



(19)

Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 043 399 B9

(12)

KORRIGIERTE EUROPÄISCHE PATENTSCHRIFT

Hinweis: Bibliographie entspricht dem neuesten Stand

- (15) Korrekturinformation:
Korrigierte Fassung Nr. 1 (W1 B1)
 Korrekturen, siehe Seite(n) 27-60
- (48) Corrigendum ausgegeben am:
08.10.2003 Patentblatt 2003/41
- (45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:
09.04.2003 Patentblatt 2003/15
- (21) Anmeldenummer: **00105929.4**
- (22) Anmeldetag: **23.03.2000**

(54) Hepatitis C Virus Zellkultursystem

Hepatitis C Virus cell culture system
 Système de culture du virus de l'hépatite C

- (84) Benannte Vertragsstaaten:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
 Benannte Erstreckungsstaaten:
AL LT LV MK RO SI
- (30) Priorität: **03.04.1999 DE 19915178**
- (43) Veröffentlichungstag der Anmeldung:
11.10.2000 Patentblatt 2000/41
- (73) Patentinhaber: **Bartenschlager, Ralf, Dr.**
55239 Gau-Odernheim (DE)
- (72) Erfinder: **Bartenschlager, Ralf, Dr.**
55239 Gau-Odernheim (DE)
- (74) Vertreter: **Rudolph, Ulrike, Dr.**
Patentanwältin
In der Schanz 10
69198 Schriesheim (DE)
- (56) Entgegenhaltungen:
WO-A-98/39031 **WO-A-99/04008**
WO-A-99/67394 **US-A- 5 851 758**
US-A- 5 874 565

- YOO ET AL: "Transfection of differentiated Huh7 with in vitro transcribed HCV RNA and establishment of a culture infected with HCV" JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, Bd. 69, Nr. 1, 1995, Seiten 32-38, XP002022696 ISSN: 0022-538X
- DASH SRIKANTA ET AL: "Transfection of HepG2 cells with infectious hepatitis C virus genome." AMERICAN JOURNAL OF PATHOLOGY, Bd. 151, Nr. 2, 1997, Seiten 363-373, XP001029512 ISSN: 0002-9440
- DATABASE EMBL [Online] 11. März 1998 (1998-03-11) retrieved from EBI Database accession no. D89815 XP002180379
- DATABASE EMBL [Online] 18. April 1997 (1997-04-18) retrieved from EBI Database accession no. D85516 XP002180641
- DATABASE EMBL [Online] 29. November 1995 (1995-11-29) retrieved from EBI Database accession no. D63922 XP002180380
- LOHMANN V ET AL: "REPLICATION OF SUBGENOMIC HEPATITIS C VIRUS RNAs IN A HEPATOMA CELLINE" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, Bd. 285, 2. Juli 1999 (1999-07-02), Seiten 110-113, XP000960693 ISSN: 0036-8075
- DATABASE EMBL [Online] 5. Juli 1999 (1999-07-05) retrieved from EBI Database accession no. AJ238799 XP002180381

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist. (Art. 99(1) Europäisches Patentübereinkommen).

- BLIGHT K.J. ET AL: 'Efficient initiation of HCV RNA Replication in cell culture' SCIENCE Bd. 290, 08 Dezember 2000, Seiten 1972 - 1974

Beschreibung

[0001] Die Erfindung betrifft ein Hepatitis-C-Virus (HCV)-RNA-Konstrukt mit der Fähigkeit zur Replikation in eukaryontischen Zellen sowie ein Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, d.h. die mit HCV-spezifischem Genmaterial transfiziert sind.

[0002] Das Hepatitis C Virus (HCV) ist eine der Hauptursachen chronischer und sporadischer Leberkrankungen weltweit. Die meisten HCV-Infektionen verlaufen ohne erkennbare klinische Symptome, allerdings werden 80-90% der Infizierten dauerhafte Virusträger und bei 50% dieser dauerhaften Virusträger kommt es zu einer chronischen Leberentzündung mit unterschiedlichen Schweregraden. Ca. 20% der chronisch Infizierten entwickeln im Laufe von 10 bis 20 Jahren eine Leberzirrhose, auf deren Basis sich ein primäres Leberzellkarzinom entwickeln kann. Die chronische Hepatitis C ist heute die Hauptindikation für eine Lebertransplantation. Eine Kausaltherapie gibt es bisher noch nicht. Die einzige derzeit verfügbare Therapie ist die hochdosierte Verabreichung von Interferon-Alpha oder eine Kombination aus Interferon-Alpha und dem Purin-Nukleosidanalogon Ribavirin. Allerdings sprechen nur ca. 60 % aller Behandelten auf diese Therapie an und bei diesen kommt es in mehr als der Hälfte aller Fälle nach dem Absetzen der Behandlung zu einer erneuten Virämie.

Aufgrund der hohen Prävalenz, gerade auch in den Industrieländern, den schwerwiegenden Folgen chronischer Infektionen und dem Fehlen einer Kausaltherapie ist die Entwicklung einer HCV-spezifischen Chemotherapie ein wesentliches Ziel der pharmazeutischen Forschung und Entwicklung. Hauptproblem hierbei ist bisher das Fehlen eines geeigneten Zellkultursystems, das ein Studium der Virus-Replikation und der Pathogenese in eukaryontischen Zellen ermöglicht.

[0003] Aufgrund der geringen Virusmengen im Blut bzw. Gewebe, dem Fehlen geeigneter Zellkultursysteme oder Tiermodelle (bis heute ist der Schimpanse das einzige mögliche Versuchstier) sowie dem Fehlen effizienter Systeme zur Produktion virus-ähnlicher Partikel, konnte die molekulare Zusammensetzung des HCV-Partikels bis heute noch nicht eingehend untersucht bzw. aufgeklärt werden. Die derzeit vorliegenden Ergebnisse lassen sich wie folgt zusammenfassen: Das HCV ist ein umhülltes Plusstrang RNA Virus mit einem Partikeldurchmesser von 50-60 nm und einer mittleren Dichte von 1,03-1,1g/ml. Es wurde erstmals 1989 molekular kloniert und charakterisiert (Choo et al., 1989; Science, 244, 359-362). Die HCV-RNA hat eine Länge von ca. 9.6 kb (= 9600 Nukleotide), eine positive Polarität und besitzt ein einziges offenes Leseraster (ORF = open reading frame), das ein lineares Polyprotein von ca 3010 Aminosäuren kodiert (siehe Rice 1996, in Virology, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp.931-960; Clarke 1997, J. Gen. Virol. 78, 2397; und Bartenschlager 1997, Intervirology 40, 378 und vgl. Fig. 1 A). Bei der Virusreplikation wird das Polyprotein durch zelluläre und virale Proteasen in die reifen und funktionell aktiven Proteine gespalten.

Innerhalb des Polyproteins sind die Proteine wie folgt angeordnet (vom Amino- zum Carboxyterminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. Das Core-Protein ist die Hauptkomponente des Nukleokapsids. Die Glykoproteine E1 und E2 sind Transmembranproteine und Hauptkomponenten der Virushülle. Sie spielen wahrscheinlich bei der Anheftung des Virus an die Wirtszelle eine wesentliche Rolle. Diese drei Proteine Core, E1 und E2 bauen den Viruspartikel auf und werden deshalb als Strukturproteine bezeichnet. Die Funktion des Proteins p7 ist noch unklar. Das Protein NS2 ist wahrscheinlich die katalytische Domäne der NS2-3 Protease, die für die Prozesierung zwischen den Proteinen NS2 und NS3 verantwortlich ist. Das Protein NS3 hat zwei Funktionen, nämlich in der aminoterminalen Domäne eine Proteaseaktivität, die für die Polyproteinprozessierung essentiell ist, und in der carboxyterminalen Domäne eine NTPase/Helikase-Funktion, die wahrscheinlich bei der Replikation der viralen RNA eine Rolle spielt. Das Protein NS4A ist ein Kofaktor der NS3-Protease. Die Funktion des Proteins NS4B ist unbekannt.

[0004] Das offene Leseraster ist an seinem 5' Ende von einer ca 340 Nukleotide langen nicht-translatierten Region (NTR = non-translated region) flankiert, die als interne Ribosomenansatzstelle (IRES = internal ribosome entry site) fungiert, und an seinem 3' Ende von einer ca. 230 Nukleotide langen NTR, die höchstwahrscheinlich für die Genomreplikation von Bedeutung ist. Eine solche 3'NTR ist Gegenstand der Patentanmeldung PCT/US 96/14033. Die Strukturproteine in dem amino-terminalen Viertel des Polyproteins werden von der Signalpeptidase der Wirtszelle gespalten. Die Nicht-Strukturproteine (NS) 2 bis (NS) 5B werden von zwei viralen Enzymen prozessiert, nämlich von der NS2-3 und der NS3/4A Proteinase. Die NS3/4A Proteinase wird für alle Spaltungen jenseits des Carboxyterminus von NS3 benötigt. Die Rolle von NS4B ist nicht bekannt. NS5A, ein hoch phosphoryliertes Protein, scheint für die Interferon Resistenz verschiedener HCV-Genotypen verantwortlich zu sein (vgl. Enomoto et al. 1995, J. Clin. Invest. 96, 224; Enomoto et al. 1996, N. Engl. J. Med. 334, 77; Gale Jr. et al. 1997, Virology 230, 217; Kaneko et al. 1994, Biochem. Biophys. Res. Commun. 205, 320; Reed et al., 1997, J. Virol. 71, 7187) und NS5B wurde als die RNA-abhängige RNA Polymerase identifiziert.

[0005] Anhand dieser Erkenntnisse wurden erste Diagnosesysteme entwickelt, die entweder auf dem Nachweis von HCV-spezifischen Antikörpern in Patientenserum oder auf dem Nachweis von HCV-spezifischer RNA mittels RT-PCR (= Reverse Transcription Polymerase Chain Reaction) beruhen, und die mittlerweile routineund/oder vorschriftsmäßig

bei allen Blutkonserven angewendet werden (müssen).

[0006] Seit der Erstbeschreibung des Genoms 1989 wurden mit Hilfe der PCR-Methode zahlreiche Teil- und Komplettsequenzen des HCV kloniert und charakterisiert. Ein Vergleich dieser Sequenzen zeigt eine hohe Variabilität des viralen Genoms, insbesondere im Bereich des NS5B-Gens, was letztendlich zu einer Einteilung in 6 Genotypen geführt hat, die selbst nochmals in Subtypen a, b, und c untergliedert sind. Die genomische Varianz ist nicht gleichmäßig über das Genom verteilt. So sind die 5'NTR und Teile der 3'NTR hoch konserviert, während bestimmte kodierende Sequenzen z.T. sehr stark variieren, vor allem die Hüllproteine E1 und E2.

[0007] Die klonierten und charakterisierten Teil- und Komplettsequenzen des HCV-Genoms wurden außerdem hinsichtlich geeigneter Angriffsziele für ein prospektives antivirales Therapeutikum untersucht. Dabei wurden drei virale Enzyme gefunden, die sich als solches Angriffsziel anbieten. Diese sind (1) der NS3/4A Proteasekomplex, (2) die NS3 Helikase und (3) die NS5B RNA-abhängige RNA Polymerase. Der NS3/4A Proteasekomplex und die NS3 Helikase konnten bereits kristallisiert und hinsichtlich ihrer dreidimensionalen Struktur aufgeklärt werden (Kim et al., 1996, *Cell*, 87, 343; Yem et al., 1998, *Protein Science*, 7, 837; Love et al., 1996, *Cell*, 87, 311; Kim et al., 1998, *Structure*, 6, 89; Yao et al., 1997, *Nature Structural Biology*, 4, 463; Cho et al., 1998, *J. Biol. Chem.*, 273, 15045); bei der NS5B RNA-abhängigen RNA Polymerase ist dies bis heute noch nicht gelungen.

Obwohl mit diesen Enzymen bedeutsame Angriffsziele für eine Therapieentwicklung der chronischen HCV-Infektion definiert sind, und obwohl sowohl mit Hilfe von 'rational drug design' als auch mit Hilfe von 'high throughput screens' weltweit intensiv nach geeigneten Inhibitoren gesucht wird, leidet die Therapieentwicklung an einem großen Defizit, nämlich dem Fehlen von Zellkultursystemen oder einfachen Tiermodellen, die es erlauben, HCV-RNA oder HCV-Antigene direkt, zuverlässig und mit einfachen laborüblichen Methoden nachzuweisen. Das Fehlen solcher Zellkultursysteme ist auch der Hauptgrund dafür, daß das Verständnis der HCV-Replikation bis heute noch sehr lückenhaft und in weiten Teilen nur hypothetisch ist.

[0008] Obwohl nach Meinung der Fachwelt eine enge evolutionäre Beziehung zwischen HCV und den Flavi- und Pestiviren besteht und für diese autonom replizierende RNAs beschrieben sind, die in verschiedenen Zelllinien ohne weiteres zur Replikation gebracht werden können und dabei relativ hohe Ausbeuten zeigen (siehe Khromykh et al., 1997, *J. Virol.* 71, 1497; Behrens et al., 1998, *J. Virol.* 72, 2364; Moser et al., 1998, *J. Virol.* 72, 5318), waren ähnliche Versuche mit HCV bisher nicht erfolgreich.

[0009] Zwar ist aus verschiedenen Publikationen bekannt, daß Zelllinien oder primäre Zellkulturen mit HCV-haltigem, hochtitrigem Patientenserum infiziert werden können, (Lanford et al. 1994, *Virology* 202, 606; Shimizu et al. 1993, *Proceedings of the National Academy of Sciences*, USA, 90, 6037-6041; Mizutani et al. 1996, *Journal of Virology*, 70, 7219-7223; M. Ikeda et al. 1998, *Virus Res.* 56, 157; Fournier et al. 1998, *J. Gen. Virol.* 79, 2376 und darin zitierte Literaturstellen, Ito et al. 1996, *Journal of General Virology*, 77, 1043-1054), diese virusinfizierten Zelllinien oder Zellkulturen erlauben jedoch nicht den direkten Nachweis von HCV-RNA oder HCV-Antigenen. Die virale RNA in diesen Zellen ist weder in einem Nothern-Blot (einem Standardverfahren zum quantitativen Nachweis von RNA) noch sind die viralen Protein in einem Western-Blot oder mittels Immunpräzipitation detektierbar. Nur mit sehr aufwendigen und indirekten Methoden ist es überhaupt gelungen, eine HCV-Replikation nachzuweisen. Diese nachteiligen Umstände zeigen klar, daß die Replikation in diesen bekannten virusinfizierten Zelllinien oder Zellkulturen absolut unzureichend ist.

[0010] Des Weiteren ist aus den Publikationen von Yoo et al. (1995, *Journal of Virology*, 69, 32-38) und von Dash et al., (1997, *American Journal of Pathology*, 151, 363-373) bekannt, daß Hepatomazelllinien mit synthetischer HCV-RNA, die mittels *in vitro* Transkription von kloniertem HCV-Genom gewonnen wurde, transfiziert werden können. In beiden Publikationen gingen die Autoren von dem Grundgedanken aus, daß das virale HCV-Genom eine Plusstrang-RNA ist, die nach dem Einschleusen in die Zelle direkt als mRNA fungiert, an die sich Ribosomen anheften und im Zuge von Translationsprozessen Virusproteine bilden, aus denen sich letztendlich neue HCV-Partikel bilden (können). Diese Virusreplikation, d.h. diese neu gebildeten HCV-Viren bzw. deren RNA wurde mittels RT-PCR nachgewiesen. Die publizierten Ergebnisse der durchgeföhrten RT-PCR sprechen jedoch dafür, daß die Effizienz der HCV-Replikation in den beschriebenen HCV-transfizierten Hepatomazellen nur sehr gering ist und jedenfalls nicht ausreicht, um Schwankungen in der Replikationsrate nach gezielter Einwirkung mit prospektiven antiviralen Therapeutika auch nur qualitativ, geschweige denn quantitativ zu messen. Außerdem ist im Stand der Technik mittlerweile bekannt (Yanagi et al., Proc. Natl. Acad. Sci. USA, 96, 2291-95, 1999), daß die hochkonservierte 3' NTR essentiell ist für die Virusreplikation, was in klarem Widerspruch zu den Behauptungen von Yoo et al. und Dash et al. steht, die für ihre Versuche in Unkenntnis des authentischen 3' Endes des HCV-Genoms ausschließlich HCV-Genome mit verkürzten 3' NTRs verwendet haben.

[0011] Rice et al. beschreiben in der WO 98 39031 die Herstellung eines infektiösen HCV Genoms und dessen Testung im Schimpansen. Die Autoren behaupten, daß solche infektiösen HCV Genome oder Teilen davon für verschiedene Zwecke, u.a. zum Aufbau von Zellkultursystemen für das HCV geeignet seien, es ist jedoch an keiner Stelle der D1 ein spezifisches, nacharbeitbares HCV-Konstrukt, das tatsächlich in Zellkultur repliziert, beschrieben.

[0012] In der WO 99/04008 ist ein kloniertes HCV Genom beschrieben, das im Schimpansen infektiös ist. Eine Infektiosität im Schimpansen erlaubt jedoch nicht den Schluß, daß das betreffende Genom auch in Zellkultur bzw. im In-vitro-Model repliziert.

- [0013] Die US 5,851,758 beschreibt von der humanen T Zelllinie H9 abgeleitete Zellklone, die mit HCV infizierbar sind, wobei nach der Infektion mit HCV zytopathische Effekte auftreten. Zur Infektion der Zellen wird Patientenserum mit nicht definiertem HCV verwendet. Im Unterschied zum Gegenstand der vorliegenden Patentanmeldung werden keine klonierten HCV-Genome verwendet und insbesondere keine klonierten HCV-Genome mit einem integrierten Selektionsmarker bzw. Reportergen, und die Zellkultur erfolgt nicht mit Hepatomzellen sondern mit humane T-Zellen.
- [0014] In der US 5,874,565 ist eine konservierte Sequenz am 3' Ende des HCV beschrieben, die für die virale Replikation essentiell ist. Es genügt jedoch nicht, die beschriebene 3'-Sequenz an ein beliebiges Genom anzuhängen und dadurch Replikation in Zellkultur zu erhalten, denn diese Sequenz ist zwar notwendig für die Replikation in Zellkultur, aber nicht hinreichend.
- [0015] Aufgabe der vorliegenden Erfindung ist die Bereitstellung eines Hepatitis-C-Virus (HCV)-RNA-Konstrukts mit der Fähigkeit zur Replikation in eukaryontischen Zellen sowie eines HCV- Zellkultursystems, bei dem die virale RNA in den transfizierten Zellen autonom und mit so hoher Effizienz repliziert, daß Schwankungen in der Replikationsrate nach gezielter Einwirkung mit virus- und insbesondere HCV-spezifischen antiviralen Therapeutika qualitativ und quantitativ und mit Hilfe gängiger, laborüblicher Meßverfahren gemessen werden können.
- [0016] Eine Lösung dieser Aufgabe besteht in der Bereitstellung eines Hepatitis-C-Virus (HCV)-RNA-Konstrukts der eingangs genannten Art, das sich dadurch auszeichnet, daß es eine Nukleotidsequenz gemäß einem der Sequenzprotokolle SEQ ID NO: 1 bis SEQ ID NO: 11 umfaßt, welche zumindest für die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) kodieren, wobei das selektierbare Markergen das Neomycinphosphotransferasegen ist und wobei anstelle des Neomycinphosphotransferasegens ein anderes Antibiotikaresistenzgen oder sonstiges Resistenzgen enthalten sein kann. Eine weitere Lösung dieser Aufgabe besteht in der Bereitstellung eines HCV-Zellkultursystems der eingangs genannten Art, bei dem die eukaryontischen Zellen humane Zellen, insbesondere Hepatomazellen sind, die vorzugsweise von einer handelsüblichen Hepatomazelllinie abstammen, aber auch aus einer entsprechenden Primärzellkultur gewonnen sein können, und bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das eine Nukleotidsequenz gemäß einem der Sequenzprotokolle SEQ ID NO: 1 bis SEQ ID NO: 11 umfaßt, welche zumindest für die HCV-spezifischen RNA-Abschnitte 5'NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) kodieren, wobei das selektierbare Markergen das Neomycinphosphotransferasegen ist und wobei anstelle des Neomycinphosphotransferasegens ein anderes Antibiotikaresistenzgen oder sonstiges Resistenzgen enthalten sein kann. "NTR" steht hier und im folgenden für "nicht-translatierte Region" und ist dem einschlägigen Fachmann als Begriff bzw. Abkürzung bekannt und geläufig. Der Begriff "HCV-RNA-Konstrukt" umfaßt hier und im folgenden sowohl Konstrukte, die das komplette HCV-Genom enthalten, als auch solche, die lediglich einen Teil davon, d.h. ein HCV-Subgenom enthalten.
- [0017] Der Begriff 5' NTR bzw. NS3 bzw. NS4A bzw. NS4B bzw. NS5A bzw. NS5B bzw. 3' NTR umfaßt im vorliegenden Zusammenhang jede Nukleotidsequenz, die im Stand der Technik als Nukleotidsequenz für den jeweils betreffenden funktionellen Abschnitt des HCV-Genoms beschrieben ist. Die Bereitstellung des erfindungsgemäßen HCV-RNA-Konstrukts ermöglicht erstmals eine detaillierte Analyse der HCV - Replikation, - Pathogenese und - Evolution in Zellkulturen. Die HCV-spezifische virale RNA kann - als vollständiges Genom oder als Subgenom - gezielt in beliebigen Mengen erzeugt werden, und es besteht die Möglichkeit, das RNA-Konstrukt zu manipulieren und damit die HCV-Funktionen auf genetischer Ebene zu untersuchen und aufzuklären. Da alle zur Zeit als Hauptangriffsziel für eine Therapie untersuchten HCV-Enzyme, nämlich die NS3/4A Protease, die NS3 Helikase und die NS5B Polymerase, in dem erfindungsgemäßen HCV-RNA-Konstrukt enthalten sind, kann es für alle entsprechenden Untersuchungen benutzt werden.
- [0018] Das in den erfindungsgemäßen HCV-RNA-Konstrukten enthaltene selektierbare Markergen (Selektionsgen), vorzugsweise ein Resistenzgen, insbesondere ein Antibiotikumresistenzgen, hat den Vorteil, daß die mit diesem Konstrukt transfizierten Zellen leicht von den nicht transfizierten Zellen selektiert werden können, indem dem Zellkulturmedium z.B. im Fall eines Antibiotikumresistenzgens das betreffende Antibiotikum zugegeben wird.
- [0019] Unter 'Antibiotikum' wird im vorliegenden Zusammenhang jede Substanz verstanden, die die nicht-transfizierten Wirtszellen oder die Zellen, in denen die HCV-RNA nur mit geringer Effizienz repliziert, am Leben oder Wachstum hindert, insbesondere Zellgifte wie z.B. Puromycin, Hygromycin, Zeocin, Bleomycin oder Blasticidin.
- [0020] Eine Alternative zu Antibiotikumresistenzgenen ist z.B. das Thymidin-Kinase-Gen, mit dem eine HAT-Selektion durchgeführt werden kann.

[0021] Bei einer alternativen Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts ist das selektierbare Markergen (Selektionsgen), insbesondere ein Antibiotikumresistenzgen, über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA bzw. der HCV-Genom- oder -Subgenomsequenz verbunden.

Damit geht der Vorteil einher, daß nach erfolgter Selektion derjenigen Zellen, in denen die HCV-RNA produktiv repliziert, in den daraus gewonnenen Zellklonen das Resistenzgen durch ribozymvennittelte Spaltung von der HCV-Subgenomsequenz abgetrennt werden kann, nämlich durch Aktivierung des einklonierten Ribozyms oder, im Fall eines Konstrukts mit einer Erkennungsstelle für ein Ribozym, durch Einschleusen des Ribozyms in die Zellen (z.B. mittels Transfektion eines Ribozymkonstrukt oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). Auf diese Weise wird ein authentisches HCV-Genom-Konstrukt ohne Resistenzgen erhalten, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist.

[0022] Eine weitere bevorzugte Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts zeichnet sich dadurch aus, daß das Konstrukt anstelle des Markergens oder zusätzlich zu dem Markergen ein integriertes Reportergen aufweist.

Unter Reportergen wird im folgenden jedes Gen verstanden, dessen Anwesenheit sich nach Überführung in einen Zielorganismus leicht und im allgemeinen mit einfachen biochemischen oder auch histochemischen Methoden nachweisen läßt, d.h. das für ein Protein kodiert, welches auch in geringen Mengen einfach und zuverlässig mit den laborüblichen Meßmethoden nachgewiesen und quantifiziert werden kann.

Diese Variante des HCV-RNA-Konstrukts hat den Vorteil, daß der Umfang der Replikation dieses Konstrukts anhand des Reportergenprodukts einfach und schnell mit laborüblichen Methoden gemessen werden kann.

Das Reportergen ist vorzugsweise ein Gen aus der Gruppe bestehend aus den Luziferasegenen, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), den GFP-Genen (green-fluorescence-protein-Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (Sezemerte-Alkalische-Phosphatase-Gen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. ELISA) bestimmt werden.

Als Reportergen kommt aber auch ein Surrogatmarkergen in Betracht. Darunter sind in diesem Zusammenhang solche Gene zu verstehen, die für zelluläre Proteine, Nukleinsäuren oder — allgemein — für solche Funktionen kodieren, die einer von der Virusreplikation abhängigen Variation unterliegen, und die infolgedessen in denjenigen Zellen, in denen sich das HCV bzw. das HCV-RNA-Konstrukt vermehrt, entweder reprimiert oder aktiviert werden. Das heißt: die Reduktion bzw. Aktivierung dieser Funktion ist ein Ersatzmarker für die Virusreplikation bzw. die Replikation des HCV-RNA-Konstrukts. Eine Variante des erfindungsgemäßen HCV-RNA-Konstrukts ist folglich dadurch gekennzeichnet, daß dessen Replikation die Expression eines (zellulären) Surrogatmarkergens beeinflußt.

[0023] Das Reportergen und das selektierbare Markergen können derart räumlich in dem Konstrukt angeordnet sein, daß sie gemeinsam ein Fusionsprotein exprimieren. Hierbei besteht die vorteilhafte Möglichkeit, daß diese beiden Gene so in dem HCV-RNA-Konstrukt angeordnet sind, daß ihre beiden exprimierten Proteine zunächst über eine Schnittstelle für eine Protease (z.B. Ubiquitin) oder über ein selbstspaltendes Peptid (z.B. das 2A-Protein der Picornaviren) fusioniert sind und erst später proteolytisch wieder getrennt werden.

[0024] Ebensogut können diese beiden Positionen aber auch derart getrennt voneinander liegen, daß beide Genprodukte separat exprimiert werden. (z.B. in der Reihenfolge: Marker- bzw. Resistenzgen — interne Ribosomenbindungsstelle — Reportergen).

Im Fall des Reportergens hat sich eine Ausführungsvariante besonders bewährt, bei der das Reportergen in das offene Leseraster des HCV-Genoms oder -Subgenoms einkloniert ist, und zwar derart, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführt wird.

[0025] Das erfindungsgemäße HCV-RNA-Konstrukt für sich genommen kann in allen seinen Variationen ebenfalls für vielfältige Zwecke eingesetzt werden. Dazu gehören vor allem:

- Die Konstruktion attenuierter Hepatitis C Viren bzw. HCV-ähnlicher Partikel und deren Produktion in Zellkulturen:

Durch zufällige oder gezielt hervorgerufene Mutationen, beispielsweise Punktmutationen, Deletionen oder Insertionen, können attenuierte HCV- oder HCV-ähnliche Partikel erzeugt werden, d.h. Viren bzw. virusähnliche Partikel mit voller Replikationskompetenz aber verringriger bzw. fehlender Pathogenität. Solche attenuierte HCV- oder HCV-ähnliche Partikel sind insbesondere als Impfstoff einsetzbar.

- Die Konstruktion von HCV-RNA-Konstrukten mit integrierten Fremdgenen, beispielsweise zur Verwendung als leberzellspezifische Genfären in der Gentherapie. Auf Grund des ausgeprägten Leberzelltropismus des HCV und der Möglichkeit, Teile des Genoms durch heterologe Sequenzen zu ersetzen, lassen sich HCV-RNA-Konstrukte herstellen, bei denen beispielsweise die Strukturproteine durch ein therapeutisch wirksames Gen ersetzt werden. Das so erhaltene HCV-RNA-Konstrukt wird in Zellen eingeschleust, vorzugsweise mittels Transfektion, die

die fehlenden HCV-Funktionen, beispielsweise die Strukturproteine, konstitutiv oder induzierbar exprimieren. Durch diese dem Fachmann unter dem Begriff der 'Transkomplementation' bekannte Technik lassen sich Viruspartikel erzeugen, in die das HCV-RNA-Konstrukt eingebaut wird. Die so erhaltenen Partikel können für die Infektion vorzugsweise von Leberzellen verwendet werden. In diesen wird das therapeutisch wirksame Fremdgen zur Expression gebracht und entfaltet damit seine therapeutische Wirkung.

- 5 • Das Auffinden permissiver Zellen, d.h. Zellen, in denen eine produktive Virusvermehrung erfolgt. Zu diesem Zweck wird entweder eines der vorgenannten HCV-RNA-Genomkonstrukte verwendet, das zur Bildung kompletter infektiöser Viren befähigt ist, oder es wird eines der vorgenannten HCV-Subgenom-Konstrukte eingesetzt, das allerdings zunächst gemäß vorgenanntem Beispiel in eine Zelllinie transfiziert wird, die die fehlenden Funktionen konstitutiv oder induzierbar exprimiert. In all diesen Fällen entstehen Viruspartikel, die zusätzlich zur HCV-Sequenz ein Resistenz- und/oder Reportergen tragen. Zum Auffinden von Zellen, in denen das HCV replizieren kann, werden diese Zellen mit den so hergestellten Viren infiziert und einer Antibiotikumselektion unterzogen oder, in Abhängigkeit vom HCV-RNA-Konstrukt, mittels Nachweis der Expression des Reportergens untersucht. Da eine Antibiotikumresistenz bzw. eine Expression des Reportergens nur dann nachweisbar ist, wenn das HCV-RNA-Konstrukt repliziert, müssen die so gefundenen Zellen permissiv sein. Auf diese Weise lassen sich nahezu beliebige Zelllinien oder primäre Zellkulturen hinsichtlich der Permissivität testen und auffinden.

10 [0026] Eine andere Variante des erfindungsgemäßen HCV-RNA-Konstrukts ist dadurch gekennzeichnet, daß das HCV-RNA-Konstrukt ein integriertes Fremdgen aufweist und dazu geeignet ist, dieses Fremdgen in eine Zielzelle einzuschleusen, die zur Expression dieses Fremdgens geeignet ist.

15 [0027] Eine weitere Variante des erfindungsgemäßen HCV-RNA-Konstrukts zeichnet sich dadurch aus, daß das HCV-RNA-Konstrukt Nukleotid- und/oder Aminosäure-Mutationen aufweist, zellkultur-adaptiert ist, mit hoher Effizienz repliziert und dadurch erhältlich ist, daß man ein erfindungsgemäßes Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein erfindungsgemäßes HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 1 bis 6 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen das HCV-RNA-Konstrukt oder Teile davon isoliert.

20 Eine Weiterbildung dieser Variante besteht darin, daß man das aus den Zellklonen isolierte HCV-RNA-Konstrukt oder Teile davon wenigstens einmal erneut passagiert, nämlich in Zellen des erfindungsgemäßen Zellkultursystems nach Anspruch 11 einschleust und diese Zellen auf/in dem dem Selektionsgen entsprechenden Selektionsmedium zu kultivieren, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen das HCV-RNA-Konstrukt oder Teile davon isoliert.

25 [0028] Ein ganz bevorzugtes HCV-RNA-Konstrukt mit hoher und sehr hoher Replikationseffizienz und infolgedessen sehr guter Eignung für die praktische Anwendung ist dadurch gekennzeichnet, daß es einen oder mehrere oder alle der in Tabelle 3 aufgelisteten Aminosäure- bzw. Nukleotidaustausche und/oder einen oder mehrere der nachfolgend aufgelisteten Aminosäureaustausche aufweist: 1283 arg -> gly , 1383 glu -> ala , 1577 lys -> arg, 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile -> val. (Die Zahlen beziehen sich auf die Aminosäurepositionen des Polyproteins des HCV-Isolats con1, siehe Tabelle 1).

30 [0029] Eine bevorzugte Variante des erfindungsgemäßen Zellkultursystems, die sich in der Praxis sehr gut bewährt hat, ist unter der Nummer DSM ACC2394 (Laborbezeichnung HuBl 9-13) bei der DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH in Braunschweig, Deutschland, hinterlegt.

35 [0030] Mit dem erfindungsgemäßen Zellkultursystem wird erstmals ein in-vitro-System bereit gestellt, in dem HCV-RNA intrazellulär, autonom und in ausreichend großen Mengen repliziert und exprimiert wird, so daß eine quantitative Bestimmung sowohl der HCV-RNA-Mengen als auch der HCV-spezifischen Proteine mit konventionellen und zuverlässig genauen biochemischen Meßmethoden durchgeführt werden kann. Das heißt: es steht erstmals ein annähernd authentisches zellgestütztes ("cell-based") HCV-Replikationssystem zur Verfügung, das für die Entwicklung und Erprobung von antiviralen Pharmazeutika dringend benötigt wird. Dieses Testsystem bietet nun die Möglichkeit, potentielle Angriffsziele für eine wirksame HCV-spezifische Therapie zu identifizieren und HCV-spezifische Chemotherapeutika zu entwickeln und zu evaluieren.

40 [0031] Die Erfindung basiert auf der überraschenden Erkenntnis, daß eine effiziente Replikation der HCV-RNA nur dann in Zellen stattfindet, wenn diese mit einem HCV-RNA-Konstrukt transfiziert wurden, das mindestens die 5' und die 3' nicht-translatierten Regionen (NTR) und die Nichtstrukturproteine (NS) 3 bis 5B umfaßt und zusätzlich ein selektierbares Markergen (Selektionsgen) aufweist. Offensichtlich sind die Strukturgene für den Ablauf der Replikation ohne wesentliche Bedeutung, während andererseits eine effiziente Replikation der HCV-RNA anscheinend nur dann stattfindet, wenn die transfizierten Zellen einem permanenten Selektionsdruck unterzogen werden, der durch das mit der HCV-RNA verbundene selektierbare Markergen (Selektionsgen) vermittelt wird. Das Markergen (Selektionsgen) scheint somit einerseits die Selektion derjenigen Zellen zu provozieren, in denen die HCV-RNA produktiv repliziert, und andererseits scheint es die Effizienz der RNA-Replikation wesentlich zu steigern.

[0032] Das erfindungsgemäße Zellkultursystem in allen seinen Variationen kann für vielfältige Zwecke eingesetzt werden. Diese umfassen:

- Das Auffinden antiviral wirksamer Substanzen. Dies können beispielsweise sein: organische Verbindungen, die unmittelbar oder mittelbar in die Virusvermehrung eingreifen (z.B. Inhibitoren der viralen Proteasen, der NS3-Helicase, der NS5B RNA-abhängigen RNA Polymerase), antisense Oligonukleotide, die an eine beliebige Zielsequenz innerhalb des HCV-RNA-Konstrukts (z.B. die 5' NTR) hybridisieren und unmittelbar oder mittelbar zu einer Beeinflussung der Virusvermehrung führen z.B. auf Grund einer Reduktion der Translation des HCV-Polyproteins oder Ribozyme, die eine beliebige HCV-RNA-Sequenz spalten und damit die Virusreplikation beeinträchtigen.
- Die Evaluierung jeglicher Art antiviral wirksamer Substanzen in Zellkultur. Solche Substanzen können beispielsweise mittels 'rational drug design' oder 'highthroughput screening' am isolierten gereinigten Enzym gefunden werden. Unter Evaluierung sind vor allem die Bestimmung der inhibitorischen Eigenschaften der entsprechenden Substanz sowie deren Wirkungsmechanismus zu verstehen.
- Die Identifikation neuer Angriffsziele, viralen oder zellulären Ursprungs, für eine HCV-spezifische antivirale Therapie. Ist beispielsweise ein zelluläres Protein essentiell für die Virusreplikation, kann mittels Hemmung dieses zellulären Proteins die Virusreplikation ebenfalls beeinflußt werden. Das Auffinden solcher auxiliären Faktoren ist mit dem erfindungsgemäßen System ebenfalls möglich.
- Der Einsatz für die Resistenzbestimmung. Es ist anzunehmen, daß auf Grund der hohen Mutationsrate des HCV-Genoms Therapieresistenzen auftreten. Solche Resistenzen, die gerade bei der klinischen Zulassung einer Substanz von großer Bedeutung sind, lassen sich mit dem erfindungsgemäßen Zellkultursystem ermitteln. Zelllinien in denen sich das HCV-RNA-Konstrukt bzw. das HCV-Genom oder - Subgenom repliziert, werden mit steigenden Konzentrationen der entsprechenden Substanz inkubiert und die Replikation der viralen RNA wird entweder anhand eines eingebrachten Reporters oder durch qualitative oder quantitative Bestimmung der viralen Nukleinsäuren oder Proteine bestimmt. Resistenz ist dann gegeben, wenn bei normaler Wirkstoffkonzentration keine Hemmung der Replikation zu beobachten ist. Durch Reklonierung der HCV-RNA (z.B. mittels RT-PCR) und Sequenzanalyse können die für Therapieresistenz verantwortlichen Nukleotid- bzw. Aminosäureaustausche ermittelt werden. Durch Einklonieren der/des entsprechenden Austausche/s in das Ursprungskonstrukt kann deren Kausalität für die Therapieresistenz bewiesen werden.
- Die Produktion von authentischen Virusproteinen (Antigene) für die Entwicklung und/oder Evaluierung von Diagnostika. Das erfindungsgemäße Zellkultursystem erlaubt auch die Expression von HCV-Antigenen in Zellkulturen. Diese Antigene können prinzipiell auch für den Aufbau diagnostischer Nachweisverfahren eingesetzt werden.
- Die Produktion von HCV Viren und virus-ähnlichen Partikeln insbesondere zur Entwicklung oder Herstellung von Therapeutika und Impfstoffen sowie für diagnostische Zwecke. Insbesondere zellkultur-adaptierte vollständige HCV-Genome, die mit dem erfindungsgemäßen Zellkultursystem hergestellt werden können, sind in der Lage, mit hoher Effizienz in Zellkulturen zu replizieren. Diese Genome besitzen alle Funktionen des HCV und sind deshalb in der Lage infektiöse Viren zu produzieren.

[0033] Die Erfindung betrifft deshalb auch die Verwendung eines erfindungsgemäßen Zellkultursystems und/oder eines erfindungsgemäßen HCV-RNA-Konstrukts zur Herstellung und/oder Evaluierung und/oder Testung von Therapeutika und/oder Diagnostika zur Behandlung von insbesondere HCV-Infektionen sowie zur Herstellung eines Impfstoffes gegen HCV-Infektionen.

[0034] Die Erfindung betrifft ferner die Verwendung eines erfindungsgemäßen HCV-RNA-Konstrukts zur Herstellung einer leberzellspezifischen Genfähre für die Gentherapie.

[0035] Das erfindungsgemäße Zellkultursystem erlaubt auch das gezielte Auffinden von HCV-RNA-Konstrukten, bei denen es auf Grund von Mutationen, die sich entweder zufällig im Rahmen der HCV-RNA-Replikation ereignen oder die gezielt in das Konstrukt eingeführt werden, zu einer Steigerung der Replikationseffizienz kommt. Solche Mutationen, die zu einer Veränderung der Replikation des HCV-RNA-Konstrukts führen, sind dem Fachmann als adaptive Mutationen bekannt.

Die Erfindung umfaßt deshalb auch die Verwendung eines erfindungsgemäßen HCV-RNA-Konstrukts nach einem der Ansprüche 1 bis 6 zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts gemäß vorstehender Beschreibung, wobei die Mutanten gegenüber dem originären HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen. Diese Verwendung ist dadurch gekennzeichnet, daß man ein Zellkultursystem gemäß Anspruch 11 oder 12, das HCV-RNA-Konstrukte gemäß einem der Ansprüche 1 bis 6 enthält, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen die HCV-RNA-Konstrukte oder Teile davon isoliert. Eine vorteilhafte Variante dieser Verwendung besteht darin, daß man die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines erfindungsgemäßen Zellkultursystems einschleust und diese auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone erntet und aus diesen Zellklonen die HCV-RNA-Konstrukte

oder Teile davon isoliert. Mit dieser Verfahrensvariante kann der Grad der adaptiven Mutationen und damit der Grad der Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten noch gesteigert werden.

Die erfindungsgemäß hergestellten zellkultur-adaptierten HCV-RNA-Konstrukte mit hoher Replikationseffizienz sind dadurch gekennzeichnet, daß sie durch Nukleotidund/oder Aminosäureaustausche von einem HCV-RNA-Konstrukt nach einem der Ansprüche 1 bis 6 ableitbar sind und daß sie mit einem der beiden vorstehend genannten Herstellungsverfahren erhältlich sind.

[0036] Die Erfindung umfaßt desweiteren die Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 1 bis 6 zur Herstellung von Mutanten eines HCV-RNA-Vollängengenoms oder eines HCV-RNA-Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-RNA-Vollängengenom oder - Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz. Diese Verwendung ist dadurch gekennzeichnet, daß man gemäß Anspruch 16 oder Anspruch 17 eine zellkultur-adaptierte Mutante des HCV-RNA-Konstrukts herstellt, daß man die Nukleotid- und Aminosäuresequenz dieser Mutante bestimmt und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Nukleotid- und Aminosäuremutationen bestimmt, und daß man diese Mutationen entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten, in ein (isoliertes) HCV-Vollängengenom oder ein HCV-Teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

[0037] Zum Nachweis bzw. zur Verifizierung derjenigen Mutationen, die tatsächlich eine Veränderung der Replikation und insbesondere eine Replikationssteigerung bewirken, kann ein Test durchgeführt werden, bei dem die bestimmten Nukleotid- und/oder Aminosäureaustausche in das ursprüngliche HCV-RNA-Konstrukt eingeführt und dieses wiederum in Zellkultur eingeschleust wird. Wenn die eingeführte Mutation tatsächlich zu einer Steigerung der Replikation führt, sollte im Fall eines HCV-RNA-Konstrukts mit selektierbarem Marker die Zahl der resistenten Zellklone bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten Konstrukt. Im Fall eines Konstrukts mit einem Reportergen sollte die Aktivität bzw. Menge des Reporters bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten.

[0038] Diese zellkultur-adaptierten HCV-RNA-Konstrukte können dazu verwendet werden, beliebige HCV-RNA-Konstrukte oder HCV-Vollängen- oder Teilgenome mit erhöhter Replikationseffizienz herzustellen. Dabei können sowohl Konstrukte mit einem selektierbaren Resistenzgen als auch Konstrukte ohne ein solches bzw. mit einem nichtselektierbaren Reportergen (z.B. Luciferase) hergestellt werden, denn aufgrund der sehr hohen Replikationseffizienz des zellkultur-adaptierten HCV-RNA-Konstrukts kann dessen Replikation auch in nicht-selektionierten Zellen nachgewiesen werden.

Die erfindungsgemäßen zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms oder eines HCV-Teilgenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, sind dadurch charakterisiert, daß sie mit einem Verfahren erhältlich sind, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 19, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

[0039] Nicht zuletzt betrifft die Erfindung auch die Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 1 bis 10 zur Erzeugung von Hepatitis C Viruspartikeln oder virus-ähnlichen Partikeln.

Besondere Eigenschaften der in den Sequenzprotokollen angegebenen Sequenzen:

[0040]

SEQ ID-NO: 1

Name: I389/Core-3'/wt

Aufbau (Nukleotidpositionen):

- 50 1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7

EP 1 043 399 B9 (W1B1)

9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
5 13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 2

10 Name: I337/NS2-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
15 3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1801-8403: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B

5. 1801-2451: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
20 6. 2452-4344: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4345-4506: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
8. 4507-5289: Nichtstrukturprotein 4B (NS4B)
9. 5290-6630: Nichtstrukturprotein 5A (NS5A)
10. 6631-8403: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
25 11. 8407-8637: HCV 3' nicht-translatierte Region

SEQ ID-NO: 3

Name: I389/NS3-3'/wt

Aufbau (Nukleotidpositionen):

- 30 1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
35 4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
40 7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
10. 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 4

45 Name: I337/NS3-3'/wt

Aufbau (Nukleotidpositionen):

- 50 1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1801-7758: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B

5. 1801-3696: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
55 6. 3697-3858: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3859-4641: Nichtstrukturprotein 4B (NS4B)
8. 4642-5982: Nichtstrukturprotein 5A (NS5A)
9. 5983-7755: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase

10. 7759-7989: HCV 3' nicht-translatierte Region

SEQ ID-NO: 5

Name: I389/NS2-3'/wt

5 Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-8418: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
5. 1813-2463: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
6. 2464-4356: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4357-4518: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
8. 4519-5301: Nichtstrukturprotein 4B (NS4B)
9. 5302-6642: Nichtstrukturprotein 5A (NS5A)
10. 6643-8415: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8419-8649: HCV 3' nicht-translatierte Region

SEQ ID-NO: 6

Name: I389/NS3-3'/9-13F

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F
5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCVNS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 7

Name: I389/Core-3'/9-13F

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase

15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 8

Name: I389/NS3-3'/5.1

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 9

Name: I389/Core-3'/5.1

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 10

Name: I389/NS3-3'/19

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3'

nicht-translatierte Region

SEQ ID-NO: 11

Name: I389/Core-3'/19

5 Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19

5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

25 [0041] Die Erfindung wird im folgenden anhand von Ausführungsbeispielen und dazugehörigen Tabellen und Figuren näher erläutert. Die erwähnten Figuren zeigen

Fig. 1 A: Die Struktur eines erfindungsgemäßen HCV-RNA-Konstrukts

Ganz oben ist eine schematische Darstellung der Struktur des kompletten parentalen HCV-Genoms gegeben mit den Positionen der Gene für die Spaltungsprodukte core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A und NS5B innerhalb des Polyproteins, und den 5' und 3' nichttranslatierten Regionen (5' NTR und 3' NTR) — als Horizontalbalken dargestellt —, und mit den beiden für die Erzeugung der Subgenom-Konstrukte ausgewählten Positionen, nämlich der Position der 'GDD-katalytischen Domäne' der NS5B RNA Polymerase (GDD) und der Position der 3' Grenze der HCV-IRES (Nukleotidpositionen 1 bis 377 bzw. 1 bis 389) — oberhalb des Genomschemas eingezeichnet —. Die Zahlen unterhalb des Genomschemas bezeichnen die entsprechenden Nukleotidpositionen.

Darunter sind schematische Darstellungen der Strukturen zweier erfindungsgemäßer, modifizierter HCV-RNA-Konstrukte (Subgenom) gezeigt, bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (Neo^R), der EMCV-IRES (E-I) und den HCV Sequenzen von NS2 bzw. NS3 bis zum authentischen 3' Ende. Die Position der das NS5B Polymerase GDD-Motiv umfassenden 10 Aminosäuren-Deletion ist jeweils mit einem Dreieck (Δ) markiert.

Fig. 1 B: Das Ergebnis einer denaturierenden Formaldehyd-Agarose-Gelelektrophorese zum Nachweis von replizierter Plusstrang-RNA in transfizierten subpassagierten Huh-7 Zellklonen. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker (M) sind links von Spur 1 angegeben.

Fig. 1 C: Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum Nachweis der Abwesenheit von integrierter Replikon-DNA in den meisten der selektierten Zellklone.

Spuren 1 und 2 zeigen die Positivkontrollen, Spur 13 die Negativ-Kontrolle. Die Zahlangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle.

Fig. 2 A: Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum sensitiven Ausschluß integrierter Replikon-DNA (Plasmid Moleküle I₃₇₇/NS3-3'/wt) in einem HCV-RNA-Konstrukt-haltigen Zellklon (9-13). Die Spuren 7 bis 11 repräsentieren das Ergebnis einer Titration von DNA-Molekülen des Konstrukts I₃₇₇/NS3-3'/wt ohne Zugabe von Gesamt-DNA des Zellklons 9-13 und die Spuren 2 - 6 repräsentieren die gleichen Plasmidmoleküle mit Zugabe von jeweils 1 μ g 9-13 DNA vor der PCR (zwecks Ausschluß eines Inhibitors der PCR in der DNA-Präparation). Spur 13 repräsentiert die Negativ-Kontrolle (PCR ohne

DNA-Sonde). Spur 1 zeigt das Ergebnis, das mit einem μ g Gesamt-DNA des Zellkons 9-13 erhalten wurde.

Fig. 2 B: Das Ergebnis eines Northern-Blot-Tests zur Quantifizierung von HCV Plus-und Minusstrang RNA. Die Pfeile markieren die Positionen von Replikon-RNA. Die "plus" und "minus" -Abgaben bezeichnen die positive (plus) bzw. negative (minus) Polarität der RNA-Kontrollen, die auf das Gel aufgetragen wurden. "Minusstrand" und "Plusstrand" bezeichnen die Spezifität der radioaktiven RNA-Sonden.

Fig. 2 C: Ergebnis einer Formaldehyd-Agarose-Gelelektrophorese nach radioaktiver Markierung der intrazellulär replizierten HCV-RNA zum Nachweis der Resistenz der HCV-RNA-Replikation gegen Dactinomycin.

Fig. 3 A: Nachweis von HCV-spezifischen Antigenen in den selektierten Zellklonen mittels Immunopräzipitation nach metabolischer Radioaktivmarkierung. Die Spuren 7 - 9 repräsentieren authentische Größenmarker (die nach transiente Expression eines HCV-RNA-Konstrukts in Huh-7-Zellen erhalten wurden); identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben.

Fig. 3 B: Ergebnisse eines Immunfluoreszenztests zum Nachweis der subzellulären Lokalisation von HCV Antigenen.

Fig. 4 : Schematische Darstellung der Struktur eines erfindungsgemäßen selektierbaren HCV-RNA-Konstrukts (komplettes Genom) bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (NeoR), einem heterologen IRES-Element, z.B. des Encephalomyocarditisvirus (E-I), dem vollständigen HCV-Leseraster und der authentischen 3' NTR.

Fig. 5: Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Antibiotikumresistenzgen (**A**) innerhalb der für das Polyprotein kodierenden Nukleotidsequenz (monocistronische RNA), und (**B**) innerhalb der 3' NTR (bicistronische RNA).

Fig. 6: Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Reportergen (**A**) als Teil eines HCV-Replikons von NS3 bis NS5B; — das Reporterprotein wird letztendlich durch virale oder durch zelluläre Proteasen aus dem Polyprotein gespalten und das selektierbare Markerogen (Selektionsgen) bzw. das Resistenzgen durch Kontransfektion in die Zellen einschleust , (**B**) als Teil eines Fusionsgens aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und green fluorescent Protein) (**C**) als Teil eines Replikons aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und das green fluorescent Protein), die über eine Nukleotidsequenz verbunden sind, welche für eine Aminosäuresequenz kodiert (schräffierter Bereich), die von einer Protease gespalten werden kann oder die über eine selbstspaltende (autokatalytische) Aktivität verfügt, (**D**) als unabhängiges Gen (hier green fluorescent protein), das von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert wird; — das Resistenzgen (hier: Neomycinphosphotransferase-Gen) wird davon unabhängig ebenfalls von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert (polycistronisches Konstrukt).

Fig. 7: Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Die dicken Linien stellen die HCV 5' und 3' NTRs dar, E-I ist eine heterologe interne Ribosomenbindungsstelle, die für die Expression des Resistenzgens notwendig ist, und das graue Quadrat stellt das Ribozym bzw. eine Erkennungsstelle für ein Ribozym dar.

Fig. 8: Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts mit Resistenzgenen und integriertem Fremdgene.

Fig. 9: Methodisches Vorgehen zum Vergleich der spezifischen Infektiosität (ausgedrückt als Anzahl gebildeter Zellkolonien) von Gesamt-RNA versus in vitro Transkripte. HCV-RNA wird mittels in vitro Transkription eines entsprechenden RNA-Konstrukts hergestellt und durch Messung der optischen Dichte bei 260 nm (OD 260 nm) quantifiziert. Eine definierte Anzahl dieser Moleküle wird mit einer bestimmten Menge Gesamt-RNA von naiven Huh-7 Zellen gemischt und diese Mischung mit Hilfe der Elektroporation in naive Huh-7 Zellen eingeschleust. Parallel dazu wird die Gesamt-RNA eines Zellklons, der mit der in Figur 1

beschriebenen Methode hergestellt wurde, mit einem im Stand der Technik bekannten Verfahren isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northernblot unter Verwendung einer HCV-spezifischen RNA-Sonde und anschließender Quantifizierung mittels Phosphoimager bestimmt. Eine definierte Menge dieser Gesamt-RNA wird analog den in vitro Transkripten in naïve Huh-7 Zellen transfiziert. Diese Zellen in beiden Ansätzen werden danach einer G418-Selektion unterzogen und die Anzahl der gebildeten Kolonien durch Auszählen nach fixieren und anfärben mit Coomassie-Brilliant-Blau bestimmt. Zur Bestimmung der Transfektionseffizienz wird jedem Transfektionsansatz 1 μ g eines Plasmids zugesetzt, das die Expression der Luziferase erlaubt. Ein Aliquot der transfizierten Zellen wird nach 24 Stunden geerntet und die Luziferaseaktivität im jeweiligen Zellysat bestimmt.

Die Anzahl der Kolonien wird jeweils auf die Luziferaseexpression normiert.

Fig. 10: Sequenzanalyse der 9-13 Klone. Gesamt-RNA des Zellklons 9-13, der durch Transfektion des HCV-RNA-Konstrukts I377/NS3-3' entstand, wurde mit einem im Stand der Technik bekannten Verfahren isoliert und das HCV-RNA-Konstrukt von Nukleotidposition 59 bis 9386 mit Hilfe der 'long-distance RT-PCR' unter Verwendung der primer S59 und A9413 amplifiziert. Die PCR-Fragmente wurden kloniert und 11 Klone (genannt 9-13 A - K) vollständig sequenziert, wobei sich die Klone D und I, E und G sowie H und J als identisch erwiesen. Die Positionen der Aminosäureunterschiede in der NS3-5B Region zwischen den reklierten HCV-RNAs und dem parentalen Konstrukt sind mit einem dicken vertikalen Strich beim jeweiligen Klon markiert. Jeder Klon wurde mit dem Restriktionsenzym *Sfi* I verdaut und das jeweilige Fragment in das parentale Konstrukt inseriert. Diese Klone wurden jeweils in Huh-7 Zellen transfiziert und die Zellen wie in Figur 1 beschrieben einer Selektion unterzogen. Die Anzahl der mit jedem Konstrukt erhaltenen Zellklone ist rechts neben dem jeweiligen Konstrukt vermerkt.

Fig. 11 A: Prinzip der Replikationsbestimmung mit Hilfe eines Reportergens. Im oberen Teil der Figur ist das HCV-DNA-Konstrukt I₃₈₉/Luc/NS3-3' dargestellt, bestehend aus der HCV 5' NTR (Nukleotidposition 1-389), dem Luziferasegen (*luc*), der IRES des Encephalomyocarditis Virus, dem HCV NS3-5B und der 3' NTR. Die Position des aktiven Zentrums der NS5B RNA-Polymerase, in das ein inaktivierender Aminosäureaustausch eingeführt wurde, ist mit 'GND' angedeutet. Die Plasmide, die für das replikationskompetente bzw. das defekte HCV-RNA-Konstrukt kodieren, werden mit dem Restriktionsenzym *Sca* I verdaut und in eine in vitro Transkription mit der T7 RNA-Polymerase eingesetzt. Nach Entfernung der Matrizen-DNA werden die jeweiligen HCV-RNA-Konstrukte mittels Elektroporation in naive Huh-7 Zellen eingeschleust und diese in regelmäßigen Abständen geerntet.

Fig. 11 B: Vergleich der Luziferaseaktivitäten in Zellen transfiziert mit dem parentalen HCV-RNA-Konstrukt I₃₈₉/Luc/NS3-3'/wt (wt) oder den folgenden Varianten: Der inaktiven RNA (318 DN), der Variante 9-I3F oder der Variante 5.1. Die Zellen wurden 6 (nicht gezeigt), 24, 48, 72, 96, 120, 144 und 168 Stunden nach der Transfektion geerntet und die Luziferaseaktivitäten luminometrisch bestimmt.

Fig. 12: Selektionierbare HCV-Vollängengenome (Konstrukte I₃₈₉/core-3'/5.1 und I₃₈₉/core-3'/9-13F).

(A) Schematische Darstellung des Vollängenkonstrukturts. Der Bereich zwischen den beiden angedeuteten Erkennungsstellen für das Restriktionsenzym *Sfi* I entspricht den Sequenzen der hoch-adaptierten RNA-Varianten 5.1. oder 9-13F.

(B) Anzahl der Kolonien die nach Transfektion von jeweils 0,1 μ g in vitro transkribierter RNA der unter A dargestellten Konstrukte I₃₈₉/core-3'/5.1 in HUH7-Zellen erhalten wurden. Angegeben ist das Ergebnis eines repräsentativen Experiments.

(C) Nachweis autonom replizierender HCV-Vollängen-RNAs in G418-resistenten Zellklonen, die nach Transfektion des entsprechenden in vitro Transkripts erhalten wurden. Die Abbildung zeigt das Autoradiogramm eines Northern Blots, der mit einer Sonde gegen das neo-Resistenzgen und der HCV 5' NTR hybridisiert wurde. Die in Spur 1 und 2 dargestellten Kontrollen entsprechen jeweils 10⁸ Molekülen der angegebenen in vitro Transkripte, gemischt mit Gesamt-RNA aus naiven Huh-7 Zellen. Die Negativkontrolle enthält ausschließlich Gesamt RNA aus naiven Huh-7 Zellen (Spur 3). Die Spuren 4-9 enthalten 3-10 μ g Gesamt-RNA aus G418-resistenten Zellklonen, die nach Transfektion von in vitro transkribierter I₃₈₉/core-3'/5.1-RNA bzw. I₃₈₉/core-3'/9-13F-RNA erhalten wurden. Die für die Selektion verwendete G418-Konzentration ist jeweils angegeben. Fünf der dargestellten Zellklone enthalten die hoch adaptierte RNA-Variante 5.1 (Spur 4-8), einer die adaptierte RNA-Variante 9-13F (Spur 9).

Fig. 13: HCV-RNA-Konstrukte mit einem Reportergen. (A) Bicistronische HCV-RNA-Konstrukte. Das Reporter-

gen wird mit Hilfe einer separaten IRES translatiert. (B) Monocistronische HCV-RNA-Konstrukte. Das Reportergenprodukt wird als Fusionsprotein mit einem HCV-Protein exprimiert. Die beiden Anteile sind über eine Erkennungssequenz für eine virale oder zelluläre Protease verbunden, die eine proteolytische Trennung der beiden fusionierten Proteinanteile erlaubt. Im gezeigten Beispiel wurden das Reportergenprodukt und das jeweilige HCV-Protein über eine Erkennungssequenz für Ubiquitin (Ub) fusioniert.

Fig. 14: Tricistronisches Vollängen HCV-RNA-Konstrukt, das zusätzlich zum Resistenzgen ein Fremdgen inseriert besitzt.

Fig. 15: Monocistronische HCV-RNA-Konstrukte, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird. Das Resistenzgen (RG) ist entweder als Fusionsprotein aktiv oder es wird so mit einer proteolytisch spaltbaren Sequenz mit dem HCV-Anteil fusioniert, daß das Resistenzgenprodukt durch eine zelluläre oder virale Protease vom HCV-Anteil abgespalten wird. Im gezeigten Beispiel wurde das Resistenzgen über die für Ubiquitin (Ub) kodierende Sequenz mit dem jeweiligen HCV-Anteil fusioniert.

Beispiel 1: Herstellung von HCV-RNA-Konstrukten

(A) Synthese und Klonierung eines vollständigen HCV-Konsensusgenoms mittels RT-PCR

[0042] Aus der Leber eines chronisch infizierten Patienten wurde das HCV-Genom, d.h die HCV-RNA wie nachfolgend beschrieben isoliert:
[0043] Aus ca. 100 mg Leber wurde die komplette RNA gemäß dem Verfahren von Chomczynski und Sacchi (1987, Anal. Biochem. 162, 156) isoliert. Mit 1 µg dieser isolierten RNA wurde eine reverse Transkription mit den Primern A6103 (GCTATCAGCCGGTTCATCCACTGC) oder A9413 (CAGGATGGCCTATTGG CCTGGAG) und dem 'expand reverse transcriptase'-System (Boehringer Mannheim, Deutschland) nach den Vorschriften des Herstellers durchgeführt. Mit den Produkten dieser reversen Transkription (RT) wurde eine Polymerase-Kettenreaktion (PCR=polymerase chain reaction) durchgeführt, und zwar unter Verwendung des 'expand long template'-Systems (Boehringer Mannheim, Deutschland), wobei der Puffer mit 2% Dimethylsulfoxid-Gehalt eingesetzt wurde. Nach einer Stunde bei 42°C wurde 1/8 dieses Reaktionsansatzes in einem ersten PCR-Durchgang mit den Primern A6103 und S59 (TGTCTTCACGCA-GAAAGCGTCTAG) oder A9413 und S4542 (GATGAGCT CGCCGCGAAGCTGTCC) eingesetzt. Nach 40 Zyklen wurde 1/10 dieses Reaktionsansatzes in einem zweiten PCR-Durchgang mit den Primern S59 und A4919 (AGCACAGC-CCGCGTCATAGCACTCG) oder S4542 und A9386 (TTAGCTCCCCG TTCATCGGTTGG) eingesetzt. Nach 30 Zyklen wurden die PCR-Produkte mittels präparativer Agarose-Gel-Elektrophorese gereinigt und die dabei eluierten Fragmente wurden in den Vektor pCR2.1 (Invitrogen) oder pBSK II (Stratagene) ligiert. Vier Klone von jedem Fragment wurden analysiert und sequenziert, und es wurde eine Konsensus-Sequenz ermittelt. Zu diesem Zweck wurden die DNA-Sequenzen miteinander verglichen. Die Positionen, an denen sich die Sequenz eines der Fragmente von den übrigen unterschied, wurde als unerwünschte Mutation betrachtet. Im Fall von Mehrdeutigkeiten der Sequenz wurden kürzere sich überlappende PCR-Fragmente der betreffenden Region amplifiziert und mehrere Klone sequenziert. Auf diese Weise konnten zahlreiche potentielle Mutationen in jedem Fragment identifiziert und somit eine isolat-spezifische Konsensussequenz etabliert werden. Diese etablierte Konsensussequenz bzw. dieses Genom gehört zum weltweit verbreiteten Genotyp 1b. Die nicht translatierte Region am 3'-Ende (=3' NTR) wurde mittels konventioneller PCR erhalten, wobei ein Antisense-Primer eingesetzt wurde, der die letzten 24 Nukleotide des im Stand der Technik bekannten 'X-tails' (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033) abdeckt. Die authentische nicht translatierte Region am 5'-Ende (=5' NTR) strangabwärts vom T7 Promotor wurde mittels PCR erzeugt, wobei zum einen ein Oligonukleotid verwendet wurde, das einem verkürzten T7 Promotor (TAA TAC GAC TCA CTA TAG) und den ersten 88 Nukleotiden von HCV entspricht, und zum anderen eines der vorgenannten Plasmide eingesetzt wurde, das eines der 5' Fragmente des Genoms trägt. Aus den subgenomischen Fragmenten mit der geringsten Anzahl an Nicht-Konsensus-Austauschen wurde ein komplettes HCV-Konsensusgenom zusammengesetzt und in einen modifizierten pBR322-Vektor insertiert. Abweichungen von der Konsensussequenz wurden mittels ortsgerichteter Mutagenese ("site-directed mutagenesis") beseitigt. Um "run-off"-Transkripte mit einem authentischen 3' Ende herzustellen, wurde die 3'-NTR der Isolate (mit dem Ende TGT) zu AGT modifiziert (gemäß der Sequenz vom Genotyp 3 = Klon 'WS' nach Kolyhalov et al., 1996, J. Virol. 70, 3363) und außerdem wurde ein zusätzlicher Nukleotidaustausch an Position 9562 vorgenommen, um die A:T Basenpaarung in der Haarnadelstruktur am 3'Ende der 3' NTR (Kolyhalov et al. ibid.) beizubehalten. Um eine interne Restriktionsstelle für das Enzym Scal zu beseitigen, wurde ferner ein sog. stiller ("silent") Nukleotidaustausch vorgenommen. Nach dem Zusammenfügen des Vollängen-Genoms mit passenden 5'- und 3' NTRen wurde die komplette HCV-Sequenz überprüft. Dabei wurde kein ungewünschter Nukleotidaustausch gefunden.

[0044] Das auf diese Weise hergestellte HCV-Genom sollte per Definition hepatotrop sein.

(B) Synthese selektierbarer HCV-Subgenom-Konstrukte

[0045] Unter Verwendung des unter (A) beschriebenen Konsensusgenoms wurden HCV-Subgenom-Konstrukte hergestellt, die das Antibiotikumresistenzgen Neomycin-Phosphotransferase (NPT) und zwei Sequenzen von internen Ribosomenbindungsstellen (IRES) enthalten. Die hierfür angewendeten biochemischen Verfahrenstechniken sind dem Fachmann bekannt und geläufig (siehe: Sambrook, J., E.F. Fritsch, T. Maniatis, 1989, Molecularcloning: a laboratory manual, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, Current Protocols in Molecular Biology, Vol. 1-3, John Wiley & Sons Inc., New York). Das Antibiotikumresistenzgen wurde unmittelbar hinter der 5' NTR insertiert, wodurch eine bicistronische RNA erhalten wurde (siehe Fig. 1 A). Ebensogut kann das Antibiotikumresistenzgen aber auch an anderer Stelle des HCV-Subgenom-Konstruktinsertiert werden, beispielsweise innerhalb der für das Polyprotein kodierenden Nukleotidsequenz, wodurch eine monocistronische RNA erhalten wird (siehe Fig. 5 A) oder in die 3' NTR (siehe Fig. 5 B). Bei den IRES-Elementen handelt es sich zum einen um eine der beiden HCV-IRES-Varianten Nukleotide 1-377 oder Nukleotide 1-389, und zum anderen um die IRES des Encephalomyocarditis Virus, die die Translation der HCV Sequenz strangabwärts von den Genen für NS2 oder NS3 bis zu dem authentischen 3' Ende des Genoms steuert.

[0046] Die beiden genannten HCV-IRES-Varianten wurden wie folgt ermittelt:

Auf der Basis von Deletionsanalysen der 3' Grenze der HCV-IRES (Reynolds et al. 1995, *EMBO J.* 14, 6010) wurden verschiedene Abschnitte der 5' NTR mit dem NPT Gen fusioniert und anhand von Kotransfektionen mit einem das T7 RNA Polymerase Gen enthaltenden Plasmid hinsichtlich der maximalen Anzahl gebildeter Kolonien analysiert. Die besten Ergebnisse wurden mit den HCV Sequenzen von 1-377 und 1-389 erhalten. Da sich das AUG-Startkodon des HCV Polyproteins an Position 342 befindet und somit in der IRES-Sequenz enthalten ist, kommt es zu einer Fusion von 12 bzw. 16 Aminosäuren des HCV-Kapsidproteins ("Core-Proteins") mit der Neomycin Phosphotransferase (siehe Fig. 1 A).

[0047] Diese modifizierten HCV-Subgenom-Konstrukte erhielten dementsprechend die Bezeichnungen I₃₇₇/NS2-3' (oder I₃₇₇/NS3-3') und I₃₈₉/NS2-3' (oder I₃₈₉/NS3-3'). Sie sind in Fig. 1A schematisch dargestellt.

[0048] Mit in-vitro-Transkripten dieser modifizierten parentalen HCV-Subgenom-Konstrukte I₃₇₇/NS2-3' (oder I₃₇₇/NS3-3') und I₃₈₉/NS2-3' (oder I₃₈₉/NS3-3') wurden verschiedene Zelllinien und Primärzellkulturen von menschlichen Hepatocyten transfiziert.

[0049] Als parallele Negativ-Kontrolle zu allen Transfektionsexperimenten wurde zu jedem modifizierten parentalen HCV-Subgenom-Konstrukt ein entsprechend modifiziertes aber defektes Subgenom konstruiert, das sich von dem parentalen dadurch unterscheidet, daß es innerhalb des Leserasters eine Deletion von 10 Aminosäuren aufweist, die das aktive Zentrum der NS5B RNA Polymerase umfaßt (Behrens et al., 1996, *EMBO J.* 15, 12; und Lohmann et al., 1997, *J. Virol.* 71, 8416).

(C) Synthese selektierbarer HCV-Genom-Konstrukte

[0050] Ein NS2-3' Subgenomkonstrukt, das am 5' Ende mit einem Fragment des Luziferasegens und der vollständigen EMCV-IRES verbunden ist, wurde mit Ncol und Spel restriktiv und mittels präparativer Agarosegelektrophorese gereinigt. Der so erhaltene Vektor wurde in einer 3-Faktor Ligation mit einem Ncol/NotI-HCV-Fragment, entsprechend den Nukleotidpositionen 342 bis 1968 des HCV-Genoms und mit einem NotI/Spel-Fragment, entsprechend den Nukleotidpositionen 1968-9605 ligiert. Das entstandene Konstrukt, bei dem das vollständige HCV-Leseraster und die 3' NTR stromabwärts dem Luziferasegenfragment und der EMCV-IRES liegen, wurde danach mit Pmel und Spel restriktiv und mit dem analog restriktierten I₃₈₉/NS3-3'/wt-Subgenomkonstrukt-Vektor ligiert. Dieses selektionierbare HCV-Genomkonstrukt ist in Fig. 4 dargestellt.

(D) Herstellung von den HCV-RNA-Konstrukten entsprechenden in-vitro-Trnaskripten

[0051] Die vorstehend beschriebenen gereinigten Plasmid DNAs wurden mit Scal linearisiert und nach Phenol/Chloroform-Extraktion und Isopropanol-Präzipitation in eine In-vitro-Trnkriptionsreaktion eingesetzt unter Verwendung der folgenden Komponenten: 80 mM HEPES, pH 7.5, 12,5 mM MgCl₂, 2 mM Spermidin, 40 mM Dithiothreitol, 2 mM von jedem NTP, 1 Einheit RNasin/µl, 50 µg/ml restriktierte DNA und ca. 2 Einheiten/µl T7 RNA Polymerase. Nach 2 Std. bei 37°C wurde die Hälfte der Menge an T7 Polymerase zugegeben und der Reaktionsansatz weitere 2h inkubiert. Zur Entfernung von DNA wurde die Mischung mit saurem Phenol extrahiert (U. Kedzierski, J.C. Porte, 1991, Bio Techniques 10, 210), mit Isopropanol präzipitiert, das Pellet in Wasser gelöst und mit DNase (2 Einheiten pro µg DNA) für 60 Min. bei 37°C inkubiert. Nach anschließender Extraktion mit saurem Phenol, saurem Phenol/Chloroform und Chloroform und Isopropanol-Präzipitation wurde die gelöste RNA mittel optischer Dichtemessungen quantifiziert und ihre

Unversehrtheit mittels Formaldehyd-Agarose-Gelelektrophorese überprüft.

Beispiel 2: Transfektionsexperimente mit der Hepatomazelllinie Huh-7

5 [0052] Bei sämtlichen Transfektionsexperimenten wurde sorgfältig darauf geachtet, daß jegliche Matrizen-DNA zuvor entfernt worden war, um zu vermeiden, daß solche DNA in transfizierte Zellen integrieren und diesen unabhängig von einer HCV-Replikation eine Neomycin-Resistenz vermittelten konnte. Deshalb wurde im Anschluß an die in-vitro-Transkription (Beispiel 1 D) die Reaktionsmischung mit 2 Einheiten DNase pro µg DNA für 60 Min. bei 37°C behandelt und mit saurem Phenol, saurem Phenol/Chloroform und Chloroform extrahiert. Vor der Verwendung für die Transfektion wurde die präzipitierte RNA mittels Formaldehyd Agarose Gel Elektrophorese analysiert.

10 [0053] Es wurden drei separate Transfektionsexperimente mit der hoch differenzierten humanen Hepatomazelllinie Huh-7 (gemäß Nakabayashi et al. 1982, Cancer Res. 42, 3858) durchgeführt. Dabei wurde jeweils 15 µg RNA in 8 x 10⁶ Huh-7-Zellen mit Hilfe der Elektroporation eingebracht und diese Zellen anschließend in Kulturschalen von 10 cm Durchmesser ausgesät. 24 Stunden nach der Aussaat wurde Neomycin (= G418) in einer Endkonzentration von 1 mg/ml zugegeben. Das Kulturmedium wurde zweimal pro Woche gewechselt. Nach 3 - 5 Wochen waren kleine Kolonien erkennbar, die isoliert und unter den gleichen Kulturbedingungen passagiert wurden.

15 [0054] Die Zellklone, die im Verlauf des ersten Experiments erhalten wurden, wurden isoliert und subpassagiert. Während dieser Prozedur starben die meisten Klone und die Endausbeute betrug nur noch 9 Klone von Zellen, die mit den parentalen HCV-Subgenom-Konstrukten transfiziert worden waren und 1 Klon (Klon 8-1) von Zellen, die mit einem defekten HCV-Genom-Konstrukt, nämlich einer defekten NS2-3' HCV-RNA transfiziert worden waren. Außer einer verkürzten Verdopplungszeit und dem gelegentlichen Auftreten von irregulär geformten Zellen wurden keine beständigen morphologischen Unterschiede zwischen diesen 9 Zellklonen und dem einen Zellklon (Klon 8-1) oder den parentalen Huh-7 Zellen gefunden.

20 [0055] Die Hauptkriterien für funktionierende HCV-Genomkonstrukte sind die Bildung von viraler RNA mit korrekter Größe und die Abwesenheit von (integrierter) Plasmid DNA, die eine G418-Resistenz übertragen bzw. vermitteln könnte.

25 [0056] Um die HCV-RNA in den Huh-7-Zellen zu bestimmen, wurde die Gesamt-RNA isoliert und mittels des gängigen Northern-Blot Verfahrens unter Verwendung einer Plusstrang-spezifischen Ribosonde (= RNA-Sonde) analysiert. Hierfür wurde von den jeweiligen Zellklonen Gesamt-RNA nach der Methode von Chomczynski und Sacchi 1987, Anal. Biochem. 162, 156 isoliert, und 10 µg RNA, was dem Gesamt-RNA-Gehalt von 0,5 - 1 x 10⁶ Zellen entspricht, mittels denaturierender Formaldehyd-Agarose-Gelelektrophorese aufgetrennt (Spuren 3 bis 12 der Fig. 1 B). Als Größenmarker mit authentischer Sequenz wurden gleichzeitig 10⁹ in-vitro-Transkripte (ivtr.), die zu den I₃₈₉/NS2-3'/wt oder den I₃₈₉/NS3-3'/wt Replikon-RNAs korrespondieren, mit aufgetrennt (Spur 1 bzw. Spur 2). Die aufgetrennte RNA wurde auf Nylon-Membranen transferiert und mit radioaktiv markierter Plusstrang-spezifischer RNA-Sonde, die komplementär zu dem kompletten NPT-Gen und der HCV-IRES von Nukleotid 377 bis Nukleotid 1 war, hybridisiert. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker sind links von Spur 1 angegeben. Die RNA Marker-Fragmente enthalten HCV-Sequenzen und hybridisieren deshalb mit der Ribosonde (= RNA-Sonde). Die Ergebnisse dieser Analyse sind in Fig. 1 B dargestellt.

30 [0057] Mit Ausnahme des mit dem defekten HCV-Genom-Konstrukt transfizierten Klons 8-1, lieferten alle Zellklone homogene HCV-RNAs korrekter Länge (ca. 8640 Nukleotide im Fall des NS2-3' und ca. 7970 Nukleotide im Fall des NS3-3' Replikons). Dieser Befund ist ein Indiz dafür, daß die funktionalen Replikons bzw. die funktionalen HCV-Genom-Konstrukte die G418 Resistenz übertragen. Um auszuschließen, daß die G418 Resistenz auf eine Plasmid-DNA zurückzuführen ist, die in das Genom der Huh-7 Wirtszelle integriert ist und unter der Kontrolle eines zellulären Promotors transkribiert wird, wurde von jedem Klon die DNA mittels einer NPT-Gen-spezifischen PCR untersucht. Hierbei wurde aus den selektierten Huh-7-Zellklonen die DNA mittels Verdau mit Proteinase K (40µg/ml, 1h, 37°C) in 10mMTris, pH7,5, 1mM EDTA, 0,5% SDS und anschließender Extraktion mit Phenol, Phenol/Chloroform und Isopropanolpräzipitation isoliert. Das DNA-Präzipitat wurde in 10 mM Tris (pH 7,5) und 1 mM EDTA gelöst und 1 Stunde mit Rnase A inkubiert. Im Anschluß an eine Phenol/Chloroform Extraktion und Ethanol Präzipitation wurde 1 µg DNA, entsprechend 4 - 8 x 10⁴ Zellen, mittels PCR unter Einsatz NPT-Gen-spezifischer Primer (5'-TCAAGACCGACCTG TCCGGTGCCC-3' und 5'-CTTGAGCCTGGCGAACAGTCGGC-3') analysiert und ein DNA-Fragment bestehend aus 379 Nukleotiden erzeugt. Die Spezifität des PCR-Produkts wurde mittels Southern Blot Verfahren nachgewiesen, wobei ein Digoxigenin-markiertes DNA Fragment eingesetzt wurde, das zu dem NPT-Gen korrespondiert. Als Positiv-Kontrollen (zum Nachweis etwa vorhandener kontaminierender Nukleinsäuren) wurde das PCR-Verfahren mit 10⁷ Plasmid Molekülen oder 1 µg DNA aus einer BHK Zelllinie, die stabil mit einem Neomycin-Resistenz-Gen transfiziert war, durchgeführt, und als Negativ-Kontrolle wurde die PCR mit denselben Reagenzien aber ohne zugesetzte DNA durchgeführt.

35 Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrollen, Spur 13 repräsentiert die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der

Nukleotid-Marker-Moleküle. Außer in Klon 7-3 (Fig. 1C, Spur 3), der von Zellen nach Transfektion mit einem NS2-3' Replikon/NS2-3'HCV-Genom-Konstrukt stammt, und in Klon 8-1 (Fig. 1C, Spur 12), der von Zellen nach Transfektion mit einem defekten HCV-Genom-Konstrukt stammt, war in keinem Zellklon eine NPT-DNA nachweisbar. Dieser Befund ist ein weiteres Indiz dafür, daß die G418 Resistenz der meisten Klone durch die replizierende HCV-RNA vermittelt wurde. Aber auch unabhängig von diesen Ergebnissen ist es unwahrscheinlich, daß HCV-RNAs mit korrekter Größe von integrierter Plasmid DNA erzeugt wird, denn die für die in-vitro-Transkription verwendeten Plasmide enthalten weder einen eukaryontischen Promotor noch ein Polyadenylierungssignal. Im Fall des Klons 7-3 ist die Resistenz deshalb höchst wahrscheinlich sowohl durch das HCV-RNA-Konstrukt bzw. die replizierende HCV-RNA als auch durch eine integrierte NPT DNA Sequenz vermittelt worden, während die Resistenz der Zellen von Klon 8-1 ausschließlich auf die integrierte Plasmid DNA zurückzuführen ist.

[0058] Um zu bestätigen, daß die G418 Resistenz von einer autonom replizierenden HCV-RNA vermittelt ist, wurde der Klon 9-13 (Fig. 1 B, Spur 11) weiteren Tests unterworfen. Klon 8-1, der integrierte Kopien des NPT-Gens trägt, wurde überall als Negativkontrolle eingesetzt. Mit dem Ziel, die Anwesenheit von NPT-DNA im Klon 9-13 rigoros auszuschließen, wurde eine PCR durchgeführt, die den Nachweis von < 1000 NPT-Gen-Kopien in ~ 40.000 Zellen erlaubt.

Das Ergebnis dieser PCR ist in Fig. 2A dargestellt. Im einzelnen wurde bei dieser PCR wie folgt verfahren: Es wurden jeweils 10^6 - 10^2 Plasmid Moleküle ($I_{377}/NS3-3'/wt$) entweder direkt (Spuren 7 - 11) oder nach Zugabe von jeweils 1 µg 9-13 DNA (Spuren 2 - 6) in dem Test eingesetzt. Die Spezifität der amplifizierten DNA Fragment wurde mittels Southern Blot unter Verwendung einer NPT-spezifischen Sonde bestimmt. Eine PCR ohne DNA-Sonde wurde als Negativ-Kontrolle durchgeführt (Spur 12).

Selbst mit dieser sensitiven Methode wurde in einem µg DNA des Zellklons 9-13 keine Plasmid DNA gefunden (Spur 1). Um die Menge an HCV Plus- und Minusstrang RNAs in diesen Zellen abzuschätzen, wurde eine Verdünnungsreihe von Gesamt-RNA mit dem Northern-Blot-Verfahren unter Verwendung einer Plus- oder Minusstrang-spezifischen radioaktiv markierten Ribosonde (= RNA-Sonde) analysiert. Hierfür wurden jeweils 8, 4 oder 2 µg Gesamt-RNA, die aus den Zellklonen 9-13 und 8-1 isoliert worden waren, parallel zu bekannten Mengen analoger in-vitro-Transkripte mit Plus- oder Minusstrang-Polarität (Kontroll-RNAs) im Northern-Blot-Verfahren analysiert und anschließend einer Hybridisierung unterworfen. Die Hybridisierung wurde mit einer Plusstrang-spezifischen Ribosonde, die das komplette NPT-Gen und die HCV-IRES abdeckte ('plusstrand', obere Bildtafel), oder mit einer Minusstrang-spezifischen RNA-Sonde, die zu der NS3-Sequenz komplementär war ('minusstrand', untere Bildtafel) durchgeführt. Die Pfeile markieren die Positionen von Replikon-RNA. Die Ergebnisse dieser Analyse sind in Fig. 2 B dargestellt.

Im Fall des Plusstrangs wurden ca. 10^8 Kopien/µg Gesamt-RNA nachgewiesen, was 1000 - 5000 HCV-RNA-Molekülen pro Zelle entspricht, während die Menge an Minusstrang-RNA 5- bis 10-fach niedriger war. Dieses Ergebnis stimmt mit der Annahme überein, daß die Minusstrang RNA die replikative Zwischenform bzw. Zwischenkopie ist, die als Vorlage für die Synthese der Plusstrang Moleküle dient.

Da die Reaktion im wesentlichen von der viralen RNA-abhängigen RNA Polymerase katalysiert wird, sollte die Synthese der HCV-RNAs resistent gegen Dactinomycin sein, einem Antibiotikum, das selektiv die RNA-Synthese von DNA-Matrizen inhibiert, nicht jedoch die RNA-Synthese von RNA-Matrizen. Um diese Vermutung zu bestätigen, wurden Zellen mit [³H] Uridin in Anwesenheit von Dactinomycin inkubiert, die radioaktiv markierten RNAs extrahiert, mittels denaturierender Agarose-Gel-Elektrophorese aufgetrennt und mit Hilfe eines handelsüblichen Bio-Imagers unter Verwendung einer [³H]-sensitiven Bildplatte analysiert. Hierfür wurden jeweils ca. 5×10^5 Zellen der Klone 9-13 und 8-1 mit 100 µCi [³H]Uridin für 16 Std. in Abwesenheit (-) oder Gegenwart (+) von 4 µg/ml Dactinomycin (Dact) inkubiert. Im Anschluß an diese Markierungsreaktion wurde die Gesamt-RNA präpariert und mittels Formaldehyd-Agarose-Gel-Elektrophorese analysiert. In den beiden ersten Spuren ist nur 1/10 der Gesamt-RNA dargestellt. Die radioaktiv markierte RNA wurden mit einem BAS-2500 Bio-Imager (Firma Fuji) sichtbar gemacht.

Die Ergebnisse dieser Analyse sind in Fig. 2 C dargestellt. In Übereinstimmung mit dem Inhibitor-Profil der NS5B Polymerase (Behrens et al., 1996, EMBOJ. 15, 12 und Lohmann et al., 1997, J. Virol. 71, 8416) war die Replikation der HCV RNA nicht durch Dactinomycin beeinflußt worden, während die Synthese von zellulärer RNA gehemmt worden war. Um die Identität der viralen RNA zu bestätigen, wurde eine RT-PCR zur Rekonstruktion der replizierten Sequenzen durchgeführt. Die Sequenzanalyse der rekonstruierten RNA zeigte, daß die RNA in dem Klon 9-13 HCV-spezifisch ist und mit dem transzribierten Transkript des HCV-Konstruktks $I_{377}/NS3-3'/wt$ übereinstimmt.

[0059] Zur Analyse der viralen Proteine wurden die betreffenden Zellen zunächst metabolisch mit [³⁵S] Methionin/Cystein radioaktiv markiert, anschließend lysiert und danach die HCV-spezifischen Proteine mittels Immunopräzipitation aus den Zell-Lysaten isoliert. Die Ergebnisse dieser Analysen sind in Fig. 3 A dargestellt. Im einzelnen wurde dabei wie folgt verfahren: Zellen der Zellklone 9-13 (wt) und 8-1 (Δ) waren durch Behandlung für 16 Stunden mit einer dem Fachmann geläufigen und im Handel erhältlichen Protein-Markierungs-Mischung (z.B. NEN Life Science) metabolisch radioaktiv markiert worden. Mittels Immunopräzipitation (IP) unter nichtdenaturierenden Bedingungen (z.B. nach Bartenschlager et al., 1995, J. Virol. 69, 7519) und unter Verwendung von drei verschiedenen Antisera (3/4, 5A, 5B, gemäß Markierung am oberen Ende der Spuren 1 bis 12) waren die HCV-spezifischen Proteine vom Zell-Lysat abgetrennt worden.. Die Immunokomplexe wurden mittels Tricine SDS-PAGE analysiert und mittels Autoradiographie

sichtbar gemacht. Um authentische Größenmarker zu erhalten, wurde das homologe Replikonkonstrukt I₃₇₇/NS3-3'/wt einer transienten Expression mit dem Vaccinia Virus T7-Hybrid System in Huh-7 Zellen unterworfen. Die dabei erhaltenen Produkte waren als Größenmarker (Spuren 7 - 9) parallel zu den Zellen der Klone 9-13 und 8-1 behandelt worden. Identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben. Es ist anzumerken, daß das verwendete NS3/4-spezifische Antiserum ('3/4') bevorzugt mit NS4A und NS4B reagiert, was zu einer Unterrepräsentation von NS3 führt.

[0060] Alle viralen Antigene waren eindeutig nachweisbar und ihre apparenten Molekulargewichte zeigten keine Abweichungen gegenüber denjenigen, die nach transienter Expression desselben bicistronischen HCV-RNA-Konstruktur in den ursprünglichen Huh-7 Zellen ermittelt wurden. Um die subzelluläre Verteilung der viralen Antigene zu bestimmen, wurde eine Immunfluoreszenz-Nachweisreaktion unter Einsatz von NS3- und NS5A-spezifischen Antiseren durchgeführt (z.B. nach Bartenschlager et al., 1995, *J. Virol.* 69, 7519). Hierfür wurden Zellen der Klone 9-13 (wt) und 8-1 (Δ) 24 Std. nach dem Aussäen auf Deckgläsern mit Methanol/Azeton fixiert und mit polyklonalen NS3- oder NS5A-spezifischen Antiseren inkubiert. Die gebundenen Antikörper wurden mit einem kommerziell erhältlichen FITC-konjuguierten Anti-Kaninchchen-Antiserum sichtbar gemacht. Zur Unterdrückung unspezifischer Fluoreszenzsignale wurden die Zellen mit dem Farbstoff 'Evans Blue' gegenfärbt.

[0061] Die Ergebnisse dieses Nachweistests sind in Fig. 3 B dargestellt. Mit beiden Antiseren war eine starke Fluoreszenz im Zytoplasma nachweisbar. Die NS5A-spezifischen Antiseren führten außerdem zu einer schwachen Zellkern-Fluoreszenz, was darauf hindeutet, daß zumindest kleine Mengen dieses Antigens auch zum Zellkern gelangen. Die generell dominierende Präsenz der viralen Antigene im Zytoplasma ist jedoch ein starkes Indiz dafür, daß die HCV-RNA Replikation im Zytoplasma stattfindet — so wie das bei den meisten RNA-Viren der Fall ist.

[0062] Diese Ergebnisse belegen klar, daß mit dem hier beschriebenen Versuchsansatz der Aufbau eines Zellkultursystems für das HCV gelungen ist, dessen Effizienz alles bisher bekannte um Größenordnungen übersteigt und erstmalig den Nachweis viraler Nukleinsäuren und Proteine mit konventionellen und bewährten biochemischen Methoden erlaubt. Erst diese Effizienz erlaubt überhaupt detaillierte Untersuchungen der HCV-Pathogenese, genetische Analysen verschiedener HCV-Funktionen und ein genaues Studium der Virus-/Wirtszellwechselwirkungen, wodurch sich neue Ansatzpunkte für die Entwicklung einer antiviralen Therapie definieren lassen.

Beispiel 3: Transfektion von Huh-7 Zellen mit HCV-Genomkonstrukten

[0063] Huh-7 Zellen werden wie in Beispiel 2 beschrieben transfiziert und selektiert, wobei hier jedoch selektionsfähige Konstrukte verwendet werden, die das vollständige Virusgenom enthalten. Die erhaltenen Zellklone werden analog dem Beispiel 2 mittels PCR auf Abwesenheit von HCV-DNA untersucht und die produktive Replikation der HCV-RNA wird danach mittels Northern Blot, [³H]Uridinmarkierung in Anwesenheit von Dactinomycin, Nachweis der viralen Proteine bzw. Antigene vorzugsweise mit Hilfe des Western Blots, der Immunopräzipitation oder der Immunfluoreszenz nachgewiesen. Im Gegensatz zu den im Beispiel 2 beschriebenen Ansätzen lassen sich mit dem hier beschriebenen Konstrukt außerdem vollständige und sehr wahrscheinlich infektiöse Viren erhalten, was bei den dort (in Beispiel 2) beschriebenen Subgenomkonstrukten nicht der Fall ist. Diese Viren, die in der Zelle und dem Zellkulturerstand vorhanden sind, werden beispielsweise mittels Ultrazentrifugation, Immunpräzipitation oder Fällung mit Polyethylenglykol konzentriert und alle exogenen, d.h. nicht im Viruspartikel eingebauten Nukleinsäuren werden mittels Inkubation mit Nukleasen (RNase, DNase, Mikrococcusnuklease) verdaut. Auf diese Weise lassen sich alle kontaminierenden Nukleinsäuren, die nicht im schützenden Viruspartikel enthalten sind, entfernen. Die geschützte virale RNA wird nach Inaktivierung der Nukleasen, beispielsweise mittels Inkubation mit Proteinase K in einem SDS-haltigen Puffer durch Extraktion mit Phenol und Phenol/Chloroform isoliert und mittels Northern Blot oder RT-PCR unter Verwendung HCV-spezifischer Primer nachgewiesen. Auch in diesem Versuchsansatz ist die Kombination des beschriebenen HCV-Konsensusgenoms mit einem Selektionsmarker entscheidend für die effiziente Produktion von viraler RNA, viralen Protein und damit von HCV-Partikeln.

Beispiel 4: Herstellung und Anwendung eines HCV-RNA Konstrukt, bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-Subgenom-Sequenz verbunden ist.

[0064] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 oder Beispiel 3 hergestellt, bei dem ein Antibiotikumresistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Solche Konstrukte sind in Fig. 7 schematisch dargestellt. Huh-7 Zellen werden wie in Beispiel 2 beschrieben mit diesem HCV-RNA-Konstrukt transfiziert. Nach der Transfektion in die Zellen erfolgt zunächst die Selektion mit dem entsprechenden Antibiotikum. In den dabei erhaltenen Zellklonen wird das einklonierte Ribozym aktiviert oder, im Fall eines Konstrukt, das eine Erkennungsstelle für ein Ribozym trägt, wird das Ribozym in die Zelle eingeschleust (z.B. mittels Transfektion eines Ribozymkonstrukt oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). In beiden Fällen wird durch die ribozymvermittelte Spaltung das Resistenzgen von der

HCV-RNA-Sequenz abgetrennt. Das Ergebnis ist im Fall des HCV-Genom-Konstrukts ein authentisches HCV-Genom ohne Resistenzgen, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist. Im Fall des HCV-Subgenom-Konstrukts entsteht ein HCV-Replikon ohne Resistenzgen.

5 Beispiel 5: Kotransfektion eines HCV-RNA-Konstruktur mit einem separaten Luziferase-Transfektionskonstrukt

[0065] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 (A) oder Beispiel 3 oder Beispiel 4 hergestellt. Parallel dazu wird ein Transfektionskonstrukt hergestellt, welches das Luziferasegen umfaßt, wobei dieses Luziferasegen vermittels einer ersten Nukleotidsequenz, die für eine HCV-Protease- (z.B. NS3-Protease-) Spaltungsstelle kodiert, mit einer zweiten Nukleotidsequenz, die für ein anderes Protein oder einen Teil eines anderen Proteins kodiert, verbunden ist. HCV-RNA-Konstrukt und Transfektionskonstrukt werden in beliebige Wirtszellen, vorzugsweise Hepatomazellen, insbesondere Huh-7-Zellen, eingeschleust. Dies kann auf die in Beispiel 2 beschriebene Art und Weise geschehen. Das Produkt des modifizierten Luziferasegens ist ein Luziferase-Fusionsprotein, in dem die Luziferase auf Grund der Fusion mit dem Fremdanteil inaktiv ist. In transfizierten Zellen mit hoher HCV-Replikation wird das Fusionsprotein, das ja eine Schnittstelle für eine HCV-Protease enthält, gespalten und damit die aktive Form der Luziferase freigesetzt, die sich durch luminometrische Messung bestimmen läßt. Wird die Replikation des HCV-RNA-Konstrukts gehemmt, wird das Fusionsprotein nicht gespalten und keine aktive Luziferase freigesetzt. Infolgedessen ist die quantitative Bestimmung der Luziferase ein Maß für die Replikation des HCV-Subgenom-Konstrukts. Anstelle des Luziferasegens kann ebensogut ein anderes Reportergen verwendet werden, das in analoger Weise modifiziert ist, so daß seine Expression von der Virusreplikation abhängt, obwohl dieses Reportergen nicht Bestandteil des HCV-Subgenom-Konstrukts ist. Es kann auch ein zelluläres Protein, welches durch die HCV-Proteine oder Nukleinsäure inaktiviert oder aktiviert wird, als sogenannter Surrogatmarker verwendet werden. In diesem Fall ist die Expression bzw. Aktivität dieses Surrogatmarkers ein Maß für die Replikation der viralen DNA.

25 Beispiel 6: Herstellung von HCV-Subgenom-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfären für die Gentherapie

[0066] Diese rekombinanten und selektionierbaren HCV-Subgenom-Konstrukte werden in trans-komplementierende Helferzelllinien transfiziert, d.h. in Zelllinien, die induzierbar oder konstitutiv die fehlenden Funktionen (beispielsweise die Strukturproteine) exprimieren. Zellklone, die ein funktionelles HCV-Subgenom-Konstrukt enthalten, lassen sich durch entsprechende Selektion etablieren. Die von der Wirtszelle exprimierten Virus-Strukturproteine erlauben die Bildung von Viruspartikeln, in die die RNA der HCV-Subgenom-Konstrukte eingeschleust wird. Das Ergebnis sind also virus-ähnliche Partikel, die ein erfindungsgemäßes HCV-Subgenom-Konstrukt einschließlich des einklonierten Fremdgens enthalten und die dieses mittels Infektion auf andere Zellen übertragen können. Ein Beispiel für ein solches Konstrukt ist in Fig. 8 dargestellt. Es besteht auch die Möglichkeit, das hier beschriebene erfindungsgemäßes HCV-Subgenom-Konstrukt mit integriertem Fremdgen direkt als Expressionsvektor einzusetzen. Dabei wird analog dem vorgenannten Verfahren vorgegangen, allerdings mit dem Unterschied, daß Zelllinien transfiziert werden, die keine trans-komplementierenden Faktoren exprimieren. In diesem Fall dient das HCV-Konstrukt also lediglich als Expressionsvektor.

40 Beispiel 7: Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte

(A) Isolationsverfahren

[0067] Für die Bestimmung adaptiver Mutationen und die Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte wurde wie folgt verfahren: Zellen wurden mit einem HCV-RNA-Konstrukt wie unter den Beispielen 1 und 2 beschrieben transfiziert und G418-resistente Zellklone hergestellt. Zur Bestimmung der Replikationskompetenz (darunter wird in diesem Zusammenhang die Anzahl G418-resistenter Zellklone verstanden, die pro Mikrogramm transfizierter HCV-RNA bzw. HCV-RNA-Konstrukt erhalten wird) wurde exemplarisch die Gesamt-RNA aus einem der Zellklone, genannt 9-13 (Fig. 1B, Spur 11) isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot wie in Fig. 2 B beschrieben bestimmt. 10 Mikrogramm der Gesamt-RNA, die ca. 10^9 Moleküle HCV-RNA enthielt, wurde anschließend per Elektroporation in naïve Huh-7 Zellen eingeschleust (Fig. 9). Parallel dazu wurden 10^9 in vitro Transkripte der analogen neo-HCV-RNA, die mit isolierter Gesamt-RNA aus naïven Huh-7 Zellen auf eine Gesamt-RNA-Menge von 10 µg aufgefüllt worden war, in naïve Huh-7 Zellen transfiziert. Nach Selektion mit G418 wurde die Anzahl der Zellkolonien, ausgedrückt in 'colony forming units (cfu) pro Mikrogramm RNA' in den beiden Ansätzen bestimmt. Bei einer Konzentration von 500 µg/ml G418 im Selektionsmedium betrug die Zahl der Kolonien, die mit der in der **isolierten** Gesamt-RNA enthaltenen HCV-RNA erhalten wurde, ca. 100.000 cfu pro Mikrogramm HCV-RNA. Dagegen wurden mit derselben Menge in vitro transkribierter HCV-RNA nur 30 - 50 Kolonien erhalten. Dieses Ergebnis belegt, daß die

spezifische Infektiosität der HCV-RNA, die aus den Zellklonen isoliert wurde, ca. 1.000 - 10.000-fach höher ist als die Infektiosität der analogen in vitro Transkripte. Das methodische Vorgehen ist in Fig. 9 dargestellt.

[0068] Mit Hilfe der 'long-distance RT-PCR' wurde die HCV-RNA aus der Gesamt-RNA der 9-13 Zellen amplifiziert, die PCR-Amplikate kloniert und zahlreiche Klone sequenziert. Ein Vergleich der Sequenzen dieser rekliierten RNAs mit der Sequenz der RNA, die ursprünglich in die naïiven Huh-7 Zellen eingeschleust wurde ergab, daß die rekliierten RNAs zahlreiche Aminosäureaustausche besaßen, die über die gesamte HCV-Sequenz verteilt waren (Fig. 10).

⁵ *Sfil*-Fragmente dieser rekliierten Mutanten wurden im Austausch gegen das analoge *Sfil*-Fragment des ursprünglichen Replikonkonstrukt in dieses eingeführt und RNAs der jeweiligen Mutanten wurden in naïve Huh-7 Zellen eingeschleust. Nach Selektion mit G418 wurde dann für jede HCV-RNA-Mutante die Zahl der gebildeten Kolonien bestimmt.

¹⁰ Während mit der Ausgangs-RNA nur 30 - 50 Kolonien pro Mikrogramm RNA erhalten wurde war die Koloniezahl bei zwei der rekliierten Varianten deutlich höher (Fig. 10). Im Fall der HCV-RNA-Konstrukte 9-13I und 9-13C betrug die spezifische Infektiosität 100 - 1.000 cfu pro Mikrogramm RNA und beim 9-13F Replikon sogar 1.000 - 10.000 cfu pro Mikrogramm RNA. Diese Ergebnisse zeigen, daß die Aminosäureaustausche in dem analysierten NS3-5B-Bereich der Mutanten 9-13I, 9-13C und insbesondere 9-13F zu einer deutlichen Erhöhung der Replikationskompetenz führten.

¹⁵ Demgegenüber waren alle anderen HCV-RNA-Konstrukte (9-13 A, B, G, H und K) nicht mehr replikationskompetent, enthielten also letale Mutationen.

[0069] Zwecks Beantwortung der Frage, welche der Aminosäureaustausche im 9-13F-Konstrukt zur Steigerung der Replikation führten, wurden die Austausche einzeln oder in Kombination in das Ausgangs-HCV-RNA-Konstrukt eingeführt und die entsprechenden RNAs in naïve Huh-7 Zellen eingeschleust. Das Ergebnis der Transfektionen mit diesen RNAs ist in Tabelle 1 zusammengefaßt. Daraus wird ersichtlich, daß im vorliegenden Beispiel die hohe Replikationskompetenz durch mehrere Mutationen bedingt ist. Den größten Beitrag leisten die Aminosäureaustausche in den HCV-RNA-Abschnitten NS5A und NS4B. Auch die einzelnen Austausche in der NS3-Region leisten einen Beitrag, der möglicherweise auf einem Synergismus dieser Einzelaustausche beruht.

²⁰ Diese Befunde belegen, daß es durch die G418-Selektion der Zellen, die mit den *neo*-HCV-RNA-Konstrukten transfiziert wurden, zur Anreicherung solcher HCV-RNAs kam, die eine deutlich höhere Replikationskompetenz hatten. Mit dem hier beschriebenen Versuchsansatz lassen sich HCV-RNA-Konstrukte mit sehr unterschiedlicher Replikationseffizienz selektieren. Je höher die Konzentration des Antibiotikums in dem Selektionsmedium ist, in/auf dem die HCV-RNA-Konstrukt-haltigen Zellen zwecks Selektion kultiviert werden, desto höher muß der Grad an adaptiven Mutationen und damit die Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten sein, damit die Zellen auswachsen können. Werden die Selektionen mit niedrigeren Antibiotikum-Konzentrationen durchgeführt, können auch solche Zellen überleben und sich vermehren, die im Vergleich geringer adaptive Mutationen und eine weniger hohe Replikationseffizienz aufweisen.

²⁵ Das bisher beschriebene HCV-RNA-Konstrukt 9-13F, das mehrere adaptive Mutationen enthielt, hatte eine erwiesenermaßen höhere Replikationseffizienz als die parentale HCV-RNA. Um HCV-RNAs mit noch höherer Replikation in Zellkultur zu erhalten, wurde die HCV-RNA, die in der Gesamt-RNA eines ausgewählten Zellklons enthalten war, mehrfach in naiven Huh-7 Zellen passagiert. Dieser ausgewählte Zellklon, genannt 5-15, wurde durch Transfektion mit dem HCV-RNA-Konstrukt I₃₈₉/NS3-3' erhalten (Fig. 1). Er entspricht weitgehend dem Zellklon 9-13, der durch Transfektion mit einem HCV-RNA-Konstrukt hergestellt wurde, das eine um 22 Nukleotide kürzere HCV-IRES besaß (I₃₇₇/NS3-3'; Fig. 1). 10 Mikrogramm Gesamt-RNA, isoliert aus dem Zellklon 5-15, wurden mittels Elektroporation in naïve Huh-7 Zellen eingeschleust und die Zellen einer Selektion mit 1 mg/ml G418 unterzogen. Aus einem der so erzeugten Zellklone wurde wiederum Gesamt-RNA isoliert, in naïve Huh-7 Zellen transfiziert und analog selektiert. Dieser Vorgang wurde insgesamt viermal wiederholt. Nach der vierten Passage wurde aus einem Zellklon die Gesamt-RNA isoliert und die *neo*-HCV-RNA mit Hilfe der 'long-distance RT-PCR' amplifiziert. Das amplifizierte DNA-Fragment wurde mit dem Restriktionsenzym *Sfil* verdaut und in das *Sfil*-restringierte Ausgangskonstrukt I₃₈₉/NS3-3' inseriert. Insgesamt ³⁰ wurden über 100 DNA-Klone erhalten und zunächst mittels Restriktionsverdau analysiert. In vitro transkribierte RNA von ca. 80 dieser Klone wurde jeweils in naïve Huh-7 eingeschleust und einer Selektion mit 500mg/ml G418 unterzogen. Von den 80 untersuchten *neo*-HCV-RNA-Varianten erwiesen sich die allermeisten als replikationsdefekt. Bei zwei Mutanten, genannt 5.1 und 19, war die spezifische Infektiosität, ausgedrückt als 'colony forming units' pro Mikrogramm RNA, jedoch sehr deutlich erhöht (Tabelle 2). Durch mehrfache Passage der RNA in Zellkultur lassen sich offensichtlich HCV-RNAs herstellen, deren Replikationseffizienz aufgrund von Mutationen (sog. "adaptiven Mutationen) mehrere Größenordnungen höher ist als die der ursprünglich aus dem Patienten klonierten RNA.

(B) Modifikationsverfahren

⁵⁵ [0070] Solche nach (A) erzeugten und identifizierten adaptiven Mutationen können in ein wenig replikationskompetentes HCV-RNA-Konstrukt übertragen werden und führen zu einer massiven Steigerung der Replikation dieses Konstrukt. Diese Steigerung ist so hoch, daß damit nachweislich HCV-RNAs in Zellkultur zur Replikation gebracht werden können, die kein selektierbares Markerprotein mehr besitzen. Fig. 12 zeigt einen Vergleich der Replikationseffi-

zienz von HCV-RNAs, die entweder der Ausgangssequenz oder den adaptierten Sequenzen 9-13F bzw. 5.1 entsprachen. Zwecks einfacher Messung wurde das *neo*-Gen entfernt und durch das Gen für die Luziferase ersetzt. Als Negativkontrolle diente wiederum ein HCV-RNA-Konstrukt, das auf Grund einer inaktivierenden Mutation der NSSB RNA-Polymerase replikationsdefekt war. Schon 24 Stunden nach der Transfektion erkennt man einen deutlichen Unterschied in der Luziferaseaktivität zwischen der defekten RNA und den 9-13F bzw. 5.1-Konstrukten während zwischen der defekten RNA (318 DN) und dem Ausgangs-RNA-Konstrukt (wt) das keine adaptiven Mutationen besaß, kaum ein Unterschied zu sehen war. Während des gesamten Beobachtungszeitraums wurde die höchste Luziferaseaktivität und damit die höchste Replikation mit der 5.1-RNA erhalten. Diese Befunde belegen nicht nur die hohe Replikationseffizienz dieser RNA, sondern zeigen auch, daß es möglich ist, mit adaptierten HCV-RNA-Konstrukten ein Zellkultursystem aufzubauen, für das die Anwesenheit eines selektierbaren Gens nicht mehr notwendig ist. Eine zusammenfassende Übersicht der Nukleotid- und Aminosäureunterschiede zwischen dem Ausgangskonstrukt und den Mutanten 9-13F, 5.1 und 19 ist in Tabelle 3 gegeben.

Beispiel 8: Herstellung zellkultur-adaptierter HCV-RNA-Vollängengenome

[0071] In den Beispielen 1 bis 7 wurde stets eine subgenomische HCV-RNA verwendet, die die gesamte Strukturproteinregion von Core bis einschließlich p7 bzw. NS2 fehlte. Im vorliegenden Beispiel 8 wird gezeigt, daß es möglich ist, mit Hilfe der adaptierten NS3-5B-Sequenz ein HCV-Vollängengenom in Zellkultur zur Replikation zu bringen. Zu diesem Zweck wurde zunächst das Sfil-Fragment der gemäß Beispiel 7 hergestellten, hoch adaptierten HCV-RNA 5.1 in ein selektionierbares HCV-Vollängengenom transferiert (Fig. 12). Dieses HCV-Genom wurde in naïve Huh-7 Zellen transfiziert und einer Selektion mit unterschiedlichen G418-Konzentrationen unterzogen. In Abhängigkeit von der Selektionsstärke (der G418-Konzentration) wurde eine unterschiedlich große Zahl an Zellklonen erhalten (Fig. 12 B). Im Vergleich dazu wurden mit dem unveränderten HCV-Vollängengenom, das keine-adaptiven Mutationen enthielt, keine Kolonien erhalten, ebenso mit der Negativkontrolle, die auf Grund einer inaktivierenden Mutation in der NS5B RNA-Polymerase replikationsdefekt war. Zum Nachweis dafür, daß die so entstandenen Zellklone tatsächlich ein autonom replizierendes HCV-Vollängenkonstrukt enthielten, wurde Gesamt-RNA aus mehreren Zellklonen isoliert und mittels Northern-Blot analysiert. In allen Zellklonen war die Vollängen HCV-RNA eindeutig nachweisbar (Fig. 12). Damit ist eindeutig belegt, daß es mit Hilfe der an Zellkulturen adaptierten HCV-Sequenzen möglich ist, ein HCV-Vollängengenom herzustellen, das mit hoher Effizienz und autonom in einer Zelllinie repliziert, d.h. es können mit dem erfindungsgemäßen System auch adaptierte HCV-Vollängengenome hergestellt werden. Da dieser Klon darüber hinaus die vollständige HCV-Sequenz besitzt, also auch die für die Viruspartikelbildung notwendigen Strukturproteine, ist es mit diesem System möglich, große Mengen infektiöser Viruspartikel in Zellkulturen herzustellen. Zum Nachweis dieser Viren werden zellfreie Überstände von Zellen, die ein replizierendes HCV-Vollängengenom tragen, auf naive Huh-7 Zellen gegeben und die so infizierten Zellen einer Selektion mit G418 unterzogen. Jeder Zellklon, der unter diesen Bedingungen auswächst, geht auf eine infizierte Zelle zurück. Die Viren in den Zellkulturüberständen von Zellen, die ein replizierendes HCV-Vollängengenom besitzen, können aber auch mit verschiedenen im Stand der Technik bekannten Verfahren wie Ultrazentrifugation oder Mikrodialyse angereichert und gereinigt werden und dann zur Infektion naiver Zellen verwendet werden. Mit diesem Verfahren ist eindeutig gezeigt, daß mit dem erfindungsgemäßen HCV-Zellkul-

20 tursystem zellkultur-adaptierte HCV-Vollängengenome hergestellt werden können, die mit hoher Effizienz in Zellen replizieren und infektiöse Viren produzieren. Diese können ebenfalls durch Infektion eines Versuchstiers, vorzugsweise dem Schimpansen, nachgewiesen werden.

Beispiel 9: Herstellung von HCV-Vollängen-Konstrukten und HCV-Subgenom-Konstrukten mit Reportergen.

[0072] Es wird ein HCV-RNA-Konstrukt hergestellt, bei dem anstelle des Antibiotikumresistenzgens ein Reportergen eingefügt wird (Fig. 13). Dabei kann die Replikation anhand der Menge bzw. der Aktivität des Reportergens bzw. Reportergenprodukts bestimmt werden. Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferase-gene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green fluorescence protein Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (sezernierte alkalische Phosphatasegen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. enzyme-linked immunosorbent assay, ELISA) bestimmt werden. Das Reportergen kann entweder von einer eigenen IRES exprimiert werden oder in Form eines Fusionsproteins, das entweder als solches aktiv ist oder mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Beispiel 10: Herstellung von HCV-Vollängen-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfahre für die Gentherapie oder als Expressionsvektor.

[0073] Das Konstrukt (Fig. 14) wird in Zellen eingeschleust und führt dort zur Bildung von HCV-Viruspartikeln, die zur Infektion weiterer Zellen verwendet werden können. Da die Viruspartikel eine RNA mit einem Fremdgen enkapsuliert haben kann dieses in den so infizierten Zellen zur Produktion des von diesem Fremdgen kodierten Proteins benutzt werden. Zellen, die mit dem Konstrukt transfiziert wurden, exprimieren ebenfalls das Fremdgen.

Beispiel 11: Herstellung von monocistronischen HCV-RNA-Konstrukten, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird.

[0074] Für bestimmte Untersuchungen ist es von Vorteil, wenn das HCV-RNA-Konstrukt kein heterologes IRES-Element besitzt. Solche Untersuchungen sind beispielsweise die Bestimmung der Interferonresistenz. Wird eine Zelle, die ein HCV-RNA-Konstrukt besitzt, mit Interferon-alpha oder -beta inkubiert, kommt es zu einer Reduktion der Replikation der HCV-RNA. Zur Aufklärung des Wirkungsmechanismus ist es notwendig, daß das HCV-RNA-Konstrukt keine heterologe IRES besitzt, da ansonsten nicht bestimmt werden kann, ob die Interferon-vermittelte Hemmung durch eine Hemmung der HCV-Replikation oder durch eine Hemmung der heterologen IRES vermittelt wird. Deshalb werden Konstrukte hergestellt, bei denen das Resistenzgen mit einem HCV-Protein fusioniert wird (Fig. 15). Entweder das Fusionsprotein ist als solches aktiv, oder das Resistenzgenprodukt wird mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV)-Protease von diesem abgespalten wird.

Tabelle 1

Spezifische Infektiositäten (cfu/µg RNA) der HCV-RNA-Konstrukte mit adaptiven Mutationen, die bei der 9-13F Mutante gefunden und in das parentale HCV-RNA-Konstrukt I ₃₈₉ /NS3-3'/wt eingeführt wurden		
Aminosäureaustausch ¹	HCV-Protein	cfu/µg RNA ²
kein		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -> ala + 1577 lys -> arg + 1609 lys -> glu)	NS3	360 - 420
1936 pro -> ser	NS4B	500 - 1000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5B	30 - 60
alle zusammen		5000

¹ Aminosäureaustausch im Polyprotein des HCV-Isolats Con-1 (EMBL-Genbank No. AJ238799); die Aminosäuren sind im Dreibuchstabenkode angegeben.

² Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500µg/ml G418.

Tabelle 2

Spezifische Infektiositäten (cfu/µg RNA) des parentalen HCV-RNA-Konstrukts I ₃₈₉ /NS3-3'/wt und der Varianten 9-13C, 9-13I, 9-13F, 5.1 und 19.	
Transfizierte RNA-Variante	cfu/µg RNA ¹
Wildtyp	30 - 50

¹ Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500µg/ml G418.

EP 1 043 399 B9 (W1B1)

Tabelle 2 (fortgesetzt)

5	Spezifische Infektiositäten (cfu/ μ g RNA) des parentalen HCV-RNA-Konstrukts I ₃₈₉ /NS3-3'/wt und der Varianten 9-13C, 9-13I, 9-13F, 5.1 und 19.	
10	Transfizierte RNA-Variante	cfu/ μ g RNA ¹
	9-13 C	100 - 1.000
	9-13 I	100 - 1.000
	9-13 F	1.000 - 10.000
15	5.1	50.000 - 100.000
	19	50.000 - 100.000

¹ Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500 μ g/ml G418.

Tabelle 3

20	Nukleotid- und Aminosäureunterschiede zwischen dem parentalen HCV-RNA-Konstrukt I ₃₈₉ /NS3-3'/wt und den Mutanten 9-13I, 9-13F, 5.1 und 19			
25	HCV Mutante	Nukleotidposition	Nukleotidaustausch	Aminosäureaustausch
	9-13 I	3685	C > T	Pro > Leu
		4933	C > T	Thr > Met
30		5249	T > C	-
		8486	C > T	-
35		8821	G > A	Trp > stop
		8991	C > G	Arg > Gly
40		9203	A > G	-
		9313	T > C	Phe > Ser
45		9346	T > C	Val > Ala
50	9-13 F	3866	C > T	-
		4188	A > G	Arg > Gly
		4489	A > C	Glu > Ala
55		4562	G > A	-
		4983	T > C	-
		5071	A > G	Lys > Arg
		5166	A > G	Lys > Glu
		6147	C > T	Pro > Ser
		6829	A > G	Glu > Gly
		7329	A > G	Lys > Glu
		7664	A > G	Ile > Val
		8486	C > T	-
		8991	C > G	Arg > Gly
	NK5.1	4180	C > T	Thr > Ile
		4679	C > T	-
		4682	T > C	-

Tabelle 3 (fortgesetzt)

Nukleotid- und Aminosäureunterschiede zwischen dem parentalen HCV-RNA-Konstrukt I ₃₈₉ /NS3-3'/wt und den Mutanten 9-13I, 9-13F, 5.1 und 19				
	HCV Mutante	Nukleotidposition	Nukleotidaustausch	Aminosäureaustausch
5		5610	C > A	Leu > Ile
		6437	A > G	-
10		6666	A > G	Asn > Asp
		6842	C > T	-
		6926	C > T	-
15		6930	T > C	Ser > Pro
		7320	C > T	Pro > Ser
		7389	A > G	Lys > Glu
20	NK19	3946	A > G	Glu > Gly
		4078	C > G	Ala > Gly
		4180	C > T	Thr > Ile
		4682	T > C	-
25		5610	C > A	Leu > Ile
		5958	A > T	Met > Leu
		6170	T > A	-
		6596	G > A	-
30		6598	C > G	Ala > Gly
		6833	C > T	-
		6842	C > T	-
		6930	T > C	Ser > Pro
35		7141	A > G	Glu > Gly
		7320	C > T	Pro > Ser
		7389	A > G	Lys > Glu
40		7735	G > A	Ser > Asn

[0075] Angegeben sind die Unterschiede der Nukleotid- und Aminosäuresequenzen zwischen der Ausgangs-HCV-RNA-Sequenz Con 1 (EMBL-Genbank No. AJ238799) und denen derzellkulturadaptierten HCV-RNAs. Die Zahlen beziehen sich auf die Nukleotid- und Aminosäurepositionen des HCV-Isolats Con1.

45

50

55

SEQUENCE LISTING

<110> Bartenschlager, Ralf
 5
 <120> Hepatitis C Virus cell culture system
 <130> ba-1
 10
 <140> 199 15 178.4
 <141> 1999-04-03
 <160> 11
 15
 <170> PatentIn Ver. 2.1
 <210> 1
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus
 20
 <400> 1
 gccagccccc gattggggc gacactccac catagatcac tccccgtga ggaactactg 60
 tcttcacgcga gaaagcgct agccatggcg ttagtatgag tgctcgccg cctccaggac 120
 cccccctccc gggagagcca tagtggctcg cggaccggt gagaatcccg gaattgccag 180
 gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccagta gtgtgggtc gcgaaaggcc ttgtgtact gcctgatagg 300
 gtgttgcga gtgccccggg aggtctcgta gaccgtgcac catgacgacg aatcctaaac 360
 25
 ctcaaagaaa aacccaaacgt aacaccaacg ggcgcgcatt gattgaacaa gatggattgc 420
 acgcagggttc tccggccgct tgggtggaga ggctattcg cgatgactgg gcacaacaga 480
 caatcggtcg ctctgatgcc gccgttcc ggctgtcgc gcaggggcgc cccgttctt 540
 ttgtcaagac cgacctgtcc ggtgcctga atgaactgca ggacgaggca ggcgcgcstat 600
 cgtggctggc cacgacgggc ttcccttgcg cagctgtct cgacgttgc actgaagcgg 660
 gaagggactg gctgctattt ggcgaagtgc cggggcagga tctcctgtca tctcacctt 720
 30
 ctcctgcccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgate 780
 cggctacccg cccattcgac cacaacgcg aacatcgcat cgagcgagca cgtactcgga 840
 tggaaaggccg tcttgcgtat cagatgtatc tggacgaaga gcatcagggg ctcgcgcag 900
 ccgaactgtt cgccaggttc aaggcgcgcg tgccgcacgg cgaggatctc gtctgaccc 960
 atggcgatgc ctgcgttccg aataatcatgg tggaaatgg cgccttttgc ggattcatcg 1020
 actgtggccg gctgggtgt gcgaccgcg atcaggacat acgcttgcg acccgtata 1080
 ttgctgaaga gcttggccgc gaaatggctg accgcttctc cgtctttac ggtatcgccg 1140
 35
 ctcccgattt gcagcgcata gccttctatc gccttcttgc cgagtttctc ttagttaaa 1200
 cagaccacaa cggtttccct ctgcggat caattccgc cctctccctc cccccccct 1260
 aacgttactg gccgaagccg ctggaaataa ggccgggtgt cgtttgtcta tatgttattt 1320
 tccaccatat tgccgttctt tgcaatgtg agggcccgaa aacctggccc tgtcttctt 1380
 acgagcattt ctagggtctt ttccctctc gccaaaggaa tgcaggatctc gtgtaatgtc 1440
 40
 gtgaaggaaag cagttctctt ggaagcttgc tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcgcg ggaacccccc acctggcgcg aggtgcctc gcggccaaaa gccacgtgt 1560
 taagatacac ctgcacaaaggc ggacacaccc ctagcgcacg ttgtgagttt gatagttgt 1620
 gaaagagtca aatggcttc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
 gtacccattt gtatggatc tgatctgggg ctcgggtgca catgcatttc atgtttag 1740
 tcgagggttaa aaaacgtcta ggccccccga accacggggc cgtgttttc ctttggaaaa 1800
 cacgataata ccatggcgcac gaaatctaaa cctcaaaaggaa aaacccaaacg taacaccaac 1860
 45
 cggcccccac aggacgtcaa gttccgggc ggtggtcaga tcgtcggtgg agtttacctg 1920
 ttgcgcgcga ggggccccag gttgggtgt cgccgcacta ggaagacttc cgagcggtcg 1980
 caacctcggt gaaaggcaca acctatcccc aaggctcgcc agcccgagg tagggcctgg 2040
 gtcagcccg ggtaccctctg gcccctctat ggcaatgagg gcttgggtg ggcaggatgg 2100
 ctccctgtcac cccgtggctc tggcttagt tggggccca cggacccccc gcttaggtcg 2160
 cgcattttgg gtaagggtcat cgataccctc acgtggggct tcgcgcattc catgggtac 2220
 attccgcctcg tggcgccccc cctagggggc gctgcccagg ccctggcgc tggcgccgg 2280
 gttctggagg acggcgtaa ctatgcacca gggaaatctgc cccgttgcgc cttttctatc 2340
 ttccctttgg ctgttgcgtc ctgtttgacc atccctcgat cccgttatgc atgtgcac 2400
 gtatccggag tgcattatgt cacaacgcg tgctccaaacg caagcattgt gtatggggca 2460
 ggggacatgaa tcatgcatac ccccggtgc gtggccctgcg ttggggagaa caactccctcc 2520
 cgcgtctggg tagcgctcac tccacgcgc gggggccaggaa acgctagcg cccactacg 2580
 acgatacgcac gccatgtcga ttgtgtcgat gggggccggctg ctctctgcgc cgcctatgtac 2640
 55
 gtgggagatc tctgcggatc tgggtttccctc gtcgcgcacg tggatccatc ctcgcctcg 2700

	cgccacgaga cagtacagga ctgcatttgc tcaatatac cccggccacgt gacaggtcac	2760
5	cgtatggctt gggatatgtat gatgaactgg tcacctacag cagcccttagt ggtatcgac	2820
	ttactccgga tcccacaacgc tgcgtggat atggggcg gggccccattt gggaggctta	2880
	gggggccttg cctactattt catggggggg aactgggcta aggttctgtat tgcgtatgcta	2940
	ctctttccg gcggtacgg gggaaacctat gtgcacagggg ggacgatggc caaaaacacc	3000
	ctcgggatta cgtccctt ttcacccggg tcataccaga aaatccagct tgtaaacacc	3060
10	aacggcagct ggcacatcaa caggactgcc ctgaaactgca atgactccct caacactggg	3120
	ttccttgcgt cgctgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatt	3180
	gccagctgca gccccatcgca cgcgttcgtc cagggggtggg gggccatcac ttacaatgag	3240
	tcacacagct cggaccagag gccttattgt tggcaactacg caccggggcc gtgcggatc	3300
	gtaccccggg cgcagggttg tggtccaggat tactgttca ccccaagccc tgcgtgggt	3360
	gggacgaccg accgggttcgg cgtccctacg tacagttggg gggagaatgaa gacggacgtg	3420
15	ctgcttctta acaacacggc gcccggcaaa ggcaactgggt ttggctgtac atggatgaat	3480
	agcaactgggt tcaccaagac gtgcgggggc ccccggtgtt acatcggggg gatcggcaat	3540
	aaaaccttga cctggccccc ggactgcttc cggaaagcacc ccgaggccac ttacaccaag	3600
	tgtggttcgg ggccttgggtt gacacccaga tgcttggtcc actaccata caggctttgg	3660
	cactaccctt gcaactgtcaa ctttaccatc ttcaaggta ggtatgtacgt gggggggagt	3720
	gagcacaggc tcgaagccgc atgcaattgg actcgaggag acggtttaa cctggaggac	3780
	aggacagat cagacttag cccgctgtc ctgtctacaa cggagtgcc ggtattgccc	3840
20	tgttccttca ccaccctacc ggctctgtcc actgggttga tccatctcca tcaaaacgtc	3900
	gtggacgtac aatacgtta cggatataggg tggcgggttgc tctccttgc aatcaaattgg	3960
	gagtatgtcc tggcttcttcg goggacgcgc gctgtctgtc ctgtctgtgg	4020
	atgatgtcgc tgatagtcgca agctggggcc gcccstagaga acctgggttgc cctcaacacgc	4080
	gcattccgtgg ccggggcga tggcatttc tcttcttcg tggcttcttcg tgcgtccctgg	4140
	tacatcaagg cggggctgtt ccctggggcg gcatatggcc tctacggcgat atggccgtt	4200
25	ctccctgtcc tgcgtgggtt accaccacga gcatacgcga tggacggggaa gatggcagca	4260
	tgcgtccggag ggcgggtttt cgttaggtctg atacttgc cttgtcacc gcactataag	4320
	ctgttccctcg ctaggctcat atgggggttca caatattttt tcaccaggcgc cgaggcacac	4380
	ttgcaagtgtt ggatccccccc cctcaacgtt cggggggggcc gcatatgttgc catcctcctc	4440
	acgtgcgcga tccacccaga gctaattttt accatcacca aaatcttgc cggcataactc	4500
30	ggtccactca tgggtcttca ggctggata accaaatgtc cgtaatttcg ggcgcacac	4560
	gggctcattc gtgcattgtc gctgggtcgaa aagggttgcg ggggttcatca tgcctaaatg	4620
	gctctcatga agttggccgc actgacaggat acgtacgtt atgaccatct caccggactg	4680
	cgggactggg cccacggggg cctacgagac cttgcgggttgc cagttgagcc cgtcgcttc	4740
	tctgtatggt agaccaaggat tatacctggg gggcagaca cccggcggt tggggacatc	4800
	atcttggggcc tggcggttc cggccggcagg gggggggaga tacatctggg accggcagac	4860
	agccttgaag ggcagggggtg ggcatttc cgccttattt cggcctactc ccaacagacg	4920
	cgaggcctac ttggctgtcat catcaactacg ctcacagggc gggacaggaa ccagggtcgag	4980
35	ggggagggtcc aagtgggttca caccgcacaa caatcttcc tggcgttgc cgtcaatggc	5040
	gtgtgttggc ctgtctatca tgggtccggc tcaaagaccc ttggccggccca aaaggggccca	5100
	atcacccaaa tgcacaccaa tggggaccat gacctgttcg gctggcaacgc gccccccggg	5160
	gcccgttccct tgcacccatg caccgtccgc agctggacc ttacttgcgat caccggat	5220
	gccgatgtca ttccgggtgc cggggggggc gacaggcagg ggagctact ctccccccagg	5280
	cccgctcttc acttgcaggg ctcttcggc ggttccactgc tctgccttc ggggcacgct	5340
40	gtgggcatcc ttcgggttcg cgtgtgcacc cgggggggttgc cgaaggcggt ggactttgtt	5400
	cccgctcgat ctatggaaac cactatgcgg tccccgggttgc tcaacggacaa ctcgtccct	5460
	ccggccgtac cgcagacattt ccagggtggc catctacacg cccctactgg tagcggcaag	5520
	agcaactaagg tgcgggttgc gtatgcggcc caagggtata aggtgttgc cttgtacccgc	5580
	tccgtccggc ccacccatgg tttccggggc tatatgtcta aggacatgg tatcgaccct	5640
	aacatcagaa ccggggtaag gaccatcacc acgggttgc ccatcacgtt cttccacctat	5700
	ggcaagtttc ttggccgacgg tgggtgttgc tggggcgctt atgacatcat aatatgttat	5760
45	gagtggccact caactgactc gaccactatc ctggcgtatc gcacagtctt ggaccaagcg	5820
	gagacggctg gaggcgtact cgtgtgttc gcccacgcgtt cgcctccggg atcggttacc	5880
	gtgccacatc caaacatcgaa ggagggtggct ctgtccaggca ctggagaaat ccccttttat	5940
	ggcaagggca tccccatcgaa gaccatcgaa gggggggaggc acctcatttt ctgcatttc	6000
	aagaagaaat gtgtatggatc cgcgcgttgcat tggccggcc tcggactcaa tgcgttagca	6060
	tattaccggg gcttgcgtat atccgtcata ccaactagcg gagacgtcat tgcgttagca	6120
50	acggacgttc taatgcggg ctttaccygc gatttcgact cgtgtatcgat ctgcataaaaa	6180
	tgtgtccatcc agacagtgc cttcaggcttgc gaccggact tcaccatttgaa gacggacgacc	6240
	gtggccacaag acgcgggttc acgctcgatc cggcgaggca ggactggtag gggcaggatg	6300
	ggcattttaca gttttgttgc tccaggagaa cggccctcgg gcatgttgc ttccctcggtt	6360
	ctgtgcgttgc gctatgacgc gggctgttgc tggtacggc tcacggccgc cggacacctca	6420
	gttaggttgc gggcttaccc aacacacca ggggttgc ctcgtccaggaa ccatctggag	6480
	ttctgggaga ggcgttttaccc aggcctcacc cacatagacg cccattttt tgcctcggact	6540
55	aaggcaggcag gagacaactt cccctacgtt gtagcataacc aggctacggt gtgcggccagg	6600
	gctcaggctc cacctccatc gtgggaccaa atgttggaaat gtctcatacg gctaaaggcct	6660

acgtgcacg ggccaacgcc cctgtgtat aggctggag ccgttcaaaa cgaggttact 6720
 accacacacc ccataaccaa atacatcatg gcatgcgtg cggctgacct ggaggtcg 6780
 acagcacct gggctgtgtt agggcggagtc ctgcgcgtc tggccgcgt 6840
 acaggcagcg tggtcattgt gggcaggatc atcttgtcc gaaaggccgc catcattccc 6900
 gacagggaaag tccttaccc ggagttcgat gagatggaaag agtgcgcctc acacccctt 6960
 tacatcgAAC agggatgcg aacttcaaac agaaggcaat cgggttgcgt 7020
 caaacagcca ccaagcaagc ggaggctgtc gtcgggtgg tggatccaa gtggcggacc 7080
 ctgcagcct tctggcggaa gcataatgtgg aatttcatca gcgggataca atatttagca 7140
 ggctgtcca ctctgcgttgg caaccccgatc atagcatcac tgatggcatt cacagccct 7200
 ataccaggcc cgctcaccac ccaacatacc ctccgttta acatctggg gggatgggt 7260
 gcccggccaaac ttgtctcc cagcgtgtc tctgcgttgc taggcgcggg catcgctgg 7320
 gcccgttgg gcagcatagg ccttggaaag gtgctgtgg atatttggc agttatgg 7380
 gcagggggtgg caggcgccgt cgtggccctt aaggctatga gcggcgagat gccctccacc 7440
 gaggacgtgg ttaacacttcc cctgtgtatc ctctccctg gcccgttgcgt 7500
 gtgtgcgcag cgatactgcg tcggacgtg gcccagggg aggggggtgt gcagtggatg 7560
 aaccggctgtc tagcgttgcg ttcgggggtt aaccacgtt ccccccacgc ctagtgcct 7620
 gagagcgacg ctgcagcagc tgcacttcg atccctctca gtcttacatc cactcagctg 7680
 ctgcaggaggc ttccacgtg gataacacgg gactgtcga cggcatgtc cggctgtgg 7740
 ctaagagatg tttgggttg gatatgcacg gtgttactg atttcaagac ctggctccag 7800
 tccaagctcc tggccggatt gcccggatc cccttcttca catgtcaacg tgggtacaag 7860
 ggagtctggc gggcgcacgg catcatgca accacctgcc catgtggagc acagatcacc 7920
 ggacatgtga aaaacgggttcc catgaggatc gtggggcttta ggacctgttag taacacgtgg 7980
 catggacat tccccatcaa cgcgtacacc acggggccctt gcacggccctc ccoggcgcca 8040
 aattattcta gggcgctgtg cgggggtggct gctgaggagt acgtggaggt tacgggggt 8100
 ggggatttcc actacgtgac gggcatgacc actgacaacg taaagtgcgc gtgtcagggtt 8160
 cccggccccc aattttcac agaagtggat ggggtgcgtg tgcacaggta cgctccagcg 8220
 tgcaaaacccc tcctacgggaa ggaggtcaca ttccgtgtc ggctcaatca atacctgggtt 8280
 ggttcacagc tcccatgcga gcccgaaccc gacgttagcag tgctcaactc catgctcacc 8340
 gacccttccc acattacggc ggagacggct aacgttaggc tggccagggg atctccccc 8400
 tccttggcca gtcatcagc tagccagct tctgcgcctt ccttgcaggc aacatgcact 8460
 acccgcatg actcccccggc egctgaccc atcgaggcga accttcgtg gggcaggag 8520
 atggggcgaaa acatcaccggc cgtggagtca gaaaataggta tagtaattttt ggactctttc 8580
 gagccgcetcc aagcggagga ggttggaggaa gaagtatccg ttccggcgga gatectgcgg 8640
 agtccagga aattccctcg aacgtatgcac atatggcac gcccggattttaaaccctcca 8700
 ctgttagagt cctggaaaggc cccggactac gtccctccag tggtaacacgg gtgtccattt 8760
 ccgcctgcca aggccctcc gataccacctt ccacggagga agaggacggt tgctctgtca 8820
 gaatctaccg tgcatttcgc ttggcgaggat ctcgcaccaa agacccctgg cagctccgaa 8880
 tcgtcgcccg tgcacagcgg cacggcaacg gcctctccgc accagccctc cgacgacggc 8940
 gacgcgggat ccgacgttgcgtc tccatgcctt cccttgcggg ggagccgggg 9000
 gatcccgtatc tcagcgtacgg gtcttgcgtc gacgtatgcac atatggcac gcccggattttaaaccctcca 8700
 gtctgtgtct cgtatgtcttca cacatggaca aactctttgc tccgttccacca caacttggtc 9180
 gaaaccaagc tgcccatcaa tgcactgacg cggcagaaga aggttacccat tgacagactg 9240
 tatgttacaa catctcgac cgcacgttgc tctaaggaga tgaaggcgaa ggctgtccaca 9300
 cagggttgcgtt acgaccata cccggacgtg gttaaggctt aacttcttcc acattcgcc 9360
 agatctaaat ttggcttatgg ggcaaaaggac gtccggaaatccatc gggcggttaac 9420
 cacatccgtt ccgtgtggaa ggacttgcgtc gaaagactg agacacactt tgacaccacc 9480
 atcatggcaaa aaaatggatgt ttctgtgtc caaccaggaga agggggggccg caagccagct 9540
 ccgccttatcg tattccaga tttgggggtt cgtgtgtcg agaaaatggc cccttacat 9600
 gtggcttccca ccctccctca ggcgtgtat ggcttccat acggattcca atactctt 9660
 ggacagcgccgg tgcagttcctt ggtgaatgcc tggaaagcgaa agaaaatggcc tatgggctt 9720
 gcatatgaca cccgtgttt tgactcaacg gtcactgaga atgacatccg tggtaaggag 9780
 tcaatctacc aatgttgta cttggccccc gaagccagac agggcataag gtgcgtcaca 9840
 gagcggcttt acatcgaaaa cccctgtact aattctaaat ggcagaactg cggctatcg 9900
 cggtgcgcgcg cgagcgtgtt actgacgacc agtgcgttgc ataccctcac atgttactt 9960
 aaggccgtcg cggcgttgcg agtgcgttgc acgttgcgttgc acgttgcgttgc 10020
 gacgacgttgc tgcgttatcg tggtaaggcg gggacccaaag aggacggaggc gggcttacgg 10080
 gccttcacgg aggctatgc tagatactt cccatgtgtt cggccatccgc cggctatcg 10140
 tacgactttgg agtggataac atcatgttcc tccatgtgtt cgtgcgttgc cgtatgtatct 10200
 ggcaaaagggt tgcgttatctt caccgtgtac cccacccccc cccttgcgtg ggttgcgtt 10260
 gagacagctt gacacacttcc agtcaattcc tggcttagggca acatcatcat gtatgcgc 10320
 accttgcgttgg caaggatgtatc cctgtatgtt catttcttcc ccatcccttca agtgcgttgc 10380
 caacttgcgtt aagcccttgcgtt tggcttagggca acatcatcat gtatgcgc 10440
 gacctacccatc agatcattca acgactccat ggccttacgg cattttactt ccatagttac 10500
 tctccaggttgc agatcaatag ggtggcttca tgcctcaggaa aacttgggtt accggccctt 10560
 cgagtcgttgc gacatcgccgc cagaagtgtc cgcgttgcgttgc tactgtccca ggggggggg 10620

gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcaact 10680
ccaatcccg ctgcgtcccc gttggattta tccagcttgt tcgttgctgg ttacagcggg 10740
5 ggagacatat atcacagcct gtctcggtcc cgaccggct ggttcatgtg gtgcctactc 10800
ctactttctg tagggtagg catctatcta ctcccccaacc gatgaacggg gagctaaaca 10860
ctccaggcca ataggccatc ctgtttttt ccctttttt ttttctttt tttttttttt 10920
ttttttttt ttttttttc tcctttttt ttcctctttt ttccttttc tttcctttgg 10980
tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040
10 agagagtgct gatactggcc tctctgcaga tcaagt 11076

10

15

20

25

30

35

40

45

50

55

<210> 2
 <211> 8637
 <212> DNA
 5 <213> Hepatitis C Virus

 <400> 2
 gccagcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgctt agccatggcg ttagtatgag tgctgtcgac cttccaggac 120
 cccccctccc gggagagcca tagtggtctg cggaaaccggt gatcacccg gaattgccag 180
 10 gacgaccggg tccttcttg gatcaacccg ctcaatgcct ggagatttg gcgtcccccc 240
 gcgagactgc tagccagta gtgtgggtc gcgaaaggcc ttgtgtact gcctgtatagg 300
 gtgtgtcgca gtccccccgg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
 ctc当地
 15 cggccgctt ggtggagagg ctattcggt atgactgggc acaacagaca atcggctgct 480
 ctgtatccgg cgtgttccgg ctgtcagcgc agggggcgccc ggttctttt gtcaagaccg 540
 acctgtccgg tgccctgaat gaactgcagg acgaggcage gcggtatcg tggctggcca 600
 cgacggcgt tccttgcga gctgtgtcg acgttgtcac tgaagcggga agggactggc 660
 tgctattggg cgaagtggcg gggcaggatc tccgtcatc tcacettgt cctgcccaga 720
 aagtatccat catggctgtat gcaatgcggc ggctgcatac gcttgcattc gctacctgco 780
 cattcgaacca ccaagcggaa catcgcatcg aycgagcacf tactcgatg gaagccggc 840
 ttgtcgatca ggatgtatcg gacgaagagc atcaggggct cgcgcagcc gaactgttgc 900
 20 ccaggctcaa ggcgcgcatg cccgacggcg aggatctcg cgtgacccat ggcgtatgcct 960
 gcttgcgaa tatcatgtt gaaaatggcc gctttctgg attcatcgac tggccggc 1020
 tgggtgtggc ggaccgttat cggacatag cgttggctac ccgtgtatatt gctgaagagc 1080
 ttggcggcga atgggtgtac cgttctctcg tgctttacgg tattcggctt cccgatcg 1140
 agcgtatcg ctttatcg ctttgcgtt agttttctcg agttttaaaca gaccacaacg 1200
 25 gtttccctct agcggatca attccggccc tcccccctcc ccccccctaa ctttactggc 1260
 cgaagccgct tggaaataagg ccgggtgtcg tttgtctata tgattttt caccatattg 1320
 ccgtctttt gcaatgttag ggcggggaa cctggccctg tcccttgc gaggatttct 1380
 aggggtcttt cccctctcg caaaggaatg caaggctctg tgaatgtcgta gaagggaaagca 1440
 gtttctctgg aagttttt aagacaaaca acgtctgtat cggccctttt caggcagcgg 1500
 aaccccccac ctggcgacag gtgcctctgc ggcggggaa cacgtgtata agatacacct 1560
 30 gcaaaggcgg cacaacccca gtggcacgtt gtgagttgaa tagttgtgaa aagagtcaaa 1620
 tggctctctt caagcgatatt caacaagggg ctgaaggatg cccagaaggat accccattgt 1680
 atgggatctg atctggggcc tcgggtgcaca tgctttatcat gtgttttagtc gaggttaaaa 1740
 aacgtctagg ccccccgaac cacggggacg tgggtttctt ttgaaaaaca cgataatacc 1800
 atgggacggg agatggcagc atcggtcgga ggcgggttt tcgttaggtt gataactctt 1860
 accttgcac cgcactataa gcttgcattc gtaggtctt tatgggtgtt acaatatttt 1920
 atcaccaggc cggaggcaca ctttgcgtt tggatcccccc ccttcaacgt tggggggggc 1980
 35 cgcgtatcg tcattcttcc ctcgtgcgcg atccacccag agctaattt taccatcacc 2040
 aaaatcttcg tcgcataact cggccactt atgggtctt cggctgttat aaccaaaatgt 2100
 ccgtacttcg tgcgcgacaca cgggtctt cgtgcacatc tgctgtgtcg gaagggttct 2160
 gggggtatt atgtccaaat ggctctcatg aagttggccg cactgacagg tacgtacgtt 2220
 tatgaccatc tcaccccaact gcccggactgg gcccacgcgg gcctacgaga ctttgcgggt 2280
 40 gcagttgagc ccgtcgctt ctctgtatag gagaccaagg ttatcacctt gggggcagac 2340
 acccgccgt gtggggacat catcttggc ctggccgtt ccgcggcgg gggggagggag 2400
 atacatctgg gaccggcaga cagccttga gggcagggtt ggcgactctt cgcgcctatt 2460
 acggcctact cccaaacagac gogaggcata cttggctgca tcatacttag cctcacaggc 2520
 cgggacagga accaggtcgat gggggagggtc caagtggctt ccaccgcaac acaatctt 2580
 ctggcgacat gctgtcaatgg cgtgtgttgg actgtctatc atgggtccgg ctcaaagacc 2640
 45 ctggccggcc caaaggggccc aatcaccatca atgtacacca atgtggacca ggacctcg 2700
 ggctggcaag cggcccccgg ggcgcgttcc ttgacaccat gcacccgtgg cagctcgac 2760
 cttaacttgg tcacgaggca tgcgtatgtc attccgggtc gccggcgcccc cggacaggcagg 2820
 gggagccatc tctcccccac gcccgttcc tacttgaagg gcttccggg cggccactgt 2880
 ctctggccct cggggcgcgcg tggggcattt tttccgggtc cctgtgtcgac cccgggggtt 2940
 gccaaggcgg tggactttt acccgtcgat tctatggaa ccaatgtcg tggccgggtc 3000
 50 ttacgggaca actcgcccc tccggccgtt ccgcagacat tccagggtggc ccatctacac 3060
 gcccctactg gtagcgccaa gagactaag gtggccgtt cgtatgcac ccaagggtat 3120
 aagggtcttgc ttctgaaccc gtcgtcgcc gcccacccat gtttccgggtc gtatatgtct 3180
 aaggcgcacatg gtatcgaccc taacatcaga accggggtaa ggaccatcac cacgggtgc 3240
 cccatcacgt actccaccta tggcaagttt cttggcgcg gtgggtgtc tggggggcc 3300
 tatgacatca taatatgtga tgagtgcac tcaactgtact cggccactat cctgggcac 3360
 ggcacagtcg tggaccaagc ggagacggct ggagcgcgc tgcgtgtgtc cggccaccgt 3420
 acgcctccgg gatcggtcac cgtgcccacat ccaaacatcg aggagggtggc tctgtccagc 3480
 55 actggagaaa tcccccttta tggcaagcc atccccatcg agaccatcaa gggggggagg 3540

5	cacctcattt tctgccattc caagaagaaa tgtgatgagc tcgcccggaa gctgtccggc 3600 ctcgactca atgctgttagc atattaccgg ggccttgatg tatccgtcat accaactagc 3660 ggagacgtca ttgtcgtagc aacggacgt ctaatgacgg gctttaccgg cgatttcgac 3720 tcagtgatcg actgaataac atgtgtcacc cagacagtcg acttcagct ggaccggacc 3780 ttcaccattg agacgacgac cgtgccacaa gacgcgggtg caccgtcgca gccccggaggc 3840 aggactggta gggcaggat gggcattac aggtttgtg cttccaggaga acggccctcg 3900 ggcatgttcg attcctcggt tctgtcgag tgctatgacg cgggctgtgc ttgtacgag 3960 ctcacgcccggc ccgagacctc agttaggtt cgggcttacc taaacacacc agggttgccc 4020 gtctgccagg accatctggg gttctgggag ggagacaact tcccctacct gtagcatac 4140 gcccatttct tgcctccagac taagcaggca ggagacaact cccacccat cgtggacca aatgtggaa 4200 caggctacgg tgcgtcccgag ggctcaggct ccacccat cctgtgtta taggctggg 4260 tgttcatac ggctaaagcc tacgctgcac gggccaacgc ccctgtgtta taggctggg 4320 gcgttcaaa acgaggttac taccacacac cccataacca aatacatcat ggcgtgcgt 4380 tcggctgacc tggaggtcg caccggcacc tgggtgctgg tagggcgagt cctagcagct 4440 ctggcccggtt attcgtcgac aacaggcacc tggtgttacc tggcaggat catctgtcc 4500 gaaaagccgg ccatcattcc cgacaggaa gggaggttgc tgatgtggaa 4560 gagtgcgcct cacacctcc ttacatcgaa cagaaggcaa tcgggttgc gcaaacagcc gttggatcca agtggcggac ctcgaagcc actctgcctg gcaaccccgcc gatagcatca 4740 acggggatac aatatttagc aggcttgcctc tgatggcat tcacagcctc tattaccaggc aacatcctgg ggggatgggt ggcgcggccaa cttgtccctc ccagcgtgc ttctgtttc 4800 gtaggcgcgg gcatcgttgg agcggctgtt gatattttgg caggttatgg agcagggttgc agcggcgaga tgccctccac cgaggacgtg ggcgccttag tcgtcggtt cgtgtcgca gagggggctg tgcagtggat gaaccggctg tcccccacgc actatgtgcc tgagagcgc actgttacca tcaactcgt gctgaagagg agtcttacca gatcgttgcgt acggccatgtt cccgtctgtt gctaagagat gatttcaaga cctggctcca gtccaaagtc tcatgttacac gtgggttacaa gggagtctgg ccatgtggag cacatgtac aggacgttac gtaacacgtg gcatggaaaca tgcacgcctt cccggcgcc aaatttattct tacgtggagg ttacgcgggtt ggggatttc gtaaagtgc cgtgtcaggt tccggcccc ttgcacaggt acgctccacgt gtgcaaaaccc gggctcaatc aatacctgg gtgctactt ccgtgttacac cgacccttcc ctggccagg gacccggattt acaaccctcc tccttgaagg caacatgtac taccgtcat aacctctgtt ggcggcggag gttagtaattt tggactctttt cgagccgtc gttccggcggtt agatccgtcg gacccggattt acaaccctcc actgttagag gtgttacacg ggtgtccatt aagaggacgg ttgtctgtc agaatctacc aagaccttcg gcaagctccga gaccagccctt ccgacgacgg cgacgccccca ccccttggagg gggagccggg gaggaggcta gtgaggacgt cgtctgtgc atcagcccat ggcgtcggtt atcagcccat ggcgtcggtt gaaaccacag ctccgttacc acaacttgg aagggttacactt ttgacagact gcaaggcttgc atgaaggcg aggcgccac agttaaggct ctgacgcccc cacattcgcc ctatcccgca agggcgatcc ccacatccgc gagacccacca ttgacaccac catcatggca aagggggccg gcaaggccagg gagaaaaatgg cccttacga tgggtctcc tacggattcc aatacttcg tggacagccgg aagaaatgc ctatggctt aatacgatcc gtgttggagga gtaatctac caggccatata ggccatataa ggtcgctcac agagccgtt gggcagaact gggcagaact gggctatcg ccgtgtccggc
10	
15	
20	
25	
30	
35	
40	
45	
50	
55	

aataccctca catgttactt gaaggccgct gcggcctgtc gagctgcgaa gctccaggac 7560
 tgcacgatgc tcgttatgcgg agacgacctt gtcggttatct gtgaaagcgc ggggacccaa 7620
 5 gaggacgagg cgagcctacg ggccttcacg gaggctatga cttagatactc tgccccccct 7680
 ggggacccgc ccaaaccaga atacgacttg gagttgataa catcatgctc ctccaatgtg 7740
 tcagtgcgc acgatgcattt tggaaaagg gtgtactatc tcaccctgtga ccccaccacc 7800
 ccccttgcgc gggctgcgtg ggagacagct agacacactc cagtaattc ctggcttaggc 7860
 aacatcatca tttatgcgcc caccttgtgg gcaaggatga tcctgtatgc tcatttttc 7920
 tccatcttc tagctcaggaa acaacttggaa aaaggccctag attgtcagat ctacggggcc 7980
 10 ttttacttca ttgagccact tgacctacatc cagatcattc aacgacttca tggccttagc 8040
 gcattttcac tccatagtta ctctccaggat gagatcaata ggggtggcttc atgcctcagg 8100
 aaacttgggg taccgcctt gcgagtctgg agacatcggg ccagaagtgt ccgcgcctagg 8160
 ctactgtccc aggggggggag ggctgccact tttggcaagt acctcttcaa ctgggcagta 8220
 aggaccaagg taaaactcac tccaaatccccg gctgcgtccc agttggattt atccagctgg 8280
 15 ttcgttgcgt gttacagcgg gggagacata tatcacagcc tttctcgtgc ccgcacccgc 8340
 tggttcatgt ggtgcctact cctactttct tttttttttt tttttttttt ctccctttttt tttcctctttt 8400
 ccatgttgcgt ggtgcctact cctactttct tttttttttt tttttttttt ctccctttttt tttcctctttt 8460
 tttttttttt tttttttttt tttttttttt tttttttttt ctccctttttt tttcctctttt 8520
 ttttcctttt ctttcctttt gttgtccat ctttagcccta gtcacggcta gttgtgaaag 8580
 20 gttttttttt tttttttttt tttttttttt tttttttttt ctccctttttt tttcctctttt 8637

25

30

35

40

45

50

55

5 <210> 3
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

 10 <400> 3
 gccagcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgctt agccatggcg ttagtatgag tgctgtcag cctccaggac 120
 ccccccctccc gggagagcca tagttgtctg cggaaaccggt gagtacaccg gaattgccag 180
 gacgacccggg tcctttcttg gatcaacccg ctaatgcct ggagatttg gcgtcccccc 240
 15 gcgagactgc tagccagta gtgttgggtc gcgaaaggcc ttgtgttact gcctgtatagg 300
 gtgttgcga gtccccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
 ctc当地agaaa aacccaaacgt aacaccaacg ggccgcgcatt gattgaacaa gatggattgc 420
 acgcagggttc tccggccgt tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 20 caatcggctg ctctgtatcc gccgtgttcc ggctgtcagc gcaggggcgc ccgggttctt 540
 ttgtcaagac cgacctgtcc ggtccccctga atgaactgca ggacgaggca ggcggctat 600
 cgtggctggc cacgacgggc gttccttgcg cagctgtgtc cgacgttgc actgaagcgg 660
 gaagggactg gctgtattt ggcgaagtgc cggggcagga tccctgtca tctcacccctt 720
 25 ctcctggcg gaaagtatcc atcatggctg atgcaatgcg gcccgtcat acgcttgcatt 780
 cggctacccgt cccattcgac caccaggcga aacatgcgt cgacgaggca cgtactcgga 840
 tggaaaggccgg tcttgtcgat caggatgatc tggacgaaaga gcatcagggg ctcgcgcagg 900
 30 cc当地actgtt cgccaggctc aaggcgcgcg tggccgacgg cgaggatetc gtctgtgacc 960
 atggcgtatgc ctgcttgcg aatatcatgg tggaaaatgg cccgttttgc ggattcatcg 1020
 actgtggccg gctgggttg gcccggcgtt atcaggacat acgcttgcgtt acccgtata 1080
 ttgtcaaga gcttggccgc gaatggctg accgttccct cgtctttac ggtatcgccg 1140
 35 ctcccgattc gcagcgcatac gccttcttgc gccttcttgc cgagtttcc tgtagttaaa 1200
 cagaccacaa cggttccct ctacgggat caattccgc cctctccctt cccccccctt 1260
 aacgttactg gccgaaggcc ctttgaataa ggccgggtgt cgtttgtcta tatgttattt 1320
 tccaccatat tgccgtctt tggcaatgtg agggcccgga aacctggccc tgcatttctt 1380
 acgagcattc cttaggggtt ttccctctc gccaaaggaa tcaaggtctt gttaatgtc 1440
 40 gtgaaggaaag cagttccctt ggaagcttct tgaagacaaa caacgtctgt acgcaccctt 1500
 tgcaggcagc ggaacccccc acctggcgac aggtgcctt cggccaaaa gccacgtgt 1560
 taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttt gatagttgt 1620
 gaaagagtca aatggcttc ctcaagcgta tcaacaagg ggctgaagga tgcccagaag 1680
 gtacccatt gtatggatc tgatctgggg ctcgggtgc catgttttac atgtgttttag 1740
 tcgaggttaa aaaacgtctt ggcggcccgaa accacggggc cgtgggtttt ctttggaaaa 1800
 45 cagataataa ccatggccgc tattacggcc tactccaaac agacgcgagg cctacttggc 1860
 tgcatacatca ctgccttac aggccgggac aggaaccagg tcgaggggg ggtccaatgt 1920
 gtctccaccg caacacaatc ttccctggcg acctgcgtca atggcgatgtt ttggactgtc 1980
 tatacatgtt ccggctcaaa gacccttgcg gccccaaagg gccaatcac ccaaatgtac 2040
 accaatgtgg accaggaccc cgtcggtctt ccacgcggcc cccggggcgc ttcccttgaca 2100
 ccatgcacct gccggcagctt ggacctttac ttggtcacga ggcataccgc tgcattccg 2160
 gtgcgcggc gggcgacag caggggggacg ctactctccc ccaggccgtt ctccctactt 2220
 aagggtcttt cgggggttcc actgtctgc ccctcggggc acgctgtggg catcttccg 2280
 50 gctggcgtgt gcacccgggg ggttgcgaag gccgtggact ttgttccctt cagatctatg 2340
 gaaaccacta tgcgttcccc ggttttgcacg gacaactcg cccctccgc cgtaccgcag 2400
 acattccagg tggcccatct acacggccct actggtagcg gcaagagcac taaggtgcgg 2460
 gctgcgtatg cagcccaagg gtataagggtt ttgttgcgtt acccgccgtt cggccgcacc 2520
 cttaggtttcg gggcgatatac gtcttaaggca catgttacatcg acccttaatc cagaacccggg 2580
 gtaaggacca tcaccacggg tgcccccattt acgtacttca cctatggca gtttcttgc 2640
 gacgggtgtt gctctgggg cgcctatgc atcataatat gtatgtatgtt ccactcaact 2700
 gactcgacca ctatcttggg catggcaca gtccctggacc aaggcgagac ggtggagcg 2760
 55 cgactcgatcg tgctcgccac cgctacgcctt ccggggatcgg tcaccctgtcc acatccaaac 2820
 atcgaggagg tggctctgtc cagcaactggaa gaaatccctt ttatggca acccatcccc 2880
 atcgagacca tcaaggggggg gaggcaccc attttctggcc atttcaagaa gaaatgtat 2940
 gagctcgccg cgaagctgtc cggccctggaa ctcaatgtcg tagcatatta cccggggcctt 3000
 gatgtatccg tcataccaaac tagggagac gtcattgtcg tagcaacggc cgcctataatg 3060
 acgggcttta cccggcgtt cgcactcgatg atcgacttca atacatgtgtt caccctggaca 3120
 gtcgacttca gcctggacc gacccatcacc attggagacca cgaccgtgc acaagacgcg 3180
 gtgtcactcgat cgcaggccggc agggaggactt ggttagggca ggtatggcat ttacagggtt 3240
 gtgactccag gagaacggcc ctggggatcg ttgttgcgtt cgggttgcgtt cagatgtat 3300
 gacggggctt gtcgttggta cgcactcgat cccggcggaga cctcgtttag gttgcgggctt 3360
 tacctaaaca caccagggtt gcccgttgc caggaccatc tggagggttgc ggagagcg 3420
 ttacaggcc tcacccacat agacggccat ttcttgcctt agactaagca ggcaggagac 3480
 aacttccctt acctggatcg ataccagggtt acgggtgtgc ccagggtctca ggctccaccc 3540

5	ccatcgtaggg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacgggcca 3600 acgccccctgc tgtataaggct gggagccgtt caaaacgagg ttactaccac acaccccccata 3660 accaaataca tcatggcatg catgtcggtt gacctggagg tcgtcacgag cacctgggtg 3720 ctggtaggcg gagtccttagc agctctggcc cggtattgcc tgacaacagg cagcgtggc 3780 attgtggca ggatcatctt gtccggaaag ccggccatca ttcccacag ggaagtcctt 3840 taccggaggt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggaa 3900 atgcagctcg ccgaacaatt caaacagaag gcaatcggt tgctgaaac agccaccaag 3960 caagcggagg ctgctgtcc cggtgtggaa tccaagtggc ggacctcga agccttctgg 4020 gctaaggata tgfgaattt catacgccgg atacaatatt tagcaggctt gtccactctg 4080 cctggcaacc ccgcgatagc atcaactgtg gcattcacag cctctatcac cagcccgctc 4140 accacccaaat ataccctctt gttaacatc ctgggggat gggtgccgc ccaacttgtc 4200 cctcccgagcg ctgtttctgc tttctgtgc gccgcatcg ctggagcggc tttggcagc 4260 ataggccttg ggaagggtct tggtatattt ggatgggtt atggagcagg gttggcaggc 4320 ggctctgtgg ccttaaggt catgagcggc gagatccctt ccaccggaga cttgggttaac 4380 ctactccctg ctatccctc ccctggcgcctt ctatgtcgtg gggtcgtgt cgccagcgata 4440 ctgcgtcgcc acgtgggccc aggggagggg gctgtcagt ggatgaaccg gctgatagcg 4500 ttcgcttcgc gggtaacca cgttcccccc acgcactatg tgcctgagag cgacgctgca 4560 gcacgtgtca ctcagatctt ctctagtctt accatcaactc agctgtgaa gaggcttac 4620 cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg 4680 gattggatat gcacgggttt gactgattt aagacctgca tccagtcacaa gtcctgcgg 4740 cgattgcccgg gagtccctt ctctcatgt caacgtgggt acaaggaggt ctggcggggc 4800 gacggcatca tgcaaacccac ctgcccattgt ggagcacaga tcacccgaca tttggaaaac 4860 ggttccatga ggatcggtgg gcctaggacc tgtagtaaca cgtggcatgg aacattcccc 4920 attaacgcgtt acaccacggg cccctgcacg ccctcccccgg cgccaaatta ttctaggcgc 4980 ctgtggcggg tggctgtca ggatgtcg gaggtagcgc ggggtgggg tttccactac 5040 gtgacggggca tgacactgtg caacgttaaaatggg tggatgggtt ggctgtgcac aggtacgc 5100 ttcacagaag tggatgggtt ggctgtgcac aggtacgc cagctgca acccctcta 5160 cggggaggagg tcacattctt ggtcgggctc aatcaatacc tggttgggtc acagctccca 5220 tgcgagcccg aaccggacgt acaactgtc acgttccatgc tcacccgaccc cttccacatt 5280 acggcggaga cggctaagcg taggctggcc aggggatctc cccctccctt ggccagctca 5340 tcagctagcc agctgtctgc gccttcctt cggacgcgtg acggcaacat gcactaccgg tcatgactcc 5400 ccggacgcgtg acctcatcga ggcaacactc tgccgtgc cggagatggg cgggaacatc 5460 acccgcgtgg agtcagaaaa taaggttaga attttgact cttcgagcc gtcacagcg 5520 gaggaggatg agaggaaatg atccgttccg gctggagatcc tgccggaggtc caggaaattc 5580 cctcgagcga tgccatatg ggcacggcccg gattacaacc ctccactgtt agagtccctgg 5640 aaggacccgg actacgtccc tccactgtt cactgggtc cttggccgc tgccaaggcc 5700 cctccgatacc cacccacggc gggatgttgc acgggtgtcc tgcaataacttccatc 5760 tctgccttgc cggagctgc cacaaggatc tgccgtgc cggatgcgc gggctgcac 5820 agcggcacgg caacggctc ttctgaccat ccctccgacg acggcgcacgc ggatccgc 5880 gtttagtgcgt actcctccat gccccccctt gaggggagc cgggggatcc cgtatctc 5940 gacgggtctt ggtctaccgt aaggcaggag gctagtggg acgtcgctg ctgctcgatg 6000 tcctacacat ggacaggcgc cctgtatcactc ccatgcgtg cggagaaac caagctgccc 6060 atcaatgcac tgagcaactc ttgtccctgt caccacaact tggcttatgc tacaacatct 6120 cgcagcgcaaa gcctcgccca gaagaagggtc acctttgaca gactcgagg cctggacgac 6180 caactaccggg acgtgtcaaa ggagatgtaa gcaaggcgt ccacagttaa ggctaaactt 6240 ctatccgtgg aggaaggctt taaagctgacg cccccacatt cggccagatc taaatttggc 6300 tatggggcaa aggacgtccg gacactatcc agcaaggccg ttaaccacat cgcctccgt 6360 tggaaaggact tgctggaaaga cactgagaca ccaatttgaca ccacccatcat ggcaaaaat 6420 gagggtttct cgcgtccaaacc agagaagggg ggcggcaacgc cagctcgct tatcgattc 6480 ccagattttgc gggatgttgc tgccgtggaa atggccctt acgtatgtgtt ctccacccctc 6540 cctcaggccg tgatgggtc ttcacatcgaa ttcaataactt ctcctggaca ggggtcgag 6600 ttcctgtgtt atgcctggaa agcgaagaaa tgccctatgg gcttccatc tgacacccgc 6660 tggttttactt caacgtcactc tgagaatgtc atccgtgtt aggagtcaat ctaccaatgt 6720 tgtgacttgg ccccccggac cagacaggcc ataagggtcgc tcacagagcg gctttacatc 6780 ggggggccccc tgactaattt taaaaggccgg aactcggtt atcgccgggt cccgcgcgagc 6840 gtgtactga cgaccacgt cggtatacc ctcacatgtt acttgaaggc cgctcgcc 6900 tgtcgagctg cgaagctcca ggactgcacg atgctcgat gggagacgca cttgtcggtt 6960 atctgtgaaa ggcggggac ccaagaggac gaggcgagcc tacggccctt cacggaggct 7020 atgacttagat actctgtcccccc ccttggggac ccggccaaac cagaatacga ctggaggtt 7080 ataacatcatc gtcctccaa tggatgtcactc ggcacgtat catctggca aagggtgtac 7140 tatctcaccat gtgaccacccac cccccccctt gtcacatgtt atcgatgtat cggccacccatc 7200 actccagtc attcctgttggt aggaacatc atccatgtat cggccacccatc gttggggac 7260 atgatctgttgc tgactcattt ctgttccatc ctccatgttgc aggaacaact tggaaaagcc 7320 cttagattgtc agatctacgg ggctgttac tccattgtac cacttgcactt acctcagatc 7380 attcaacgc tccatggcct tagcgttactt tcactccata gttactctcc aggtgagatc 7440 aatagggtgg cttcatgtccct caggaaactt ggggttccatc ctttgcgagc ctttgcgagatc 7500
10	
15	
20	
25	
30	
35	
40	
45	
50	
55	

cgggcagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
5 tcccagttgg atttatccag ctgggtcggt gctggttaca gcgggggaga catatatcac 7680
agcctgtctc gtgcccgacc ccgcgtggtc atgtggtgc tactcctact ttctgttaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatcctgtt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860
ttttctcctt tttttttctt cttttttcc ttttcttcc tttgggtggct ccatcttagc 7920
cctagtcacg gctagctgt aaaggtccgt gagccgcttg actgcagaga gtgctgatac 7980
10 tggcctctct gcaatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 4
 <211> 7989
 <212> DNA
 5 <213> Hepatitis C Virus

 <400> 4
 gccagcccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgctgtcgag cctccaggac 120
 ccccccctccc gggagagcca tagtggctcg cggAACCGGT gagaTCACCG gaATTGCCAG 180
 10 gacgacGGGG tcctttcttg gatacAccCG ctcaatgcct ggAGATTTGG gcgtGCCCCC 240
 gcgagactgc tagccgagta gtgtgggtc gcgAAAGGCC ttgtgtact gcctgatAGG 300
 gtgttgcga gtGCCCCGGG aggtctcgta gaccgtgcac catgacacg aatcctaAac 360
 ctcAAAGAAA aaccAAAGGG cgccgcatga ttGAACAAGA tggattgcac gcaggTTCTC 420
 cgccgctt ggtggagagg ctattcggct atgactggc acaacagaca atcggctgtc 480
 15 ctgatGCCGc cgttccgg ctgtcagcgc aggggCGCC ggttctttt gtcAAGACCG 540
 acctgtccgg tgcctgaat gaaACTGCAGG acgaggcAGC gCGGCTATCG tgctggca 600
 cgacggcgt tccttgcgca gctgtgctcg acgttgcac tgaAGCGGA aggactgge 660
 tgctattggg cgaagtGCCG gggcaggatc tcctgtcatc tcaccttgc cctGCCAGA 720
 aagtatccat catggctgtat gcaatgcggc ggctgcatac gcttgcattc gctacctGCC 780
 cattcgacca ccaAGCGAAA catcgcatcg acgcgacacg tactcgatg gaAGCCGGTC 840
 20 ttgtcgatca ggatgatctg gacgaagacg atcaggggct cgcgcAGCC gaactgttcg 900
 ccaggctcaa ggccgcgtat cccgacGGCG aggatctcg ctgtacccat ggcgtatgcct 960
 gcttgcgaa tatcatggtg gaaaatggcc gcttttctgg attcatgcac tgTGGCCGGC 1020
 tgggtgtggc ggaccgtat caggacatag ctgtggctac ccgtgatatt gctgaagAGC 1080
 ttggcggcga atgggctgac cgttccctcg tgctttacgg tATCggcgtt cccgattcgc 1140
 agcgcatecgc ctttatcgat ctttcgtac agtttctcg agtttAAACA gaccacaACG 1200
 gtttccctct acggggatca attccccc tctccctccc cccccctaa cgttactggc 1260
 25 cgaagccgct tggaaataagg cccgggtgcg tttgtctata tgttatTTTc caccatattg 1320
 ccgttctttt gcaatgttag gggccggaa cctggccctg tcttcttgc gaggcattct 1380
 aggggtcttt cccctctcgc caaAGGAATG caaggctctg tgaatgtcgta gaAGGAAGCA 1440
 gttcctctgg aagttcttg aagacAAACA acgtctgtac cgtaccctt caggcagcgg 1500
 aaccccccac ctggcgcacag gtgcctctgc ggccAAAGC cacgtgtata agatacacct 1560
 gcaaaggcgg cacaACCCCA gtccacgtt gtgagttgaa tagttgtgaa aagagtcaaa 1620
 30 tggctctcct caagcgtatt caacaAGGGG ctgaaggatg cccagaAGGT accccattgt 1680
 atgggatctg atctggggcc tcgggtcaca tgctttacat gtgttagtc gaggttaaaa 1740
 aacgtctagg ccccccgaac caccgggacg tggttttctt ttgaaaaaca cgataatacc 1800
 atggcgccta ttacggccta ctcccaacag acgcgaggcc tactgggtg catcatca 1860
 agcctcacag gcccggacag gaaccagggtc gaggGGGGGG tccaatgtgt ctccaccgc 1920
 acacaatctt tcttggcgc tctgtcaat ggcgtgtgtt ggactgtcta tcatggtgcc 1980
 35 ggctcaaaga cccttgcgg cccaaAGGGC ccaatCACCC aaatgtacac caatgtggac 2040
 caggacctcg tcggctggca agcGCCCCC ggggcgcgtt ctttgcacacc atgcacctgc 2100
 ggcagctcg acctttactt ggtcacgagg catgcccgtt tcattccgtt ggcggccgg 2160
 ggcgacagca gggggagcct actctcccc aggcccgtt cctacttggaa gggcttctcg 2220
 ggcggtccac tgctctgccc ctggggcac gctgtggca tctttgggc tgccgtgtgc 2280
 40 acccgaggGGG ttgcgaaggc ggtggacttt gtaccgtcg agtctatgga aaccactatg 2340
 cggccccccg tcttcacgca caactcgtcc cctccggccg taccgcagac attccagggt 2400
 gcccacatc acgccccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 2460
 gcccaagggt ataagggtct tgctctgaaac ccgtccgtcg ccgcacccct aggtttcggt 2520
 ggttatatgt ctaaggcaca tggtatcgac cctaacatca gaACCCGGGT aaggaccatc 2580
 accacgggtg cccccatcac gtactccacc tatggcaagt ttcttgcga cgtgggttgc 2640
 45 tetggggcgc cctatgacat cataatatgt gatgagtgc actcaactga ctgcaccact 2700
 atctggggca tcggcacagt ctgggaccaa gcccggacgg ctggagcggc atctgtcg 2760
 ctcgcaccccg ctacgcctcc gggatcggtc accgtccac atccaaacat cgaggagggt 2820
 gctctgtcca gcaactggaga aatccccctt tatggcaag ccattcccat cgagaccatc 2880
 aaggGGGGGG ggcacctcat ttctgtccat tccaagaaga aatgtatgaa gctcggccgc 2940
 aagctgtccg gcctcgact caatgtgtc gcatattacc gggcccttga tttatccgtc 3000
 50 ataccaacta gcccggacgt cattgtcgta gcaacggacg ctctaatgac gggctttacc 3060
 ggcgatttcg actcagtgtat cgactgtca acatgtgtca cccagacagt cgacttcgc 3120
 ctggacccga ctttcacccat tgagacgacg accgtccac aagacgcggt gtcacgcgtc 3180
 cagcggcgcag gcaggactgg tagggcagg atggcattt acaggtttgt gactccaggaa 3240
 gaacggccct cgggcatgtt cgattcctcg gttctgtcg agtgcata ggcgggctgt 3300
 gcttggtagc agctcacgca cgcggagacc tcagtttgt tgccggctta cttAAACACA 3360
 ccagggttgc cctgtccca ggaccatctg gagttctggg agagcgtctt tacaggcctc 3420
 55 accccacatag acgccccattt ctgtccca gactaaggcagg caggagacaa cttccccctac 3480
 ctggtagcat accaggctac ggtgtgcgc agggtcagg ctccacccatc atcgtggac 3540

	ccaatgtgga	agtgtctcat	acggctaaag	cctacgctgc	acggggcaac	gccctgtcgt	3600
5	tataggcgtgg	gagccgttca	aaacgagggt	actaccacac	accccataac	caaatacatc	3660
	atggcatgca	tgtcggtca	cctggaggtc	gtcacagaca	cctgggtgt	ggtaggcgga	3720
	gtccttagcag	ctctggccgc	gtattgcctg	acaacaggca	gcgtggtcat	tgtgggcagg	3780
	atcatcttgt	ccggaaagcc	ggccatcatt	cccgcacagg	aagtcttta	ccgggagttc	3840
10	gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacaggaaat	gcagctcgcc	3900
	gaacaattca	aacagaaggc	aatcgggtt	ctgcaaacag	ccaccaagca	acgggaggct	3960
	gctgctcccg	tgggtgaaatc	caagtggcg	accctcgaa	ccttctggc	gaagcatatg	4020
	tgaatttca	tcagcggat	acaatatta	gcaggcttg	ccactctgc	tgcacaaaa	4080
15	gcgatagcat	cactgtatgc	attcacagcc	tctatccaca	gcccgcac	cacccaacat	4140
	accctccctgt	ttaacatctt	ggggggatgg	gtggccccc	aacttgcet	tccacgcgt	4200
	gcttctgtt	tgcgtggc	ccgcattcg	ggagcgct	ttggcagcat	aggccttgg	4260
	aagggtgtt	tggatattt	ggcaggttat	ggagcagg	tggcaggc	gctgtggc	4320
	ttaaggtca	tgagcggcga	gatgcctcc	accgaggacc	tggtaac	actccctgct	4380
20	atccctctccc	ctggcgcct	agtgcgtgg	gtcgtgtgc	cagcgatact	gcgtcgac	4440
	gtgggccccag	gggagggggc	tgtcagtgg	atgaaccgc	tgatagcgtt	cgcttcg	4500
	ggtaaccacg	tctcccccac	gcactatgt	cctgagagc	acgctcg	acgtgtact	4560
	cagatcctct	ctagtcttac	catcaactcg	ctgctgaaga	ggcttcacca	ttggatcaac	4620
25	gaggactgt	ccacgcct	ctccggctcg	tggctaagag	atgtttgg	ttggatatgc	4680
	acgggtgttga	ctgatttcaa	gacctggctc	cagtccaa	tcctggc	attgcgg	4740
	gtcccttct	tctcatgtca	acgtgggtac	aaggagtc	ggcggggc	cggcatatg	4800
	caaaccacct	gcccatgtt	agcacagatc	accggacat	taaaaacgg	ttccatgagg	4860
30	atcggtgggc	ctagggact	tagtacacg	tggcatgg	cattccccat	taacgcgt	4920
	accacggggcc	cctgcaccc	ctccccggc	ccaaattatt	ctagggcg	ttgggggt	4980
	gtgtgttgg	agtacgtt	ggttacgg	gtggggatt	tccactac	gacgggc	5040
	accactgaca	acgtaaatg	ccctgtc	tttgcggcc	ccgaatttctt	cacaagatg	5100
35	gtgggggtgc	ggttgcacag	gtacgctca	gcgtcaaa	ccctcttac	ggaggaggtc	5160
	acattcttgg	tcgggctcaa	tcaatacct	gttgggtc	agctccat	cgagccgaa	5220
	ccggacgt	cagtgc	ttccatgtc	accgaccct	cccacattac	ggcggagacg	5280
	gctaagcgta	ggctggcc	ggatctcc	ccctcctt	ccagctc	actagccag	5340
40	ctgtctgc	cttcctgaa	ggcaacatgc	actacccgt	atgactcccc	ggacgctgac	5400
	ctcatcgagg	ccaaacctt	gtggcggc	gagatgg	ggaacatc	ccgcgtgg	5460
	tcagaaaata	aggtgtaat	tttgactt	ttcgagcc	tccaagc	ggaggatg	5520
	agggaaat	ccgttccgg	ggagatct	cggaggtc	ggaaatttcc	tcgagcgat	5580
45	cccatatgg	cacggccg	ttacaacc	ccactgtt	agtcctgg	ggacccgg	5640
	tacgccctc	cagtgtt	cgggtgtt	ttgcgcct	ccaaaggccc	tccgatacc	5700
	cctccacgg	ggaagagg	ggttgc	tcagaatct	ccgtgtt	tcccttgg	5760
	gagctcgca	caaagac	ccgcgtc	gaatctcg	ccgtc	cgcacgg	5820
50	acggcctctc	ctgaccag	ctccgac	ggcgcacgg	gatccgac	tgatcg	5880
	tcctccatgc	ccccctt	ggggagcc	ggggatccc	atctcag	cggtctt	5940
	tctaccgtt	gcfaggagg	tagtgg	gtcgtct	gtcgtat	ctacacat	6000
	acaggcgcc	tgtacac	atgcgt	gaggaaac	agctccat	caatgcact	6060
55	agcaactt	tgctcc	ccacaactt	gtctatgt	caacatct	cagcga	6120
	ctgcggcaga	agaagg	cttgc	ctgcagg	tggac	ctacggg	6180
	gtgtctcaagg	agatgaa	gaaggcgt	acagtt	ctaaactt	atccgtgg	6240
	gaagcctgt	agctgac	cccacatt	gccagat	aatttgc	tggggcaaa	6300
	gacgtccgg	acctatcc	caaggcgtt	aaccat	gtccctgt	gaaggactt	6360
	ctggaaagca	ctgagac	aattgac	accat	caaaaaatg	gttttct	6420
	gttcaacc	agaagg	ccgca	gtcgc	tcttccc	agatttgg	6480
	gttcgtgtt	gcfggaaa	ggccctt	gtgtgtt	ccacccccc	teaggccgt	6540
	atgggctt	catacg	ccatact	cttgc	gggtc	aggctt	6600
	gcctggaaatg	cgaagaaat	ccctatgg	ttcgc	atatg	acacccgt	6660
	acggtact	agaatgac	ccgtgtt	gagtca	accaatgtt	tgacttgg	6720
	cccgaa	gacagg	ccat	acagag	tttacat	gggccccct	6780
	actaattct	aagg	cgct	ccggcgt	cgcg	agcg	6840
	accagctcg	gtataacc	cacatgtt	ttgaagg	ctgcgg	ctcgac	6900
	aagctccagg	actgcac	gtctgtat	ggagac	ttgtc	ttatctgt	6960
	gcggggacc	aagg	gacg	ccggc	tttgc	gttgcgt	7020
	tctggccccc	ctggg	gaccc	ggagg	tttgc	gttgcgt	7080
	tcctccaaat	tgtc	gtcgt	gat	ccaccc	tttgc	7140
	gaccggcc	cccc	tttgc	gggg	tttgc	gttgcgt	7200
	tcctggctag	gcaacat	catgtat	gggg	tttgc	gttgcgt	7260
	actcatttct	tctccat	tctagct	gggg	tttgc	gttgcgt	7320
	atctacgggg	cctgttact	cattgag	gggg	tttgc	gttgcgt	7380
	catggcctt	gcfgat	tttgc	tttgc	tttgc	gttgcgt	7440
	tcatgcct	ggaaactt	tttgc	tttgc	tttgc	gttgcgt	7500

gtccgcgcta	ggctactgtc	ccaggggggg	agggctgcca	cttggca	gtaccttc	7560
aactggcag	taaggacca	gctaaactc	actccaatcc	cggctgcgtc	ccagttggat	7620
ttatccagct	ggttcggtc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	7680
gcccccccc	gctggttcat	gtgggccta	ctcctacttt	ctgttaggggt	aggcatatat	7740
ctactcccc	accgatgaac	ggggagctaa	acactccagg	ccaataggcc	atccctgtttt	7800
tttccctttt	tttttttctt	ttttttttt	ttttttttt	ttttttttt	ttctcctttt	7860
tttttcctct	tttttcctt	ttcttcctt	tggtggctcc	atcttagccc	tagtcacggc	7920
tagctgtgaa	aggccgtga	gccgcgtac	tgcagagagt	gctgatactg	gcctctctgc	7980
agatcaagt						7989

15

20

25

30

35

40

45

50

55

5 <210> 5
 <211> 8649
 <212> DNA
 5 <213> Hepatitis C Virus
 <400> 5
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tttcacgca gaaagcgct agccatggcg ttagtatgag tgctgtgcag ctcaggac 120
 ccccccctccc gggagagcca tagtggctg cggaccggt gactacccg gaattgccag 180
 10 gacgaccggg tcctttctt gataaccccg ctcaatgcct ggagatttg gcgtgcccc 240
 gcgagactgc tagccgagta gtgtgggtc gcgaaaggcc ttgtgtact gcctgatagg 300
 gtgttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcag aatccctaaac 360
 ctcaaagaaa aaccaaacgt aacaccaacg ggcgcggcat gatggacaa gatggattgc 420
 acgcagggtc tccggccgt tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 15 caatcgctg ctctgatgcc gccgttcc ggctgtcagc gcaggggcgc cccgttctt 540
 ttgtcaagac cgacctgtcc ggtgccttga atgaactgca ggacgaggca ggcggctat 600
 ctgttgcggc cacgacgggc gttcttgcg cagctgtgtc cgacgttgc actgaagcgg 660
 gaaggactg gctgtattt ggcgaaagtgc cggggcagga ttcctgtca ttcacccctt 720
 ctccgtccga gaaagtatcc atcatggctg atgaatgcg gcggctgcat acgcttgc 780
 cgcttacctg cccattcgac caccacgcg aacatcgcat cgagcagca cgtactcgga 840
 tggaaaggccg tcttgcgtat caggatgatc tggacgaa gcatcagggg ctgcgcggc 900
 20 ccgaactgtt cgccaggctc aaggcgcgca tggccgcacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgttgcgg aatatcatgg tggaaaaatgg cgcgtttctt ggattcatcg 1020
 actgtggccg gctgggtgtc gcggaccgt atcaggacat agcgttgc acccggtata 1080
 ttgtcaaga gcttggccgc gaatgggtct accgttctt cgtgtttac ggtatcgcc 1140
 25 ctccccgattc gcagcgcata gccttctatc gccttcttga cgagtcttc ttagttaaa 1200
 cagaccacaa cgggttccct ctgcgggat caattccggc cctccccc ccccccctt 1260
 aacgttactg gccgaaggcg ctttgaataa ggccgggtgt cgtttgtata tatgttattt 1320
 tccaccat tgcgttctt tggcaatgtg agggccggg aacctggccc tgcgttctt 1380
 acgacattt ctaggggtt ttccctctc gccaaggaa tgcaaggatc gttagatgtc 1440
 gtaaggaag cagttctt ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcagc ggaacccccc acctggcgac aggtgcctt gcggccaaaa gcccgtgt 1560
 30 taagatacac ctgcaaaaggc ggacacaaccc cagtgccacg ttgtgagttt gatagttgt 1620
 gaaagagtca aatggcttc ctcaagcgta ttcaacaagg ggctgaagga tggccagaag 1680
 gtacccatt gtatggatc tgatctgggg cctcggtgca catgtttac atgtgtttag 1740
 tcgaggttaa aaaacgtcta ggccccccga accacggggc cgtgttttc ctttgaaaaa 1800
 cacgataata ccatggaccg ggagatggca gcatcggtc gaggccggc tttcgtaggt 1860
 ctgtactct tgccttgc accgcactat aagctgttcc tgcgttgcgcatatgggt 1920
 ttacaatatt ttatcaccat ggcggaggca cacttgcacg tggatccc cccctcaac 1980
 35 gttcggggg gcccgatgc cgtcatctc ctgcgtgca cgatccccc agagtaatc 2040
 ttaccatca caaaaatctt gctgcctata ctgcgttccat tcatgttgc ccaggctgg 2100
 ataaccaaaatg tgcgtactt cgtgcgcgca cagggtcttca atggctctca tgaagttggc cgcactgaca 2160
 cggaaagggtt ctgggggtca ttatgtccaa atggctctca tggatccc cccctcaac 2220
 ggtacgtacg ttatgacca ttcacccca ctgcggact gggccacgc gggctacga 2280
 gaccttgcgg tggcaggta gcccgtgca ttctctgata tggagaccaa gttatcacc 2340
 40 tggggggcag acaccggcgt gttggggac atcatcttg gcctgcccgt ctccggccgc 2400
 agggggaggg agatacatct gggaccggc gacagcctt aagggcagggt gtggcgactc 2460
 ctgcgcctta ttacggctt ctccaaacag acgcgggccc tacttggctg catcatact 2520
 agcctcacag gcccggacag gaaccaggc gaggggagg tccaaatgtt ctccaccgca 2580
 acacaatctt tccggccgac ctgcgtcaat ggcgtgttt ggactgtcta tcatgttgc 2640
 ggctcaaaga cccttgcgg cccaaaggcc ccaatcaccc aaatgtacac caatgtggac 2700
 caggacctcg tcggctggca agcggccccc gggcgccgtt cttgcacacc atgcaccc 2760
 ggcagctcg gaccttactt ggtcacgagg catggcgatg tcattccggc gcgcggccgg 2820
 ggcgacagca gggggaggtt actctccccc aggccgtct cctacttggaa gggcttctcg 2880
 ggcggccatc tgctctggccctt ctcggggac gctgtggca tcttcgggc tgcgtgtgc 2940
 acccgagggtt ttgcaggac ggttggactcg atgtatgg aaccactatg 3000
 45 cygtccccgg tcttcacggca caactcgatc ctcggccggc tacccggcgtt attccagggt 3060
 gcccacatctac acggccctac tggtagcgcc aagagcacta aggtgcggc tgcgtatgc 3120
 gcccaagggtt ataagggtct tggctctgaa cctgcgtgca cccgcacccctt aggtttcg 3180
 gctgtatgtt ctaaggcaca tggatcgac cctaaacatca gaaagggtt aaggaccatc 3240
 accacgggtt ccccccacatctac gtactccacc tatggcaatg ttcttcggca cgggtgggtgc 3300
 tctggggccg cctatgacat cataatatgt gatgagtgc actcaactga ctcgaccact 3360
 atccctggccatc tcggcacatc ctcggacaa gcccggacgg ctggagcgcg actcgctgt 3420
 50 ctcggccacccg ctacgcctcc gggatcggtc accgtgcaccc atccaaacat cgaggagggt 3480
 gctctgtcca gcaactggaga aatccctt tatggcaaaag ccatccccat cgagaccatc 3540

	aaggggggga	ggcacctcat	tttctgccc	tccaagaaga	aatgtatgt	gctcgccgcg	3600
5	aagctgtccg	gcctcgact	caatgtgt	gcatattacc	ggggcctt	tgtatccgtc	3660
	ataccaacta	gcccggacgt	cattgtgt	gcaacggac	ctctaatt	gggtttacc	3720
	ggcgatttcg	actca	gtgt	acatgtgt	ccca	gacagt	3780
	ctgacc	ccat	tgagac	gacg	accgt	gcccgt	3840
	cagcggcgag	gcaggact	gg	atgggcattt	acagg	ttgt	3900
	gaacggccct	cggc	catgtt	cgatt	cctcg	gtgt	3960
	gcttggtagc	agc	tac	cgcc	gagacc	tcagg	4020
	ccagggttgc	cgt	tgc	ggaccat	tcgtt	tgccc	4080
10	accacatag	acgccc	attt	cttgc	acta	aggaca	4140
	ctgttagcat	accagg	ctac	gtgt	cgcc	aggg	4200
	caa	atgtgg	atgt	ctcat	acgg	cctgc	4260
	tataaggctgg	gagc	cggttca	aaac	gagg	actacc	4320
	atggcatgca	tgtcg	ggct	cctgg	gagg	gtc	4380
	gtcctagcag	ctctgg	ccgc	gtatt	gc	gtgt	4440
15	atcatcttgt	ccgg	aaag	ggccat	ccc	gacagg	4500
	gatgagatgg	aag	agt	tcac	cctt	acat	4560
	gaacaattca	aac	agaagg	atc	gggtt	ctgaaac	4620
	gctgctcccg	tgg	ttaa	caat	ttt	ggtt	4680
	tggaaattca	tcag	cggtt	acaat	at	tttt	4740
	gcatagcat	cact	gtat	tttt	tttt	tttt	4800
20	accctcttgt	ttaa	cat	tttt	tttt	tttt	4860
	gcttctgttt	tgc	tag	tcg	ggct	tttgc	4920
	aagg	gtt	tttt	ggc	agg	tttt	4980
	tttaagg	tg	acg	cc	agg	tttt	5040
	atcc	tctccc	ctgg	ccct	atgt	tttt	5100
	gtgggccc	gg	gggg	tttt	tttt	tttt	5160
25	ggtaaccacg	tct	cccc	ccat	gcac	tttt	5220
	cagatcctct	ctag	tct	cat	act	tttt	5280
	gaggactgct	ccac	ccat	ctt	ctt	tttt	5340
	acgtgttga	ctg	attt	caat	gggtt	tttt	5400
	gtcc	ccat	tc	acgt	gggtt	tttt	5460
	caa	acc	ct	tttt	tttt	tttt	5520
30	atcg	gggtt	ccat	tttt	tttt	tttt	5580
	accacggg	cct	gac	tttt	tttt	tttt	5640
	gct	ctt	gggg	tttt	tttt	tttt	5700
	accactgaca	ac	gtt	tttt	tttt	tttt	5760
	gat	gggtt	tttt	tttt	tttt	tttt	5820
	acatt	tttt	tttt	tttt	tttt	tttt	5880
35	ccgg	act	tttt	tttt	tttt	tttt	5940
	gcta	gg	tttt	tttt	tttt	tttt	6000
	ctgt	cc	tttt	tttt	tttt	tttt	6060
	ctc	at	tttt	tttt	tttt	tttt	6120
	tca	aaa	tttt	tttt	tttt	tttt	6180
	agg	taa	tttt	tttt	tttt	tttt	6240
	ccgt	ttt	tttt	tttt	tttt	tttt	6300
40	ccca	at	tttt	tttt	tttt	tttt	6360
	tac	gtt	tttt	tttt	tttt	tttt	6420
	cct	cc	tttt	tttt	tttt	tttt	6480
	gag	ctt	tttt	tttt	tttt	tttt	6540
	acgg	cc	tttt	tttt	tttt	tttt	6600
45	tctt	ccat	tttt	tttt	tttt	tttt	6660
	tctt	ccat	tttt	tttt	tttt	tttt	6720
	acagg	cc	tttt	tttt	tttt	tttt	6780
	aga	actt	tttt	tttt	tttt	tttt	6840
	ctg	cc	tttt	tttt	tttt	tttt	6900
	gt	cc	tttt	tttt	tttt	tttt	6960
	gaag	ctgt	tttt	tttt	tttt	tttt	7020
50	gac	gtt	tttt	tttt	tttt	tttt	7080
	ctg	gg	tttt	tttt	tttt	tttt	7140
	gtt	cc	tttt	tttt	tttt	tttt	7200
	atgg	ttt	tttt	tttt	tttt	tttt	7260
	gc	cc	tttt	tttt	tttt	tttt	7320
	acgg	ttt	tttt	tttt	tttt	tttt	7380
55	ccca	gg	tttt	tttt	tttt	tttt	7440
	acta	ttt	tttt	tttt	tttt	tttt	7500

accagctgcg gtaataccct cacatgttac ttgaaggccg ctgcggcctg tcgagctg 7560
 aagctccagg actgcacgat gctcgatgc ggagacgacc ttgtcggtat ctgtgaaagc 7620
 gcggggaccc aagaggacga ggcgagccta cgggccttca cgaggactat gactagatac 7680
 tctgcccccc ctggggaccc gcccaaacca gaatacgact tggagtttat aacatcatgc 7740
 tcctccaatg tgtcagtgc gcacgatgca tctggcaaaa gggtgtacta tctcaccctg 7800
 gaccccacca ccccccattgc gcgggctgcg tgggagacag cttagacacac tccagtcaat 7860
 tcctggctag gcaacatcat catgtatgcg cccaccttgt gggcaaggat gatcctgtatg 7920
 actcatttct tctccatcct tctagcttag gaacaacttg aaaaagccct agattgtcag 7980
 atctacgggg cctgttactc cattgagcca cttgacctac ctcagatcat tcaacgactc 8040
 catggcctta gogcattttc actccatagt tactctccag gtgagatcaa tagggtggct 8100
 tcatgcctca gaaaaacttg ggtaccgccc ttgcgagttt ggagacatcg ggccagaagt 8160
 gtcccgctca ggctactgtc ccaggggggg agggctgcca cttgtggcaa gtaccttcc 8220
 aactggcag taaggaccaa gctcaaactc actccaaatcc cggctgcgtc ccagttggat 8280
 ttatccagct ggtcgttgc tggttacagc gggggagaca tatatcacag cctgtctcg 8340
 gcccgaaccc gctggttcat gtggtgccta ctcctacttt ctgtaggggt aggcatctat 8400
 ctactccccca accgatgaac ggggagctaa acactccagg ccaataggcc atcctgtttt 8460
 tttccctttt tttttttctt tttttttttt tttttttttt tttttttttt ttctcctttt 8520
 tttttcctct ttttttcctt ttcttcctt tggtgctcc atcttagccc tagtcacggc 8580
 tagctgtgaa aggtccgtga gccgcttgac tgcaagagat gctgatactg gctctctgc 8640
 agatcaagt 8649

25

30

35

40

45

50

55

<210> 6
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

<400> 6

```

gccaactgtt cggcaggctc aaggcgcgca tggccgcacgg cgaggatctc gtcgtgacc 960
atggcgatgc ctgtttcccg aatatcatgg tggaaaatgg ccgttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcccggcgt atcaggacat agcgtggct acccggtata 1080
ttgctgaaga gcttggccgc gaatgggtgt accgcttctc cgtctttac ggtatcgccg 1140
ctccccgattc gcagcgcatac gccttctatac gccttcttgc cgagtcttc tgagttaaa 1200
cagaccacaa cggtttccct cttagggat caattccggcc cctctccct cccccccct 1260
aacgttactg gcccgaaggcg cttgaaataa ggccgggtgt cgtttgcata tatgttattt 1320
tccaccatata tggcgttcc ttggaaatgtg agggccggaa aacctggccct tgcttcttgc 1380
acgacgatcc cttaggggtt tttccctc gccaaaggaa tgcaaggatct gttagatgtc 1440
gtgaaggaag cagtttccct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcgcg ggaacccccc acctggcgac aggtgcctt gcccggccaa gccacgtgt 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgcacgc ttgtgagttt gatagttgt 1620
gaaagagtca aatggcttcc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatggatc tgatctgggg cctcggttca catgctttac atgtgttttag 1740
tcgaggttaa aaaacgtcta ggccccccga accacggggc cgtgggtttc ctttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
tgcattcatca cttagcctcac aggccgggac aggaaccagg tcgagggggc ggtccaagt 1920
gtctccacccg caacacaatc tttccctggcg acctcggttca atggcggtgt ttggactgtc 1980
tatcatgttg cccgctaaa gacccttgcg ggcccaaagg gcccataatcac cccaaatgtac 2040
accaatgtgg accaggacct cgtcggttgcgaa gggccggccg ttcccttgaca 2100
ccatgcaccc tggccggcgtc ggaccctttac ttggtcaacgc ggcatggccg tgcattccg 2160
gtgcggccggc gggggcggacag cagggggagc ctactctccc ccagggccgt ctccctacttgc 2220
aagggtctt cggccgggtcc actgtctgc ccctcgggc atgtgttggg catctttcg 2280
gctggcggtgt gcacccggg ggttgcgaag gccgggtggact ttgttccctt ctagtctatg 2340
gaaaccacta tgcggtcccc ggttccacgc gacaactctgt cccctccggc cgtaccgcag 2400
acattccagg tggcccatct acacgcccct actggtagcg gcaagagac taaggtgccc 2460
gctgcgtatg cagcccaagg gtataagggtt ctgttgcata acccgccgt cggccggccacc 2520
ctaggtttcg gggcgatata tgcataaggca catggatctcg accctaaatcat cagaaccggg 2580
gtagggacca tcaccacggg tgccccatc acgtactcca cctatggca gtttcttgc 2640
gacgggtgtt gctctggggc cgccatgtac atcataatat gtatgtgtt ccactcaact 2700
gactcgacca ctatctggg categgcaca gtcctggacc aagcgagac ggttggagcg 2760
cgactcgatcg tgctcgccac cgctacgcctt ccggatcgcc tcaccgttgc acatccaaac 2820
atcgaggagg tggctctgtc cagactcgatc gaaatccctt tttatggca gaccatcccc 2880
atcgcgacca tcaagggggg gaggcacctt attttctgcc atttcaagaa gaaatgtgt 2940
gagctcgcccg cgaagcttac cggcctcgga ctcaatgtcg tagcatatta ccggggccctt 3000
gatgtatccg tcataccaac tagggggac gtcattgtcg tagcaacggc cgctctaatg 3060
acggggctta cggcgattt cgactcgatg atcgacttgc atacatgtgt cacccagaca 3120
gtcgacttca gcctggaccc gacccatcacc attgagacca cgaccgttgc acaagacgcg 3180
gtgtcacgtt cgcagccggc aggaggact ggttagggca ggtatggcat ttacagggtt 3240
gtgactccag gagaacggcc ctggggcatg ttgcattctt cgggttctgt cgagtgtat 3300
gacgggggtt gtgcgttgc ctagactcgatc cccggggaga cctcagtttgc gtttggggct 3360
tacctaaaca caccaggctt gcccgttgc caggaccatc tggagggttgc ggagagcg 3420
tttacaggcc tcacccatcat agacgcccattt ttcttgc tccatggca ggcaggagac 3480
aacttccctt acctggtagc ataccaggctt acgggttgcg ccagggttca ggctccaccc 3540

```

	ccatcggtggg	accaaatgtg	ggagtgtctc	atacggctaa	agcctacgct	gcacgggcca	3600
5	acccccctgc	tgtataggct	gggagccgtt	caaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcggt	gacctggagg	tgttcacgag	cacctgggtg	3720
	ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgccc	tgacaacagg	cagcgtggc	3780
	atgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccgacag	ggaagtccctt	3840
	taccggaggt	tcatgtggat	ggaagagtgc	gcctcacacc	tcccttacat	cgaacaggga	3900
10	atgcagctcg	ccgaacaatt	caaacagaag	gcaatcggtt	tgctcaaacc	agccaccaag	3960
	caagcggagg	ctgctgtcc	cgtgtggaa	tccaaagtggc	ggaccctcga	agcctctgg	4020
	gcgaagcata	tgtgaaattt	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
	cctggcaacc	ccgcgatagc	atcaactgt	gcattcacag	cctctatcac	cagccgcctc	4140
15	accacccaac	ataccctctt	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgtt	4200
	cctcccagcg	ctgcttctgc	ttcgttaggc	ccggcgcac	ctggagcggc	tgttggcagc	4260
	ataggccttg	ggaagggtct	tgtgtatatt	ttggcagggtt	atggagcagg	gttggcaggc	4320
	gcgctgtgg	cctttaaggt	catgagcggc	gagatgcctt	ccaccgagga	cctgggttac	4380
20	ctactccctg	ctatccttc	ccctggcgcc	ctagtcgtc	gggtcggtt	cgacgcgtat	4440
	ctgcgtggc	acgtggggcc	aggggagggg	gtgtgcagt	ggatgaacc	gctgtatagcg	4500
	ttocttcgtc	gggtaacca	cgtctcccc	acgcactatg	tgtctgagag	cgacgcgtca	4560
	gcacgtgtca	ctcagatct	ctctagtctt	accatcaact	agctgtgaa	gaggcttcac	4620
	cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggtaag	agatgtttgg	4680
25	gattggatat	gcacgggtt	gactgatttc	aagacctggc	tccagtccaa	gctcctgccc	4740
	cgattggcg	gagtccctt	cttctcatgt	caacgtgggt	acaaggaggt	ctggggggc	4800
	gacggcatca	tgcaaaccac	ctgcccattgt	ggagcacaga	tcacccgaca	tgtaaaaaac	4860
	ggttccatga	ggatcggtgg	gcctaggacc	tgttagtaaca	cgtggcatgg	aacattcccc	4920
30	attaacgcgt	acaccacgg	ccccgtac	ccctccccc	cgccaaattt	tcttagggcg	4980
	ctgtggcggg	tggctgtca	ggagtagcgt	gaggtacgc	gggtggggga	tttccactac	5040
	gtgacgggca	tgaccactga	caacgtaaag	tgcccgtgt	aggttccgc	ccccgaattt	5100
	ttcacagaag	tggatgggt	gccccgtcac	aggtacgc	cagcgtgaa	accctctcta	5160
35	cgggaggagg	tcacattctt	ggtcgggctc	atcaataacc	tgttgggtc	acagctccca	5220
	tgcgggcccc	aaccggacgt	agcagtgc	acttccatgc	tcacccgacc	ctccacattt	5280
	acgggggaga	cggctaagcg	taggtggcc	aggggatcte	cccccttctt	ggccagctca	5340
	tcagctagcc	agctgtctgc	gccttcctt	aaggcaacat	gcactaccc	tcatgactcc	5400
40	ccggacgtcg	acccatcgta	ggccaaaccc	ctgtggcgcc	aggagatggg	cggaacatc	5460
	acccgcgtgg	agtcagaaaa	taaggttagt	atttggact	cttcgagcc	gctccaagcg	5520
	gaggaggatg	agagggaaat	atccgttccg	gccccggatcc	tgccggaggtc	caggaaattt	5580
	cctcgagcga	tgcccatatg	ggcacgccc	gattacaacc	ctccactgtt	agagtccctgg	5640
45	aaggaccgg	actacgtccc	tccagtggta	cacgggtgtc	cattggcc	tgccaaggcc	5700
	cctccgatac	cacctccac	gagggagagg	acgggtgtcc	tgtcagaatc	taccgtgtct	5760
	tctgccttgg	cgagacgtcg	cacaaagacc	ttcggcagat	cegaatcg	ggccgtcgac	5820
	agcggcacgg	caacggcctc	tcctgaccag	ccctccgac	acggcgcac	ggatccgac	5880
50	gttgagtctgt	actcctccat	gccccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
	gacgggtctt	ggtctaccgt	aagcgaggag	gtctgttgg	acgtcgctg	ctgctcgatg	6000
	tcc tacatcat	ggacaggcgc	cctgtatc	ccatgcgt	cgagggaaac	caagctgccc	6060
	gtcaatgcac	tgagcaactc	tttgcctcgt	caccacaact	tggtctatgc	tacaacatct	6120
	cgcagcgcaa	gcctgcggca	gaagaagg	acctttgaca	gactcgagg	cttggacgac	6180
55	cactaccggg	acgtgtctca	ggagatgaag	gcaaggcgt	ccacagttaa	ggctaaaactt	6240
	ctatccgtgg	aggaacgcgt	taagctgac	ccccccat	ggccagatc	taattttggc	6300
	tatggggcaa	aggacgtccg	gaaacctatcc	agcaaggcc	ttaaccat	ccgtccctgt	6360
	tggaaggact	tgctggaaaga	caactgagaca	ccaattgaca	ccaccatcat	ggaaaaaaat	6420
	gaggttttct	gcgtccaaacc	agagaagggg	ggccgca	cagctcgct	tatcgattt	6480
	ccagattttgg	gggttcgtgt	gtgcgagaaa	atggccctt	acgatgtgtt	ctccaccctc	6540
	cctcaggccg	tgatggctc	ttcatacgg	ttcaataact	ctccctggaca	gggggtcgag	6600
	ttccctggta	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacaccgc	6660
55	tgttttact	caacggtcac	tgagaatgt	atccgtgtt	aggagtcaat	ctaccaatgt	6720
	tgtgacttgg	cccccgaaac	cagacaggcc	ataaggtc	tcacagagcg	gctttacatc	6780
	ggggggcccc	tgactaattt	taaaggccag	aactcgccgt	atcgccgtt	ccgcgcgagc	6840
	ggtgtactga	cgaccagctg	cgtaatacc	ctcacatgtt	atttgaaggc	cgctcgccg	6900
	tgtcgagctg	cgaagcttca	ggactgcac	atgctcgat	gcccggacga	ccttgcctg	6960
	atctgtgaaa	gcgcggggac	ccaagggac	gaggcgcac	tacggcc	cacggaggct	7020
	atgactagat	actctgtccc	ccctggggac	ccggccaaac	cagaatacga	cttggagttt	7080
	ataacatcat	gtcttccaa	tgttgcgt	gcccacat	catctggca	aagggtgtac	7140
	tatctcaccc	gtgacccac	cacccccc	gcgcgggc	cgtggagac	agctagacac	7200
	actccagtca	atccctggct	aggaacatc	atcatgtat	cgcccac	gttggcaagg	7260
	atgatccctga	tgactcattt	cttctccatc	ttcttagctc	aggaacaact	tggaaaagcc	7320
	ctagattgtc	agatctacgg	ggctgttac	tccatttgc	cacttgacat	acccatcgatc	7380
	attcaacgac	tccatggct	tagcgctt	tcactccata	gttactctcc	aggtgagatc	7440
	aataggggtgg	tttcatgtct	caggaaactt	ggggtaacc	ccttgcqagt	cttggagacat	7500

cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacatct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
5 tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
agcctgtctc gtgcccggacc ccgctggttc atgtggtgcc tactcctact ttctgttaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatcctgtt ttttccctt ttttttttc tttttttttt tttttttttt tttttttttt 7860
ttttctcctt tttttttcctt tttttttcc ttttttttcc tttggtggtt ccatcttagc 7920
cctagtcacg gctagctgtg aaaggtccgt gagccgcttg actgcagaga gtgctgatac 7980
10 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 7
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

<400> 7

```

 5      gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
      tcttcacgca gaaagcgct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
      cccccctccc gggagagcca tagtggctg cggAACCGGT gagtacaccg gaattgccag 180
      gacgaccggg tccttcttg gatcaacccg ctaatgcct ggagatttg gcgtgcccc 240
      gcgagactgc tagccgaga gtgtgggtc gcgAAAGGCC ttgtggtaact gcctgtatagg 300
      gtgcttgcga gtccccccc aggttcgtc gaccgtcac catgagcac aatccataac 360
      ctcaaagaaa aaccaaactg aacaccaacg ggcgcgcccatt gattgaacaa gatggattgc 420
      acgcagggtc tcggccgtc tgggtggaga gctatccgg ctatgactgg gcacaacaga 480
      caatcggtc ctctgttgc gccgtttcc gctgtcagg gcaggggcgc ccggttctt 540
      ttgtcaagac cgacctgtcc ggtggctga atgaactgca ggacgaggca ggcgcgctat 600
      cgtggctggc cacgacgggc gttccttgcg cagctgtgcg cgacgttgc actgaagcgg 660
      gaaggggactg gctgttattt ggcgaagtgc cggggcagga ttcctgtca ttcacccctt 720
      ctctggccgaa aaggtatccc atcatggctg atgcaatgcg gggctgcattt acgcttgcatt 780
      cggctacccg cccatttcgac caccaagcga aacatcgcat cgagcgagca ctgtactcgga 840
      tggaaagccgg tcttgtcgtat caggatgtatc tggacgaaga gcatcagggg ctgcgcggc 900
      ccgaactgtt cgccagggtc aaggcgcgcgca tgcccgcgg cgaggatctc gtcgtgaccc 960
      atggcgatgc ctgcgttgcg aatatcatgg tggaaaatgg ccgcgttttgcg 1020
      actgtggccg gctgggtgtg gcgaccgtt atcaggacat acgcttggctt acccggtata 1080
      ttgtcaaga gcttggcggc gaatgggtg accgcttccat cgtgttttgcg 1140
      ctcccatttc gcagcgcattt gccttctatc gccttcttgcg cgagtcttc tgagttaaa 1200
      cagaccacaa cggtttccctt cttagcggat caattccgcgc cctctccctt ccccccctt 1260
      aacgttactg gccgaagccg cttggataaa ggcgggtgtg cgtttgtcta tatgttattt 1320
      tccacccat tgcgttccat tggcaatgttgg acggccgggaa aacctggccc tgcgttccat 1380
      acgaggattt cttaggggtt tttcccttc gccaaaggaa tgcaggatgtt gttgaatgtc 1440
      gtgaagggaa gatgttccat ggaagcttct tgaagacaaa caacgtctgt acgcaccctt 1500
      tgcaggcgcg ggaacccccc acctggcgac aggtgcctt cggccaaaa gccacgtgt 1560
      taagatacac ctgcaaaaggc ggcacaaccc cagtgccac cttgtgatgg gatagttgt 1620
      gaaagagtca aatggcttccat ctcacggttca ttcaacaagg ggctgaagga tgcccagaag 1680
      gtacccattt gtatggatc tgatctgggg ctcgggtgcg catgttttgcg 1740
      tcgaggatccaaa aacacgttca ggcggccggc accacggggg cgtgttttgcg 1800
      cacgataata ccatgggcac gaatctaaa ctcacaaagaa aaaccaaactg taacaccaac 1860
      cggcccccac aggacgttca gttccgggc gttgttcaga tcgtcggtgg agtttacctt 1920
      ttggcgcgcgaa gggggccccc gttgggtgtg cgcgcgacta ggaagacttc cgacgggtcg 1980
      caacctcggtgaa agggcgacaa acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
      gctcggcccg ggtacccctt gccccttat gcaatggggg gcttgggtg ggcaggatgg 2100
      ctccgttcac cccgtggctc tggccttagt tggggccccc cggacccccc gctgtgtcg 2160
      cgcacattgg gtaaggatctt cgtatccctt acgtggggatc tegccgatctt catggggatc 2220
      attccgtcg tggcgcggcc cctagggggc gttggccagggg ccctggcgca tggcgtccgg 2280
      gttctggagg acggcgatggaa ctatgcaaca gggaatctgc cgggttgcgc 2340
      ttcttttgcgat tttgttgcgtc ctgttttgcgat atccagttt ccgttatgcg 2400
      gtatccggat gttaccatgt cacgaacgcgatc tgctccaaacg caagcattgtt gttgtggc 2460
      gcgacatgat tcatgcatac cccgggtgc gtgcctgcg ttggggaa caactccctt 2520
      cgtgtctggg tagcgctcac tcccacgcgc gggcccgaggaa acgctagcgatccctt 2580
      acgatacgac gccatgtcgat tttgtctcgat gggggggctg ctctctgcgc cgctatgtac 2640
      gtggggatc tctgcggatc ttttttgcgat gtcggccaggc ttttttgcgat 2700
      cggcacgaga cagtagggatc ctgcaattgc tcaatataatc cggccacgt gacaggtcac 2760
      cgtatggctt gggatatgtt gatgtactgg tcaatccatc cagccctatgtt ggtatcgatgg 2820
      ttactccggaa tccccacaatc tgcgtggat atgggtgggg gggccattt gggagttctt 2880
      gcccccttgcgat cttactattt catgtggggg aactgggtca agtttgcgtat ttttttgcgat 2940
      ctcttttgcgat cgttgcgttgcg gggggccatc ttttttgcgat gtttttgcgat 3000
      ctcgggatcat ctttttgcgat ttttttgcgat gtttttgcgat 3060
      aacggcagctt ggcacatcaa caggactgcgc tcaatccatc aatccatcgtt ttttttgcgat 3120
      ttccttgcgat cgttgcgttgcg ttttttgcgat gtttttgcgat 3180
      gccagctgcgat gggccatcgcg ttttttgcgat gtttttgcgat 3240
      tcacacagctt cggaccaggatc gccttattgtt ttttttgcgat gtttttgcgat 3300
      gtaccccgccgat cgcaggatgtt gggggccatcgcg ttttttgcgat gtttttgcgat 3360
      gggacgaccgcgat accgggttccat cgttgcgttgcg ttttttgcgat gtttttgcgat 3420
      ctgttttgcgat acaacacgcgc gggccatcgcg ttttttgcgat gtttttgcgat 3480
      agcaactgggtt ttttttgcgat gggggccatcgcg ttttttgcgat gtttttgcgat 3540
  
```

aaaaccccttga cctggccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
 tgggttcgg ggccttgggt gacacccaga tgcttggtcc actaccata caggctttgg 3660
 5 cactaccctt gcaactgtcaa cttaaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
 gagcacaggc tcaagggccgc atgcaatttg actcgaggag agegttggaa cctggaggac 3780
 agggacagat cagagcttag cccgctgtcg ctgttacaa cggagtggca ggtattggccc 3840
 tggcttcca ccacccttacc ggetctgtcc actgggttga tccatcttca tcagaacgtc 3900
 gtggacgtac aatactgtta cgttataggg tggcggttgc aatcaaattgg 3960
 gagatgttcc tggcttctg gcggacgcgc gctgtctgtc ctgttgg 4020
 atgatgtgc tgatagctca agctgaggcc gcccataaaaa acctgggttgc 4080
 10 gcatccgtgg cggggggcgca tggcattctc tccttctcg tggcttctg tgctgcctgg 4140
 tacatcaagg gcaggctggt ccctggggcg gcatatgccc tctacggcgt atggccgcta 4200
 ctctgctcc tgcgtggcggtt accaccacga gcatacgcacca tggaccggga gatggcagca 4260
 tcgtgcggag ggcgggtttt ctaggtctg atactcttgc ctttgttacc gcactataag 4320
 ctgttctcg ctaggtctat atgggtggta caatattttt tcaccaggcgc cgaggcacac 4380
 ttgcaagtgt ggtatcccccc cctcaacgtt cggggggggcc gcgatgccgt catccttctc 4440
 15 acgtgcgcga tccaccacaa gctaattttt accatcacca aaatcttgc cggcataactc 4500
 ggtccactca tggtgcgttca ggctgggtata accaaagtgc cgtacttgc ggcgcacac 4560
 gggcttattt gtcgtgcgt gctggcggg aagggtgtc ggggttattt tgccaaatgg 4620
 gctctcatga agttggccgc actgacaggt acgtacgtt atgaccatct cacccactg 4680
 cgggacttggg cccacgggg cttacagacat cttgcgttgc cagttgagcc cgtcgttcc 4740
 tctgatatgg agacaaatgg tttttttttt gggcagaca cccggcgttgc tggggacatc 4800
 20 atcttggggcc tggccgttcc cggccgcagg gggaggggaga tacatctggg accggcagac 4860
 agccttgaag ggcagggttgc ggcacttcc ggccttattt cggcttactc ccaacagac 4920
 cgaggccatc ttggctgtcat catcaacttgc ctcacaggcc gggacaggaa ccaggtcgag 4980
 gggaggttcc aagtggtctc caccgcaaca caatcttcc tggcgttcc cgtcaatggc 5040
 gtgtgttggc ctgttctatca tggtgcgttgc taaaagaccc ttggccggcc aaaggggccca 5100
 atcaccctaaa tggtacacaa tggtggaccag gacacgttgc gctggcaagc gccccccggg 5160
 25 ggcgttccct tgacaccatg cacctgcggc agctcgacc tttacttggt cacgaggcat 5220
 gccgatgtca ttccgggtcgcc cggggggggc gacagcagggg ggaggctact ctccccccagg 5280
 cccgttctctt acttggatggg ctcttggggc ggttccactgc tctggccctc gggcatgtct 5340
 gtgggcattt ttccgggttgc cgttgcacc cgggggggtt cgaaggccgtt ggactttgtt 5400
 cccgtcgagt ctatggaaac cactatgcgg tccccgggttgc tcaacggaaat ctcgtccct 5460
 cccggccgtac cgcagacatc ccagtggttcc catctacacg cccctactgg tagcggcaag 5520
 30 agcaactaagg tggccgttgc gtatgcagcc caagggtata aggtgtttgtt cctgaaccgg 5580
 tccgtcgccg ccaccctagg tttccggggc tatatgttca aggacatgg tattggaccct 5640
 aacatcagaa cgggggttggg gaccatcacc acgggtggcc ccatcacgtt ctcacccat 5700
 ggcgaagtttcc ttggccacgg tgggtgtctt gggggcgcctt atgacatcat aatatgttgc 5760
 gagtgccact caactgactc gaccactatc ctgggcatcg gcacagtcc gggaccaagcg 5820
 gagacggctg gacgcgttgc gtcgttgc gccaccgttgc cgcctccggg atcggttacc 5880
 35 gtgccacatc caaacatcgat ggggtggcgtt ctgttccatgc ctggagaaat ccccttttt 5940
 ggcggccatc tcccatcgatc gaccatcaag gggggggggc accttcatttt ctggccattcc 6000
 aagaagaaat gtgtgtatgc cggccgttgc ctatccggcc tggacttcaaa tgctgttagca 6060
 tattaccggg gcttgcgttgc atccgttcaaa ccaacttagcg gacgttgc tggctgttagca 6120
 acggacgttc taatgacggg cttaaccggc gattttgcact cgttgcgttca ctgcacatata 6180
 tgggtgttgc agacagtgc ttcacccgttgc gacccgttgc ttcacattttt gacgttgcacc 6240
 40 gtggccacaag acgcgggttgc acgttgcgttgc cggcgaggca ggactgttag gggcaggatg 6300
 ggcattttaca ggtttgttgc tccaggagaa cggcccttgc gcatgttgc ttcctcggtt 6360
 ctgttgcgttgc gctatgttgc gggctgttgc tgggttgc gtcgttgc ttcacccgttgc 6420
 gtaggttgc gggcttacctt aaacacacca gggcttgc tctggccatgg ccatctgggg 6480
 ttctggggaga gcttgcgttgc accgttccatc cacaatggcc cccattttt ttcctcggtt 6540
 aggcaggccatc gagacaaatcc ccccttacatgc gtagtacatcc aggttacgggtt gtcgtccagg 6600
 gtcaggccatc cacccttccatc gttttttttt atgtggggatgtt gtcgttgc gtcgttgc 6660
 45 acgttgcaccc ggcggccatc ccccttacatgc gttttttttt atgtggggatgtt gtcgttgc gtcgttgc 6720
 accacacacc ccataacaa atacatcatgc gcatgttgc ttcacccgttgc 6780
 acggacgttgc gggcttgcgttgc aggttgcgttgc ttcacccgttgc 6840
 acaggccatc tgggttgcgttgc gggcttgcgttgc atcttgcgttgc gaaaggccggc catccatcc 6900
 gacaggccatc tcccttaccc gggcttgcgttgc gggcttgcgttgc acacccgttgc 6960
 tacatcgaaatc agggatgttgc gtcgttgcgttgc gggcttgcgttgc 7020
 50 caaacagccatc ccaacacacc ccataacaa atacatcatgc gcatgttgc ttcacccgttgc 7080
 ctcgttgcgttgc tggccatcc gcatgttgc ttcacccgttgc 7140
 ggcttgcgttgc ctcttgcgttgc caaccccgcc atacatcatgc ttcacccgttgc 7200
 atcaccacacc cgcgttgcgttgc gggcttgcgttgc atacatcatgc ttcacccgttgc 7260
 gccggccatc ttggccatcc cggccgttgc ttcacccgttgc 7320
 gccggcttgcgttgc gacgttgcgttgc ccccttacatgc gggcttgcgttgc 7380
 gcaaggccatc gggccgttgc ttcacccgttgc 7440
 gaggacgttgc ttcacccgttgc 7500

5	gtgtgcgcacg cgatactcgcc tcggcacgtg ggcccagggg agggggctgt gcagtgatgc 7560 aaccggctga tagcgttcgc ttgcggggt aaccacgtct cccccacgca ctatgtgtct 7620 gagagcgacg ctgcagcagc tgtactcgat atccctctcta gtcttaccat cactcagctg 7680 ctgaagaggc ttaccaggat gatcaacgag gactgctcca cgccatgctc cggctcggtt 7740 ctaagagatg tttgggattt gatatgcacg gtgttgcactg atttcaagac ctggctccag 7800 tccaagctcc tgccgcgatt gccggagtc cccttcttcatgtcaacg tgggtacaag 7860 ggagtctggc ggggcgacgg catcatgcaa accacctgcc catgtggagc acagatcacc 7920 ggacatgtga aaaacggttt catgaggatc gtggggccta ggacctgttag taacacgtgg 7980 catggAACAT tccccattaa cgctacacc acggggccct gcacgccc cccggcgcca 8040 aattattcttta gggcgctgtt gcgggtggct gctgaggagt acgtggaggt tacgcccgggt 8100 ggggatttcc actacgtgac gggcatgacc actgacaacg taaagtgcctt gtgtcaggtt 8160 ccggccccccg aattcttccat agaagtggat ggggtgcgtt tgacacaggta cgctccagcg 8220 tgcaaacccttccat tccatcggtt ggaggtcaca ttccctgtcg ggctcaatca atacctgggtt 8280 gggtcacagc tcccatcggtt gcccgaaccg gacgttagcag tgctcacttc catgtctcacc 8340 gacccttccc acattacggc ggagacggct aacgttaggc tggccagggg atctccccc 8400 tccttgccca gctcatcagc tagcoagctg tctgcgcctt ctttgcaggc aacatgcact 8460 acccgtcatg actcccccggc cgctgacccatc atcgaggcca acctcctgtg gcggcaggag 8520 atggccgggaa acatcaccccg cgtggaggatc gaaaataagg tagtaatttt ggactctttc 8580 gagccgccttcc aagcgaggaga ggatgagagg gaagtatccg ttccggcgga gatcctcggtt 8640 aggtccaggaa aattccctcg agcgatgccc atatgggcac gccccgatta caaccctcca 8700 ctgttagagt cctggaggaa cccggactac gtccctccag tggtagacccg gtgtccattt 8760 ccgcctgcca aggccccctcc gataccacct ccacggaggag agaggacgggt tgcctgtca 8820 gaatctaccg tgccttctgc cttggcggtt ctcgcccacaa agacccctgg cagctccgaa 8880 tcgtccggccg tgcacagcgg cacggcaacg ggctctccgtt accagccctc cgacgacggc 8940 gacgccccat cccgacgttgc gtcgtactcc tccatgcccc cccttgcagggg ggagccgggg 9000 gatcccgatc ttagcgacgg gtcttggctt accgttaacg aggaggctag tgaggacgtc 9060 gtctgtctgtt ctagtgcctt ccatggaca ggcgcctgtt acgtccatc tgacgccttgc 9120 gaaaccaaggc tgcccgtaa tgcactgagc aactcttgc tccgtcacca caacttggtc 9180 tatgtaccaa catctcgca cgcacccctg cggcagaaga aggtcacctt tgacagactg 9240 cagggtctgg acgaccacta cccggacgtt ctcaggaga tgaaggcgaa ggcgtccacca 9300 gttaaggcttcc aacttctatc cgtggaggaa gctgttaacg tgacccccc acattcggtt 9360 agatctaaat ttggctatgg ggcaaggac gtccggaaacc tatccagcaaa ggcgttaac 9420 cacatccgct ccgtgtggaa ggacttgcg gaagacactg agacaccaat tgacaccacc 9480 atcatggcaaa aaaatgaggat tttctgtgtt caaccaggaga agggggccg caagccagct 9540 cgcccttatcg tattcccgat tttgggggtt cgtgtgtgcg agaaaatggc cctttacgat 9600 gtggcttccat ccctccctca ggccgtgtt ggtcttccat acggattcca atactctct 9660 ggacagcgccg tcgagttcc tggtaatgcc tggaaagcga agaaatggcc tatgggcttc 9720 gcatatgaca cccgtgttt tgactcaacg gtcactgaga atgacatccg tggtagggag 9780 tcaatctacc aatgttgtga cttggccccca gaaaggccagc aggcataag gtcgtccaca 9840 gagccgtttt acatcggggg cccctgtact aattctaaag ggcagaactg cggctatcgc 9900 cggtccgcgcg cgagccgtgtt actgacgacc agctcggtt ataccctcac atgttattt 9960 aaggccgttgcg cggccgtgtc agctcgacag ctccaggact gcacatgtct cgtatgcgg 10020 gacgaccctt tgcgttatctg tggaaagcgc gggacccaaag aggacgaggc gagccatacg 10080 gccttcacgg aggctatgac tagatactt gccccccctg gggacccggc caaaccagaa 10140 tacgacttgg agttgataac atcatgttcc tccaaatgtgtt cagtcgcgc cgtatgcac 10200 ggcaaaagggt tgcgttatctt caccgtgac cccaccaccc cccttgcgcg ggctgcgtgg 10260 gagacagacta gacacactcc agtcaattcc tggcttaggca acatcatcat gtatgcgccc 10320 accttggggg caaggatgat cctgtatgact cattttcttccatccat agctcaggaa 10380 caacttgaaa aagccctaga ttgtcagatc tacggggcct gttactccat tgagccactt 10440 gacccatctc agatcattca acgactccat ggccttagcg cattttccat ccatagtttac 10500 tctccagggtt agatcaatag ggtggcttca tgcctcaggaa aacttggggt acggcccttg 10560 cgagtcgttgc gacatcggtt cagaagtgtt cgcgttaggc tactgtccca gggggggagg 10620 gctgccactt gtggcaaggta cctttcaac tggcaggtaa ggaccaagct caaactcact 10680 ccaatccggg ctgcgtccca gttggattt tccagctggt tcgttgcgtt ggcttccatc 10740 ggagacatatac acacagctt gtctcgttgc cggccctgtt ggcttccatc gtgcctactc 10800 ctactttctgtt taggggttagg catctatcta ctcccaacc gatgaacggg gagctaaaca 10860 ctccaggcca ataggccatc ctgtttttt ccctttttt ttttctttt ttttctttt 10920 ttttttttt ttttttttccctttttt ttcctttttt tttccttttccctttt 10980 tggctccatc ttagccctag tcacccgttgc ctgtgaaagg tccgtgagcc gcttgcactc 11040 agagagtgttccatc gatactggcc tctctgcaga tcaagt 11076
---	---

5 <210> 8
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

10 <400> 8
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 ttttcacgca gaaagcgct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
 ccccccctccc gggagagcca tagtggtct cgaaaccgt gagtacaccg gaattgccag 180
 gacgaccggg tcctttctt gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccagata gtgtgggtc gcgaaaggcc ttgtgtact gcctgatagg 300
 gtgcttgcga gtgccccggg aggtctcgta gaccgtcac catgagcaca aatcctaaac 360
 ctcggaaa aaccaaacgt aacccaacg ggcgcgcat gattgaacaa gatggattgc 420
 acgcagggttc tccggccgt tgggtggaga ggcttattcg ctatgactgg gcacaacaga 480
 caatcggtcg ctctgtatcc gccgtgtcc ggctgtcagc gcaggggcgc cccgttctt 540
 ttgtcaagac cgacctgtcc ggtggccctga atgaactgca ggacggggca ggcggctat 600
 cgtggctggc cacgacgggc gttecttgcg cagctgtgtc cgacgttgc actgaagcgg 660
 gaaggggactg gctgtatgg ggcgaagtgc cggggcaga ttcctgtca ttcacctt 720
 ctctggccga gaaagtatcc atcatggctg atcaatgcg gcgctgtcat acgcttgatc 780
 cggctacccgt cccattcgac caccacgca aacatcgcat cgagcggca cgtactcgga 840
 tgaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgcgg 900
 ccgaactgtt cgccaggtc aaggcgcga tgcccggcgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgcttccg aatatcatgg tggaaaatgg cgcgtttctt ggattcatcg 1020
 actgtggccg gctgggtgt gcgaccgct atcaggacat agcgttggct acccgtata 1080
 ttgtctgaaga gcttggcgcc gaatgggtct accgcttctt cgtcttac ggtatcgccg 1140
 15 ctcggattc gcagcgcatac gccttctatc gccttcttgc cgagttttc tgagttaaa 1200
 cagaccacaa cggttttccctt ctggggat caattccggc cctcccttc cccccccct 1260
 aacgttactg gccgaagccg ctggaaataa ggccgggtgt cggttgcata tatgttattt 1320
 tccaccatata tggctgtttt tggcaatgtg agggggccga aacctggccc tgcgttctt 1380
 acgaggattt ctgggggtct ttcccttc gccaaaggaa tgcaagggtt gtgaatgtc 1440
 gtgaagggaaag cagttctctt ggaagttct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcgc ggaacccccc acctggcgc acgtgcctt gcgccaaaa gcccacgtgt 1560
 20 taagatacac ctgcaaaaggc ggacacaaccc ttcaacaagg ggctgaagga tgcccagaag 1620
 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
 gtacccatt gtatgggatc tgatctgggg ctcgggtgca catgttttac atgtgtttag 1740
 tcgagggttaaaa acgtcta ggccccccga accacgggga cgtgttttc ctttggaaaa 1800
 cacgataata ccatggcgcc tattacggcc tactccaaac agacgcgagg cctacttggc 1860
 tgcatcatca ctggctcac aggccgggac aggaaccagg tcgaggggg ggtccaagtg 1920
 25 gtctccaccg caacacaatc ttctctggcg acctgcgtca atggcggtg ttggactgtc 1980
 tatcatgggt cgggctcaa gacccttgcg gggccaaagg gcccacatcac ccaaatgtac 2040
 accaatgtg accaggaccc ctgcggctgg caagcgcccc cggggcgcg ttccttgaca 2100
 ccatgcacct gccgcagtc ggaccttac ttgtcaca ggcacggca tgcattccg 2160
 gtgcgcggc ggggcacag cagggggagc ctactctcc ccaggccgt ctcttactt 2220
 aagggcttcc cggcggtcc actgtctgc ccctcggggc acgctgtgg catcttccg 2280
 30 gctggcggt gcacccggg ggttgcgaag ggggtggact ttgttccgt cgagtctatg 2340
 gaaaccacta tgcggcccc ggttccacg gacaactctg cccctccggc cgtaccgcag 2400
 acattccagg tggcccatct acacggccct actggtagcg gcaagagcac taaggtccg 2460
 gctcgatgtc cagcccaagg gtataagggt cttgtctga acccgccgt cgccgcacc 2520
 35 ctaggtttcg gggcgatata tgcataaggca catggatcg accctaatacat cagaatcg 2580
 gtaaggacca tcaccacggg tgccccatc acgtactcca cctatggca gtttcttgcc 2640
 gacgggtgggt gctctgggg cgccatgac atcataatat gtgatgatgtg ccaactcaact 2700
 gactcgacca ctatccctggg catcgccaca gtcctggacc aagcgagac ggctggagcg 2760
 cgactcgatc tgctccac cgtacgcct cgggatcg tcaccgtgc acatccaaac 2820
 atcgaggagg tggctctgtc cagactggaa gaaatccct ttatggca agccatcccc 2880
 atcgagacca tcaagggggg gaggcaccc attttctgc attccaagaa gaaatgtat 2940
 40 gagctcgccg cgaagctgtc cggccctcgga ctcaatgtcg tagcatatta cccggggctt 3000
 gatgtatcc tcataccaac tagcgagac gtcattgtcg tagcaacggc cgctctaatg 3060
 acgggcttta cgggtgactt cgactcgatg atcgactgca atacatgtgt caccacgaca 3120
 gtcgacttca gcctggaccc gacccatcattt gatggacgca cgaccgtgcc acaagacgcg 3180
 gtgtcacgtc cgcagccggc aggcaggact ggtggggca ggtatggcat ttacagggtt 3240
 gtgactccag gagaacggcc ctggggatc ttgcattctt cgggttgcg cgagtctat 3300
 gacgcgggct gtgcttggta cgagctcacg cccggccgaga cctcgtttag gttgcgggct 3360
 45 tacctaaaca caccagggtt gccccatc caggaccatc tggagttctg ggagagcg 3420
 tttacaggcc tcaccacatc agacgccccat ttcttgcctt agactaagca ggcaggagac 3480
 50 aactccctt acctggtagc ataccaggct acgggtgtcg ccagggtca ggctccac 3540

5 ccatcggtgg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacgggcca 3600
 acgcccctgc tgtataggtct gggagccgtt caaaaacgagg ttactaccac acaccccaata 3660
 accaaataca tcatggcatg catgtcggt gacctggagg tcgtcaccgag caccctgggtg 3720
 ctggtaggcg gagtccttagc agctctggcc gctattgccc tgacaacagg cagcgtggtc 3780
 attgtggca ggatcatctt gtccggaaag ccggccatca ttcccgacag ggaagtcctt 3840
 taccggaggt tcgtatggat ggaagagtgc gcctcacacc tcccttacat cgaacaggaa 3900
 atgcagctcg ccgaacaatt caaaacagaag gcaatcggtg tcgtcacaac agccaccaag 3960
 caagcggagg ctgctgtcc cgtgtgggg tccaatgtgg ggaccatcg agccttctgg 4020
 gcaagcata tggaaattt catcagcggg atacaattt tagcaggtt gtccactctg 4080
 10 cctggcaacc ccgcgatagc atcactgttgcattcacag cctctatcac cagcccgctc 4140
 accacccaaac ataccctctt gtttaacatc ctggggggat gggtggccgc ccaacttgct 4200
 cctcccaagcg ctgctctgc tttcgtaggc gccggcatcg ctggagccgc tggggcagc 4260
 ataggccttgc ggaagggtgt tggatatt ttggcagggtt atggagcagg ggtggcaggc 4320
 ggcgtcggtt ctttaaggt catgagccgc gagatgccct ccaccgagga cctggtaac 4380
 ctactccctg ctatccctc ccctggcgcc ctatcggtcg gggtcgtgtg cgacagcgata 4440
 ctgcgtcgcc acgtggggcc aggaggagggt gctgtgcagt ggtgaaccg gctgatagcg 4500
 ttgcgttgcg gggtaacca cgtctcccccc acgcactatg tgcctgagag cgacgctgca 4560
 gcacgtgtca ctcaaatctt ctctgttcc accatcaactc agctgtgaa gaggcttcac 4620
 cagtggatca acgaggactg ctccacgcg tgctccggct cgtggctaag agatgtttgg 4680
 gattggatca gacgggtt gactgttttcc aagacctggc tccagttccaa gcttctggcg 4740
 15 cgattggccg gatgtccctt ctctcatgt caacgtgggtt acaaggaggat ctgggggggc 4800
 gacggcatca tgcaaacac ctggccatgt ggggcacaga tcaccggaca tggaaaaaac 4860
 ggttccatga ggatcggtt ggcttagggc ttgttagtaca cgtggcatgg aacattcccc 4920
 attaacgcgt acaccacggg cccctgcacg ccctccccgg cgccaaattt ttctaggcg 4980
 ctgtggcggtt tggctgtca ggatgtacgtt gagggttacgc ggggggggaa ttccactac 5040
 gtgacggggca tgaccactga cgacgtaaag tgccctgtgtc aggttccggc ccccaattt 5100
 ttacagaag tggatgggtt gcggttgcac aggtacgttgc cagcgtgaa accccttcta 5160
 20 cgggaggagg tcacattctt ggtggggctc aatcaatacc tgggtgggtc acagctccca 5220
 tgcgagcccg aaccggatgt agcagtgtc acttccatgc tcaccgaccc ctccacatt 5280
 acggcggaga cggctaagcg taggtggcc aggggatctc ctccccctt ggccagctca 5340
 tcagctagcc agctgtctgc gccttcctt aaggcaacat gcaactacccg tcatgactcc 5400
 ccggacgctg acctcatcgaa ggccaaacctc ctgtggcgcc aggagatggg cggaacatc 5460
 acccgctgtt agtcagaaaa taaggtgtt attttggact ctttcgagcc gctccaagcg 5520
 25 gaggaggatg agatgggttccg gcgaggatcc tgccgggtt cagggaaattt 5580
 cctcgagatc tgcccatat ggcaccccccc gattacaacc ctccactgtt agatcttgg 5640
 aaggaccctt actacgtccc tccagtggta cacgggtgtc cattggccg tgccaaggcc 5700
 cctccgatac caccttcacg gaggaaaggg acgggtgtcc tgcagaatcc taccgtgtct 5760
 tctgccttgg cggagctcgcc cacagagacc ttccggcagct ccgaatcgcc ggcgtcgac 5820
 agcggcacgg caacggcctc ttctgaccag ccctccgacg acggcgacgc gggatccgac 5880
 gttgagtcgt actcctccat gccccccctt gaggggggggc cgggggatcc cgatctcagc 5940
 30 gacgggttcc ggttacatgtt aaggcaggag gctagtgttgc acgtcgctg ctgctcgatg 6000
 tcttacacat ggacaggcgcc cctgtatcactg ccattcgctg cggaggaaac caagctgccc 6060
 atcaatgcac tgacgttccgtt caccacaact tggtctatgc tacaacatct 6120
 cgcagcgcaa gcttgcggca gaagaagggtt acctttgaca gactcgatg cctggacgac 6180
 cactaccggg acgtgtctca gggatgttgc gcgaaaggcgcc ccacagttaa ggctaaactt 6240
 35 ctatccgtgg aggaaggctt taatgttgcacccccatcc cggccagatc taaatttggc 6300
 tatggggcaaa aggacgttccg aaacctatcc agaaggccg ttaaccatcc cgcctccgt 6360
 tgaaggact tgctggaaaga cactgagaca ccaattgtaca ccaccatcat gggaaaaaaat 6420
 gaggttttctt gcttccaaacc agagaagggg ggccgcacggc cagctcgctt tategttattc 6480
 ccagattttgg ggggttgcgtt gtgcgagaaaa atggccctttt acgtatgttgc tccaccctc 6540
 cctcaggccg tggatggctt ttccatcggta ttccaaatact ctccctggaca ggggttgcag 6600
 ttccctggta atgcctggaa agcgaagaaaa tggccctatgg gtttcgcata tgacacccgc 6660
 40 tgttttgcgtt caacgggttcc tggaaatgttcc acgtcgatccat ctaccaatgt 6720
 tgtgacttgg ccccccggaa cagacaggcc ataagggttgc tcacagacgc gctttacatc 6780
 gggggccccccca tgactaattt taaagggttgc aactgtcggtt atcgccgtt ccggcgacgc 6840
 ggtgtactgtt cggaccatgtt cggatataacc ctcacatgtt acgttgcaggc cgtgtcggtt 6900
 ttgcgatgtt cggatgttccg ggtactgcacg atgtcgatg tggggacgc ctttgcgtt 6960
 atctgtgaaa ggcggggggc ccaaggaggac gggggggcc tacggccctt cacgggggtt 7020
 45 atgactatgtt actctggccccc ccctggggac ccggccaaac cagaatgttca cttggagttt 7080
 ataacatcat gcttcccaa tggatgttgc ggcacgttgc catctggcaaa aagggttgc 7140
 tatctcaccctt gtttttttttcc caccggccctt ggcggggctt cgtggggatc agctagacac 7200
 actccagtcac ttccctggctt agggatgttgc atcatgttgc cggccacccctt gtggggcaagg 7260
 atgatccctgtt tgactcattt ctctccatc ttcttagtgc aggaacaact tggaaaaagcc 7320
 cttagattgttcc agatctacgg ggcctgttac tccattgttgc cacttgcaccc acctcagatc 7380
 50 attcaacgcac tccatggccct tagcgcattt tcaactccata gttactctcc aggtgagatc 7440
 aatagggtgg cttcatgcctt caggaaactt ggggttccgc ctttgcgtt 7500

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgcgc 7560
aagtacctct tcaactggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
agcctgtctc gtgcccgacc ccgctgggtc atgtggtgcc tactcctact ttctgttaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatccgtt ttttccctt ttttttttc tttttttttt tttttttttt tttttttttt 7860
ttttccctt tttttccctt cttttttcc ttttctttcc tttggtggct ccatcttagc 7920
10 cctagtcacg gctagctgt aaagggtccgt gagccgcttg actgcagaga gtgctgatac 7980
tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 9
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

5 <400> 9
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgct agccatggcg ttagtatgag tgctgtgcg cctccaggac 120
 cccccctccc gggagagcca tagtggctcg cgaaaccggt gагtacaccg gaattgccag 180
 10 gacgaccggg tcctttctt gatcaaccccg ctaatgcct ggagatttg gcgtgcccc 240
 gcgagactgc tagccgagta gtgtgggtc cgaaaggcc ttgtgtact gcctgatagg 300
 gtgcttgcgta gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
 ctc当地agaaa aaccaaacgt aacaccaacg ggcgcgcat gattgaacaa gatggattgc 420
 acgcagggttc tccggccgt tgggtggaga ggctattcg ctatgactgg gcacaacaga 480
 15 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggcgc cccggttctt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca ggcggctat 600
 cgtggctggc cacgacgggc gttccttgcg cagctgtgtc cgacgttgc actgaagcgg 660
 gaaggactg gctgttattt ggcgaagtgc cggggcagga tctccgtca ttcaccc 720
 ctctccgtca gaaagtatcc atcatggctg atgaatgca gcccgtcat acgcttgcata 780
 eggtactcg cccattcgac caccatcgca aacatcgat cgagcggca cgtactcgga 840
 tggaaagccgg tcttgcgtat caggatgatc tggacgaa gcatcagggg ctcgcggccag 900
 20 cggactctc cgccagggtc aaggcgcga tgccgcagg cgaggatctc gtctgtgaccc 960
 atggcgatgc ctgcttgcg aatatcatgg tggaaaatgg ccgttcttgcgattcatcg 1020
 actgtggccg gctgggtgt gggggccgt atcaggatc acgcttgcgt acccggtata 1080
 ttgtcaaga gcttggccgc gaatgggtcg accgttctt cgtgttcttgcgatcc 1140
 ctcccgttcc gcaatcgatc gccttcttgcgatcc gccttcttgcgatcc 1200
 cagaccacaa cgggttccct ctagcggatc caattccccc cctctccctc cccccc 1260
 25 aacgttactg gccgaaggccg ctttggataa ggcgggtgt cgtttgtcta tatgttattt 1320
 tccaccatat tgccgttcc tggcaatgtg agggcccgaa aacctggccc tgcgttcttgcg 1380
 acgagcattc cttaggggtt ttcctctc gccaaaggaa tgcaaggatc gttgaatgtc 1440
 gtgaaggaag cagttcctt ggaagatct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcagc ggaacccccc acctggcgac aggtgccttgcg gcccggccaa gccacgtgt 1560
 taagatacac ctgcaaggc ggcacaaccc cagtgccacg ttgtgagttt gatagttgt 1620
 30 gaaagagtca aatggcttc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
 gtacccatt gtatggatc tgatctgggg cctcggtgc catgcttac atgtgttttag 1740
 tcgaggttaa aaaacgtcta ggcggcccgaa accacggggc cgtgggttcc ctttggaaaa 1800
 cacaatggcata ccatgggcac gaatcttggaa ctc当地aaacggc taacaccaac 1860
 cggcccccac aggacgttca gttccgggc ggtggtcaga tcgtcggtgg agtttacctg 1920
 ttgcccgcga gggggcccgat gttgggtgt cgcgcgacta ggaagacttc cgagcgggtcg 1980
 35 caacctcgta gaaggcgaca acctatcccc aaggctcgcc agcccgggg tagggcctgg 2040
 gtcagcccg ggtacccctg gccccttat ggcaatgggg gcttgggtg ggcaggatgg 2100
 ctccctgtcac cccgtggctc tcggctagt tggggccca cggacccccc gcttaggtcg 2160
 cgc当地ttgg gtaaggatcat cgataccctc acgtcggttgc tgcgcgatct catggggat 2220
 atcccgctcg tcggccccc ccttaggggc gctgcccagg cctggcgca tggcgtccgg 2280
 40 gttctggagg acggcgtgaa ctatgcaaca gggaaatctgc cccgttgc ttttctatc 2340
 ttccctttgg ctttgcgtc ctgtttgacc atccctatc cgc当地tgcgcaac 2400
 gtatccggag tgtaccatgt cacgaacgc tgctccaacg caagoattgt gtatgaggca 2460
 ggggacatga tcatgcatac cccgggtgc tgccctgcg ttcgggagaa caactccctc 2520
 cgtctgtggg tagcgctcac tcccacgctc gggccaggaa acgctactcg ccccaactac 2580
 acgatcgcac gccatgtcg tttgctcgat gggccggctg ctctctgtc cgctatgtac 2640
 45 gttggagatc tctcggtatc tgtttgcgtc tgcccccggc tggcccccggc tgc当地tgc 2700
 cggcacaatggc cagtagatc ctgc当地tgc tcaatatac cggccacatc gacagggtc 2760
 cgtatggctt gggatatgtat gatgaaactgg ttc当地ccggatc tccatcgatc 2820
 ttactccggatc tcccacaatgc tgc当地tgc tggatggccgg gggcccatgg gggaggatc 2880
 gggggcccttgc cctactattc catgggtgggg aactgggtca aggttctgtat tgc当地tgc 2940
 ctctttgcgg gctgtgacgg gggaaacctat gttacagggg ggacgatggc caaaaacacc 3000
 50 ctgggattatc cgtcccttgc ttc当地ccggatc tccatcgatc aaatccatgt tgtaaacacc 3060
 aacggcagct ggcacatcaa caggactgccc tgc当地tgc tcaactgcgca atgactccct caacactggg 3120
 ttc当地ccggatc cgtctgttca cgtgcacaatgc gccc当地tgc cgc当地tgc tggatggccgg agagcgcata 3180
 gccagctgca gccc当地tgc cgc当地tgc tcaactgcgca atgactccct caacactggg 3240
 tcacacagct cggaccagag gccttattgt tggcactacg caccggccgtc gtgc当地tgc 3300
 gtaccccgccgg cgc当地tgc tggatggccgg tggatggccgg tggatggccgg tgc当地tgc 3360
 gggacgaccgg accggatcgatc cgtcccttgc tcaatatac cggccacatc gacagggtc 3420
 55 ctgcttgcgatc acaacacgcg gccggccgaa ggacgatgtt gtc当地tgc tggatggccgg atggatgaat 3480
 agactgggtt tcaccaagac gtc当地tgc tggatggccgg tggatggccgg tgc当地tgc 3540

	aaaaccccttga	cctgccccac	ggactgcttc	cggaagcacc	ccgaggccac	ttacaccaag	3600
	tgtggttcgg	ggccttgggt	gacacccaga	tgcttggtcc	actaccata	caggctttgg	3660
5	cactaccct	gcactgtcaa	cttattaccatc	ttcaaggtta	ggatgtacgt	ggggggagtg	3720
	gagcacaggc	tcgaagccgc	atgcaattgg	actcgaggag	agcgttggta	cctggaggac	3780
	aggacagat	cagagcttag	cccgtctgtg	ctgtctacaa	cgagtgccg	gttattggcc	3840
	tgttccttca	ccaccctacc	ggctctgtcc	actggtttg	tccatctcca	tcagaacgtc	3900
	gtggacgtac	aatacctgtta	cggataggg	tcggcggtt	tctccttgc	aatcaaattgg	3960
	gagtatgtcc	tgttgcttctt	ccttcttctg	gcccgcgc	gctgtctgtc	ctgcttgg	4020
10	atgatgtgc	tgatagctca	agctgaggcc	gccctagaga	acctgggtt	cctcaacgcg	4080
	gcattccgtgg	ccggggcgca	tggcattctc	tccttctcg	tgttcttctg	tgctgcctgg	4140
	tacatcaagg	gcaggctgtt	ccctggggcg	gcatatgccc	tctacggcgt	atggccgcta	4200
	ctcctgtcc	tgctggcgtt	accaccacga	gcatacgcga	tggaccggga	gatggcagca	4260
	tcgtcgaggag	gcgcgggttt	cgtaggtctg	atactcttga	ccttgcacc	gcactataag	4320
	ctgttcctcg	ctaggctcat	atggtggta	caatatttt	tcaccaggc	cgaggcacac	4380
	ttgcaagtgt	ggatcccccc	cctcaacgtt	cgggggggcc	gcatgtccgt	catccttcctc	4440
15	acgtgcgcga	tccaccacaga	gctaattttt	accatcacca	aatcttgc	cgccataactc	4500
	gttccactca	tggtgcctca	ggctgttata	accaaagtgc	cgtacttgc	gcgcgcacac	4560
	gggtcttccat	gtgtgttgcgg	aagggtgt	gggggttata	tgttccaaatg	4620	
	gctctcatga	agtggccgc	actgacaggt	actgttgtt	atgaccatct	cacccactg	4680
	cgggactggg	cccacgcggg	cctacgagac	tttgcgttgg	caatttgc	cgctgtcttc	4740
	tctgatatgg	agaccaaggt	tatcacctgg	ggggcagaca	ccgcggcgt	tggggacatc	4800
20	atcttggggcc	tgcggcttc	cgccgcagg	gggagggaga	tacatctgg	acggcagac	4860
	agccttgaag	ggcagggggt	gcatcttc	gcccctatta	cggcctactc	ccaacagacg	4920
	cgaggcctac	ttggctgcat	catcaactagc	ctcacaggcc	gggacaggaa	ccaggtcgg	4980
	ggggaggtcc	aagtggtctc	caccaaca	caatctttc	tggcgcac	cgtaatggc	5040
	gtgtgttgg	ctgtctatca	tggtgcggc	tcaagaccc	ttgcggccc	aaaggccc	5100
	atcacccaaa	tgtacacaa	tgtgaccag	gacctgtcg	gttggcaac	gcgggggg	5160
25	gcgcgttcc	tgacaccatg	cacctggggc	agctcgacc	tttacttgg	cacgaggcat	5220
	gcccgtgtca	ttccgggtgc	ccggcggggc	gacagcagg	ggagctact	ctcccccagg	5280
	cccgcttcc	acttgaaggg	ctcttgcggc	ggtccactgc	tctgcctc	ggggcacgt	5340
	gttggcatct	tccgggtgc	cgtgtcacc	cgagggttgc	cgaaggcggt	ggactttgt	5400
	cccgctcgagt	ctatggaaac	caactatgcgg	tcccggtct	tcacggacaa	ctcgcccc	5460
	ccggccgtac	cgcagacatt	ccagtgccgc	catctacacg	ccctactgg	tagcggcaag	5520
30	agactaagg	tgcggctgc	gtatgcagcc	caagggtata	aggtgttgc	cctgaaccc	5580
	tccgtcgccg	ccaccctagg	tttgcgggc	tatatgtcta	aggcacatgg	tatcgacc	5640
	aacatcagaa	tggggtaag	gaccatcacc	acgggtgcc	ccatcacgt	ctccacctat	5700
	ggcaagttt	ttggcgacgg	tgggtgtct	gggggcctt	atgacatcat	aatatgtat	5760
	gagtgcact	caactgactc	gaccactatc	ctgggcac	gcacagtct	ggaccaagcg	5820
	gagacggctg	gagcgcgact	cgtgtgtc	gccaccgt	cgccctccgg	atcggtacc	5880
35	gtgccacato	caaacatcg	ggaggtggc	ctgtccagca	ctggagaat	cccttttat	5940
	ggcaaagcc	tcccatcg	gaccatca	ggggggaggc	acccat	ctgcccattc	6000
	aagaagaaat	gtgtgatgc	cgccgcga	ctgtccggc	tcggactca	tgtctgt	6060
	tattaccggg	gccttgcgt	atccgtata	ccaactagc	gagacgtcat	tgtctgt	6120
	acggacgtc	taatgacggg	ctttaaccgt	gacttcact	cagtgtatcg	ctgcaata	6180
	tgtgtcacc	agacagtca	tttcagctg	gaccgcac	tcaccattga	gacgacgacc	6240
40	gtgccacaag	acgcgggtgc	acgctcgc	cgccgaggca	ggacttgcgt	ggcaggatg	6300
	ggcattttaca	ggtttgcac	tccaggagaa	cgcccctcg	gcatgttgc	ttccctcggt	6360
	ctgtgcgt	gctatgcac	gggtgtgt	tggtacgac	tcaccccgc	cgagacct	6420
	gttaggtgc	gggttac	aaacacacca	gggttcccg	tctgcagg	ccatctggag	6480
	ttctgggaga	gcgttcc	aggcctc	cacatagac	cccat	gtcccagact	6540
	aagcaggcag	gagacaactt	cccttac	gtagcata	aggtacacgt	gtgcgc	6600
	gctcaggc	cacccatc	gtggacca	atgttgc	gtctcata	gttgcgc	6660
45	acgtgcac	ggccaaac	cctgtgt	aggctggag	ccgttcaaa	cgagggtact	6720
	accacacacc	ccataacaa	atacatcg	gcatgtat	cggtgcac	ggaggtcgtc	6780
	acgagcac	gggtgttgc	aggcgagtc	ctagcgtc	tggccgcgt	ttgcctgaca	6840
	acaggcagcg	tggtattgt	ggcaggatc	atcttgc	gaaaggccgc	catcattccc	6900
	gacagggaa	tccttacc	ggagttcgat	gagatggaa	agtgcgc	acaccc	6960
50	tacatcyaac	agggaaat	gctgcgc	caattca	agaaggca	cggttgc	7020
	caaacagcc	ccaagca	ggaggct	gctcc	tggatcc	gtggcggacc	7080
	atcgaagcc	tctggc	gcata	aat	gcgggata	atatttgc	7140
	ggcttgc	ctctgc	caaccc	atag	tgatgc	cacagcc	7200
	atcaccagcc	cgctcacc	ccaa	at	acat	ctgg	7260
	gccc	ttgtc	cac	ac	ggatgggt	7320	
55	gccc	ttgtc	cac	ac	ggatgggt	7380	
	gccc	ttgtc	cac	ac	ggatgggt	7440	
	gaggac	ttaac	ccct	at	ggatgggt	7500	

5 gtgtgcgcag cgatactgcg tcggcacgtg gccccagggg agggggctgt gcagtggatg 7560
 aaccgctga tagcgttgc ttgcggggt aaccacgtct ccccaacgca ctatgtgcct 7620
 gagagcgacg ctgcagcacf tgtactcgat ccctctcta gtcttaccat cactcagctg 7680
 ctgaagaggc ttcaccagtg gatcaacgag gactgctcca cggcatgctc cggctcgtag 7740
 ctaagagatg ttgggattt gatatgcacg gtgttactg atttcaagac ctggctccag 7800
 tccaaactcc tgccgcgatt gcccggatgc cccttcttct catgtcaacg tgggtacaag 7860
 ggagtctggc ggggcacgg catcatgcaa accacctgcc catgtggggc acagatcacc 7920
 ggacatgtga aaaacggttc catgaggatc gtggggccta ggacctgttag taacacgtgg 7980
 10 catgaaacat tcccccattaa cgcgtacacc acggggccct gcacgcccctc cccggcgcca 8040
 aattattcta gggcgctgtg gcgggtggct gctgaggagt acgtggaggt tacgcgggtg 8100
 ggggatttcc actacgtgac gggcatgacc actgacgacg taaagtgcctt gtgtcaggtt 8160
 cccggcccccg aattcttccac agaagtggat ggggtgcggg tgcacaggta cgctccagcg 8220
 tgcaaaccctt ccctacggga ggaggtcaca ttccctggctg ggctcaatca atacctgggtt 8280
 gggcacagc tcccatgcca gcccgaacccg gatgttagcag tgctcacttc catgctcacc 8340
 gacccttccc acattacggc ggagacggct aagcgttaggc tggccagggg atctcctccc 8400
 cccttggcca gctcatcagc tagccagctg tctgcgcctt ctttgaaggc aacatgcact 8460
 acccgcatg actcccgga cgctgacccctc atcgaggcca acctccctgtg gcggcaggag 8520
 atggccggga acatcaccgg cgtggagtca gaaaataagg tagtaattttt ggactctttc 8580
 gagccctcc aagcggagga ggatgagagg gaagtatccg ttccggcgga gatcctgcgg 8640
 20 aggtccagga aattccctcg agcgatgccc atatgggcac gcccggatta caaccctcca 8700
 ctgttagagt ccttggaaaggga cccggactac gtccctccag tggtagacgg gtgtccattt 8760
 ccgcctgcca aggccctcc gataccaccc tcaacggagga agaggacggt tgcctgtca 8820
 gaatctaccg tgccttctgc cttggcgag ctgcacccacag agaccttcgg cagctccgaa 8880
 tcgtccggccg tcgacagcgg cacggcaacg gcctctccctg accagccctc cgacgacgac 8940
 gacggggat ccgacgttga gtcgtactcc tccatgcccc cccttggaggg ggagccgggg 9000
 25 gatcccgatc tcaacggacgg gtcttggctt accgttaacgg aggaggctag tgaggacgtc 9060
 gtctgtctgt ctagtgccttta cacaatggaca ggcgccttgc tcacgcccattt cgctgcggag 9120
 gaaaccaaggc tgcccatcaa tgcactgagc aactcttttc tccgtcacca caacttggtc 9180
 tatgctacaa catctcgacgg cgcaacccctg cggcagaaga aggtcacctt tgacagactg 9240
 caggctctgg acgaccacta ccgggacgtg ctcaaggaga tgaaggcgaa ggcgtccaca 9300
 30 gtttaaggcta aacttctatc cgtggaggaa gcctgttaacg tgacgcccccc acattcgccc 9360
 agatctaaat ttggcttatgg ggcaaggagac gtccggaaacc tatccagcaaa ggcgttaac 9420
 cacatccgct ccgttgaaaa ggacttgcgtg gaagacactg agacaccaat tgacaccacc 9480
 atcatggcaa aaaatggat tttctgcgtc caaccagaga agggggggccg caagccagct 9540
 cgccttatcg tattcccaaaa tttgggggtt cgtgtgtgcg agaaaatggc ctttacgt 9600
 gtggcttcca ccctccctca ggccgtatgt ggctcttcat acggattcca atactctct 9660
 35 ggacacgggg tcgagttcctt ggtgaatgcc ttggaaagcga agaaaatgccc tatggcttc 9720
 gcatatgaca cccgcgttt tgactcaacg gtcactgaga atgacatccg tggtagggag 9780
 tcaatctacc aatgttgtga cttggcccccc gaaaggccagac aggccataag gtcgtccaca 9840
 gagcggctttt acatcggggg cccctctactt aattctaaag ggcagaactg cggctatcgc 9900
 cgggtcccgccg cgagcgggtt actgacgacc agctgcggta ataccctcac atgttactt 9960
 40 aaggccgctg cggcctgtcg agctgcgaag ctccaggact gcacatgtct cgtatgcgg 10020
 gacgaccccttgc tgcgttatctg tgaaagcgcg gggacccaaag aggacgaggc gaggctacgg 10080
 gccttcacgg aggctatgac tagatactctt gccccccctg gggacccggcc caaaccagaa 10140
 tacgacttgg agttgataac atcatgtctcc tccaaatgtt cagtcgcgc cgtatgtct 10200
 ggcaaaaagggt tgcgttatctt caccgtgac cccaccaccc cccttgcgcg ggctgcgtgg 10260
 gagacagacta gacacactcc agtcaattcc tggcttaggca acatcatcat gtatgcgcc 10320
 45 accttggggg caaggatgtt cctgtatgact catttcttccat ccatttttttctt agctcaggaa 10380
 caacttgaaa aagccctaga ttgtcagatc tacggggcctt gttactccat tgagccactt 10440
 gacctacccctc agatcattca acgactccat ggccttagcg cattttctactt ccatagtttac 10500
 tctccagggtt agatcaatag ggtggcttca tgcctcagaa aacttggggt accggccctt 10560
 cgagtcgttgcgacatcgccg cagaagtgtc cgcgcctactt tactgtccca ggggggggggg 10620
 gctgccactt gtggcaagttt ccttcaac tggggcgttta gggaccaagct caaactctact 10680
 50 ccaatcccggtt ctgcgtccca gttggattta tccagctgtt tcgttgcgttgg ttacagcg 10740
 ggagacatataatc acacagactt gtctcgttgc cggcccttgc gtttcatgtt gtgcctactc 10800
 ctactttctgtt taggggttagg catctatcta ctcccaacc gatgaacggg gagctaaaca 10860
 ctccaggccca ataggccatc ctgtttttt ccctttttt tttttttttt tttttttttt 10920
 tttttttttt tttttttttt ccctttttt ttcctttttt tttcccttttcc tttcccttttgg 10980
 tggctccatc ttagcccttag tcacggcttag ctgtgaaagg tccgtgagcc gcttgactgc 11040
 55 agagagtgttgc gatactggcc tctctgcaga tcaagt 11076

5 <210> 10
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

 <400> 10
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 ttttacgc gaaagcgct agccatggcg ttatgtatgag tgctgtgcag cttccaggac 120
 cccccctccc gggagagca tagtggctg cgaaaccgg gatgtacaccg gaattgccag 180
 .gacgaccggg tcctttctt gataaccccg ctcaatgcct ggagatttg gctgtcccc 240
 gcgagactgc tagccgagta gtgtgggtc gcaaaaggcc ttgtgtact gctgtatagg 300
 gtgttcgca gtccccccgg aggtctcgta gaccgtgcac catgagcacy aatcctaaac 360
 ctc aaagaaa aaccaaacs aacaccaacg ggcgcgcctt gattaaaca gatggattgc 420
 acgcgggtc tccggccgt tgggggaga ggtatccgg ctatgactgg gacaaacaga 480
 caatcgctg ctctatgtcc gccgttcc ggtgtcagc gacggggcgc cccgttctt 540
 ttgtcaagac gcacccgtcc ggtggccctt atgaaactgca ggacgaggca ggcggctat 600
 cgtggctggc cacgacggc gttcccttgc cagctgtctt cgcgttgc actaaggcg 660
 gaaggactg gctgtattt ggcgaaatgc cggggcagga ttcctgtca ttcacccctt 720
 ctccgtccg gaaagtatcc atcatggctg atgaaatgcg gcccgtcat acgcttgatc 780
 cggttacctg cccattcgac caccacggca aacatcgat cggcggagca cgtactcgga 840
 tggaaaggccg tcttgcgtat caggatgtc tggacgaa gcatcagggg ctcggccag 900
 ccgaactgtt cgccaggctc aaggcgcgc a tggccgcacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgttgcgg aatatcatgg tggaaaatgg cgcgtttctt ggattcatcg 1020
 actgtggccg gctgggtgt gcccggcgt atcaggatc acgttggct acccggtata 1080
 ttgtcaaga gcttggccgc gaatgggtg accgttctt cgtgtttac ggtatcgccg 1140
 ctcccgattt cgcgcgcata gccttcttgc gccttcttgc cgagttttc tgatgtttaa 1200
 cagaccacaa cgggttccctt ctgcggat caatccgc cctcccttcc cccccccccc 1260
 aacgttactg gccgaaggccg cttggataa ggcgggtgt cgtttgttgc tatgttattt 1320
 tccaccat tggcgtttt tggcaatgtg aggccccgg aacctggccc tttttttttt 1380
 acgacattt ctgggtt tttcccttgc gccaaggaa tgcgggtt gttttttttt 1440
 gtaaggaag cgttccctt ggaaccttgc tgaagacaaa caacgttgc tgcggccctt 1500
 tgaggcgc ggaacccccc acctggcgc aggtgcctt gcccggaaa gcccacgttgc 1560
 taagatacac ctgcaaggc ggcacaccc cagtgccacg ttgtgagtt gatagttgtt 1620
 gaaagagtca aatggcttc ctcaagcgta ttcaacaagg ggctgaaggaa tggccagaag 1680
 gtacccatt gtatggatc tgatctgggg ctcgggtca catgttttac atgtgttttag 1740
 tcgaggttaa aaaacgtcta ggccccccgaa accacggggc cgtgggtttt ctttggaaaa 1800
 cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
 tgcattatca ctggccttac aggggggac aggaaccagg tcgagggggaa ggttcaagg 1920
 gtcctccaccg caacacaatc tttccctggcg acctggcgtca atggcggtt tggactgtc 1980
 tatcatggt cccggctcaaa gacccttgc ggcggaaagg gcccacatc ccaaatgtac 2040
 accaatgtgg accaggaccc cgtgggttgc caagcgcccc cggggggcgc ttccgtaca 2100
 ccatgcacct gggcgaccc ggaccccttac ttgggtcaca ggcacgttgc tgcatttccg 2160
 gtgcggccgc gggggcagac caggggggcacttacttcc cctctggggc acgtgttggg catctttccg 2220
 aagggtctt cggggcggttcc actgttgc tttcccttgc cccctggggc acgtgttggg catctttccg 2280
 gtcggcgtgt gcaaccggg ggttgcgaag gcccggact ttgttcccttgc agtacttatg 2340
 ggaaccacta tgcggtcccc ggttccatc gacaacttgtt cccctccggc ctttccgttgc 2400
 acattccagg tggcccatc acacggccctt actggtagcg gcaagagcacttcaacttca 2460
 gtcgtatc caggccaaagg gtataagggtt cttgttccatc acccgccgtt cggccggccacc 2520
 cttagttcg gggcgatata gtttgcgtatc catgttacccg accctaaatc cagaatcggtt 2580
 gtaaggacca tcaccacggg tgccccatc acgtacttca cctatggca gtttcttgc 2640
 gacgggtgtt gtcgtggggg cgcctatgc atcataat gtttgcgtatc ccactcaactt 2700
 gactcgacca tttatccggg catggcaca gtcctggacc aagcggagac ggctggagcg 2760
 cgaactcgatc tgctccac cgcgttgcctt ccgggatccg tcacccgttgc acatccaaac 2820
 atcgaggagg tggctctgtc cagactgttgc gaaatccctt tttatggca agccatcccc 2880
 atcgagacca tcaagggggg gaggcaccc tttttcttgc atttccaaatc gaaatgttgc 2940
 gagctcgccg cgaaggatgc cggccctcgga gatgtatccg tcatccaaatc tagccggggcctt 3000
 acgggcttta cccggcgactt cgcactcgatc gtcgtatc gtcattgtcg tagcaacggc cgctctaaatg 3060
 gtcgtatc gctggggacc gacccatc gtttgcgtatc atcgacttgc atacatgttgc caccggacca 3120
 gtttgcgtatc cgcacggcg aggccggactt gtttgcgtatc attgagacca cgcggccgtt acacggacgc 3180
 ggttggggcc ggtatggccat ttcagggtt cttgttccatc cgggttccatc gtttgcgtatc 3240
 gtttgcgtatc cgggttccatc cgggttccatc cgggttccatc cgggttccatc gtttgcgtatc 3300
 gacggccggct gtttgcgtatc cgggttccatc cgggttccatc cgggttccatc gtttgcgtatc 3360
 tacctaaaca caccagggtt gcccgttgc cggaccatc tggagggttgc gggagacgc 3420
 ttttacaggcc tcaccacat agacggccat ttcttgc ttttgc ttttgc ttttgc 3480
 aacttccctt acctggtagc ataccaggct acgggttgc cccgggttgc ggcaggac 3540

5 cgggcccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgca 7620
tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
agcctgtctc gtgccccgacc ccgctgggttc atgtgggtgcc tactctact ttctgttaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggc当地atagg 7800
ccatccctgtt tttttccctt ttttttttcc tttttttttt tttttttttt 7860
ttttctccctt tttttttccctt ctttttttcc ttttttttcc tttgggtggct ccatcttagc 7920
10 cctagtcacg gctagctgtt aaagggtccgt gagccgcttg actgcagaga gtgctgatac 7980
tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

5 <210> 11
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

<400> 11

```

gccaagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtcag cctccaggac 120
ccccctccc gggagagccca tagtggctcg cggAACCGGT gagtacaccg gaattgccag 180
gacgaccggg tcctttctt gatcaacccg ctcaatgcct ggagatttg gctgtccccc 240
gcgagactgc tagccgagta gtgtgggtc gcgaaaggcc ttgtggtaact gcctgatagg 300
gtgttgcga gtgcggccgg aggttcgtca gaccgtcag catgacacg aatcctaaac 360
ctcaaagaaa aaccaaaccg aacaccaacg ggcgcgcccccat gattgaacaa gatggattgc 420
acgcagggtc tcggccogct tgggtggaga gcttattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgtatcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttctt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgcg ggacgaggca ggcgcgctat 600
cggtggctggc cacgacggg gttccttgcg cagctgtgtcgt cgacgttgc actgaagcgg 660
gaagggactg gctgctattt ggcgaagtgc cggggcagga ttcctgtca ttcacccctt 720
ctctgtccga gaaagtatcc atcatggctg atgcaatgcg gggctgcattt acgcttgcattt 780
cggttacccgt cccattcgac caccaagcga aacatcgcat cgagcggca cgtactcgg 840
tggaaaggccg tcttgcgtat caggatgtc tggacgaaga gcatcagggg ctgcgcggcag 900
ccgaactgtt cggccagggtc aaggcgcga tgccccacgg cgaggatctc gtcgtgaccc 960
atggcgtatgc ctgttgcgg aatatcatgg tggaaaatgg cgccttttctt ggttcatcg 1020
actgtggccg gctgggtgtg gcggaaccgat atcaggacat acgcttgcgtt acccgtgtata 1080
ttgtgttgcgaa gcttggccgc gaatgggctg accgcttccct cgtgttttac ggtatcgccg 1140
ctcccgattt cgcagcgcattt gccttcttgc gccttcttgc cgaggatctt ttagttaaa 1200
cagaccacaa cggtttccctt cttagggatg caattccggcc cctctccctt ccccccctt 1260
aacgttactg gcccggccg ctttgaataaa ggccgggtgtg cgttttgcattt tattttttt 1320
tccaccatattt tgccgttccctt tggcaatgtg aggccccggg aacctggccc tttttttt 1380
acgaggatcc cttaggggttctt tttcccttc gccaaggaaa tgcagggttgc ttgtatgtc 1440
gtgaaggaaag cagttccctt ggaagcttctt tgaagacaaa aacgtctgt acgcaccctt 1500
tgcaggcagc ggaacccccc acctggcgac aggtgcctctt cggccaaaa gccacgtgt 1560
taagatacac ctttgcggccgaaatccctt ctttgcggccgaaatccctt tttttttt 1620
gaaagagtca ctgcaaggcc ggcacaaccc cagtgccacg ttgtgatgtt gatagttgt 1680
gtaccccaattt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgttttt 1740
tcgagggttaa aaaacgtcta ggccccccga accacggggg cgtgttttccctt tttttttt 1800
cacgataataa ccatgggcac gaatcttaaa cctcaaaagaa aaacccaaacg taacaccaac 1860
cgccggcccac aggacgtcaaa gttccggggc ggtggtcaga tcgtcggtgg agtttacctg 1920
ttggcgcgca gggggcccaag gttgggtgtg cgcgcgacta ggaagacttc cgagcgggtcg 1980
caacctcggtg gaaggcgcaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccg ggttcccttgc gccccttat gcaatgggg gcttgggtg ggcaggatgg 2100
ctctgttcac cccgtggctc tcggctagt tggggccccca cggacccccc gcttaggtcg 2160
cgaattttgg gtaagggtcat cgataccctt acgtcggttgc tcggccatctt catgggtac 2220
atcccgctcg tcggccccc cttagggggc gtcggcagggg ccctggcgca tggcgtccgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaaatctgc cgggttgcctt cttttttt 2340
ttccctttgg ctttgcgttgc ctgtttgcacc atcccgatctt ccgttatgtt agtgcgttac 2400
gtatccggag ttttgcgttgc ctgtttgcacc atcccgatctt ccgttatgtt agtgcgttac 2460
gcccggatgtca tcatgcatac cccgggtgc gtcggcccttgc ttccggggaa caactccctt 2520
cgctgttggg tagcgctcac tcccaacgctc gcccggccaggaa acgctagctt ccccaactacg 2580
acgataacgac gccatgtcgaa tttgcgttgc ttccggccgtt ctctgttgc cgtatgtac 2640
gtggggagatc tctgcggatc ttttgcgttgc ttccggccgtt ccgttatgtt agtgcgttac 2700
cgccacgaga cagttacggaa ctgcaatttgc tcaatataatcc cggccacgt gacaggttac 2760
cgtatggctt gggatgtat gatgaactgg tcacccatag cagccctatgtt ggtatcgatgg 2820
ttactccggaa tcccaacaaatcc ttttgcgttgc ttccggccgtt ccgttatgtt agtgcgttac 2880
gcggggcccttgc ctttgcgttgc ttccggccgtt ccgttatgtt agtgcgttac 2940
ctctttggccg gctgttgcgg gggaaacctat gtcggccgggg ggacgttgc caaaaacacc 3000
ctcggttgcgg ctttgcgttgc ttccggccgtt ccgttatgtt agtgcgttac 3060
aacggcgtatcc ggcacatcaa caggactggc ctgaaactgc atgactccctt caacactggg 3120
ttcccttgcgtc ctttgcgttgc ttccggccgtt ccgttatgtt agtgcgttac 3180
gcacgtcgaa gcccatgtcgaa ttttgcgttgc ttccggccgtt ccgttatgtt agtgcgttac 3240
tcacacagatc cggaccaggatc gtcggccgtt ccgttatgtt agtgcgttac 3300
gtaccccgccg cgcagggtgtg tggccgttgc ttccggccgtt ccgttatgtt agtgcgttac 3360
ggggacgaccg accgggttccgg ctttgcgttgc ttccggccgtt ccgttatgtt agtgcgttac 3420
ctgttttgcgttca acaacacgac gccggccgtt ccgttatgtt agtgcgttac 3480
aggactgggtt tcaccaagac gtcggccgtt ccgttatgtt acatcgccggg gatcgccgtt 3540

```

	aaaaccttga	cctgccccac	ggactgcttc	cgaaagcacc	ccgaggcac	ttacaccaag	3600
	tgtggttcgg	gcccttggtt	gacacccaga	tgcttggtcc	actaccata	caggctttgg	3660
5	caactaccct	gcactgtcaa	cttattaccat	tcaggttta	ggatgtacgt	ggggggagtg	3720
	gagcacaggc	tcgaagccgc	atgcaatttg	atcgaggag	agcgtttaa	cctggaggac	3780
	aggacagat	cagagcttag	cccgtgtcg	ctgtctacaa	cggagtggca	gttattggcc	3840
	tgttccttca	ccaccctta	ggctgttcc	actggttta	tccatctcca	tcagaacgtc	3900
	gtggacgtac	aatacgtta	cgtttaggg	tcggcggtt	tctccttgc	aatcaaattgg	3960
	gagtatgtcc	tgttgcctt	ccttcttctg	gccccgcgc	gctgttgtc	ctgttgtgg	4020
10	atgatgtcgc	tgtatgtca	agctgaggcc	gccctagaga	acctgggtt	cctcaacgcg	4080
	gcatccgtgg	ccggggcgca	tggcattctc	tccttcctcg	tgttcttctg	tgctgcctgg	4140
	tacatcaagg	gcaggctgtt	ccctggggcg	gcataatgccc	tctacggcgt	atggccgcta	4200
	ctcctgttcc	tgctggcggt	accaccacga	gcatacgcga	tggaccggga	gatggcagca	4260
	tcgtcgaggag	gccccgggtt	cgttaggtctg	atactcttga	ccttgcacc	gcactataag	4320
	ctgttcctcg	ctaggctcat	atgggttta	caatattttt	tcaccagggc	cgaggcacac	4380
	ttgcaagtgt	ggatcccccc	cctcaacgtt	ccccggggcc	gcatgtccgt	catttccttc	4440
15	acgtgcgcga	tccaccacaa	gctaattttt	accatccaca	aaatctgtc	cgccataactc	4500
	ggtccactca	tggtcttca	ggctgttata	acccaaatgc	cgtacttcgt	gccccgcacac	4560
	gggcttccat	gtgtatgtc	gtgtgtcg	aagggtgtct	gggggttatta	tgtccaaatg	4620
	gctctcatga	atgtggccgc	actgacagg	acgtacgttt	atgaccatct	cacccactg	4680
	cgggactgggg	cccacgcggg	cctacgagac	cttgcgttgc	cagttgagcc	cgtcgttcc	4740
	tctgatatgg	agaccaaggt	tatcacctgg	ggggcagaca	ccggccgtg	tggggacatc	4800
20	atcttggggcc	tgcccgctc	cgcccgagg	gggaggggaga	tacatctggg	acccgcagac	4860
	agccttgaag	ggcagggggt	gcatcttc	gccccttata	cggctactc	ccaaacagacg	4920
	cgaggccctac	ttggctgtat	catcaactagc	ctcacaggcc	gggacaggaa	ccagggtcgag	4980
	ggggaggtcc	aagtggtctc	caccgaaca	aatctttcc	tggcaccctg	cgtaatggc	5040
	gtgtgttgg	ctgtctatca	tggtgcggc	tcaaagaccc	ttggcggccc	aaagggccca	5100
	atcacccaaa	tgtacaccaa	tgtgaccag	gacctgtcg	gctggcaagc	cccccccg	5160
25	ggcgcttcc	tgacaccat	cacctgcggc	agctcgacc	tttacttgg	cacgaggcat	5220
	gcccgttca	ttcccggtcg	ccccgggggc	gacagcagg	gggcctact	ctccccagg	5280
	cccgcttct	acttggagg	cttctgggc	gttccactgc	tctgcccctc	ggggcacgt	5340
	gtgggcattct	ttcgggctgc	cgtgtgcacc	cgaggggtt	cgaaggcggt	ggactttgt	5400
	cccgctcgagt	ctatggaaac	cactatgcgg	tcccccgtt	tcacggacaa	ctcgccccct	5460
	ccggcgatc	cgcagacatt	ccaggtggcc	catctacacg	cccctactgg	tagcggcaag	5520
30	agcactaagg	tgccggctgc	gtatgcaggc	caagggtata	agggtcttgc	cctgaaccgg	5580
	tccgtcgccg	ccaccctagg	tttccgggc	tatatgtcta	aggcacatgg	tatcgaccct	5640
	aacatcagaa	tcgggtaag	gaccatcacc	acgggtgccc	ccatcacgt	ctccacctat	5700
	ggcaagttt	ttgcccacgg	tggttgtct	gggggcgcct	atgacatcat	aatatgtat	5760
	gagtgcact	caactgactc	gaccactatc	ctgggcacatc	gcacgttct	ggaccaagcg	5820
	gagacggctg	gagcgcgact	cgtctgtc	gccccgccta	ccgctccggg	atcggtcacc	5880
	gtgccacatc	caaacatcga	ggaggtggct	ctgttccaga	ctggagaaat	ccccctttat	5940
35	ggcaagggca	tcccccattga	gaccatcaag	ggggggaggc	acccatcttt	ctggcattcc	6000
	aagaagaaat	gtgatgacgt	cgccgcgaag	ctgtccggcc	tcggactcaa	tgtctgtac	6060
	tattaccggg	gccttgcattgt	atccgtcata	ccaaactagcg	gagacgtcat	tgtctgtac	6120
	acggacgctc	taatgacggg	ctttaccggc	gacttcgact	cagtgtatcga	ctgcaataca	6180
	tgtgtcacc	agacagtcga	cttcagcctg	gaccgcacct	tcaccattga	gacgacgacc	6240
	gtgccacaag	acgcgggtgc	acgctcgcag	ccggcgaggc	ggactgttag	gggcaggatg	6300
40	ggcattttaca	ggtttgcac	tccaggagaa	ccggccctcg	gcatgttgc	ttcctcggtt	6360
	ctgtgcgagat	gctatgcacgc	gggcgtgtct	tggtaacgac	tcacgcggc	cgagacctca	6420
	gttaggttgc	gggcttacat	aaacacacca	gggttccccc	tctgcacca	ccatctggag	6480
	ttctggggaga	gcgttcttac	aggcctacc	cacatagacg	cccattttctt	gtcccgagact	6540
	aaggcaggcag	gagacaactt	cccttacctg	gtacgtatc	aggctacgg	gtgcgcagg	6600
	gctcaggcctc	cacccatcgt	ttggggacaa	atgtggaaat	gttccatc	gtctaaaggct	6660
45	acgctgcacgc	ggccaaacgc	cctgtgtat	aggctgggg	ccgttcaaaa	cgagggtact	6720
	accacacaccc	ccataaccaa	atacatcatg	gcatgtatgt	cggctgaccc	ggaggtcg	6780
	acgagcacct	gggtgcttgt	aggcggagtc	ctagcagctc	tggcccgct	ttgcctgaca	6840
	acaggcagcg	ttgtcattgt	gggcaggatc	atcttgcctc	gaaagccggc	catcatcccc	6900
	gacagggaaag	tccttaccg	ggagttcgat	gagatggaa	agtgcgcctc	acacccctt	6960
	tacatcgaac	agggaaatc	gctcgccaa	caattcaaa	agaaggcaat	cggttgcgt	7020
50	caaacagccca	ccaagcaagc	ggaggctgt	gtcccggtt	tggatccaa	gtggcgacc	7080
	atcgaagctt	tctggggaa	gcatatgtgg	aatttcatca	ggggatata	atatttagca	7140
	ggcttgcata	ctctgcattgg	caaccccgcg	atagcatcac	tgtatggcatt	cacaggctct	7200
	atcaccagcc	cgctcaccac	ccaacatacc	cttctgttta	acatccctgg	gggatgggtg	7260
	gcccggccaaac	ttgtcttcc	cagcgtgtct	tctgtttcg	taggcggcc	catcgcttgc	7320
	gcccgtgttgc	gcagcatagg	ccttggggaa	gtgttgttgc	atattttggc	agggttatgg	7380
55	gcaagggttgg	caggcgcgt	cgtggccctt	aaggtcatga	gccccggat	gccctccacc	7440
	gaggaccttgc	ttaacctact	ccctgtatc	ctctccccct	gcccctttagt	cgtcgggggt	7500

5 gtgtgcgcag cgatactgcg tcggcacgtg ggcccagggg agggggctgt gcagtggatg 7560
 aaccggctga tagcgttcgc ttgcgggggt aaccacgtct ccccccacgca ctatgtgcct 7620
 gagagcgcacg ctgcagcag agtcaactcg atccctctcta gtcttaccat cactcagctg 7680
 ctgaagaggc ttaccaggc gatcaacgag gactgtccca cggccatgctc cggctcgtgg 7740
 ctaagagatg tttgggattt gatatgcacg gtgttgactg atttcaagac ctggctccag 7800
 tccaagctcc tgccgcgatt gccgggagtc cccttcttct catgtcaacg tgggtacaag 7860
 ggagtctggc ggggcgacgg catcatgcaa accacctgcc catgtggagc acagatcacc 7920
 ggacatgtga aaaacggttc catgaggatc gtggggccta ggacctgttag taacacgtgg 7980
 10 catggAACAT tccccattaa cgcgtacacc acggggccct gcaccccctc cccggcgcca 8040
 aattattctt aggcgctgtg gcgggtaggt gctgaggagt acgtggaggt tacgcgggtg 8100
 ggggatttcc actacgtgac gggcatgacc actgacaacg taaagtgcctt gtgtcaggtt 8160
 cgggccccccg aattttcac agaagtggat ggggtgcggg tgcacaggta cgctccagcg 8220
 tcaaaaaaaa ttctacgggaa ggaggtcaca ttcttgcggg ggctcaatca atacctgggtt 8280
 gggtcacagc tcccatgcga gcctgaaccg gatgttagcag tgctcacttc catgctcacc 8340
 gaccctcccc acattacggc ggagacggct aagcgttaggc tggccagggg atctcccccc 8400
 cccttggcca gtcatcagc tagccagctg ttcgcgcctt ctttgcggg aacatgcact 8460
 acccgcatg actccccggc cgctgacccctc atcgaggcca acctctgtg gcggcaggag 8520
 atgggggggaa acatcaccgg cgtggagtca gaaaataagg tagtaatttt ggactctttc 8580
 gagccgtcc aagcggagga ggatgagagg ggagtatccg ttccggcgg gatcctgcgg 8640
 aggtccagga aattccctcg agcgatgccc atatgggcac gcccggatttta caaccctcca 8700
 ctgttagagt ctttggaaaggc cccggactac gtccctccag tggtagacccg gtgtccattt 8760
 cccgcgtccca agggccctcc gataccaccc tcaacggagga agaggacggt tgcctgtca 8820
 gaatctaccg tgcgttctgc cttggcgggat ctcggccacag agacccctgg cagctccgaa 8880
 tcgtcgccg tcgacagcgg cacggcaacg gcctctccctt accagccctc cgacgacggc 8940
 gacgcgggat ccgcgttgc gtcgtactcc tccatgcccc cccttgggggg ggagccgggg 9000
 gatcccgatc tcaagcgacgg gtcttggctt accgtaaacg aggaggctag tgaggacgtc 9060
 gtctgtgtct cgtatgccta cacatggaca ggccgcctga tcacccatg cgctgcggag 9120
 gaaaccacggc tccccatcaa tgcactgagc aactctttgc tccgtccacca caacttggc 9180
 tatgctacaa catctcgacg cgcaaacctg cggcagaaga aggtcacctt tgacagactg 9240
 caggtccctgg acgaccacta ccgggacgtg ctcaggaga tgaaggcggaa ggctccacca 9300
 30 gttaaggcta aacttctatc cgtggaggaa gcctgttaacg tgaccccccc acattcgccc 9360
 agatctaaat ttggctatgg ggcaaaaggac gtccggaaacc tatccagcaa ggccgttaac 9420
 cacatccgct ccgtgtggaa ggacttgctg gaagacactg agacaccaat tgacaccacc 9480
 atcatggcaa aaaaatgggt tttctgtgtc caaccagaga agggggggccg caagccagct 9540
 cgccttatcg tattccaga tttgggggtt cgtgtgtgcg agaaaatggc ctttacgat 9600
 gtggcttccca ccctccctca ggccgtgatg ggctcttcat acggatttcca atactctct 9660
 35 ggacagcggg tcgagttctt ggtgaatgcc tggaaagcga agaaaatgccc tatgggcttc 9720
 gcatatgaca cccgtgttt tgactcaacg gtcactgaga atgacatccg tggtagggag 9780
 tcaatctacc aatgttgta cttggccccc gaagccagac aggccataag gtcgtccaca 9840
 gagcggcttt acatcggggg cccctgact aattctaaag ggcagaactg cggctatcgc 9900
 cggtgcgcgc cgagcggtgt actgacgacc agctgcggta ataccctcac atgttactt 9960
 40 aaggccgctg cggcctgtcg agctgcgaag ctccaggact gcacgatgct cgtatgcgg 10020
 gacgaccttgc tcgttatctg taaaagcgcg gggacccaaag aggacgaggc gggctacgg 10080
 gccttacgg aggctatgac tagatactt cccccccctg gggacccggc caaaccagaa 10140
 tacgacttgg agttgataac atcatgtcc tccaatgtgt cagtcgcgcg cgtatgcac 10200
 ggcaaaaagggt tgactatctt caccgtgac cccaccaccc cccttgcggc ggctgcgtgg 10260
 gagacagcta gacacactcc agtcaattcc tggctaggca acatcatcat gtatgcgc 10320
 45 accttgggg caaggatgtt cctgtatgact catttcttccatccatccat agctcaggaa 10380
 caacttggaaa aagcccttaga ttgtcagatc tacggggctt gttactccat tgagccactt 10440
 gacccatctc agatcattca acgactccat ggccttagcg cattttccat ccatagttac 10500
 tctccagggt agatcaatacg ggtggcttca tgcctcaggaa aacttggggt accggccctt 10560
 cgagtcgtgg aacatcgggc cagaagtgtc cggcgttaggc tactgtccca gggggggagg 10620
 gtcgcactt gtggcaagta cctcttcaac tggcaggtaa ggaccaagct caaactcact 10680
 50 ccaaatccgg ctgcgtccca gttggattta tccagctggt tcgttgcgtgg ttacagcggg 10740
 ggagacatatac atcacagcc tgcgttgcctt cggcccttgcgtt gttcatgtt gtcctactc 10800
 ctactttctg taggggtagg catctatcta ctccccaacc gatgaacggg gagctaaaca 10860
 ctccaggcca ataggccatc ctgtttttt ccctttttt tttttttttt tttttttttt 10920
 tttttttttt tttttttttt tccctttttt ttcctttttt ttccttttc ttcctttttt 10980
 tggctccatc ttagccctag tcacgcttag ctgtgaaagg tccgtgagcc gcttgacttc 11040
 55 agagagtgtt gatactggcc tctctgcaga tcaagt 11076

Patentansprüche

1. Hepatitis-C-Virus (HCV)-RNA-Konstrukt mit der Fähigkeit zur Replikation in eukaryontischen Zellen, **dadurch gekennzeichnet**,
 5 **daß** es eine Nukleotidsequenz gemäß einem der Sequenzprotokolle SEQ ID NO: 1 bis SEQ ID NO: 11 umfaßt, welche zumindest für die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NSSB und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) kodieren, wobei das selektierbare Markergen das Neomycinphosphotransferasegen ist und wobei anstelle des Neomycinphosphotransferasegens ein anderes Antibiotikaresistenzgen oder sonstiges Resistenzgen enthalten sein kann.
 10
2. HCV-RNA-Konstrukt nach Anspruch 1, **dadurch gekennzeichnet**,
daß es anstelle des Markergens oder zusätzlich zu dem Markergen ein integriertes Reportergen aufweist.
 15
3. HCV-RNA-Konstrukt nach Anspruch 2, **dadurch gekennzeichnet**,
daß das Reportergen ausgewählt ist aus der Gruppe bestehend aus den Luziferasegenen, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidase-Gen), den GFP-Genen (green-fluorescence-protein-Genes), dem GUS-Gen (Glukuronidasegen) und dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen).
 20
4. HCV-RNA-Konstrukt nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, **daß** dessen Replikation die Expression eines (zellulären) Surrogatmarkergens beeinflußt.
 25
5. HCV-RNA-Konstrukt nach einem der Ansprüche 1 bis 4, **dadurch gekennzeichnet**, **daß** das Reportergen und das selektierbare Markergen derart räumlich in dem Konstrukt angeordnet sind, daß sie gemeinsam ein Fusionsprotein exprimieren.
 30
6. HCV-RNA-Konstrukt nach einem der Ansprüche 1 bis 5, **dadurch gekennzeichnet**, **daß** es ein integriertes Fremdgen aufweist und dazu geeignet ist, dieses Fremdgen in eine Zielzelle einzuschleusen, die zur Expression dieses Fremdgens geeignet ist.
 35
7. HCV-RNA-Konstrukt nach einem der Ansprüche 1 bis 6, **dadurch gekennzeichnet**, **daß** es Nukleotid- und/oder Aminosäure-Mutationen aufweist, zellkultur-adaptiert ist, mit hoher Effizienz repliziert und dadurch erhältlich ist, daß man ein Zellkultursystem gemäß Anspruch 11, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 1 bis 6 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen das HCV-RNA-Konstrukt oder Teile davon isoliert.
 40
8. HCV-RNA-Konstrukt nach Anspruch 7, **dadurch gekennzeichnet**, **daß** man das aus den Zellklonen isolierte HCV-RNA-Konstrukt oder Teile davon wenigstens einmal erneut passagiert, nämlich in Zellen des Zellkultursystems gemäß Anspruch 11 einschleust und diese Zellen auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen das HCV-RNA-Konstrukt oder Teile davon isoliert.
 45
9. HCV-RNA-Konstrukt gemäß einem der Ansprüche 1 bis 8, **dadurch gekennzeichnet**,
daß es einen oder mehrere der nachfolgend aufgeführten Aminosäureaustausche aufweist, nämlich 1283 arg -> gly und/oder 1383 glu -> ala und/oder 1577 lys -> arg und/oder 1609 lys -> glu und/oder 1936 pro -> ser und/oder 2163 glu -> gly und/oder 2330 lys -> glu und/oder 2442 ile -> val aufweist, zellkultur-adaptiert ist und mit hoher Effizienz repliziert.
 50
10. HCV-RNA-Konstrukt gemäß einem der Ansprüche 1 bis 9, **dadurch gekennzeichnet**,
daß es einen oder mehrere der in Tabelle 3 aufgeführten Nukleotid- und/oder Aminosäureaustausche aufweist, wobei Tabelle 3 Bestandteil dieses Anspruchs ist.
 55
11. HCV Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, **dadurch gekennzeichnet**, **daß** die eukaryontischen Zellen humane Hepatomazellen sind und daß das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 1 bis 10 ist.

12. Zellkultursystem nach Anspruch 11, **dadurch gekennzeichnet**,
daß die das HCV-RNA-Konstrukt enthaltenden Zellen bei der DSMZ, Braunschweig, BRD, unter der Hinterlegungsnummer DSM ACC2394 (Laborbezeichnung HuBl 9-13) hinterlegt sind.
- 5 13. Verwendung eines Zellkultursystems nach einem der Ansprüche 11 oder 12 und/oder eines HCV-RNA-Konstrukts
nach einem der Ansprüche 1 bis 10 zur Herstellung und/oder Evaluierung von Therapeutika und/oder Diagnostika
für HCV-Infektionen.
- 10 14. Verwendung eines Zellkultursystems nach einem der Ansprüche 11 oder 12 und/oder eines HCV-RNA-Konstrukts
nach einem der Ansprüche 1 bis 10 zur Herstellung eines Impfstores gegen HCV-Infektionen.
- 15 15. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 1 bis 10 zur Herstellung einer leberzellspezifischen Genfähre für die Gentherapie.
- 20 16. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 1 bis 6 zur Gewinnung von zellkultur-adaptierten Mutanten dieses HCV-RNA-Konstrukts, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, **dadurch gekennzeichnet**,
daß man ein Zellkultursystem gemäß Anspruch 11 oder 12, das HCV-RNA-Konstrukte gemäß einem der Ansprüche 1 bis 6 enthält, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen die HCV-RNA-Konstrukte oder Teile davon isoliert.
- 25 17. Verwendung eines HCV-RNA-Konstrukts nach Anspruch 16, **dadurch gekennzeichnet**,
daß man die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 11 oder 12 einschleust und diese auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone erntet und aus diesen Zellklonen die HCV-RNA-Konstrukte oder Teile davon isoliert.
- 30 18. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 1 bis 6 zur Herstellung von Mutanten eines HCV-Vollängengenoms oder eines HCV-Teilgenoms oder eines beliebigen HCV-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, **dadurch gekennzeichnet**,
- 35 - daß man gemäß Anspruch 16 oder Anspruch 17 eine zellkultur-adaptierte Mutante des HCV-RNA-Konstrukts herstellt,
- daß man die Nukleotid- und Aminosäuresequenz dieser Mutante bestimmt und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Nukleotid- und Aminosäuremutationen bestimmt,
40 - und daß man diese Mutationen entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten, in ein (isoliertes) HCV-Vollängengenom oder ein HCV-Teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.
- 45 19. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 1 bis 10 zur Erzeugung von Hepatitis C Viruspartikeln oder virus-ähnlichen Partikeln.

Claims

1. Hepatitis-C-virus (HCV) RNA construct with the ability to replicate in eukaryotic cells, **characterized in**
that it comprises a nucleotide sequence in accordance with one of the sequences listed as SEQ ID NO: 1 to SEQ ID NO: 11, which codes at least for the HCV specific RNA-segments 5'NTR, NS3, NS4A, NS4B, NS5A, NS5B and 3'NTR and additionally for a marker gene for selection (selection gene), whereby the marker gene for selection is the neomycin phosphotransferase gene and whereby another antibiotic resistance gene or other selection gene may be contained in place of the neomycin phosphotransferase gene.
- 55 2. HCV RNA construct according to claim 1, **characterized in**
that it comprises instead of the marker gene or in addition to the marker gene an integrated reporter gene.

3. HCV RNA construct according to claim 2, **characterized in**
that the reporter gene is a gene from the group consisting of the luciferase genes, the CAT-gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP genes (green fluorescence protein genes), the GUS gene (glucuronidase gene) and the SEAP gene (secreted alkaline phosphatase gene).
- 5
4. HCV RNA construct according to claims 1 to 3, **characterized in**
that its replication influences the expression of a (cellular) surrogate marker gene.
- 10
5. HCV RNA construct according to claim 1 to 4, **characterized in**
that the reporter gene and the marker gene for selection are spatially arranged in the construct in such a way that they are expressed together as a fusion protein.
- 15
6. HCV RNA construct according to claim 1 to 5, **characterized in**
that it comprises an integrated foreign gene and can be used to introduce this foreign gene into a target cell which is suited for expressing this foreign gene.
- 20
7. HCV RNA construct according to claim 1 to 6, **characterized in that** it comprises nucleotide and/or amino acid mutations, that it is cell culture-adapted, that it replicates with high efficiency and that it is obtained by using a cell culture system according to claim 11, in which the transferred HCV specific gene material is an HCV RNA construct with a selection gene according to one of the claims 1 to 6, by culturing on/in the selection medium that is appropriate for the particular selection gene, by harvesting the grown cell clones, and by isolating the HCV RNA construct or parts thereof from these cell clones.
- 25
8. HCV RNA construct according to claim 7, whereby the isolated HCV-RNA-construct or parts thereof are subjected to at least one further passage, i.e. are introduced into cells of a cell culture system according to claim 11, these cells are cultivated on/in the selection medium that is appropriate for the particular selection gene, the grown cell clones are harvested and the HCV RNA construct or parts thereof are isolated from these cell clones.
- 30
9. HCV RNA construct according to one of the claims 1 to 8, **characterized in**
that it comprises one or several of the following specified amino acid exchanges, namely 1283 arg -> gly and/or 1383 glu -> ala and/or 1577 lys -> arg and/or 1609 lys -> glu and/or 1936 pro -> ser and/or 2163 glu -> gly and/or 2330 lys -> glu and/or 2442 ile -> val, that it is cell culture-adapted and that it replicates with high efficiency.
- 35
10. HCV RNA construct according to claim 1 to 9, **characterized in**
that it comprises one or several of the nucleotide and/or amino acid exchanges specified in table 3, whereby table 3 is constituent of this claim.
- 40
11. HCV cell culture system, composed essentially of eukaryotic cells, which contain transferred HCV specific genetic material, **characterized in that** the eukaryotic cells are human hepatoma cells and that the transferred HCV specific genetic material is an HCV RNA construct in accordance with one of the claims 1 to 10.
- 45
12. Cell culture system according to claim 11, **characterized in**
that the cells containing the HCV RNA construct are deposited at the DSMZ, Braunschweig, FRG, under the deposit number DSM ACC2394 (laboratory name HuBl 9-13).
- 50
13. A use of a cell culture system according to claim 11 or 12 and/or a HCV RNA construct according to one of the claims 1 to 10 for the production and/or evaluation of therapeutics and/or diagnostics for HCV infections.
14. A use of a cell culture system according to claim 11 or 12 and/or a HCV RNA construct according to one of the claims 1 to 10 for the production of a vaccine against HCV infections.
- 55
15. A use of an HCV RNA construct according to one of the claims 1 to 10 for the production of a liver-cell-specific vector in gene therapy.
16. A use of an HCV RNA construct according to one of the claims 1 to 6 for the production of cell culture-adapted mutants of this HCV RNA construct, whereby the mutants replicate with increased efficiency in comparison with the HCV RNA construct, **characterized in that** one cultivates a cell culture system in accordance with claim 11 or 12, which contains HCV RNA constructs in accordance with one of the claims 1 to 6 on/in the selection medium

that is appropriate for the particular selection gene, that one harvests the grown cell clones, and that one isolates the HCV RNA constructs or parts thereof from these cell clones.

5 17. A use of an HCV RNA construct according to claim 16, **characterized in**

that one performs at least one further passage with the isolated HCV RNA constructs, i.e. one introduces the isolated HCV RNA constructs into cells of a cell culture system according to claim 11 or 12 and cultivates these cells on/in the selection medium that is appropriate for the particular marker gene, one harvests grown cell clones and isolates the HCV RNA constructs or parts thereof from these cell clones.

10 18. Use of an HCV RNA construct according to one of the claims 1 to 6 for the production of mutants of an HCV full length genome or an HCV partial genome or any HCV construct having an increased replication efficiency in comparison with the original HCV full length genome or partial genome or HCV RNA construct, **characterized in**

- 15 - **that** one creates a cell culture-adapted mutant of the HCV RNA construct in accordance with claim 16 or claim 17,
- **that** one analyzes the nucleotide and amino acid sequence of this mutant and by comparison with the nucleotide and amino acid sequence of the original HCV RNA construct determines the type, number and positions of the nucleotide and amino acid mutations in which the original HCV RNA construct differs from this mutant,
- 20 - and **that** one introduces these mutations found in the mutant either by direct mutagenesis or by exchange of fragments of the HCV RNA construct, which contain the mutations concerned into an (isolated) HCV full length genome or an HCV partial genome or any HCV RNA construct.

25 19. A use of an HCV RNA construct according to one of the claims 1 to 10 for the production of hepatitis C virus particles or virus-like particles.

Revendications

30 1. Construction d'ARN du virus de l'hépatite C (VHC) avec la capacité de réPLICATION dans des cellules eucaryotes, **caractérisée en ce qu'elle comprend une séquence de nucléotides selon l'un des protocoles de la séquence SEQ ID N° : 1 à SEQ ID N° : 11, codant pour au moins les sections d'ARN spécifiques du VHC 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B et 3'NTR et en outre, un gène marqueur sélectionnable (gène de sélection), le gène marqueur sélectionnable étant le gène de la néomycine phosphotransférase et un autre gène résistant aux antibiotiques ou un autre gène résistant pouvant être contenu à la place du gène de néomycine-phospho-transférase.**

35 2. Construction d'ARN du VHC selon la revendication 1, **caractérisée en ce qu'elle présente, à la place du gène marqueur ou en complément du gène marqueur, un gène rapporteur intégré.**

40 3. Construction d'ARN du VHC selon la revendication 2, **caractérisée en ce que** le gène reporter est sélectionné dans le groupe constitué des gènes de la luciférase, du gène CAT (gène de la chloramphénicol acétyltransférase), du gène lacZ (gène bêta-galactosidase), des gènes GFP (gène de la protéine vert fluorescent), du gène GUS (gène de la glucuronidase) et du gène SEAP (gène de la phosphatase alcaline sécrétoire)

45 4. Construction d'ARN du VHC selon l'une des revendications 1 à 3, **caractérisée en ce que** sa réPLICATION influe sur l'expression d'un gène marqueur substitut (cellulaire).

50 5. Construction d'ARN du VHC selon l'une des revendications 1 à 4, **caractérisée en ce que** le gène reporter et le gène marqueur sélectionnable sont agencés de telle façon dans l'espace de la construction qu'ils expriment ensemble une protéine de fusion.

55 6. Construction d'ARN du VHC selon l'une des revendications 1 à 5, **caractérisée en ce qu'elle présente un gène étranger intégré et est apte à introduire ce gène étranger dans une cellule cible qui est apte à exprimer ce gène étranger.**

7. Construction d'ARN du VHC selon l'une des revendications 1 à 6, **caractérisée en ce qu'elle présente des mutations de nucléotide et / ou d'acides aminés, elle est adaptée à la culture cellulaire, elle se réplique avec une efficacité élevée et elle peut être obtenue en cultivant un système de culture cellulaire selon la revendication 11, dans lequel le matériel génétique spécifique au VHC introduit est une construction d'ARN du VHC avec un gène**

de sélection selon l'une des revendications 1 à 6, sur / dans le milieu de sélection correspondant au gène de sélection, en récoltant les clones cellulaires cultivés et en isolant de ces clones cellulaires la construction d'ARN du VHC ou des parties de celle-ci.

- 5 8. Construction d'ARN du VHC selon la revendication 7, **caractérisée en ce que** l'on effectue au moins une fois de nouveau un passage de la construction d'ARN du VHC isolée à partir des clones cellulaires ou des parties de celle-ci, à savoir que l'on introduit ladite construction d'ARN du VHC dans des cellules du système de culture cellulaire selon la revendication 11 et que l'on cultive ces cellules sur / dans le milieu de sélection correspondant au gène de sélection, que l'on récolte les clones cellulaires cultivés et que l'on isole de ces clones cellulaires la construction d'ARN du VHC ou des parties de celle-ci.
- 10 9. Construction d'ARN du VHC selon l'une des revendications 1 à 8, **caractérisée en ce qu'elle présente un ou plusieurs échanges des acides aminés suivants, à savoir 1283 arg-> gly et / ou 1383 glu-> ala et / ou 1577 lys-> arg et / ou 1609 lys-> glu et / ou 1936 pro-> ser et / ou 2163 glu-> gly et / ou 2330 lys-> glu et / ou 2442 ile-> val, qu'elle est adaptée à la culture cellulaire et elle se réplique avec une efficacité importante.**
- 15 10. Construction d'ARN du VHC selon l'une des revendications 1 à 9, **caractérisée en ce qu'elle présente un ou plusieurs échanges de nucléotides et / ou d'acides aminés figurant dans le tableau 3, le tableau 3 faisant partie intégrante de cette revendication.**
- 20 11. Système de culture cellulaire du VHC qui comprend essentiellement des cellules eucaryotes contenant un matériel génétique spécifique au VHC introduit, **caractérisé en ce que** les cellules eucaryotes sont des cellules d'hépatome humain et **en ce que** le matériel génétique spécifique du VHC introduit est une construction d'ARN du VHC selon l'une des revendications 1 à 10.
- 25 12. Système de culture cellulaire selon la revendication 11, **caractérisé en ce que** les cellules contenant la construction d'ARN du VHC sont enregistrées auprès du DSMZ, Braunschweig, RFA, sous le numéro d'enregistrement DSM ACC2394 (code laboratoire HuBl 9-13).
- 30 13. Utilisation d'un système de culture cellulaire selon l'une des revendications 11 ou 12 et / ou d'une construction d'ARN du VHC selon l'une des revendications 1 à 10 pour la production et / ou l'évaluation de produits thérapeutiques et / ou diagnostiques pour les infections au VHC.
- 35 14. Utilisation d'un système de culture cellulaire selon l'une des revendications 11 ou 12 et / ou d'une construction d'ARN du VHC selon l'une des revendications 1 à 10 pour la production d'un vaccin contre les infections au VHC.
- 40 15. Utilisation d'une construction d'ARN du VHC selon l'une des revendications 1 à 10 pour la production d'un vecteur génétique spécifique des cellules hépatiques pour la thérapie génique.
- 45 16. Utilisation d'une construction d'ARN du VHC selon l'une des revendications 1 à 6 pour obtenir des mutants de cette construction d'ARN du VHC adaptés à la culture cellulaire, les mutants présentant par rapport à la construction d'ARN du VHC une efficacité de réplication accrue, **caractérisée en ce que** l'on cultive un système de culture cellulaire selon la revendication 11 ou 12 contenant la construction ARN du VHC selon l'une des revendications 1 à 6, sur / dans le milieu de sélection correspondant au gène de sélection, que l'on récolte les clones cellulaires cultivés et que l'on isole de ces clones cellulaires des constructions d'ARN du VHC ou des parties de celles-ci.
- 50 17. Utilisation d'une construction d'ARN du VHC selon la revendication 16, **caractérisée en ce que** l'on effectue au moins une fois de nouveau un passage de la construction d'ARN du VHC isolée, à savoir que l'on introduit ladite construction d'ARN du VHC dans les cellules d'un système de culture cellulaire selon la revendication 11 ou 12 et que l'on cultive celles-ci sur / dans le milieu de sélection correspondant au gène de sélection, que l'on récolte les clones cellulaires cultivés et que l'on isole de ces clones cellulaires la construction d'ARN du VHC ou des parties de celle-ci.
- 55 18. Utilisation d'une construction d'ARN du VHC selon l'une des revendications 1 à 6, pour la production de mutants d'un génome entier du VHC ou d'un génome partiel du VHC ou d'une construction quelconque du VHC avec, par rapport au génome entier du VHC ou au génome partiel ou à la construction d'ARN du VHC d'origine, une efficacité de réplication accrue, **caractérisée en ce que**

➤ l'on produit selon la revendication 16 ou 17 un mutant adapté à la culture cellulaire de la construction d'ARN du VHC,

5 ➤ que l'on détermine la séquence de nucléotides et d'acides aminés de ce mutant et que l'on détermine, par comparaison avec la séquence de nucléotides et d'acides aminés de la construction d'ARN du VHC d'origine, la nature, le nombre et les positions des mutations des nucléotides et des acides aminés,

➤ et que l'on introduit ces mutations soit par une mutagenèse ciblée, soit par l'échange de sections de séquences contenant les mutations concernées, dans un génome entier du VHC (isolé) ou un génome partiel du VHC ou une construction d'ARN du VHC quelconque.

10 **19.** Utilisation d'une construction d'ARN du VHC selon l'une des revendications 1 à 10 pour la production de particules virales de l'hépatite C ou de particules de type viral.

15

20

25

30

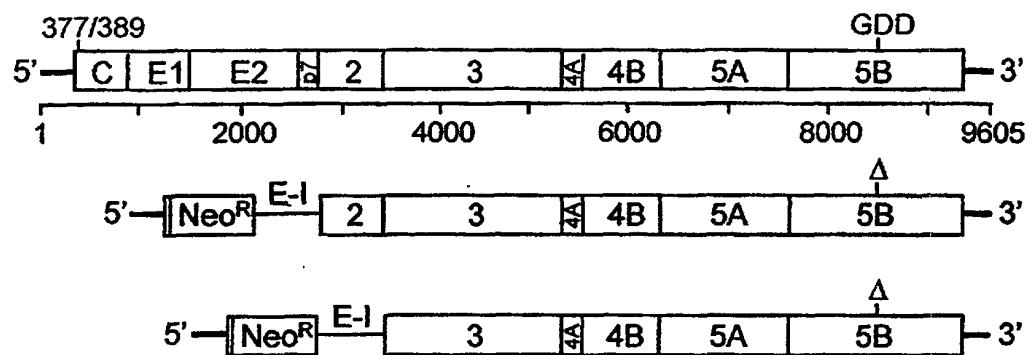
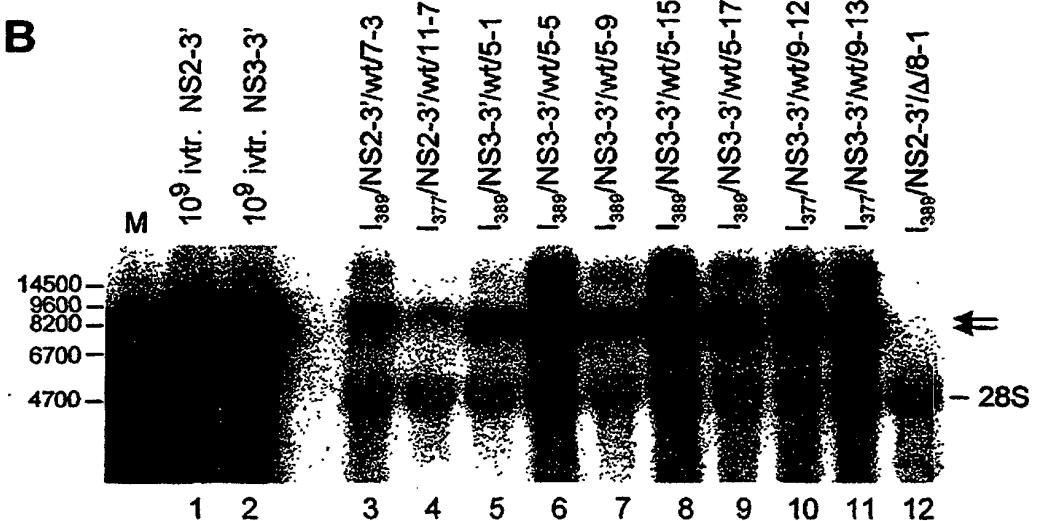
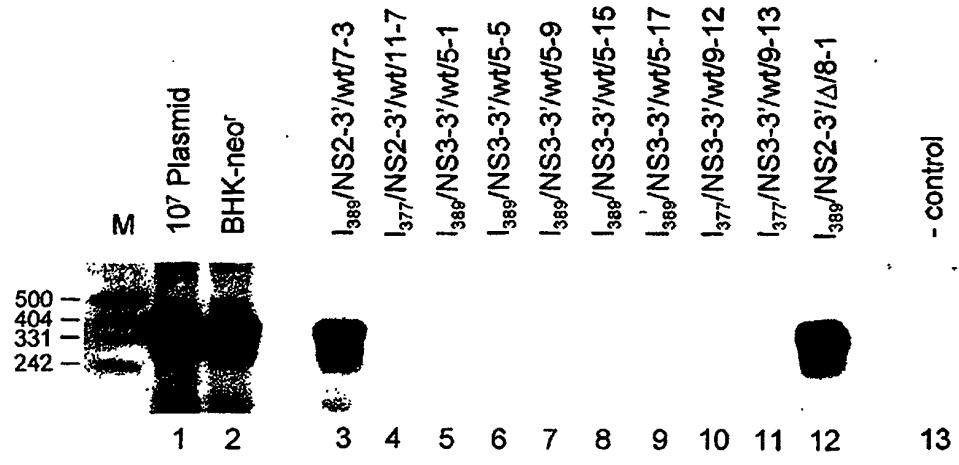
35

40

45

50

55

A**B****C****Fig. 1**

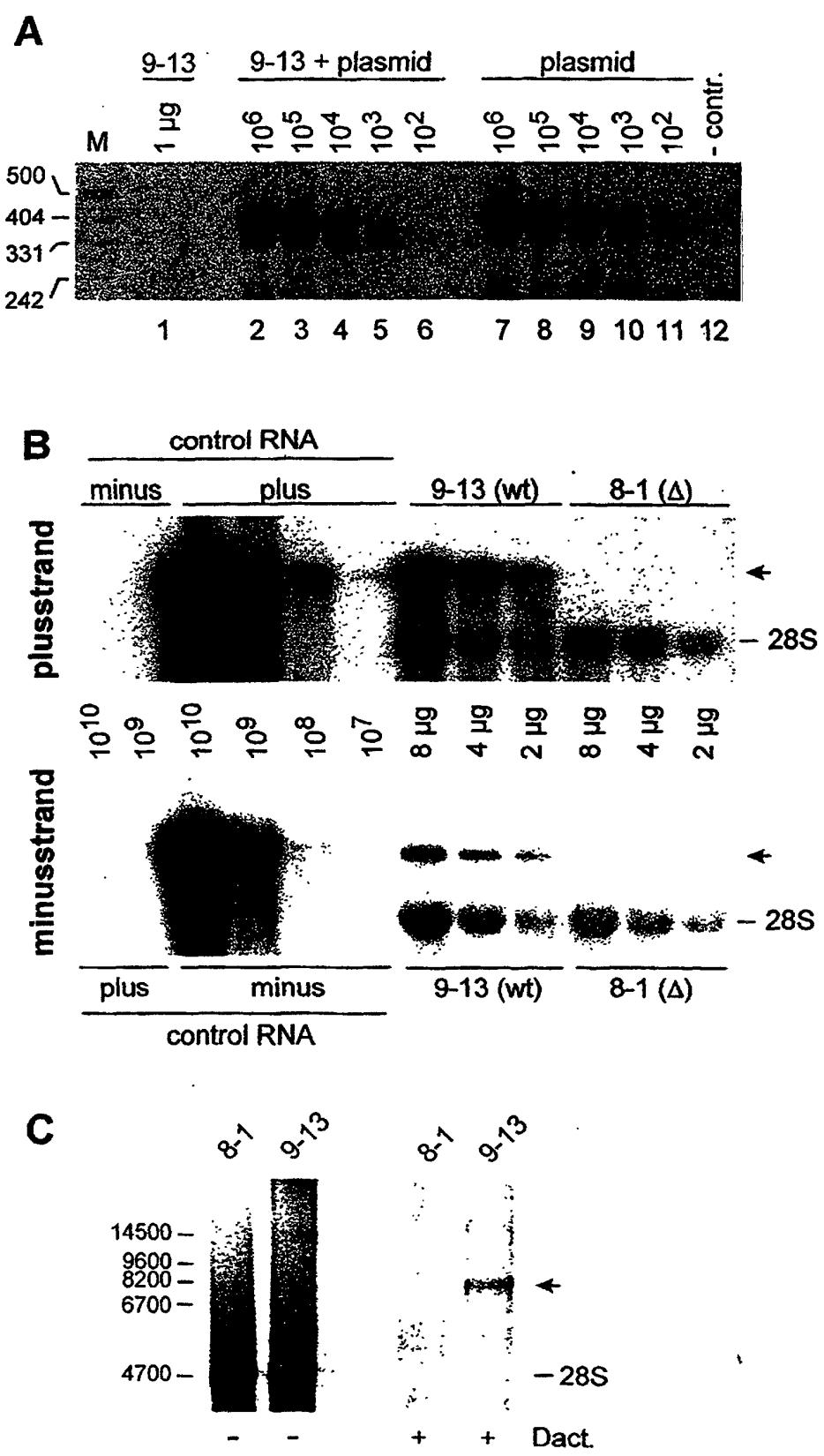
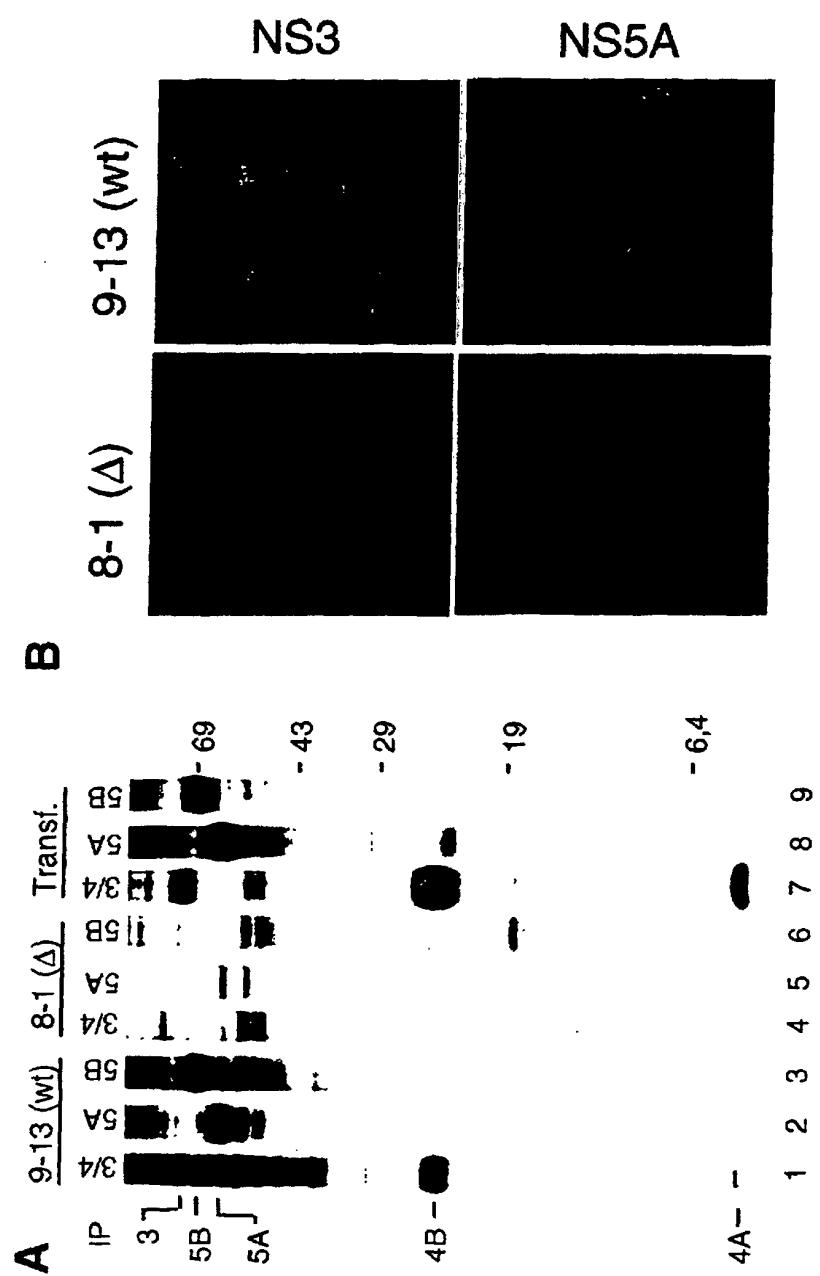


Fig. 2

FIG. 3



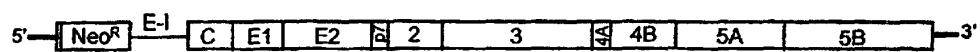
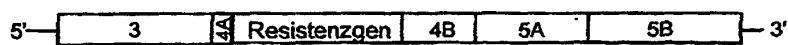


Fig. 4

A



B

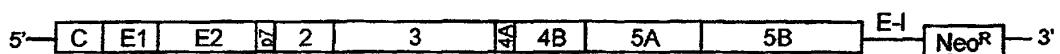
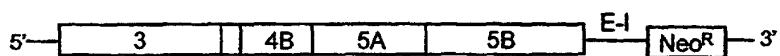
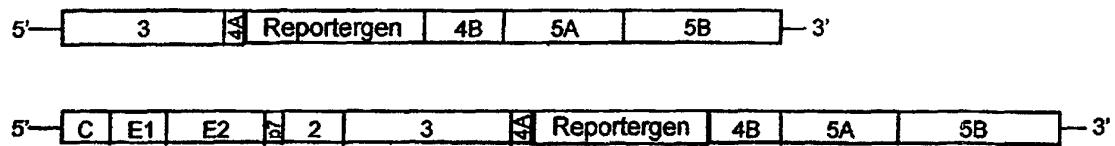
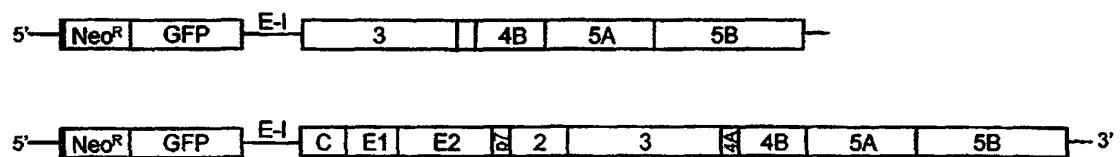


Fig. 5

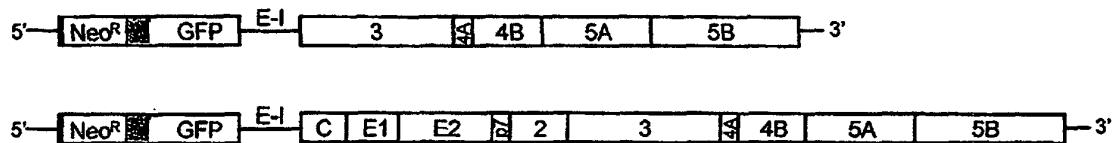
A



B



C



D

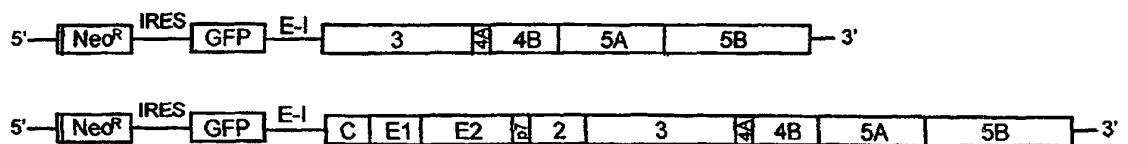


Fig. 6

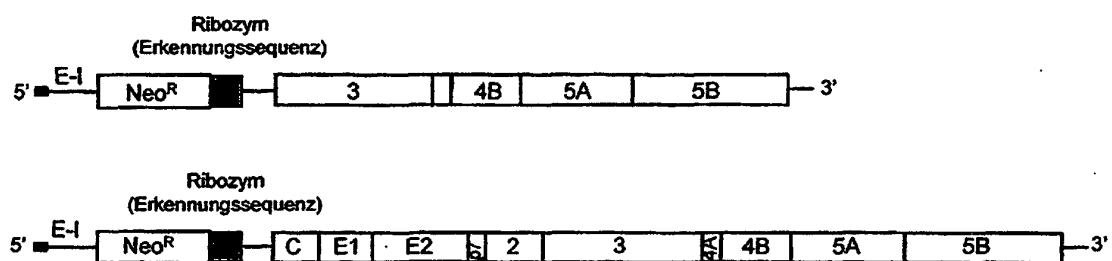


Fig. 7

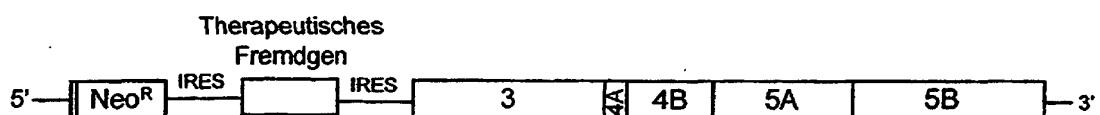


Fig. 8

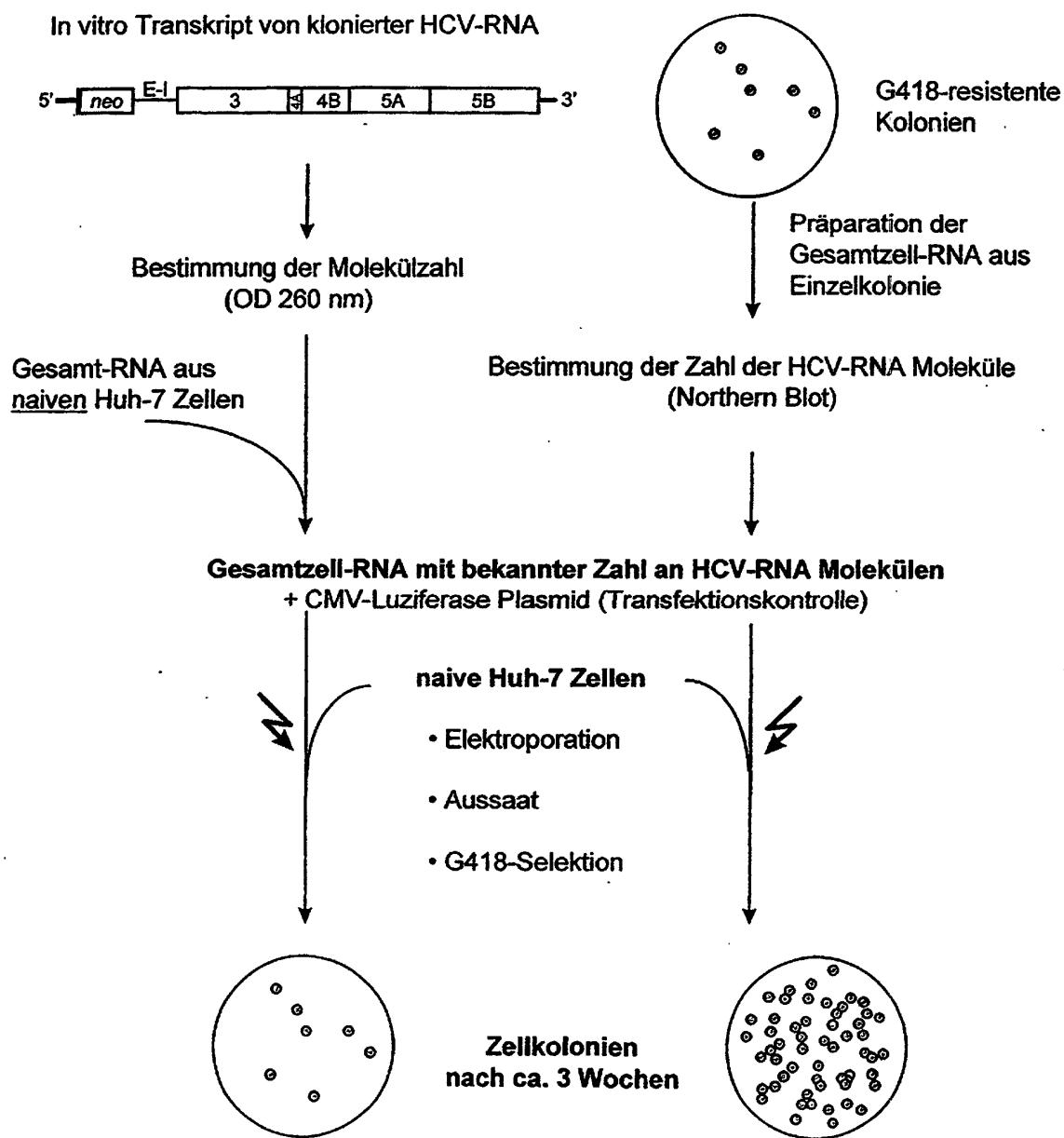


Fig. 9

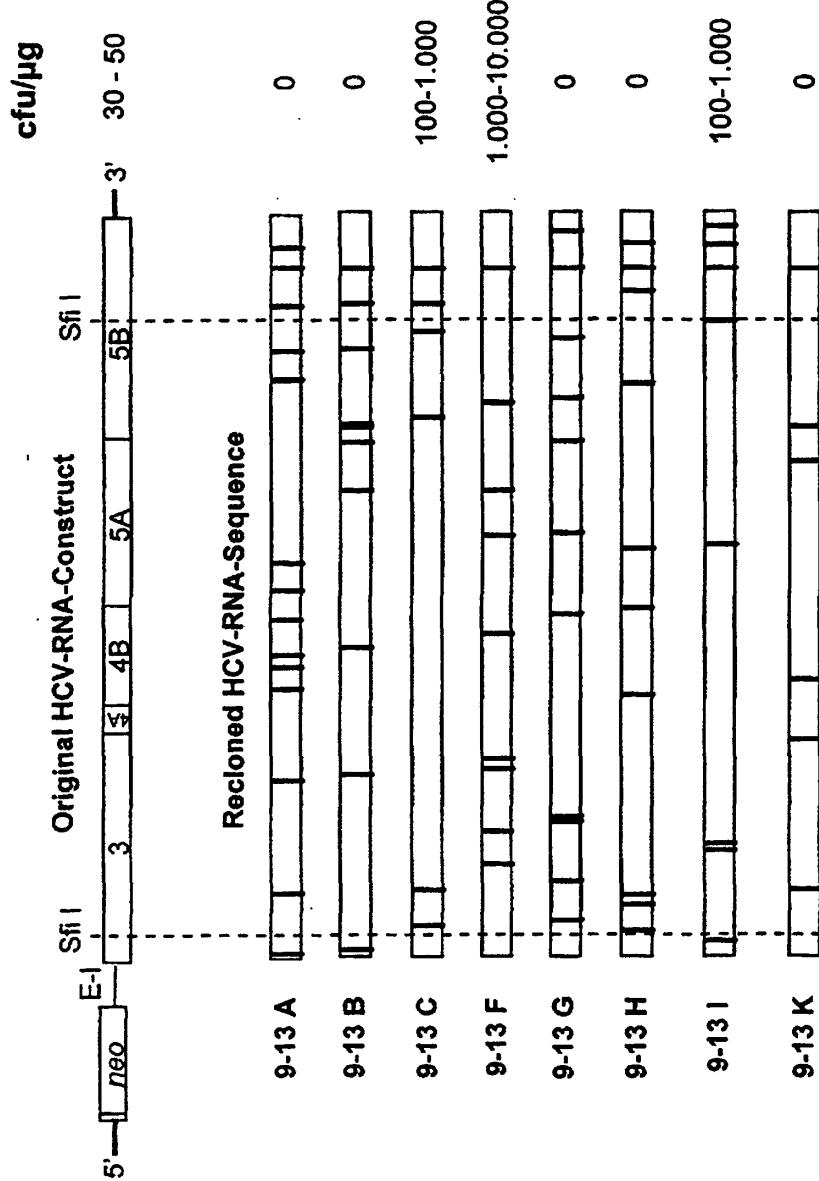
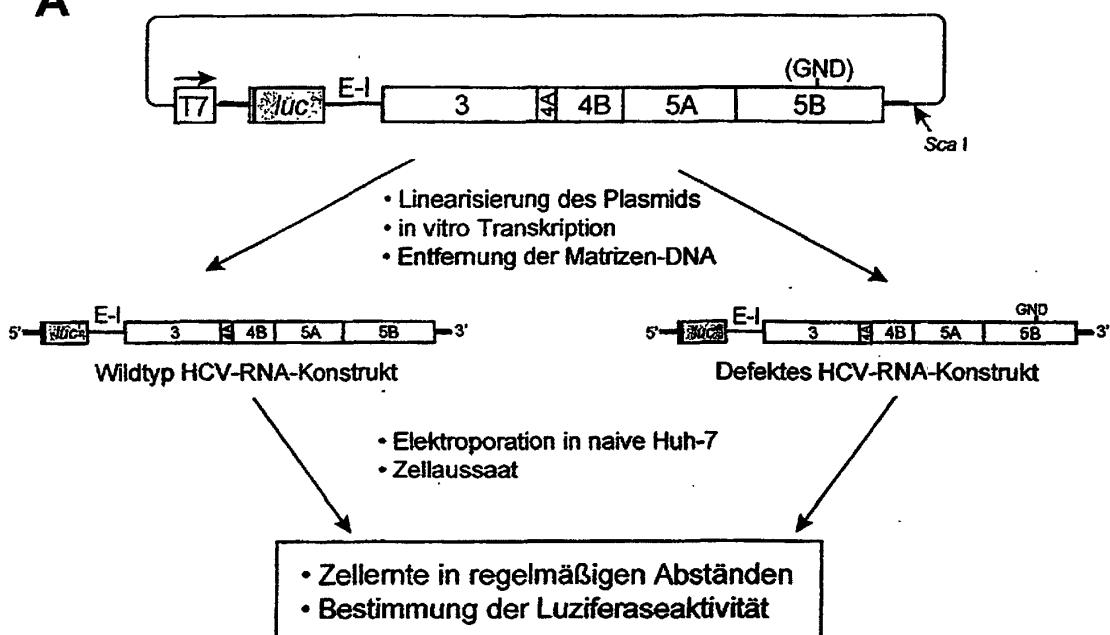


Fig. 10

A



B

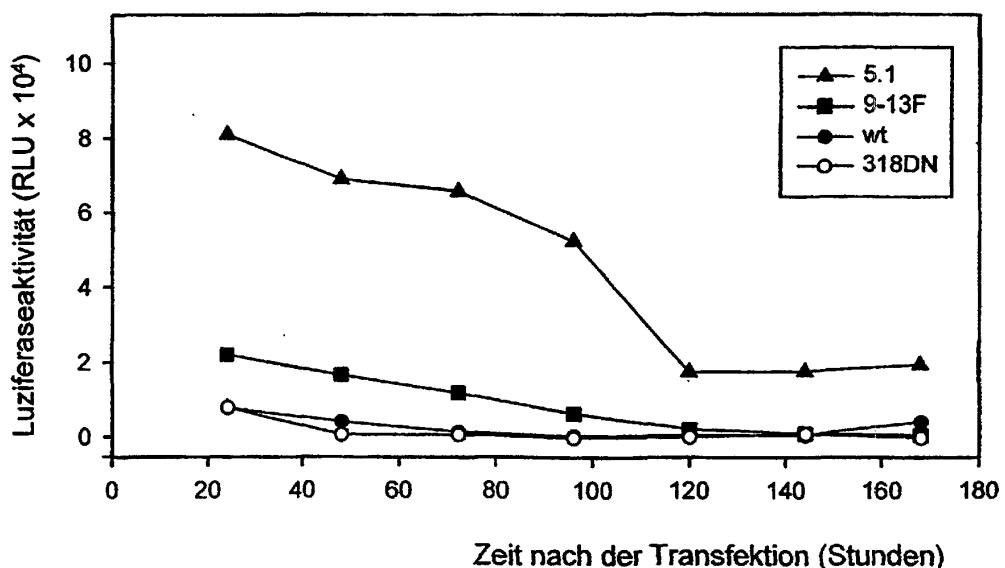
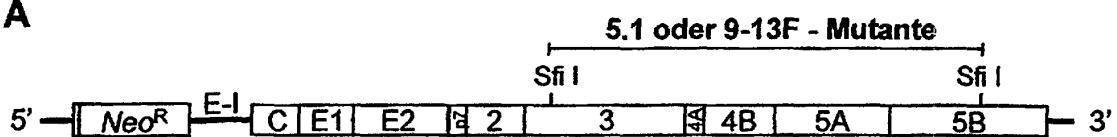


Fig. 11

A



B

G418-Konzentration:	500 µg/ml	250 µg/ml	100 µg/ml
Anzahl erhaltener Kolonien:	1	23	75

C

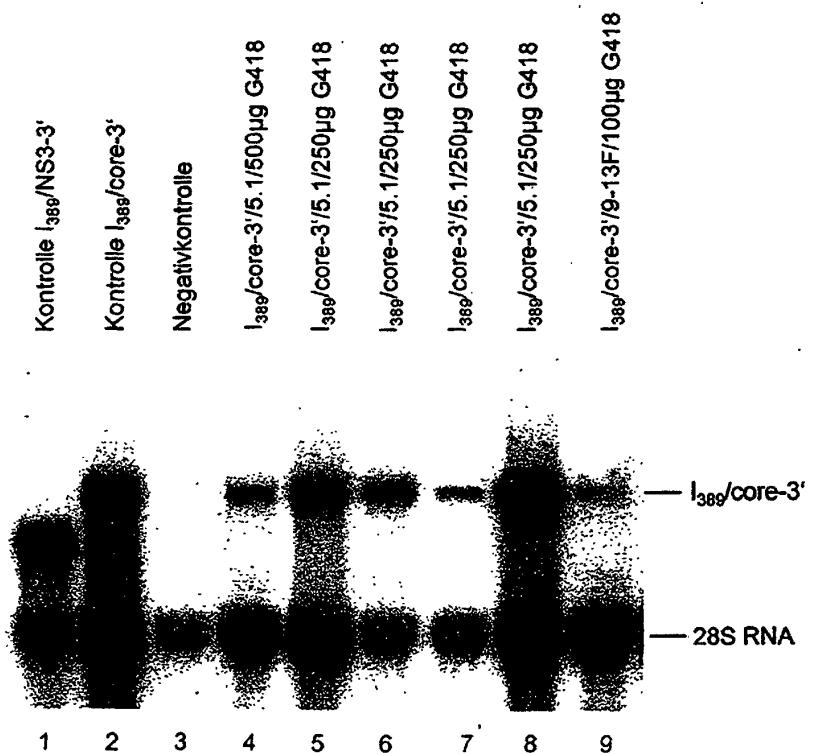


Fig. 12

Fig. 13

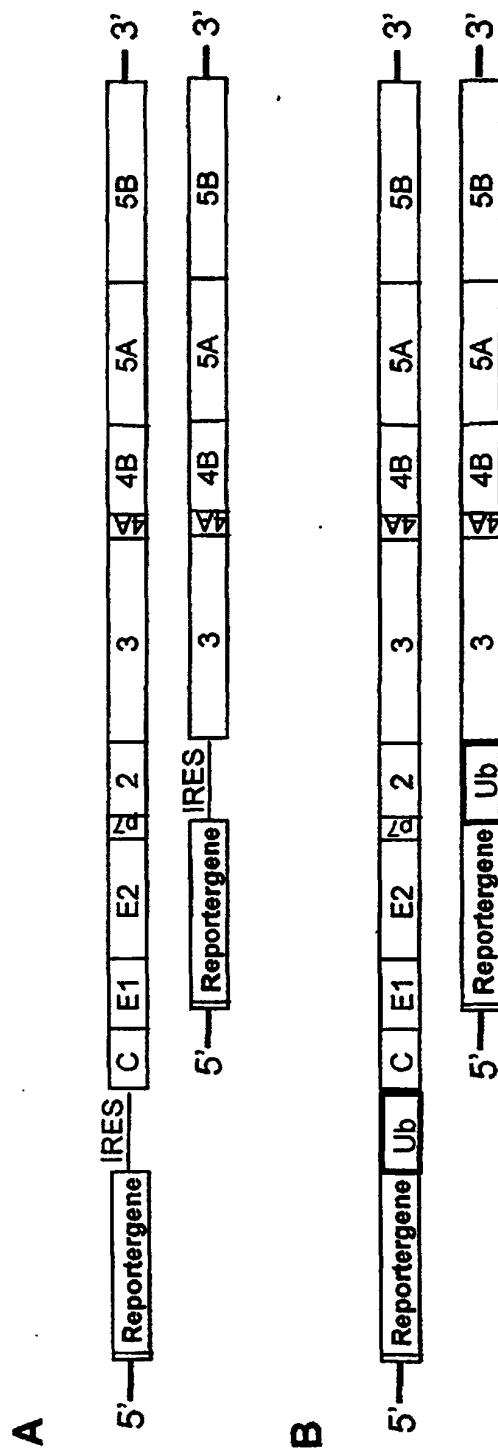


Fig. 14

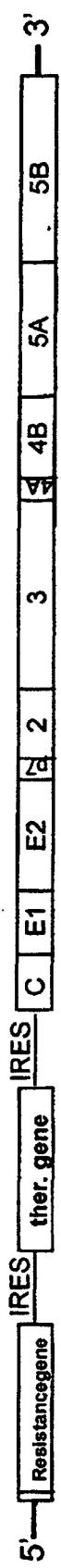


Fig. 15

