Virol* 2005; 79: 4977-4990
72. Bruen TC, Bryant D. A Simple and Robust Statistical Test for Detecting the Presence of Recombination. *Genetics Society of America* 2006; 127: 2665-81
73. Bon F, Fascia P, Dauvergne M, Tenenbaum D, Planson H, Petion AM, Pothier P, Kohli E. Prevalence of group A rotavirus, human calicivirus, astrovirus, and adenovirus type 40 and 41 infections among children with acute gastroenteritis in Dijon, France. *J. Clin. Mikrobiol.* 1999; 37: 3055-3058
74. Oh DY, Gaedicke G, Schreier E. Viral agents of acute gastroenteritis in German children: prevalence and molecular diversity. *J Med Virol* 2003; 71: 82-93
75. Meyerhans A, Maier R, Vartanian JP, Wain-Hobson S. The non-clonal and transitory nature of HIV in vivo. *swiss medical weekly* 2003; 133: 451-454

76. Iritani N, Seto Y, Kubo H, Murakami T, Haruki K, Ayata M, Ogura H. Prevalence of Norwalk-Like Virus Infections in Case of viral Gastroenteritis among Children in Osaka City, Japan. *J. Clin. Mikrobiol.* 2003; 41: 1756-1759
77. Marshall JA, Salamone S, Yuen L, Catton MG, Wright JP. High level excretion of Norwalk-like virus following resolution of clinical illness. *Pathology* 2001; 33: 50-52
78. Posada D. Evaluation of methods for detecting recombination from DNA sequences: empirical data. *Mol. Biol. Evol.* 2002; 19: 708-717
79. Posada D, Crandall KA. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci USA* 2001 98: 13757-13762
80. Burke DS. Recombination in HIV: an important viral evolutionary strategy. *Emerg. Infect. Dis.* 1997; 3: 253-259
81. Robertson DL, Sharp PM, McCutchan FE, Hahn BH. Recombination in HIV-1. *Nature* 1995; 374: 124-126
82. Prasad BVV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK. X-ray Crystallographic Structure of the Norwalk Virus Capsid. *Science* 1999; 286: 287-290
83. Nagy PD, Bujarski JJ. Genetic recombination in brome mosaic virus: effect of sequence and replication of RNA on accumulation of recombinants. *J Virol* 1992; 66: 6824-6828
84. White KA, Morris TJ. Nonhomologous RNA recombination in tombusviruses: generation and evolution of defective interfering RNAs by stepwise deletions. *J Virol* 1994; 68: 14-24
85. White KA, Morris TJ. Recombination between defective tombusvirus RNAs generates functional hybrid genomes. *Proc Natl Acad Sci USA* 1994 91: 3642-3646
86. Nagy PD, Simon AE. In vitro characterization of late steps of RNA recombination in turnip crinckle virus. Role of motif1-hairpin structure. *Virology* 1998; 249: 379-392
87. Nagy PD, Ogiela C, Bujarski JJ. Mapping Sequences active in homologous RNA recombination in brome mosaic virus: prediction of recombination hot spots. *Virology* 1999; 254: 92-104
88. Nagy PD, Bujarski JJ. Homologous RNA recombination in brome mosaic virus: AU-rich sequences decrease the accuracy of crossovers. *J Virol* 1996; 70: 415-426
89. Lopman BA, Reacher MH, Van Duijnhoven Y, Hanon FX, Brown D, Koopmans M. Viral gastroenteritis outbreaks in Europe 1995-2000. *Emerg. Infect. Dis.* 2003; 9: 90-96

## 7 ANHANG

### 7.1 ABBILDUNGSVERZEICHNIS

Abbildung 1:	Maximum likelihood Phylogram zur Darstellung der Verwandtschaftsverhältnisse innerhalb der Caliciviridae und zu den Picornaviridae	2
Abbildung 2:	Schematische Darstellung von Genomorganisation und Benutzung der Open-Reading-Frames (ORF) und subgenomischer RNAs anhand exemplarischer Caliciviren der einzelnen Genera	9
Abbildung 3:	Ikosaeder aus unterschiedlichen Perspektiven	11
Abbildung 4:	Elektronenmikroskopische Darstellung des Capsids	12
Abbildung 5:	Geometrie des Virus	12
Abbildung 6a, b, c:	Das Capsidprotein mit Darstellung der variablen und konservierten Regionen	14
Abbildung 7a, b:	a: Elektronenmikroskopische Darstellung des Norovirus Capsid Proteins, b: Zentraler Querschnitt eines rVLPs	15
Abbildung 8:	Schematische Darstellung des NV-Genoms	53
Abbildung 9:	UV-Licht-Photo der RT-PCR Proben in der Gel-Elektrophorese	54
Abbildung 10:	UV-Licht-Photo der Gel-Elektrophorese der nPCR. Dresden174 amplifiziert mit unterschiedlichen Primern	55

Abbildung 11:	UV-Licht-Photo der Gel-Elektrophorese nach Aufreinigung	56
Abbildung 12 :	Evolutionäre Verwandtschaftsverhältnisse innerhalb der NV-GG II/4 basierend auf der Nukleotidabfolge der gesamten Capsidsequenz	59
Abbildung 13:	Durch SplitsTree anhand der Sequenzen erstellter phylogenetischen Baum	62
Abbildung 14:	Similarity Plot des rekombinanten Stamms Hamburg180/1997/GE mit den Parentalstämmen Hamburg189/1997/GE und Bochum024/1998/GE	63
Abbildung 15:	ML-Bäume der rekombinanten Stämme Hamburg180/1997/GE und Hamburg/137/1997/GE	64
Abbildung 16:	Aminosäuren-Alignment der vollständigen Capsidsequenzen	66
Abbildung 17:	Charakterisierung der Zusammensetzung der Sequenz und der RNA-Sekundärstruktur der wahrscheinlich rekombinanten Region	68

## 7.2 TABELLENVERZEICHNIS

Tabelle 1:	Länge der amplifizierten Fragmente in Kilobasen	53
Tabelle 2:	GeneBank accession numbers der von uns amplifizierten und sequenzierten Stämme	57

Tabelle 3:	Sawyer's Test	61
Tabelle 4:	Breakpoint Analyse der Rekombinanten und ihrer wahrscheinlichen Parentalstämme	65

### 7.3 ABKÜRZUNGSVERZEICHNIS

µM	Mikromolar
A, mA	Ampère, Milliampère
aa	amino acids
BriV	Bristol-Virus
C	Konzentration
CamV	Camberwell-Virus
CarV	Carlton-Virus
ChiV	Chiba-Virus
CSG	Caliciviridae Study Group
CV	Calicivirus
ddNTPs	Didesoxyribonucleosidtriphosphate
DIC	Disseminated Intravasal Coagulation
DMSO	Dimethyl Sulfoxide
DNA	Desoxyribonucleinsäure
dNTPs	Desoxyribonucleosidtriphosphate
DSV	Desert-Shield-Virus
EM	Electron Microscopy
EpR	Eppendorf-Röhrchen
et al.	et alii
GE	Gastroenteritis
h	Stunde
HCl	Salzsäure
HEL	Helicase

HesV	Hesse-Virus
HilV	Hillingdon-Virus
HuCV	Human Calicivirus
HV	Hawaii-Virus
IEM	Immuno-Elektronenmikroskopie
ITCV	International Committee on the Taxonomy of Viruses
JBV	Jena-Bovine-Virus
kb	Kilobasen
KCl	Kalium-Chlorid
kDa	kilo Dalton
KEM	Kryo-Elektronenmikroskopie
LB-Medium	Luria-Bertani-Medium
LeeV	Leeds-Virus
LorV	Lordsdale-Virus
LR	Likelihood Ratio
M	Molar
MeIV	Melksham-Virus
min.	Minute
ML-Methode	Maximum-Likelihood-Methode
Mm	Millimolar
Mt-Keira-Virus	MtKV
MusV	Musgrove-Virus
MX	Mexico-Virus
nm	Nanometer
NorV	Norwalk-Virus
nPCR	nested PCR
NS	Nukleinsäure
nt	Nukleotide
NTPase	Nucleosidtriphosphatase
NV	Norovirus
ORF	Open-Reading-Frame

PEC	Porcine-Enteric-Calicivirus
Pol	RNA-abhängige RNA-Polymerase
Pro	Proteinase
QAV	Queens-Arms-Virus
RNA	Ribonucleinsäure
rpm	rounds per minute
RT	Raumtemperatur
SeaV	Seacroft-Virus
sec	Sekunde
SinV	Sindlesham-Virus
SMV	Snow-Mountain-Virus
SouV	Southampton-Virus
SSIIRT	SuperScript II H-Reverse Transcriptase
SV	Sapovirus
Tm	Schmelztemperatur
TorV	Toronto-Virus
TSR	Template Suppressing Reagent
V	Virus
VLP	Virus Like Particles
VPg	Virus Protein
WinV	Winchester-Virus
z. B.	zum Beispiel

## Danksagung

Ich danke Herrn Prof. Rethwilm für das Thema meiner Dissertationsarbeit und das in mich gesetzte Vertrauen. Auf seine fachliche Betreuung konnte ich mich zu jeder Zeit verlassen, so dass ich in angenehmer Atmosphäre lernen konnte wissenschaftlich zu denken und zu arbeiten. Darüber hinaus bin ich sehr dankbar für die überaus rasche und präzise Korrektur meiner Dissertationsschrift.

Ich danke Jacques Rohayem für die freundschaftliche Zusammenarbeit, sein Engagement und die stetige Bereitschaft meine Fragen und Anregungen sinnvoll mit mir zu diskutieren. Wir waren ein gutes Team!

Zudem danke ich allen Mitarbeitern der Virologie der TU Dresden, ohne deren entgegenkommende Hilfsbereitschaft diese Arbeit nicht möglich gewesen wäre.

Ich danke Mama, Papa, Johannes und Oma, weil sie mich bestärkt haben und mir auch heute mit Rat und Tat zur Seite stehen. Meiner Mutter danke ich für das Lesen meiner Dissertationsschrift und weil sie für mich immer ein offenes Ohr hat. Mein Vater ist stets zur Aufmunterung parat, Oma ist die vernünftigste alte Lady, die ich kenne und Johannes und ich befinden uns füreinander ohnehin 24/7 in Rufbereitschaft ;-).

Ich danke Euch.

Von meinen Freunden danke ich vor allem Daniel, der mir während der gesamten Arbeit zur Seite stand, und ein Freund ist, auf den ich mich immer verlassen kann, und Anusha, die mich am Tag meines Staatsexamens bereits daran erinnert hat, dass *jetzt* der Zeitpunkt sei, meine Dissertation zu Ende zu schreiben ;-); ich danke Dir für die jahrelange intensive Freundschaft, die uns verbindet.

Ich danke Raikie für seine stetige Bereitschaft zur freien Meinungsäußerung, Regina, für ihre Geradlinigkeit und Hilfsbereitschaft, Jan, für einen guten Start in Hamburg, Gereon für die vielen interessanten Gespräche und Diskussionen und Amy, Alex, Dan, Matt, Greg und Shilen, mit denen ich eine tolle Zeit in London hatte.

And last, but not least, thank you Steve, for your loving patience and support.

Ohne Euch hätte ich es vielleicht auch geschafft, aber es wäre nicht so schön gewesen  
- danke, danke, danke!!





