

**Zeitschrift:** Schweizer Archiv für Tierheilkunde SAT : die Fachzeitschrift für Tierärztinnen und Tierärzte = Archives Suisses de Médecine Vétérinaire ASMV : la revue professionnelle des vétérinaires

**Herausgeber:** Gesellschaft Schweizer Tierärztinnen und Tierärzte

**Band:** 138 (1996)

**Heft:** 2

**Artikel:** Identifikation und Charakterisierung von neuen und unbekanntem Coronaviren mit Hilfe von RT-PCR und degenerierten Primern

**Autor:** Tobler, K. / Ackermann, M.

**DOI:** <https://doi.org/10.5169/seals-589671>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

**Download PDF:** 30.01.2025

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

## Identifikation und Charakterisierung von neuen und unbekanntem Coronaviren mit Hilfe von RT-PCR und degenerierten Primern

K. Tobler und M. Ackermann

### Zusammenfassung

Um das Genom des Epizootischen Virus Diarrhöe Virus (EVDV) untersuchen zu können, wurde eine spezielle Methode der reversen Transkription und Polymerase Kettenreaktion (RT-PCR) entwickelt. Die publizierten Sequenzen von verschiedenen Coronaviren wurden miteinander verglichen und auf Regionen untersucht, welche bei allen Vertretern vorkommen. Solche Stellen wurden auf der Ebene der Aminosäuresequenz im Gen für das Nichtstrukturprotein 1b (ORF1b), sowie in den Genen für die Hauptstrukturproteine (S, M und N) gefunden. Aufgrund der Degeneriertheit des genetischen Codes kann eine einzige Aminosäure durch verschiedene Nukleinsäure-Triplets kodiert werden. Bei der Synthese von Primern für die RT-PCR wurde dieser Degeneriertheit Rechnung getragen. Es wurden Primergemische hergestellt, welche alle möglichen Varianten in der Nukleinsäuresequenz berücksichtigten. Mit Hilfe dieser Methode wurde insgesamt ein circa 4000 Basen grosses Stück des Genoms von EVDV durch RT-PCR amplifiziert und kloniert. Die Analyse der Genomsequenz ergab, dass EVDV eine interessante Stelle zwischen dem Virus der Transmissiblen Gastroenteritis (TGE) der Schweine und dem Humanen Coronavirus 229E einnimmt. Es ist anzunehmen, dass die hier eingeschlagene Strategie auch für die Analyse anderer, bisher unbekannter Coronaviren und anderer nicht züchtbarer Viren erfolgreich sein könnte.

**Schlüsselwörter:** Coronavirus – EVD-Virus – RT-PCR – degenerierte Primer

### Identification and characterisation of new and unknown coronaviruses by using RT-PCR and degenerated primers

A modified method of the reverse transcription followed by polymerase chain reaction (RT-PCR) was developed in order to examine the genome of a recently discovered virus, the porcine epidemic diarrhoea virus (PEDV), which resembled morphologically the coronaviruses. The published sequences of the genomes of various coronaviruses were compared. On the level of the amino acid sequence, conserved regions, common to all coronaviruses, were found in the gene encoding the nonstructural protein 1b as well as in the genes coding for the major structural proteins (S, M, and N). Due to the degeneration of the genetic code, some amino acids may be encoded by different nucleotide triplets. In order to compensate for this degeneration, mixtures of primers were synthesized, containing a variety of nucleotide sequences which together represented all possible codons for the conserved amino acid sequences. This method allowed to amplify and clone approximately 4000 base pairs of the genome of PEDV. An analysis of the genomic sequences revealed that PEDV holds an interesting intermediate position between human coronavirus 229E and Transmissible Gastroenteritis virus. We postulate that the method presented in this contribution may be useful to study and characterize other unknown viruses, especially viruses for which no cell cultures for propagation are available.

**Key words:** coronavirus – PED virus – RT-PCR – degenerated primers

### Einleitung

Das Virus der Epizootischen Virus Diarrhöe (EVD, EVDV) wurde Ende der 70er Jahre erstmals als Erreger von akuten Durchfallerkrankungen bei Schweinen allen Alters beschrieben (Pensaert und Debouck, 1978). Vor allem bei neugeborenen Ferkeln nahm die Infektion oft einen fatalen Verlauf. Das Krankheitsbild erinnerte stark

an dasjenige der Transmissiblen Gastroenteritis (TGE), welche von einem Coronavirus, dem TGEV, verursacht wird. Elektronenmikroskopische Aufnahmen des EVD-Virus zeigten tatsächlich coronavirus-ähnliche Strukturen. Andererseits konnte keine serologische Verwandtschaft zwischen TGEV und dem EVD-Virus gezeigt werden. Es dauerte danach fast 10 Jahre, bis ein Zellkultur-System für die Vermehrung von EVDV *in vitro* gefunden

wurde (Hofmann und Wyler, 1988). Aus diesem Grunde wurden lange Zeit keine Fortschritte in der Forschung über EVDV erzielt.

Die Familie der Coronaviridae umfasst zwei Genera (Coronavirus und Torovirus), wobei das Genus Arterivirus zu den coronaähnlichen Viren gezählt wird. Das Genus der Coronaviren wird in vier Gruppen gegliedert, wobei die Coronaviren der Gruppen 1 und 2 die Vertreter umfassen, welche bei Säugetieren und beim Menschen vorkommen. Wichtige Vertreter der Gruppe 1 sind TGEV und das Virus der Felinen Infektiösen Peritonitis (FIPV), aber auch das Humane Coronavirus 229E (HCV229E). In der Gruppe 2 sind das Bovine Coronavirus (BCV), das Maushepatitis Virus (MHV) und das Humane Coronavirus OC43 (HCVOC43) zu nennen. Die Gruppeneinteilung beruht auf antigenetischer Verwandtschaft und auf Ähnlichkeiten im Aufbau des Genoms. Um EVDV innerhalb der Coronaviren näher zu klassieren und zugleich eine Basis für weitere Untersuchungen zu bekommen, schien es sinnvoll, die Nukleinsäuresequenz des Genoms von EVDV zu bestimmen.

Das Genom der Coronaviren besteht aus einer einzelsträngigen, polyadenylierten Ribonukleinsäure (RNS) mit positiver Polarität. Seine Grösse von 25000 bis 30000 Basen (25 bis 30 Kilobasen, kb) macht es zu einem der längsten RNS-Moleküle, die in der Natur vorkommen. Der typische Aufbau eines Genoms der Coronaviren mit ORF1a (ORF = Open Reading Frame, offenes Leseraster) und ORF1b am 5' Ende und den Strukturprotein-Genen (S, M und N) am 3' Ende ist in Abbildung 1 dargestellt. Zwischen den einzelnen Genen gibt es sogenannte intergene (IG) Sequenzen. Bei der Replikation dienen diese IG-Sequenzen als Startpunkte für die Synthese von subgenomischer RNS. Die Synthese wird dann bis zum 3' Ende des Genoms fortgesetzt. Auf diese Weise entsteht ein sogenanntes «nested set» von RNS, wobei alle das 3' Ende mitbeinhalten, während auf der 5' Seite, je nach der Stelle, an der die Synthese begonnen hat, mehr oder weniger Gene vorhanden sind (Abbildung 1). Zu Beginn unserer Studie war von der Nukleinsäuresequenz von EVDV überhaupt nichts bekannt. Von anderen Coronaviren hingegen war recht viel Information verfügbar. Die Genome von zwei Coronaviren (MHV und Infektiöse Bronchitis der Hühner, IBV) waren vollständig durchsequenziert worden. Von etwa acht anderen waren mindestens Sequenzinformationen im Bereich des 3' En-

des publiziert worden (Tobler et al., 1994 und Referenzen darin).

Unser Plan bestand darin, die publizierten Sequenzen der verschiedenen Coronaviren miteinander zu vergleichen und nach Regionen zu suchen, die bei allen Vertretern vorkamen. Aufgrund dieser konservierten Sequenzen wollten wir Oligonukleotide synthetisieren und als Primer für die reverse Transkription und die Polymerase-Kettenreaktion (PCR) verwenden. Die amplifizierten Stücke, welche Teile des Genoms von EVDV repräsentierten, sollten dann kloniert und sequenziert werden. Im vorliegenden Manuskript wird beschrieben, wie diese Strategie zur Klassifizierung von EVDV innerhalb der Familie der Coronaviridae führte.

## Material und Methoden

**Vergleich der publizierten Coronavirussequenzen:** Mit dem GCG (Genetic Computing Group; University of Wisconsin)-Programm wurden bereits publizierte Sequenzen von Coronaviren und verwandten Viren miteinander verglichen. Aus den konservierten Regionen wurden degenerierte Primer für RT-PCR-Amplifikationen abgeleitet. Die Primersequenzen sind in Tabelle 1 dargestellt, während die Lokalisation der Primer in Abbildung 1 zu sehen ist. Sämtliche verwendeten Primer wurden von der Firma Microsynth (Balgach) synthetisiert.

**RNS-Isolation:** Es wurde virale RNS aus gereinigten Viren oder isolierte total RNS aus infizierten Zellen verwendet. Die RNS wurde nach Kingston (1991) mit Guanidin-thiocyanat extrahiert, wie es andernorts bereits beschrieben wurde (Bridgen et al., 1993; Duarte et al., 1994).

**Reverse Transkription:** Die Reverse Transkription (RT) wurde in zwei Schritten nach dem Protokoll von Wirth et al. (1991) durchgeführt. In einem Primerbindungsschritt wurde die RNS von zirka 1.5 Mio. infizierten Zellen mit 500 ng Primer 24 oder 25(dT) für eine Stunde bei 42 °C inkubiert (10 µl Reaktionsvolumen). Um die RNS vor Degradation durch RNase zu schützen, wurde das RNS/Primer-Gemisch mit 20 Einheiten RNasin (Promega) ergänzt. Danach wurden 20 Einheiten AMV-reverse Transkriptase (Pharmacia), dNTPs und Puffer zugegeben und für weitere zwei Stunden bei 42 °C inkubiert.

Tabelle 1: Primer, die für RT-PCR verwendet wurden.

Primer Nr.	Sequenz	Bemerkungen
23	AAGCTTTTACTAYTTDGGHACAGGACC	18-fach degeneriertes 27-mer
24	CTCGAGCGACCCAGAMGACWCCKTC	8-fach degeneriertes 25-mer
25	GACTAGTTGGTGGAGWTTTAAAYCCWGA	8-fach degeneriertes 27-mer
25(dT)	GACTAGTTGGTGGAGWTTTAAAYCCWGA (T) 20-30	8-fach degeneriert
50	CCTCTACAAGCAATGTACC	nicht degeneriertes 19-mer
55	GGAKAAGGTKAATGARTGYGT	16-fach degeneriertes 21-mer
56	CCAKACVTACCAWGGCCAYTT	24-fach degeneriertes 21-mer

Oligos, welche die selbe Orientierung wie das Genom haben, sind mit ungeraden, solche mit komplementärer Orientierung mit geraden Nummern bezeichnet. Degenerierte Stellen, welche durch verschiedene Basen besetzt sein können, wurden nach dem Code der «International Union of Biochemistry» dargestellte (R: A oder G; Y: C oder T; H: A, C oder T; V: A, C oder G; M: A oder C; K: G oder T; B: C, G oder T; D: A, G oder T; S: G oder C; W: A oder T).

Die entstandene komplementäre DNS (cDNS) wurde anschliessend mit Phenol/Chloroform extrahiert und in 20 µl Wasser gelöst. Davon wurden jeweils 2 µl für die Amplifikation in der PCR verwendet.

**Reverse Transkription des 3' Endes der viralen RNS:** Um das polyadenylierte 3' Ende des viralen Genoms zu amplifizieren, wurde das RACE-Protokoll (**R**apid **A**mplification of **c**DNA **E**nds; Frohman et al., 1988) benutzt. Dabei wird zur Erhöhung der Spezifität bei der späteren PCR und um gleichzeitig die dem 3' Ende des Genoms komplementäre DNS (cDNS) mit einer bestimmten Sequenz zu markieren, für die reverse Transkription ein beliebiger Primer mit bekannter Sequenz durch Terminale Desoxyribonukleotidyl Exotransferase (TDT) mit dT am 3' Ende verlängert. Dieser modifizierte Primer wird zur cDNS-Synthese verwendet, wobei er mit seinem poly(dT)-Teil an den poly(A)-Teil der RNS bindet. Während der Synthese entsteht dann eine einzelsträngige cDNS, die am 5' Ende unmittelbar vor dem poly(dT) die bekannte Primersequenz aufweist.

In unserer Anwendung wurde der Primer 25 mit TDT (Boehringer Mannheim) für 30 Minuten bei 37 °C behandelt. Der mitgelieferte Reaktionspuffer wurde mit 1.5 mM Kobaltchlorid ergänzt. Das molare Verhältnis von Primer und dTTP in der Reaktion war 1:50. Unter den gewählten Reaktionsbedingungen verlängerte sich der Primer um jeweils 20–30 Tymidine. Der modifizierte Primer wurde als 25(dT) bezeichnet.

**PCR-Amplifikationen:** Das Amplikon mit den Primern 24 und 25 wurde unter folgenden Reaktionsbedingungen produziert: 38 Zyklen mit 50 Sekunden bei 94 °C (Denaturierung), 1 Minute bei 48 °C (Bindung der Primer) und 1 Minute bei 72 °C (Elongation). Für die PCR-Amplifikation der cDNA im RACE-Experiment wurden die Primer 23 und 25 (jetzt ohne poly(dT)-Verlängerung) verwendet. Die Reaktion erfolgte in 40 Zyklen mit 50 Sekunden bei 94 °C, 1 Minute bei 47 °C und 2.5 Minuten bei 72 °C. Die Reaktionsgemische setzten sich standardmässig aus je 10 µM Primer, 208 µM dNTP und 0.1 Einheiten *Taq*-Polymerase (Perkin Elmer Cetus) zusammen. Die Reaktion mit einem Gesamtvolumen von 30 µl hatte eine Pufferzusammensetzung von 50 mM Kaliumchlorid, 10 mM Tris pH 8.3, 1 mM Magnesiumchlorid, 10 µg/ml Gelatine, 0.045% NP-40 und 0.045% Tween20. Diese beiden Reaktionen wurden in einem Hybaid Intelligent Heating Block (Model IHB 2024) durchgeführt.

Die Amplifikation mit den Primern 55 und 56 wurde mit *Pfu*-Polymerase (Stratagene) und dem mitgelieferten Puffer unter folgenden Bedingungen erfolgreich durchgeführt: 1 Zyklus mit 5 Minuten bei 95 °C, 1 Minute bei 45 °C und 5 Minuten bei 74 °C und anschliessend 39 Zyklen mit 50 Sekunden bei 95 °C, 1 Minute bei 45 °C und 2.5 Minuten bei 74 °C. Für die PCR-Reaktionen mit *Pfu*-Polymerase wurden dieselben Primerkonzentrationen verwendet. Als Reaktionspuffer wurde der mitgelieferte Puffer verwendet. Es wurden 2, 5 Einheiten Enzym pro Reaktion eingesetzt. Das Reaktionsvolumen war 30 µl.

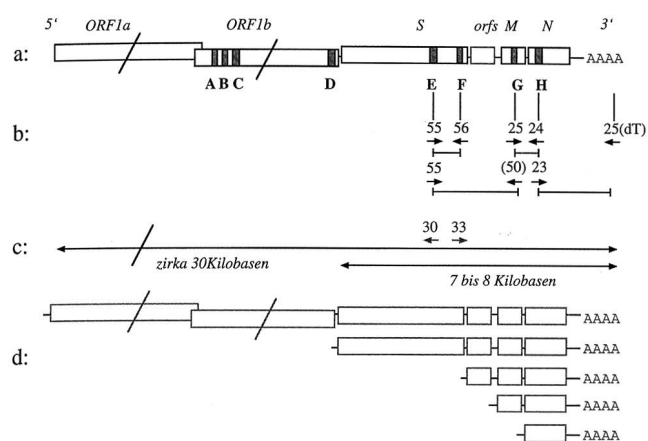
Die Amplikons wurden mit gebräuchlichen Methoden in die *EcoRV*-Stelle des Plasmids Bluescript KS(+) (Stratagene) kloniert und mit dem Sequenase 2 Kit (United States Biochemical Corporation, USB, Cleveland, Ohio) sequenziert. Es wurden mindestens drei Stränge von zwei Klonen aus zwei verschiedenen PCR-Reaktionen sequenziert, um mögliche Fehler der Reversen Transkriptase oder der DNS-Polymerase zu erkennen.

## Resultate

Zu Beginn des Projektes waren keine Sequenzinformationen zu EVDV verfügbar. Es lagen nicht einmal Untersuchungen über die genomische Struktur vor, und Plasmide mit klonierter cDNS des viralen Genoms existierten nicht. Um die Erarbeitung dieser Daten möglichst gezielt vorzubereiten, wurden zunächst die verfügbaren Datenbanken nach Nukleotid- und Proteinsequenzen anderer Coronaviren und coronaähnlicher Viren durchsucht. Auf diesem Weg wurden (Teil-)Sequenzen von 12 Coronaviren gefunden. Weil im Bereich der ORF1-Region nur gerade zwei Coronaviren sequenziert worden waren, wurden hier die publizierten Sequenzen coronaähnlicher Viren (Equines Arteritis Virus und Bern Virus) in die Analysen miteinbezogen (Referenzen betreffend der verwendeten Sequenzen in Tobler et al., 1994). Die vergleichenden Analysen wurden auf dem Niveau der Proteine, der Aminosäuresequenzen, durchgeführt. Die gefundenen homologen Regionen sind als dunkle Balken in den offenen Leserastern in Abbildung 1 dargestellt und mit den Buchstaben A bis H bezeichnet. Die konservierten Stellen in den Genen für ORF1b, das Spikeprotein (S), das Membranprotein (M) und das Nukleoprotein (N) dienten als Grundlage für die Synthese von Primern für die reverse Transkription und die PCR.

Im genetischen Code steht jedes Triplet von Nukleotiden für ein bestimmtes Signal oder eine bestimmte Aminosäure. Die einzelnen Aminosäuren können jedoch entweder von einem einzigen oder aber von mehreren, maximal sechs verschiedenen Triplets kodiert werden. Bei der Synthese der Primer wurde diesem Umstand Rechnung getragen, indem sogenannte degenerierte Primer hergestellt wurden. Entsprechend der Primersequenz einer Aminosäure, welche durch unterschiedliche Triplets kodiert wird, so synthetisierten wir alle möglichen Formen als Gemisch. Dadurch entstanden Primer mit einem Degenerationsgrad von achtzehn bis vierundzwanzig. Die Länge der Primer wurde mit vierundzwanzig Nukleotiden (Basen) konstant gehalten. Die verwendeten degenerierten Primer, welche zu einer erfolgreichen Amplifikation führten, sind in Abbildung 1 an den betreffenden Bindungsstellen dargestellt. Mit den Primern 24 und 25 konnten 0.7 Kilobasenpaare (kbp) der cDNS amplifiziert werden. Das amplifizierte Stück stellte etwa die Hälfte des M-Gens und einen Teil des N-Gens dar. Um das 3' Ende des Genoms als cDNS-Klon zu bekommen, wurde die RACE-Methode von Frohman (siehe Material und Methoden) angewendet. Am 3' Ende des





**Abb. 1a:** Schematische Darstellung eines typischen Coronavirusgenoms. Die offenen Leseraster sind als Rechtecke dargestellt. Oberhalb der Darstellung sind die Gene bezeichnet: Polyprotein (ORF1a und ORF1b; nicht massstabgetreu gezeichnet), Spikeprotein (S), diverse ORFs (orfs), Membranprotein (M) und Nukleokapsidprotein (N). In der ganzen Familie der Coronaviren konservierte Regionen sind mit den Buchstaben A bis H bezeichnet: **A:** GDD-Motiv, **B:** cysteinreiche Region, **C:** Helikase-Motiv, **D:** konservierte Region, **E:** zirka 300 aa lange konservierte Region in S, **F:** konservierte Region im S-Protein mit Konsensussequenz YVKWPKYVWL, **G:** konservierte Region im M-Protein mit Konsensussequenz SWWSFNPE, **H:** konservierte Region im N-Protein: GYW-(4aa)-R-(4-7aa)-G-(10aa)-FYYLGTGP-(11aa)-DGV(V/F)WVA. Die konservierten Regionen E, F, G und H wurden verwendet, um degenerierte Primer für RT-PCR zu entwickeln. **1b:** Die Primer wurden in der Skizze mit den entsprechenden Nummern eingezeichnet. Die resultierenden PCR-Produkte wurden ebenfalls eingezeichnet. Primer 25(dT) am 3' Ende des Genoms wurde für das RACE-Experiment gebraucht. Primer 50 ist ein nicht degenerierter Primer, der aufgrund von ersten Sequenzanalysen synthetisiert werden konnte. Die PCR-Produkte decken etwa 4 Kilobasen am Ende des EVD-Genoms ab. (**1c:** Massstab) **1d:** Darstellung der subgenomischen mRNA, die während des Infektionszyklus in den infizierten Zellen transkribiert werden. Die Anzahl der mRNA unterscheidet sich zwischen den verschiedenen Coronaviren.

Primers 25 wurde ein poly(dT) Schwanz angehängt und komplementäre DNS (cDNS) wurde auf der viralen RNS-Vorlage synthetisiert. Auf dieser Vorlage wurde mit den Primern 23 und 25 der restliche Teil des N-Gens sowie die nicht translatierte Region des genomischen 3' Endes bis zum poly(A) Schwanz amplifiziert. Es entstand ein 1,6 kbp grosses PCR-Produkt. Mit den Primern 55 und 56, deren Bindungsstellen im S-Gen lokalisiert sind, konnte ein weiteres 0,7 kbp Produkt amplifiziert werden. Dies gelang allerdings erst, nachdem die *Taq*-Polymerase durch *Pfu*-Polymerase ersetzt und zugleich eine neue PCR Maschine in Betrieb genommen worden war. Um

die Lücke zwischen den Bindungsstellen der Primer 56 und 25 zu schliessen, wurde schliesslich ein normaler, nicht degenerierter Primer synthetisiert (Primer 50), welcher aufgrund der ersten Sequenzierresultate entwickelt werden konnte. Die Primer 30 und 33 waren ebenfalls in der RT-PCR verwendet worden, aber es gelang nicht, damit spezifische Produkte zu amplifizieren, obwohl die Bedingungen variiert wurden.

Die mit PCR amplifizierten Stücke des Genoms wurden anschliessend in Plasmiden kloniert und sequenziert. Da mit einer geringen Prozentzahl von Ablesefehlern der reversen Transkriptase und der thermophilen DNS-Polymerase bei der RT-PCR zu rechnen war, wurden Klone aus mindestens zwei verschiedenen PCR-Produkten sequenziert. Die Sequenz wurde mindestens dreimal bestimmt, wobei beide Ableserichtungen berücksichtigt wurden. Insgesamt wurde die Sequenz von etwa 4000 Basenpaaren am 3' Ende des Genoms von EVDV bestimmt.

Detaillierte Analysen der Sequenzresultate sind in anderen Zeitschriften bereits publiziert worden (Bridgen et al., 1993, und Duarte et al., 1994). Das Genom von EVDV erwies sich als typisch für Coronaviren der Gruppe 1. Der engste Verwandtschaftsgrad bestand zum Humanen Coronavirus 229E, etwas schwächer erwies er sich zum Virus der Transmissiblen Gastroenteritis. Zu den anderen Coronaviren fiel die Homologie kontinuierlich ab. Relativ zum Maus-Hepatitis-Virus und zum Virus der Infektiösen Bronchitis des Huhns lagen noch Homologien im Bereich von 15 bis 20% vor. Die Gene für das N-Protein, das M-Protein, das sM-Protein und Teile des S-Proteins von EVDV, HCV229E und TGEV wiesen eine Identität in der Aminosäuresequenz von zwischen 30% und 60% auf. Im Genomabschnitt zwischen dem S-Gen und dem sM-Gen wurde ein möglicher Offener Leseraster gefunden, der in den verschiedenen sequenzierten cDNS-Klonen verschiedene Deletionen und Mutationen aufwies. In diesem Genomabschnitt unterscheiden sich auch andere Coronaviren markant voneinander.

Anschliessend an die Sequenzbestimmung wurde ein Vergleich der degenerierten Primer mit den tatsächlichen Sequenzen auf dem viralen Genom durchgeführt. Die Ergebnisse sind in Abbildung 2 dargestellt. In der Mehrzahl der Fälle ergab die Degeneration in mindestens einem Primer des Gemisches ein passendes Nukleotid. Es zeigte sich jedoch, dass eine Anzahl von nicht aufeinander passenden Nukleotiden (mismatches) toleriert wurde, ausser wenn sich diese zu nahe am 3' Ende des Primers befanden, wo die Elongation stattfinden sollte. Ein anderer Primer, der das Nukleotid Analog Inosin enthielt, das im Prinzip mit allen Nukleotiden paaren sollte, führte nicht zu einer erfolgreichen Amplifikation.

**Abbildung 2: Primer, die in RT und PCR verwendet wurden. Die Primersequenzen (Primer) werden mit den mit der Sequenz des EVDV-Genoms (EVDV-Genom) und der davon abgeleiteten Aminosäuresequenz (EVDV aa Sequenz) verglichen. Die Nukleotide des Genoms, welche zu denjenigen des Primer passten, sind schwarz schattiert dargestellt, um die Bindung zu veranschaulichen. Die Aminosäuresequenzen (EVDV) werden zudem den Consensussequenzen (Consensus) aus dem Vergleich mit den anderen Coronaviren gegenübergestellt. Die bindende Genomsequenz komplementär gerichteter Primer sind in einer zusätzlichen Linie (invers EVDV) umgekehrt dargestellt, um den Vergleich übersichtlicher zu gestalten. In Klammern wird angegeben, ob der Primer in dem RT-PCR-Experiment erfolgreich war. Die Primer 30 und 33 haben in den Experimenten nicht funktioniert.**

<p>23: AAGCTTTTACTA (T/C) TT (T/A/G) GG (A/C/T) ACAGGACC  <b>T T T C T A C T A T A C</b> C C T C C G G A A C A G G A C C T            F Y Y L G T G P            F Y Y L G T G P            (F)</p>	<p>Primer im N-Gen (PCR)            EVDV Genom            EVDV aa Sequenz            Consensus</p>
<p>24: CTCGACGACCCAGA (C/A) GAC (A/T) CC (G/T) TC  <b>C T T T A G C A A C C C A G A A A C A C C C T C</b>            G A G G G T G T T T T C T G G G T T G C T A A A G            E G V F W V A            D G V F W V A            (E) (I) V</p>	<p>Primer im N-Gen (PCR)            invers EVDV            EVDV Genom            EVDV aa Sequenz            Consensus</p>
<p>25: GACTAGTTGGTGGAG (T/A) TTTAA (T/C) CC (A/T) GA  <b>A C A T T C T T G G T G G T C T T T T A A T C C A T C A</b>            S W W S F N P E            S W W S F N P E            (F) (W)</p>	<p>Primer im M-Gen (PCR)            EVDV Genom            EVDV aa Sequenz            Consensus</p>
<p>30: CGGATCCAGTT (C/A) TACCATA (T/G) A (T/G) ATT (G/A) TA  <b>T A C A A C C A C A C T T A C C A T T A C A A A C A T A</b>            T A T G T T T G T A A T G G T A A C T C T C G T T G T A            Y V C N G N S R C            Y V C N G N P R C            (F) G</p>	<p>Primer im S-Gen (PCR)            invers EVDV            EVDV Genom            EVDV aa Sequenz            Consensus</p>
<p>33: GGAATTCAA (G/A) TGGCCTTGGT (A/G) TGT (G/I) TGG  <b>A T A C A T C A A C T G G C C T T G G T G T G T T G G</b>            K W P W W V W            K W P K Y V W</p>	<p>Primer im S-Gen (PCR)            EVDV Genom            EVDV aa Sequenz            Consensus</p>
<p>55: GGA (G/T) AAGGT (G/T) AATGA (A/G) TG (C/T) GT  <b>C A A A A G G T T A A T G A C T G C T G T</b>            K V N E C V            K V N E C V</p>	<p>Primer im S-Gen (PCR)            EVDV Genom            EVDV aa Sequenz            Consensus</p>
<p>56: CCA (G/T) AC (A/C/G) TACCA (A/T) GGCCA (T/C) TT  <b>C C A A A C C A C C A C C G G C C A C T T</b>            T C A A A G T G G C C G T G G T G G T T T G G T T T G            K W P W W V W L            Y V K W P K Y V W L</p>	<p>Primer im S-Gen (PCR)            invers EVDV            EVDV Genom            EVDV aa Sequenz            Consensus</p>

## Diskussion

Im vorliegenden Manuskript wird ein Beispiel beschrieben für die Möglichkeit, mit Hilfe von publizierten Sequenzen, davon abgeleiteten degenerierten Primern und der Polymerase-Kettenreaktion (PCR) das Genom eines unbekanntes Virus zu analysieren. Mit der Synthese und dem Einsatz von spezifischen Primern in der PCR vollzieht sich eine Selektion der im Experiment amplifizierbaren Sequenzen. Im Gegensatz zu konventionellen cDNS-Klonierungen, bei denen die Selektion auf der Ebene des Experimentes stattfindet, erstreckt sich bei unserer Methode die Selektion auch auf eine theoretische Ebene. Die experimentelle Durchführung wird dadurch einfacher. Die theoretische Vorarbeit ist aber aufwendiger. Eine Analyse von publizierten Coronavirussequenzen liess konservierte Regionen bereits auf der Ebene der genomischen RNS-Sequenzen erkennen. Der genetische Code lässt zu, dass einzelne Aminosäuren durch bis zu sechs verschiedene Basentriplets kodiert werden. Dies wird als Degeneriertheit des genetischen Codes bezeichnet. Aufgrund dieses Phänomens waren die Homologien auf der Ebene der Aminosäuresequenz besser erkennbar. Gleichzeitig bedeutete der Vergleich der Aminosäuresequenzen anstelle der Nukleinsäuresequenzen eine Datenreduktion, welche die Analyse erleichterte.

Um zu den Nukleinsäuresequenzen zu gelangen, die in der RT-PCR als Primer zu verwenden waren, erfolgte eine theoretische Rücktranslation der gefundenen Aminosäuresequenzen in Nukleinsäuresequenzen. Dabei musste wiederum die Degeneriertheit des genetischen Codes berücksichtigt werden. Die konservierten Aminosäuren der homologen Regionen sind in den verglichenen Coronaviren zum Teil durch verschiedene Basentriplets kodiert. Für die Auswahl der Primer innerhalb von homologen Regionen wurden Aminosäuren bevorzugt, die nach dem genetischen Code nur durch ein einziges Basentriplett codiert werden. Tryptophan (abgekürzt W) ist eine solche Aminosäure, die nur durch das Triplet TGG codiert wird. Serin (S) und Arginin (R) können hingegen durch je sechs verschiedene Triplets dargestellt werden. Bei der Evaluation der Primersequenzen sollten auch die Möglichkeiten anderer Basenpaarungen ausser der üblichen Watson-Crick-Paarung berücksichtigt werden. Unter gewissen Umständen kann eine unübliche Paarung zwischen A und C oder zwischen G und T erfolgen. In einer Doppelhelix ist eine falsche Paarung zwischen C und T eher möglich, als eine mit G und A, weil sich die beiden Purinringe von A und G gegenseitig stark behindern. Um allen verfügbaren RNS-Sequenzen Genüge zu leisten, wurden deshalb degenerierte Oligonukleotide als Primer verwendet. Mit diesem Ansatz und diesen Ein-

schränkungen wurden unsere Primer ausgewählt und in der RT-PCR verwendet.

Bei einem konventionellen Ansatz wird zunächst möglichst die gesamte zelluläre RNS isoliert, mit Hilfe von reverser Transkriptase in cDNS umgeschrieben und anschliessend kloniert. Eine sogenannte cDNS-Genbank enthält also die Gesamtmenge der RNS in Form von cDNS. Diese Genbank muss nach den gewünschten Klonen selektioniert werden. Die Selektion kann auf der Stufe der Nukleinsäure mittels Hybridisierung erfolgen. Falls die Klonierung der DNS in sogenannte Expressionsvektoren erfolgte, kann auch mit Antikörpern, die gegen das Genprodukt der zu klonierenden Nukleinsäure gerichtet sind, selektioniert werden.

In unserem Experiment lieferte der Vergleich von publizierten Sequenzen von ähnlichen Viren alle nötigen Sequenzinformationen zur Entwicklung von Primern für die RT-PCR. Für die Klonierung von Amplifikaten aus der PCR mit degenerierten Primern waren deshalb weder spezifische Antikörper noch Sonden für die Hybridisierung nötig.

Die Primer 30 und 33 zeigten in den RT-PCR-Experimenten nicht die gewünschten Resultate. Dies kann verschiedene Gründe haben. Die Sequenz von Primer 30 war in EVDV wohl konserviert, aber in der Nähe des 3' Endes des Primers, wo die Elongation der Polymerase stattfinden sollte, traten Unterschiede zum EVDV-Genom auf. Bereits die dritte Base vom 3' Ende her war degeneriert. Die vierte Base führte zu einer falschen Paarung. Das 3' Ende war zudem reich an Paarungen zwischen A und T. Adenosin und Tymidin sind bei der Paarung jedoch über nur zwei Wasserstoffbrücken verbunden. Im Gegensatz dazu bilden Cytosin und Guanosin bei der Paarung drei Wasserstoffbrücken aus. Deshalb blieb die Bindung zwischen dem 3' Ende des Primers und der Vorlage im Fall von Primer 30 relativ schwach. Es könnte sehr wohl sein, dass die Verwendung dieses Primers aus einem dieser Gründe nicht zu einer Amplifikation führte.

Die Primer 56 und 33 zeigten einen interessanten Gegensatz auf. Sie repräsentierten die gleiche Stelle auf dem viralen Genom, waren sich in der Sequenz jedoch komplementär und sollten deshalb die Initiation der DNS-Synthese in gegensätzliche Richtungen ermöglichen. Primer 56 war erfolgreich, während Primer 33 jedoch ein negatives Resultat lieferte. Ein Grund für das Fehlschlagen von Primer 33 kann im Inosin liegen, das in diesem Primer Verwendung fand. Es war der einzige Primer, der Inosin als Alternative für degenerierte Basen enthielt. Man weiss, dass gewisse Polymerasen dieses Basenanalog nicht als Substrat akzeptieren (Knittel und Picard, 1993). Andererseits spielt natürlich auch der jeweils zweite Primer eine wichtige Rolle für das Gelingen der PCR.

Die Amplifikation des Fragmentes im S-Gen mit den Primern 55 und 56 gelang erst, nachdem von *Taq*-Polymerase auf *Pfu*-Polymerase gewechselt worden war und nachdem ein neuer Thermocycler angeschafft werden konnte. Die Gründe für das Gelingen können auf verschiedenen Ebenen gesucht werden. Zum Beispiel ist *Pfu*-Poly-

merase thermostabiler als *Taq*-Polymerase und verfügt zudem über eine sogenannte «proof reading» Aktivität. Durch die höhere Stabilität verbraucht sich das Enzym langsamer. Beim neuen Gerät kamen kleinere und dünnere Reaktionsgefässe zum Einsatz, und der Wärmeaustausch erfolgte wahrscheinlich präziser.

Für die konservierten Sequenzen im Bereich von ORF1b waren auch degenerierte Primer synthetisiert und für die RT-PCR eingesetzt worden. Da jedoch nur gerade zwei Sequenzen für ORF1 von Coronaviren zur Verfügung gestanden hatten, wurden die publizierten Sequenzen der anderen, coronaähnlichen Viren miteinbezogen. Diese Ausweitung, aber auch die Tatsache, dass das 5' Ende der viralen RNS aufgrund der Replikation (nested set, vgl. Einleitung) in der infizierten Zelle weniger häufig vorkommt, kann für das negative Resultat mit den ORF1b-Primern verantwortlich gemacht werden. Nachdem die Amplifikation mit den Primern in den Strukturgenen (S, M und N) so erfolgreich verlief, wurde auf weitere Abklärungen bezüglich der ORF1b-Primer verzichtet.

Mit der hier für EVDV beschriebenen Methode sollte es auch möglich sein, die entsprechenden Genomabschnitte von anderen bisher unbekanntem Viren, speziell Coronaviren, zu amplifizieren. Ähnlich EVDV lassen sich verschiedene Coronaviren nur schwer in Zellkulturen vermehren. So gibt es Hinweise aus der Literatur, dass nicht züchtbare coronavirusähnliche Partikel in Patientenmaterial von Menschen und Tieren gefunden werden (Payne et al., 1987, und Referenzen darin). Aufgrund von serologischen Untersuchungen haben wir bereits postuliert, dass möglicherweise dem EVDV ähnliche Viren beim Menschen und bei der Katze existieren (Utiger et al., 1995). Eventuell lässt sich unsere Methode auch als breites diagnostisches Werkzeug für den Nachweis von Coronaviren allgemein anwenden. In einem ersten Schritt würde mit RT-PCR nachgewiesen, ob in einer bestimmten Probe überhaupt Coronaviren vorhanden seien. In weiteren Schritten müsste dann die Spezifizierung eines solchen Coronavirus erfolgen. Dies könnte mittels Restriktionsanalysen des Amplikons oder mit weiteren PCR-Amplifikationen mit spezifischen Primern geschehen. Eine einfache Diagnostik zur Unterscheidung der Coronaviren des Schweins (TGE und EVD) wäre sicherlich sinnvoll.

## Literatur

- Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M. (1993): Sequence determination of the nucleocapsid protein gene of the porcine epidemic diarrhoea virus confirms that this virus is a coronavirus related to human coronavirus 229E and porcine transmissible gastroenteritis virus. *J. Gen. Virol.* 74, 1795-1804.
- Bridgen A., Tobler K., Ackermann M. (1994): Identification of coronaviral conserved sequences and application to viral genome amplification. In: «Coronaviruses: molecular biology and virus-host interactions» (H. Laude and J. V. Vautherot). New York, Plenum.
- Duarte M., Laude H. (1994): Sequence of the spike protein of the porcine epidemic diarrhoea virus. *J. Gen. Virol.* 75, 1195-1200.

Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M., Laude H. (1994): Sequence analysis of the porcine epidemic diarrhoea virus genome between the nucleocapsid and the spike protein genes reveals a polymorphic ORF. *Virology* 198, 466-476.

Frobman M.A., Dush M.K., Martin G.R. (1988): Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci., U. S. A.* 85, 8998-9002.

Hofmann M., Wylter R. (1988): Propagation of virus of porcine epidemic diarrhoea in cell culture. *J. Clin. Microbiol.* 26, 2235-2239.

Kingston R.E. (1991): *Current Protocols in Molecular Biology*. New York, Greene Publishing Associates.

Knittel T., Picard D. (1993): PCR with degenerate primers containing deoxyinosine fails with Pfu DNA polymerase. *PCR Methods and Applications* 2, 346-347.

Pensaert M.B., Debouck P. (1978): A new coronavirus-like particle associated with diarrhoea in swine. *Arch. in Virology* 58, 243-247.

Tobler K., Bridgen A., Ackermann M. (1994): Sequence analysis of the nucleocapsid protein gene of porcine epidemic diarrhoea virus. *In:*

«Coronaviruses: molecular biology and virus-host interactions» (H. Laude and J. V. Vautherot). New York, Plenum.

Utiger A., Frei A., Carvajal A., Ackermann M. (1995): Studies on the *in vitro* and *in vivo* host range of porcine epidemic diarrhoea virus. *In:* «Corona- and Related Viruses» (P.J. Talbot and G. A. Levy). New York, Plenum.

Wirth U.V., Vogt B., Schwyzer M. (1991): The three major immediate-early transcripts of bovine herpesvirus 1 arise from two divergent and spliced transcription units. *J. Virol.* 65, 195-205.

## Dank

Die vorliegende Arbeit wurde vom Bundesamt für Veterinärwesen (Projekt Nr. 012.91.7) sowie vom Schweizerischen Nationalfonds (Projekt Nr. 31.37418.93) unterstützt. Die Autoren bedanken sich bei Dr. Felix R. Homberger für die Kommentare zum Manuskript.

*Korrespondenzadresse: Mathias Ackermann, Institut für Virologie der Vet.-med. Fakultät der Universität Zürich, Winterthurerstrasse 266a, CH-8057 Zürich*