



(72) BARTENSCHLAGER, RALF, DE

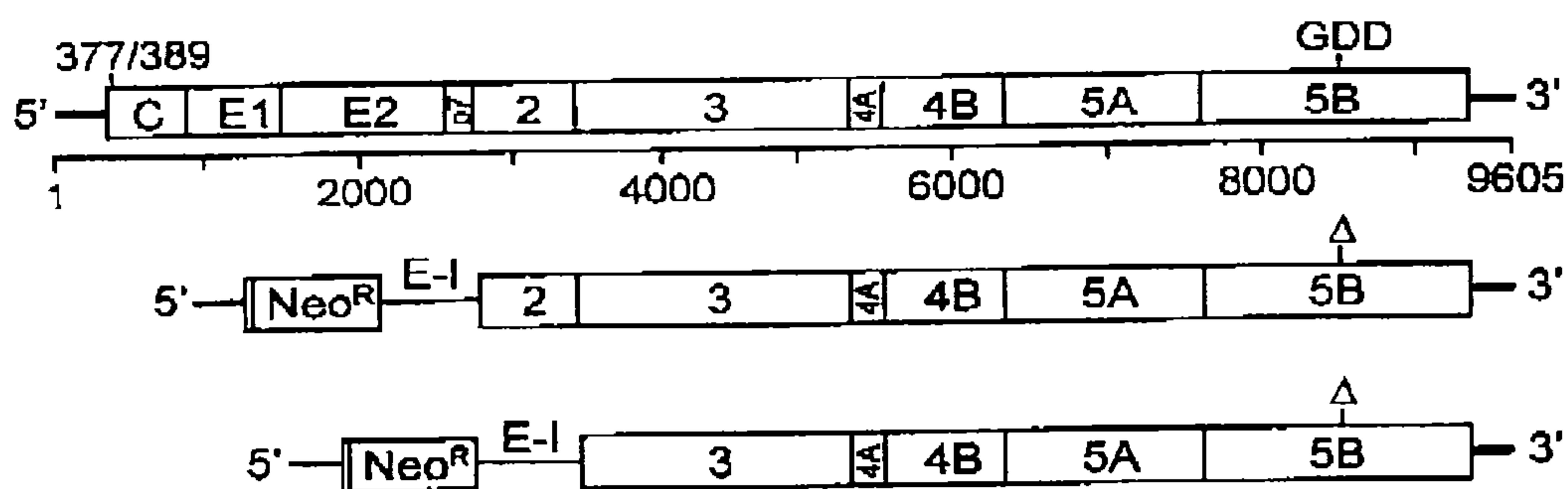
(71) BARTENSCHLAGER, RALF, DE

(51) Int.Cl.<sup>7</sup> C12N 7/02, C12N 15/51, C12N 5/10

(30) 1999/04/03 (199 15 178.4) DE

(54) **SYSTEME DE CULTURE DU VIRUS DE L'HEPATITE C**

(54) **HEPATITIS C VIRUS CELL CULTURE SYSTEM**



(57) The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).



## Hepatitis C Virus Cell Culture System

5

### ABSTRACT

The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises  
10 the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).

**Hepatitis C Virus Cell Culture System**DESCRIPTION

5 The invention relates to a hepatitis C virus (HCV) cell culture system, which comprises mainly eukaryotic cells containing transfected HCV specific genetic material, which means they are transfected with HCV specific genetic material.

10

The hepatitis C virus (HCV) is one of the main causes worldwide of chronic and sporadic liver diseases. The history of most HCV infections does not involve any obvious clinical signs, but 80 - 90 % of the infected people become

15 chronic carriers of the virus and 50 % of these chronic carriers of the virus develop chronic hepatitis with different degrees of severity. Approx. 20 % of the chronically infected develop a cirrhosis of the liver over 10 to 20 years, based on what a primary hepatocellular

20 carcinoma can develop. Nowadays chronic hepatitis C is the main indication for liver transplantation. A specific therapy does not exist until now. The only therapy currently available is high-dose administration of Interferon alpha or a combination of Interferon alpha and

25 the purine nucleoside analogue Ribavirin. However, only approx. 60 % of all treated persons respond to this therapy and with these, a new viraemia occurs in more than half of all cases after the discontinuation of the treatment.

30 Due to the high prevalence, especially in industrialized countries, the serious effects of chronic infections and the non-existence of a specific therapy, the development of a HCV specific chemotherapy is an important goal of pharmaceutical research and development. The main problem

35 lies in the previous lack of a suitable cell culture system, which enables the study of virus replication and pathogenesis in eukaryotic cells.

Due to the small amount of virus in blood or tissue, the lack of suitable cell culture systems or animal models (the chimpanzee is still the only possible experimental animal) as well as the lack of efficient systems for producing virus-like particles, it was not possible up to now, to analyze the molecular composition of the HCV particle in-depth and to solve it. The information currently available can be summarized as follows: HCV is an enveloped plus-strand RNA virus with a particle diameter of 50 - 60 nm and a medium density of 1.03 - 1.1 g/ml. It was molecularly cloned and characterized for the first time in 1989 (Choo et al., 1989, *Science*, 244, 359 - 362). The HCV-RNA has a length of approx. 9.6 kb (= 9600 nucleotides), a positive polarity and comprises one open reading frame (ORF), which encodes a linear polyprotein of approx. 3010 amino acids (see 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; and Bartenschlager 1997, *Intervirology* 40, 378 and see Fig. 1 A). During the replication of the virus the polyprotein is cleaved into the mature and functionally active proteins by cellular and viral proteases.

Within the polyprotein the proteins are arranged as follows (from the amino- to the carboxy terminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. The core protein is the main component of the nucleocapsid. The glycoproteins E1 and E2 are transmembrane proteins and the main components of the viral envelope. They probably play an important role during the attachment of the virus to the host cell. These three proteins core, E1, and E2 constitute the viral particle and are therefore called structural proteins. The function of the protein p7 is still not clear. The protein NS2 is probably the catalytic domain of the NS2-3 protease, which is responsible for the processing between the proteins NS2 and NS3. The protein NS3 has two functions, one is a

protease activity in the amino terminal domain, which is essential for the polyprotein processing, and the other a NTPase/helicase function in the carboxy terminal domain, which is probably important during the replication of the viral RNA. The protein NS4A is a co-factor of the NS3 protease. The function of the protein NS4B is unknown.

The open reading frame is flanked on its 5' end by a non-translated region (NTR) approx. 340 nucleotides in length, which functions as the internal ribosome entry site (IRES), and on its 3' end by a NTR approx. 230 nucleotides in length, which is most likely important for the genome replication. A 3' NTR such as this is the object of patent application PCT/US 96/14033. The structural proteins in the amino terminal quarter of the polyprotein are cleaved by host cell signal peptidase. The non-structural proteins (NS) 2 to (NS) 5B are processed by two viral enzymes, namely the NS2-3 and the NS3/4A protease. The NS3/4A protease is required for all cleavages beyond the carboxy terminus of NS3. The function of NS4B is unknown. NS5A, a highly phosphorylated protein, seems to be responsible for the Interferon resistance of various HCV genotypes (see 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), and NS5B has been identified as the RNA-dependent RNA polymerase.

First diagnostic systems have been developed from these findings, which are either based on the detection of HCV specific antibodies in patient serum or the detection of HCV specific RNA using the reverse transcription polymerase chain reaction (RT-PCR), and which are (must be) routinely used with all blood and blood products and/or according to the regulations.

Since the first description of the genome in 1989 several partial and complete sequences of the HCV have been cloned and characterized using the PCR method. A comparison of these sequences shows a high variability of the viral genome in particular in the area of the NS5B gene, which eventually resulted in the classification of 6 genotypes, which are again subdivided into the subtypes a, b, and c.

The genomic variance is not evenly distributed over the genome. The 5'NTR and parts of the 3'NTR are highly conserved, while certain encoded sequences vary a lot, in particular the envelope proteins E1 and E2.

The cloned and characterized partial and complete sequences of the HCV genome have also been analyzed with regard to appropriate targets for a prospective antiviral therapy. In the course of this, three viral enzymes have been discovered, which may provide a possible target. These include (1) the NS3/4A protease complex, (2) the NS3 Helicase and (3) the NS5B RNA-dependent RNA polymerase. The NS3/4A protease complex and the NS3 Helicase have already been crystallized and their three-dimensional structure determined (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). It has not been successful until now with the NS5B RNA-dependent RNA polymerase.

Even though important targets for the development of a therapy for chronic HCV infection have been defined with these enzymes and even though a worldwide intensive search for suitable inhibitors is ongoing with the aid of rational drug design as well as high throughput screening, the development of a therapy has one major deficiency, namely the lack of cell culture systems or simple animal models, which allow direct, reliable identification of HCV-RNA or

HCV antigens with simple methods which are common in the laboratory. The lack of these cell culture systems is also the main reason that to date the comprehension of HCV replication is still incomplete and mainly hypothetical.

5

Although according to the experts, a close evolutionary relationship exists between HCV and the flavi- and pestiviruses, and self-replicating RNAs have been described for these, which can be used for the replication in different cell lines with a relatively high yield, (see 10 Khromykh et al., 1997, *J. Virol.* 71, 1497; Behrens et al., 1998, *J. Virol.* 72, 2364; Moser et al., 1998, *J. Virol.* 72, 5318), similar experiments with HCV have not been successful to date.

15

Although it is known from different publications that cell lines or primary cell cultures can be infected with high titre patient serum containing HCV, (Lanford et al. 1994, *Virology* 202, 606; Shimizu et al. 1993, *Proceedings of the* 20 *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 and bibliographical references quoted in here; Ito et al. 1996, *Journal of* 25 *General Virology*, 77, 1043 - 1054), these virus-infected cell lines or cell cultures do not allow the direct detection of HCV-RNA or HCV antigens. The viral RNA in these cells can not be detected in a Northern Blot (a standard method for the quantitative detection of RNA) or 30 the viral protein in a Western Blot or with immunoprecipitation. It has only been possible to detect HCV replication with very costly and indirect methods. These disadvantageous facts show that obviously the replication in these known virus-infected cell lines or 35 cell cultures is completely insufficient.

Furthermore it is known from the publications of Yoo et al. (1995, *Journal of Virology*, 69, 32 - 38) and of Dash et al., (1997, *American Journal of Pathology*, 151, 363 - 373) that hepatoma cell lines can be transfected with synthetic HCV-RNA, which are obtained through in vitro transcription of the cloned HCV genome. In both publications the authors started from the basic idea that the viral HCV genome is a plus-strand RNA functioning directly as mRNA after being transfected into the cell, permitting the synthesis of viral proteins in the course of the translation process, and so new HCV particles are (could be) formed. This viral replication, which means these newly formed HCV viruses and their RNA, have been detected through RT-PCR. However the published results of the RT-PCR carried out indicate, that the HCV replication in the described HCV transfected hepatoma cells is not particularly efficient and is not sufficient to measure the quality, let alone the quantity of the fluctuations in the replication rate after an targeted action with prospective antiviral treatments. Furthermore it is prior art (Yanagi et al., *Proc. Natl. Acad. Sci. USA*, 96, 2291-95, 1999), that the highly conserved 3' NTR is essential for the virus replication. This knowledge strictly contradicts the statements of Yoo et al. and Dash et al., who used for their experiments only HCV genomes with shorter 3' NTRs since they did not know the authentic 3' end of the HCV genome.

The object of the present invention is to provide a HCV cell culture system, where the viral RNA self-replicates in the transfected cells with such a high efficiency that the quality and quantity of the fluctuations in the replication rate can be measured with common methodologies usually found in the laboratory after a targeted action with virus and prospective HCV specific antivirals in particular.

The solution to this problem is to provide a cell culture system of the prior mentioned type, where the eukaryotic

cells are human cells, in particular hepatoma cells, which are preferably derived from a normal hepatoma cell line, but can also be obtained from an appropriate primary cell culture, and where the transfected HCV specific genetic material is a HCV-RNA construct, which essentially comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR preferably in the order mentioned as well as a minimum of one marker gene for selection (selection gene).

10

Here and in the following "NTR" stands for "non-translated region" and is a known and familiar term or abbreviation to the relevant expert.

15

Here and in the following the term "HCV-RNA construct" comprises constructs, which include the complete HCV genome, as well as those, which only include a part of it, which means a HCV subgenome.

20

A preferred variation of the cell culture system according to the invention, which had proven to be worthwhile in practice, is lodged at the DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (German collection of Microorganisms and Cell Cultures) in Braunschweig, Germany under the number DSM ACC2394 (laboratory name HuBl 9-13).

25

With the cell culture system according to the invention an in vitro system is provided for the first time, where HCV-RNA is self-replicated and expressed intracellularly and in a sufficient amount, so that the quantity of the amounts of HCV-RNA as well as the HCV specific proteins can be determined with conventional and reliably precise biochemical measuring methods. This means an almost authentic cell-based HCV replication system is available for the first time, which is urgently needed for the development and testing of antiviral drugs. This test system provides the possibility of identifying potential

30  
35

targets for an effective HCV specific therapy and developing and evaluating HCV specific chemotherapeutics.

The invention is based on the surprising finding that  
5 efficient replication of the HCV-RNA only occurs in cells  
if they have been transfected with an HCV-RNA construct,  
which comprises at least the 5' and the 3' non-translated  
regions (NTR) and the non-structural proteins (NS) 3 to 5B  
and additionally a marker gene for selection (selection  
10 gene). The structural genes are obviously without great  
importance for replication, whereas efficient replication  
of the HCV-RNA apparently only occurs if the transfected  
cells are subject to permanent selection pressure, which is  
imparted by the marker gene for selection (selection gene)  
15 linked to the HCV-RNA. Consequently the marker gene  
(selection gene) seems on one hand to provoke the selection  
of those cells, where the HCV-RNA replicates productively,  
and it seems on the other hand to considerably increase the  
efficiency of the RNA replication.

20

An object of the invention is also a cell-free HCV-RNA  
construct, characterized in that it comprises the HCV  
specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B,  
and 3' NTR, preferably in the order mentioned, as well as a  
25 marker gene for selection (selection gene).

In the present context the terms 5' NTR and NS3 and NS4A  
and NS4B and NS5A and NS5B and 3' NTR comprise each  
nucleotide sequence, which is described in the state of the  
30 art as the nucleotide sequence for each functional segment  
of the HCV genome.

By providing a HCV-RNA construct such as this, a detailed  
analysis of the HCV replication, pathogenesis and evolution  
35 in cell culture is possible for the first time. The HCV  
specific viral RNA can specifically be created as a  
complete genome or subgenome in any amount, and it is

possible, to manipulate the RNA construct and consequently to examine and identify the HCV functions on a genetic level.

5 Because all HCV enzymes identified as a main target for a therapy at the moment, namely the NS3/4A protease, the NS3 helicase and the NS5B polymerase, are included in the HCV-RNA construct according to the invention, it can be used for all relevant analyses.

10

An embodiment of the HCV-RNA construct, which has proven to be worthwhile in practical use, stands out by the fact that it comprises the nucleotide sequence according to the sequence protocol SEQ ID NO:1.

15

Further embodiments with similar good properties for practical use are characterized in that they comprise a nucleotide sequence either according to sequence protocol SEQ ID NO:2 or according to sequence protocol SEQ ID NO:3  
20 or according to sequence protocol SEQ ID NO:4 or according to sequence protocol SEQ ID NO:5 or according to sequence protocol SEQ ID NO:6 or according to sequence protocol SEQ ID NO:7 or according to sequence protocol SEQ ID NO:8 or according to sequence protocol SEQ ID NO:9 or according to  
25 sequence protocol SEQ ID NO:10 or according to sequence protocol SEQ ID NO:11.

It is possible to provide the HCV subgenomic construct with a 3' NTR, which has a nucleotide sequence so far unknown in  
30 the state of art, a nucleotide sequence, which has been selected from the group of nucleotide sequences (a) to (i) listed in the following:

(a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTAGCTTT  
35 TTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCCCAGACTGGAG  
T

- (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTGTCT  
TTTTTTTTTC TTTTTTTGA GAGAGAGAGT CTCACTCTGT TGCCCAGACT  
GGAGC
- 5 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTAATCTTT  
TTTTTTTTCT TTTTTTTGA GAGAGAGAGT CTCACTCTGT TGCCCAGACT  
GCAGC
- (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTAGTC  
10 TTTTTTTTTT TTTTTTTTTT TGAGAGAGAG AGTCTCACTC TGTTGCCAG  
ACTGGAGT
- (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTGTCT  
TTTTTTTTTT TTTTTTTTTT TGAGAGAGAG AGTCTCACTC TGTTGCCAG  
15 ACTGGAGT
- (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTGTCT  
TTTTTTTTTT TTTTTTTTTT TTGAGAGAGA GAGTCTCACT CTGTTGCCA  
GACTGGAGT  
20
- (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTGTCT  
TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA GTCTCACTCT GTTGCCCAGA  
CTGGAGT
- 25 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTTAAT  
CTTTTTTTTT TTTTCCTTT TTTTGAGAGA GAGAGTCTCA CTCTGTTGCC  
CAGACTGGAG T
- (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTAATC  
30 TTTTTTTTTT TTTCTTTTT TTTTGAGAG AGAGAGTCTC ACTCTGTTGC  
CCAGACTGGA GT

The marker gene for selection (selection gene) included in  
the HCV-RNA constructs according to the invention is  
35 preferably a resistance gene, in particular an antibiotic  
resistance gene.

This has the advantage that the cells transfected with this construct can easily be selected from the non-transfected cells by adding for example the appropriate antibiotic to the cell culture medium in the case of an antibiotic  
5 resistance gene.

In the present context 'antibiotic' means any substance, which impedes the non-transfected host cells or the cells, where the HCV-RNA is not replicating efficiently,  
10 continuing to live or grow, especially the cell poison Puromycin, Hygromycin, Zeocin, Bleomycin or Blasticidin.

A preferred marker gene for selection (selection gene) and resistance gene, which has proven to be worthwhile in  
15 practice, is the neomycin phosphotransferase gene.

An alternative for the antibiotic resistance genes is for example the thymidine kinase gene, which can be used to carry out a HAT selection.  
20

The marker gene for selection (selection gene), the preferred resistance gene and the most preferred antibiotic resistance gene is preferably positioned in the HCV-RNA construct after the HCV 5' NTR, which means downstream from  
25 the 5' NTR and upstream from the HCV reading frame. However, an insertion in the area of the 3' NTR or another site of the HCV genome or subgenome, for example within the polyprotein, is also conceivable.

30 In another embodiment of the HCV-RNA construct according to the invention the marker gene for selection (selection gene), in particular an antibiotic resistance gene, is linked to the HCV-RNA or HCV genomic or subgenomic sequence via a ribozyme or a recognition site for a ribozyme.  
35

This has the advantage, that after the selection of the cells, in which the HCV-RNA is replicating productively,

the resistance gene in the obtained cell clones can be separated from the HCV subgenomic sequence through a ribozyme-dependent cleavage, namely by activating the inserted ribozyme or in the case of a construct with a recognition site for a ribozyme, by transfecting the ribozyme into the cells (for example through the transfection of a ribozyme construct or infection with a viral expression vector, into which the appropriate ribozyme has been inserted). By this means an authentic HCV genomic construct can be obtained without a resistance gene, which can then form authentic infectious virus particles.

Another preferred embodiment of the HCV-RNA construct according to the invention is characterized in that the construct has at least one integrated reporter gene.

In the following a reporter gene means any gene, whose presence can be easily detected with, in general, simple biochemical or also histochemical methods after being transferred into a target organism, which means a gene, that encodes for a protein, which can be easily and reliably detected and quantified in small amounts with the common measuring methods in the laboratory.

This variation of the HCV-RNA construct has the advantage that the extent of the replication of this construct can be easily and quickly measured with the methods common in the laboratory using the reporter gene product.

The reporter gene is preferably a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) or the SEAP gene (secreted alkaline phosphatase gene). This reporter gene and its products, namely the relevant reporter proteins, can be

detected for example using fluorescence, chemiluminescence, colorimetric measurements or by means of immunological methods (for example ELISA).

5 A surrogate marker gene can also be considered as a reporter gene. In this context it includes those genes, which encode for cellular proteins, nucleic acids or generally for those functions, which are subject to variation depending on the replication of the virus, and  
10 which consequently are either suppressed or activated in the cells, in which the HCV or the HCV-RNA construct multiplies. This means, the suppression or activation of this function is a surrogate marker for the replication of the virus or the replication of the HCV-RNA construct.

15

The positions of the reporter genes and the marker gene for selection (selection gene) can be selected in such a way, that a fusion protein made from both genetic products will be expressed. This has the advantage that these two genes  
20 can be arranged in such a way in the HCV-RNA construct that their two expressed proteins are fused via a recognition sequence for a protease (for example ubiquitin) or via a self-cleaving peptide (for example the 2A protein of the Picornaviruses) at first and will be separated  
25 proteolytically later.

These two positions might as well lie apart from each other in such way, that both genetic products are separately expressed (for example in the order: marker or resistance  
30 gene - internal ribosome binding site - reporter gene).

In the case of the reporter gene one embodiment has proven to be particularly worthwhile, where the reporter gene is cloned into the open reading frame of the HCV genome or  
35 subgenome in such a way that it will only be transferred to an active form after proteolytic processing.

The cell culture system according to the invention can be used for various purposes in each of its embodiments. These comprise:

- 5 • The detection of antiviral substances. This can include for example: organic compounds, which interfere directly or indirectly with viral growth (for example inhibitors of the viral proteases, the NS3 helicase, the NS5B RNA-dependent RNA polymerase), antisense oligonucleotides,  
10 which will hybridize to any target sequence in the HCV-RNA construct (for example the 5' NTR) and will have an direct or indirect influence on the virus growth for example due to a reduction of translation of the HCV polyprotein or ribozymes, which cleave any HCV-RNA  
15 sequence and consequently impair virus replication.
- The evaluation of any type of antiviral substances in the cell culture. These substances can be detected on the isolated purified enzyme for example with 'rational  
20 drug design' or 'high-throughput screening'. Evaluation means mainly the determination of the inhibitory features of the respective substance as well as its mode of action.
- 25 • The identification of new targets of viral or cellular origin for a HCV specific antiviral therapy. If for example a cellular protein is essential for viral replication, the viral replication can also be influenced by inhibiting this cellular protein. The  
30 system according to the invention also enables the detection of these auxiliary factors.
- The determination of drug resistance. It can be assumed that resistance to therapy occurs due to the high  
35 mutation rate of the HCV genome. This resistance, which is very important for the clinical approval of a substance, can be detected with the cell culture system

according to the invention. Cell lines, in which the HCV-RNA construct or the HCV genome or subgenome replicates, are incubated with increasing concentrations of the relevant substance and the replication of the viral RNA is either determined by means of an introduced reporter gene or through the qualitative or quantitative detection of the viral nucleic acids or proteins. Resistance is given if no or a reduced inhibition of the replication can be observed with the normal concentration of the active substance. The nucleotide and amino acid replacements responsible for the therapy resistance can be determined by recloning the HCV-RNA (for example by the means of RT-PCR) and sequence analysis. By cloning the relevant replacement/s into the original construct its causality for the resistance to therapy can be proven.

- The production of authentic virus proteins (antigens) for the development and/or evaluation of diagnostics. The cell culture system according to the invention also allows the expression of HCV antigens in cell cultures. In principle these antigens can be used as the basis for diagnostic detection methods.

- The production of HCV viruses and virus-like particles, in particular for the development or production of therapeutics and vaccines as well as for diagnostic purposes. Especially cell culture adapted complete HCV genomes, which could be produced by using the cell culture system according to the invention, are able to replicate in cell culture with high efficiency. These genomes have the complete functions of HCV and in consequence they are able to produce infectious viruses.

The HCV-RNA construct according to the invention by itself can also be used for various purposes in all its embodiments. This includes first of all:

- The construction of attenuated hepatitis C viruses or HCV-like particles and their production in cell cultures:

5

Attenuated HCV or HCV-like particles can be created by accidental or purposefully introduced mutations, such as point mutations, deletions or insertions, which means viruses or virus-like particles with complete ability to replicate, but reduced or missing pathogenicity. These attenuated HCV or HCV-like particles can be used in particular as vaccine.

10

- The construction of HCV-RNA constructs with integrated foreign genes, used for example as liver cell specific vector in gene therapy. Due to the distinctive liver cell tropism of the HCV and the possibility of replacing parts of the genome by heterologous sequences, HCV-RNA constructs can be produced, where for example the structural proteins can be replaced by a therapeutically effective gene. The HCV-RNA construct obtained in this way is introduced into cells preferably by means of transfection, which express the missing HCV functions, for example the structural proteins, in a constitutive or inducible way. Virus particles, carrying the HCV-RNA construct, can be created by means of this method known to the expert under the term 'transcomplementation'. The particles obtained can preferably be used for the infection of liver cells. Within these the therapeutically effective foreign gene will be expressed and will consequently develop its therapeutic effect.

20

25

30

- The detection of permissive cells, which means cells, in which a productive virus growth occurs. For this purpose either one of the HCV-RNA genomic constructs previously mentioned, which is able to form complete infectious viruses, or one of the HCV subgenomes previously

35

mentioned, which according to the previously mentioned example will be transfected in a cell line first, which expresses the missing functions in a constitutive or inducible way, is used. In each case virus particles are created, which carry a resistance and/or reporter gene apart from the HCV sequence. In order to detect cells, where the HCV is able to replicate, these cells are infected with viruses generated in this way and subject to an antibiotic selection or they are examined depending on the HCV-RNA construct by means of determining the presence of the expression of the reporter gene. Because an antibiotic resistance or reporter gene expression can only be established, when HCV-RNA construct replicates, the cells detected in this way must be permissive. Almost any cell line or primary cell culture can be tested in regard to the permissivity and detected in this way.

20 The cell culture system according to the invention also permits targeted discovery of HCV-RNA constructs for which there is an increase in the efficiency of replication due to mutations. This occurs either by chance, in the context of HCV-RNA replication, or by targeted introduction into the construct. These mutations, leading to a change in the replication of the HCV-RNA construct, are known to experts as adaptive mutations. The invention therefore also includes a method for obtaining cell culture adapted mutants of a HCV-RNA construct according to the invention following the above description, in which the mutants have increased replication efficiency compared to the original HCV-RNA construct. It further includes a method for the production of mutants of a HCV-RNA full-length genome or of a HCV-RNA subgenome or of any HCV-RNA construct with increased replication efficiency compared to the original HCV-RNA full-length genome or subgenome or HCV-RNA construct, as well as cell culture adapted mutants of HCV-

RNA constructs, HCV-RNA full-length genomes and HCV subgenomes with increased replication efficiency compared to the original constructs, subgenomes or full-length genomes.

5

The method according to the invention for the production of cell culture adapted mutants of a HCV-RNA construct according to the invention, in which the mutants have increased replication efficiency compared to the HCV-RNA construct, is characterised in that a cell culture system according to claim 1, in which the transfected HCV specific genetic material is a HCV-RNA construct with a selection gene according to claims 4 to 19, is cultivated on/in the selection medium corresponding to the selection gene, that  
10 the cultivated cell clones are collected and that the HCV-RNA construct is isolated from these cell clones.

In an advantageous extension of this production method, the isolated HCV-RNA constructs are passaged at least one more  
20 time, that is they are transfected in cells of a cell culture system according to claim 1, the thus obtained cell culture system according to claim 1, in which the transfected HCV specific genetic material is the isolated HCV-RNA construct with a selection gene, is cultivated  
25 on/in the selection medium corresponding to the selection gene, the cultivated cell clones are collected and the HCV-RNA constructs are thus isolated.

Using this process variation, the quantity of adaptive  
30 mutations and hence the degree of replication efficiency in the relevant HCV-RNA constructs can be increased even further.

The method according to the invention for the production of  
35 mutants of a HCV-RNA full-length genome or of a HCV-RNA subgenome or of any HCV-RNA construct with increased replication efficiency compared to the original HCV-RNA

full-length genome or subgenome or HCV-RNA construct, has the following features. Using one of the two production methods presented above, a cell culture adapted mutant of a HCV-RNA construct is produced, isolated from the cells, 5 cloned using state of the art known methods and sequenced. By comparing with the nucleotide and amino acid sequence of the original HCV-RNA construct, the type, number and position of the mutations is determined. These mutations are then introduced into an (isolated) HCV subgenome or 10 full-length genome or any HCV-RNA construct, either by site-directed mutagenesis, or by exchange of DNA fragments containing the relevant mutations.

A test can be carried out to determine or verify which 15 mutations actually are responsible for an alteration of replication efficiency, particularly an increase in replication. In this test the corresponding nucleotide and/or amino acid changes are introduced into the original HCV-RNA construct and the modified construct is then 20 transfected in cell culture. If the introduced mutation actually leads to an increase in replication, then for a HCV-RNA construct with a selectable marker gene, the number of resistant cell clones in the artificially mutated construct should be noticeably higher compared to the 25 untreated construct.

In the case of a construct with a reporter gene, the activity or quantity of the reporter should be noticeably higher for the artificially mutated construct compared to 30 the untreated one.

The cell culture adapted HCV-RNA constructs with high replication efficiency according to the invention are characterized in that, through nucleotide or amino acid 35 exchanges, they are derivable from a HCV-RNA construct according to one of the claims 4 to 19, and that they are

obtainable using one of the two production processes presented above.

5 These cell culture adapted HCV-RNA constructs can be used to produce any HCV-RNA constructs or HCV full-length or subgenomes with increased replication efficiency. Both constructs with a selectable resistance gene and constructs without one or with a non-selectable reporter gene (e.g. luciferase) can be produced in this way, since replication  
10 of cell culture adapted HCV-RNA constructs can also be demonstrated in non-selected cells due to their high replication efficiency.

The cell culture adapted mutants of a HCV-RNA construct or  
15 HCV-RNA full-length genome or HCV subgenome with high replication efficiency compared to the original HCV-RNA construct or the original HCV full-length genome are characterized in that they are obtainable by a method in which the type and number of mutations in a cell culture  
20 adapted HCV-RNA construct are determined through sequence analysis and sequence comparison and these mutations are introduced into a HCV-RNA construct, particularly a HCV-RNA construct according to one of the claims 4 to 19, or into an (isolated) HCV-RNA full-length genome, either by site-  
25 directed mutagenesis, or by exchange of DNA fragments containing the relevant mutations.

A group of preferred HCV-RNA constructs, HCV full-length genomes and HCV subgenomes with high and very high  
30 replication efficiency, which are consequently highly suitable for practical use is characterised in that it contains one, several or all of the amino acid or nucleic acid exchanges listed in table 3 and/or one or several of the following amino acid exchanges: 1283 arg -> gly , 1383  
35 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile

-> val. (The numbers refer to the amino acid positions of the polyprotein of the HCV isolate con1, see Table 1).

5 **Special features of the nucleotid sequences according to the sequence listings:**

**SEQ ID-NO: 1**

Name: I389/Core-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 10 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 15 4.1813-10842: HCV Polyprotein from Core up to nonstructural protein 5B
  5. 1813-2385: HCV Core Protein; structural protein
  6. 2386-2961: envelope protein 1 (E1); structural protein
  - 20 7. 2962-4050: envelope protein 2 (E2); structural protein
  8. 4051-4239: Protein p7
  9. 4240-4890: nonstructural protein 2 (NS2); HCV NS2-3 Protease
  - 25 10.4891-6783: nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
  - 11.6784-6945: nonstructural protein 4A (NS4A); NS3 Protease cofactor
  - 12.6946-7728: nonstructural protein 4B (NS4B)
  - 30 13.7729-9069: nonstructural protein 5A (NS5A)
  - 14.9070-10842: nonstructural protein 5B (NS5B); RNA-dependent RNA-polymerase
  - 15.10846-11076: HCV 3' non-translated region

**SEQ ID-NO: 2**

Name: I337/NS2-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 5 2.342-1181: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1190-1800: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 10 4.1801-8403: HCV Polyprotein from nonstructural protein 2 up to nonstructural protein 5B
  5. 1801-2451: nonstructural protein 2 (NS2); HCV NS2-3 Protease
  6. 2452-4344: nonstructural protein 3 (NS3); HCV NS3
  - 15 Protease/Helicase
  7. 4345-4506: nonstructural protein 4A (NS4A); NS3 Protease cofactor
  8. 4507-5289: nonstructural protein 4B (NS4B)
  9. 5290-6630: nonstructural protein 5A (NS5A)
  - 20 10.6631-8403: nonstructural protein 5B (NS5B); RNA-dependent RNA-polymerase
  - 11.8407-8637: HCV 3' non-translated region

**SEQ ID-NO: 3**

25 Name: I389/NS3-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 30 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-7767: HCV Polyprotein from nonstructural protein 3 up to nonstructural protein 5B

5. 1813-3708: nonstructural protein 3 (NS3); HCV NS3  
Protease/Helicase
6. 3709-3870: Nonstructural protein 4A (NS4A); NS3  
Protease Cofactor
- 5 7. 3871-4653: Nonstructural protein 4B (NS4B)
8. 4654-5994: Nonstructural protein 5A (NS5A)
9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-  
dependent RNA-Polymerase
10. 7771-8001: HCV 3' non-translated Region

10

**SEQ ID-NO: 4**

Name: I337/NS3-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 15 2.342-1181: HCV Core Protein-Neomycin Phosphotransferase  
fusion protein; selectable Marker
- 3.1190-1800: internal ribosome entry site from  
encephalomyokarditis virus; directs translation of the  
downstream located HCV open reading frame
- 20 4.1801-7758: HCV Polyprotein from Nonstructural protein 3  
up to Nonstructural protein 5B
  5. 1801-3696: Nonstructural protein 3 (NS3); HCV NS3  
Protease/Helicase
  6. 3697-3858: Nonstructural protein 4A (NS4A); NS3  
Protease Cofactor
  - 25 7. 3859-4641: Nonstructural protein 4B (NS4B)
  8. 4642-5982: Nonstructural protein 5A (NS5A)
  9. 5983-7755: Nonstructural protein 5B (NS5B); RNA-  
dependent RNA-Polymerase
  - 30 10. 7759-7989: HCV 3' non-translated Region

**SEQ ID-NO: 5**

Name: I389/NS2-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from  
5 encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-8418: HCV Polyprotein from Nonstructural protein 2 up to Nonstructural protein 5B
5. 1813-2463: Nonstructural protein 2 (NS2); HCV NS2-3  
10 Protease
6. 2464-4356: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
7. 4357-4518: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
8. 4519-5301: Nonstructural protein 4B (NS4B)  
15
9. 5302-6642: Nonstructural protein 5A (NS5A)
- 10.6643-8415: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
11. 8419-8649: HCV 3' non-translated Region

20

**SEQ ID-NO: 6**

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

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 25 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 30 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 9-13F
5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
- 35 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3

Protease Cofactor

7. 3871-4653: Nonstructural protein 4B (NS4B)  
 8. 4654-5994: Nonstructural protein 5A (NS5A)  
 9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-  
 5 dependent RNA-Polymerase  
 7771-8001: HCV 3' non-translated Region

**SEQ ID-NO: 7**

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

10 Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region  
 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase  
 fusion protein; selectable Marker  
 3.1202-1812: internal ribosome entry site from  
 15 encephalomyokarditis virus; directs translation of the  
 downstream located HCV open reading frame  
 4.1813-10842: HCV Polyprotein from Core up to Nonstructural  
 protein 5B of the cell culture-adapted mutant 9-13F  
 5. 1813-2385: HCV Core Protein; structural protein  
 20 6. 2386-2961: envelope protein 1 (E1); structural  
 protein  
 7. 2962-4050: envelope protein 2 (E2); structural  
 protein  
 8. 4051-4239: Protein p7  
 25 9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3  
 Protease  
 10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3  
 Protease/Helicase  
 11.6784-6945: Nonstructural protein 4A (NS4A); NS3  
 30 Protease Cofactor  
 12.6946-7728: Nonstructural protein 4B (NS4B)  
 13.7729-9069: Nonstructural protein 5A (NS5A)  
 14.9070-10842: Nonstructural protein 5B (NS5B); RNA-  
 dependent RNA-Polymerase  
 35 15.10846-11076: HCV 3' non-translated Region

**SEQ ID-NO: 8**

Name: I389/NS3-3'/5.1

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 5 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase  
fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from  
encephalomyokarditis virus; directs translation of the  
downstream located HCV open reading frame
- 10 4.1813-7767: HCV Polyprotein from Nonstructural protein 3  
up to Nonstructural protein 5B of the cell culture-  
adapted mutant 5.1
5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3  
Protease/Helicase
- 15 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3  
Protease Cofactor
7. 3871-4653: Nonstructural protein 4B (NS4B)
8. 4654-5994: Nonstructural protein 5A (NS5A)
9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-  
20 dependent RNA-Polymerase
- 7771-8001: HCV 3' non-translated Region

**SEQ ID-NO: 9**

Name: I389/Core-3'/5.1

25 Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase  
fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from  
30 encephalomyokarditis virus; directs translation of the  
downstream located HCV open reading frame
- 4.1813-10842: HCV Polyprotein from Core up to Nonstructural  
protein 5B of the cell culture-adapted mutant 5.1
5. 1813-2385: HCV Core Protein; structural protein

6. 2386-2961: envelope protein 1 (E1); structural protein
7. 2962-4050: envelope protein 2 (E2); structural protein
- 5 8. 4051-4239: Protein p7
9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3, Protease
10. 4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
- 10 11. 6784-6945: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
12. 6946-7728: Nonstructural protein 4B (NS4B)
13. 7729-9069: Nonstructural protein 5A (NS5A)
14. 9070-10842: Nonstructural protein 5B (NS5B); RNA-
- 15 dependent RNA-Polymerase
15. 10846-11076: HCV 3' non-translated Region

**SEQ ID-NO: 10**

Name: I389/NS3-3'/19

- 20 Composition (Nucleotide positions):
1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
3. 1202-1812: internal ribosome entry site from
- 25 encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
4. 1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 19
- 30 5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
6. 3709-3870: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
7. 3871-4653: Nonstructural protein 4B (NS4B)
- 35 8. 4654-5994: Nonstructural protein 5A (NS5A)

9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

7771-8001: HCV 3' non-translated Region

5 **SEQ ID-NO: 11**

Name: I389/Core-3'/19

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker

3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame

4.1813-10842: HCV Polyprotein from Core up to Nonstructural protein 5B of the cell culture-adpated mutant 19

5. 1813-2385: HCV Core Protein; structural protein

6. 2386-2961: envelope protein 1 (E1); structural protein

7. 2962-4050: envelope protein 2 (E2); structural protein

8. 4051-4239: Protein p7

9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3 Protease

10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase

11.6784-6945: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor

12.6946-7728: Nonstructural protein 4B (NS4B)

13.7729-9069: Nonstructural protein 5A (NS5A)

14.9070-10842: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

15.10846-11076: HCV 3' non-translated Region

The invention is described in detail in the following by way of examples of embodiments and respective diagrams. These are as follows

5 **Fig. 1 A:** The structure of a HCV-RNA construct according to the invention

10 On the top a diagram of the structure of the complete parental HCV genome is given with the positions of the genes for the cleavage products core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B within the polyprotein, and the 5' and 3' non-translated regions (5' NTR and 3' NTR) - shown as a thick horizontal line -, and with the  
15 two positions selected for the creation of the subgenomic constructs, namely the position of the 'GDD catalytic domain' of the NS5B RNA polymerase (GDD) and the position of the 3' end of the HCV-IRES (nucleotide positions 1 to 377 and 1 to 389) -  
20 - drawn above the diagram of the genome. The numbers below the diagram of the genome indicate the respective nucleotide positions.

25 Diagrams of the structure of two modified HCV-RNA constructs (subgenome) according to the invention are shown below, consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (Neo<sup>R</sup>), the EMCV-IRES (E-I) and the HCV sequences of NS2 or NS3 up to the authentic 3' end. The position of  
30 the 10-amino acid deletion comprising the NS5B polymerase GDD motive is marked with a triangle ( $\Delta$ ).

35 **Fig. 1 B:** The result of a denaturing formaldehyde-agarose gel electrophoresis for the detection of replicating plus-strand RNA in transfected subpassaged Huh 7 cell clones.

5 The positions of HCV specific RNAs (arrows) and the 28S rRNA are specified to the right of lane 12, the size (number of nucleotides) of the RNA marker (M) is specified to the left of lane 1.

10 **Fig. 1 C:** The result of a PCR test with subsequent Southern Blot to demonstrate the absence of integrated replicon DNA in most of the selected cell clones.

The lanes 1 and 2 show the positive control, lane 13 the negative control. The figures to the left of lane 1 indicate the size of the nucleotide marker molecules.

15 **Fig. 2 A:** The result of a PCR test with subsequent Southern Blot for the detection of the sensitive exclusion of integrated replicon DNA (plasmid molecule I<sub>377</sub>/NS3-3'/wt) in a cell clone containing a HCV-RNA construct (9-13).

20 The lanes 7 to 11 represent the result of a titration of DNA molecules of the construct I<sub>377</sub>/NS3-3'/wt without addition of total DNA of the cell clone 9-13, and the lanes 2 - 6 represent the result obtained with the same plasmid molecules with the addition of 1 µg 9-13 DNA each prior to the PCR (for the purpose of excluding an inhibitor of the PCR in the DNA preparation). Lane 13 represents the negative control (PCR without DNA template). Lane 1 shows the result, which was achieved with one µg total DNA of the cell clone 9-13.

35 **Fig. 2 B:** The result of a Northern Blot test to quantify the amounts of HCV plus- and minus-strand RNA.

The arrows mark the positions of replicon RNA.

The "plus" and "minus" details indicate the positive (plus) and negative (minus) polarity of the RNA controls, which have been applied to the gel. "Minus-strand" and "Plus-strand" indicate the specificity of the radioactive RNA probes.

5

10

**Fig. 2 C:** The result of a formaldehyde-agarose gel electrophoresis after radioactive labeling of the intracellular replicated HCV-RNA to demonstrate the resistance of HCV-RNA replication to dactinomycin.

15

**Fig. 3 A:** The detection of HCV specific antigens in the selected cell clones by means of immunoprecipitation after metabolic radioactive labeling.

20

The lanes 7 - 9 represent authentic size marker (which have been obtained in Huh 7 cells after the transient expression of a HCV-RNA construct); identified HCV proteins are labeled on the left edge of lane 1, the molecular weights (in Kilodalton) are specified to the right of lane 9.

25

**Fig. 3 B:** Results of an immunofluorescence test to establish the subcellular localisation of HCV antigens.

30

**Fig. 4 :** Diagram of the structure of a selectable HCV-RNA construct according to the invention (complete genome) consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (NeoR), a heterologous IRES element, for example from the encephalomyocarditis virus (E-I), the complete HCV open reading frame and the authentic 3' NTR.

35

- 5 **Fig. 5:** Diagram of the structure of HCV-RNA constructs with inserted antibiotic resistance gene **(A)** within the nucleotide sequence encoding the polyprotein (monocistronic RNA construct), and **(B)** within the 3' NTR (bicistronic RNA construct).
- 10 **Fig. 6:** Diagram of the structure of HCV-RNA constructs with inserted reporter genes **(A)** as part of a HCV replicon from NS3 to NS5B, - in the end the reporter protein is cleaved by viral or cellular proteases out of the polyprotein or the marker gene for selection (selection gene) or the resistance gene are transfected into the cells through co-transfection -, **(B)** as part of a fusion gene composed of a resistance and reporter gene (for example for the neomycin phosphotransferase and green fluorescent protein), **(C)** as part of a replicon composed of a resistance and reporter gene (for example the neomycin phosphotransferase and green fluorescent protein) bound via a nucleotide sequence, which encodes for an amino acid sequence (area with hatches), which can be cleaved by a protease or has a self-cleaving (autocatalytic) activity, **(D)** as independent gene (in this case the green fluorescent protein), which is expressed from its own internal ribosome binding site (IRES); - the resistance gene (in this case the neomycin phosphotransferase gene) is also expressed from its own internal ribosome binding site (IRES) (polycistronic construct).
- 15
- 20
- 25
- 30
- 35 **Fig. 7:** Diagram of the structure of a HCV-RNA construct, where the resistance gene is linked to the HCV-

RNA sequence via a ribozyme or a recognition site for a ribozyme.

5 The thick lines illustrate the HCV 5' and 3' NTRs, E-I is a heterologous internal ribosome binding site, which is required for the expression of the resistance gene, and the grey square illustrates the ribozyme or a recognition site for a ribozyme.

10

**Fig. 8:** Diagram of the structure of a HCV-RNA construct with resistance gene and integrated foreign gene.

**Fig. 9:** Method for comparing the specific infectivity (expressed as number of cell colonies obtained) of total RNA against in vitro transcripts. HCV-RNA is generated by in vitro transcription of a corresponding RNA construct and quantified by measurement of the optical density at 260 nm (OD 260 nm). A defined number of these molecules is mixed with a specified amount of total RNA from naive Huh-7 cells and this mixture is transfected into naive Huh-7 cells with the aid of electroporation. At the same time the total RNA of a cell clone, produced by the method described in Figure 1, is isolated using a known state of the art method and the amount of HCV-RNA contained therein is determined by means of Northern Blot using a HCV specific RNA probe and subsequent quantification via phosphoimaging. A defined amount of this total RNA is analogously transfected in naive Huh-7 cells. These cells from both the cultures are then subjected to a G418 selection and the number of colonies created is determined by counting after fixing and staining with Coomassie Brilliant Blue. For the determination of transfection efficiency 1µg of a

20

25

30

35

plasmid allowing the expression of luciferase is added to each transfection culture. An aliquot of the transfected cells is collected after 24 hours and the luciferase activity determined in the respective cell lysates. The number of colonies is always normed to the luciferase expression.

**Fig. 10:** Sequence analysis of the 9-13 clone. Total RNA of the 9-13 cell clone, resulting from transfection of HCV-RNA construct I377/NS3-3', was isolated using a known state of the art method and the HCV-RNA construct amplified from nucleotide position 59 to 9386 with the aid of 'long-distance RT-PCR' using primer S59 and A9413. The PCR fragments were cloned and 11 clones (9-13 A - K) completely sequenced. Clones D and I, E and G as well as H and J turned out to be identical, respectively. The positions of the amino acid differences in the NS3-5B region between the recloned HCV-RNAs and the parental construct are marked with a thick vertical mark for each clone. Each clone was digested with restriction enzyme *SfiI* and the respective fragment inserted in the parental construct. These clones were each transfected in Huh-7 cells and the cells subjected to selection as described in Figure 1. The number of cell clones obtained with each construct is noted next to the respective construct on the right.

**Fig.11 A:** Principle of determination of replication with the aid of a reporter gene. In the upper part of the figure, the HCV-DNA construct I<sub>389</sub>/Luc/NS3-3' is shown. It consists of the HCV 5' NTR (nucleotide positions 1-389), the luciferase gene (*luc*), the IRES of the encephalomyocarditis virus, the HCV NS3-5B and the 3' NTR. The

position of the active centre of the NS5B RNA polymerase, into which a deactivating amino acid exchange was introduced, is indicated by 'GND'. The plasmids, which code for the HCV-RNA construct which is able to replicate or is defective, are digested with the restriction enzyme *ScaI* and added to an in vitro transcription reaction with T7 RNA polymerase. After removal of the template DNA, the respective HCV-RNA constructs were transfected in naive Huh-7 cells by means of electroporation and the latter collected at regular intervals.

**Fig.11 B:** Comparison of luciferase activity in cells transfected with the parental HCV-RNA construct I<sub>389</sub>/Luc/NS3-3'/wt (wt) or the following variants: inactive RNA (318 DN), variants 9-13F or variant 5.1. The cells were collected 6 (not shown), 24, 48, 72, 96, 120, 144 and 168 hours after transfection and luciferase activities determined by luminometric measurement.

**Fig. 12:** Selectable HCV full-length genomes (constructs I<sub>389</sub>/core 3'/5.1 and I<sub>389</sub>/core 3'/9-13F). (A) Diagram of the full-length construct. The region between both indicated recognition sites for the restriction enzyme *SfiI* corresponds to the sequences of the highly adapted RNA variants 5.1. or 9-13F. (B) Number of colonies which were obtained after transfection of 0.1 µg in vitro transcribed RNA of the construct I<sub>389</sub>/core-3'/5.1. described under A into HUH7-cells. The result of a representative experiment is given. (C) Demonstration of autonomously replicating HCV full-length RNAs in G418 resistant cell clones which were obtained after transfection of the

corresponding in vitro transcript. The illustration shows the autoradiogram of a Northern Blot, which was hybridised with a probe against the neo-resistance gene and the HCV 5' NTR. The controls shown in lanes 1 and 2 each correspond to  $10^8$  molecules of the indicated in vitro transcripts, mixed with total RNA from naive Huh-7 cells. The negative control contains only total RNA from naive Huh-7 cells (lane 3). Lanes 4-9 contain 3-10  $\mu$ g total RNA from G418 resistant cell clones, which were obtained after transfection by in vitro transcribed I<sub>389</sub>/core 3'/5.1 RNA or I<sub>389</sub>/core 3'/9-13F RNA. The G418 concentration used for the selection is indicated in each case. Five of the cell clones shown contain the highly adapted RNA variant 5.1 (lanes 4-8), one contains the adapted RNA variant 9-13F (lane 9).

**Fig. 13:** HCV-RNA constructs with a reporter gene. (A) Bicistronic HCV-RNA constructs. The reporter gene is translated with the aid of a separate IRES. (B) Monocistronic HCV-RNA constructs. The reporter gene product is expressed as fusion protein with a HCV protein. Both portions are linked via a recognition sequence for a viral or cellular protease, which permits a proteolytic separation of the two fused protein portions. In the example shown the reporter gene product and respective HCV protein was fused through a recognition sequence for ubiquitin (Ub).

**Fig. 14:** Tricistronic full-length HCV-RNA construct, that in addition to the resistance gene possesses an inserted foreign gene (ther. gene).

**Fig. 15:** Monocistronic HCV-RNA constructs, for which the resistance gene product is expressed as a fusion protein with HCV portion. The resistance gene (RG) is either active as a fusion protein or it is fused with the HCV portion via a proteolytically cleavable sequence in such a way that the resistance gene product is split from the HCV portion by a cellular or viral protease. In the example shown the resistance gene was fused with the respective HCV portion through the sequence coding for ubiquitin (Ub).

**Example 1: Production of HCV-RNA constructs**

**(A) Synthesis and cloning of a complete HCV consensus genome by means of RT-PCR**

The HCV genome, which means the HCV-RNA, was isolated from the liver of a chronically infected patient as described in the following:

The total RNA was isolated from approx. 100 mg of liver according to the method described by Chomczynski and Sacchi (1987, Anal. Biochem. 162, 156). Using 1 µg of this isolated RNA a reverse transcription was carried out with the primers A6103 (GCTATCAGCCGGTTCATCCACTGC) or A9413 (CAGGATGGCCTATTGG CCTGGAG) and the expand reverse transcriptase system (Boehringer Mannheim, Germany) according to the manufacturer's recommendations. A polymerase chain reaction (PCR) was carried out with the products of this reverse transcription using the expand long template system (Boehringer Mannheim, Germany), in which the buffer containing 2 % of dimethyl sulfoxide was used. After one hour at 42 °C, 1/8 of the reaction mixture was used for a first PCR with primers A6103 and S59

(TGTCTTCACGCAGAAAGCGTCTAG) or A9413 and S4542 (GATGAGCT  
CGCCGCGAAGCTGTCC). After 40 cycles, 1/10 of this reaction  
mixture was used for a second PCR with primers S59 and  
A4919 (AGCACAGCCCGCGTCATAGCACTCG) or S4542 and A9386  
5 (TTAGCTCCCCG TTCATCGGTTGG. After 30 cycles the PCR products  
were purified by means of preparative agarose gel  
electrophoresis and the eluted fragments were ligated into  
the vector pCR2.1 (Invitrogen) or pBSK II (Stratagene).  
Four clones of each fragment were analyzed and sequenced,  
10 and a consensus sequence was established. For this purpose  
the DNA sequences were compared to each other. The  
positions, where the sequence of one fragment was different  
from the others, were considered as undesired mutations. In  
the case of ambiguities of the sequence, shorter  
15 overlapping PCR fragments of the respective region were  
amplified and several clones sequenced. By this means  
several potential mutations could be identified in each  
fragment and consequently an isolate specific consensus  
sequence could be established. This established consensus  
20 sequence or genome belongs to the worldwide spread genotype  
1b. The non-translated region at the 3' end (=3' NTR) was  
obtained by means of a conventional PCR, whereby an  
antisense primer, which covers the last 24 nucleotides of  
the 'X-tail' known from the state of the art was used  
25 (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215,  
744; und Rice, PCT/US 96/14033). The authentic non-  
translated region on the 5' end (=5' NTR) downstream of the  
T7 promoter was created by means of PCR, whereby an  
oligonucleotide, which corresponds to a shortened T7  
30 promoter (TAA TAC GAC TCA CTA TAG) and the first 88  
nucleotides of HCV, was used on one hand and one of the  
previously mentioned plasmids, which carries one of the 5'  
fragments of the genome, was used on the other hand. A  
complete HCV consensus genome was assembled from the  
35 subgenomic fragments with the smallest number of non-  
consensus replacements and inserted into a modified pBR322  
vector. Deviations from the consensus sequence were

eliminated by means of site-directed mutagenesis. In order to produce run-off transcripts with an authentic 3' end, the 3' NTR of the isolates (with the end TGT) was modified to AGT (according to the sequence of the genotype 3 = clone 5 'WS' according to Kolykhalov et al., 1996, J. Virol. 70, 3363) and an additional nucleotide replacement was carried out at position 9562, to preserve the A:T pairing in the hairpin structure at the 3' end of the 3' NTR (Kolykhalov et al. ibid.). In order to eliminate an internal restriction 10 site for the Scal enzyme, a silent nucleotide replacement was further carried out. After joining the full-length genome with the appropriate 5' and 3' NTRs the complete HCV sequence was analyzed. No undesired nucleotide replacement was detected.

15

The HCV genome produced in this way should be hepatotropic according to the definition.

(B) Synthesis of selectable HCV subgenomic constructs

20

By means of the consensus genome described under (A), HCV subgenomic constructs were created, which include the antibiotic resistance gene neomycin phosphotransferase (NPT) and two sequences of internal ribosome entry sites 25 (IRES). The biochemical procedures used for this are known and familiar to the expert (see: 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 30 Protocols in Molecular Biology, Vol. 1-3, John Wiley & Sons Inc., New York). The antibiotic resistance gene was inserted immediately after the 5' NTR, obtaining a bicistronic RNA as a result (see Fig. 1 A). However, the antibiotic resistance gene might as well be inserted at 35 another site of the HCV subgenomic construct, for example within the nucleotide sequence encoding for the polyprotein, obtaining a monocistronic RNA as a result (see

Fig. 5 A), or in the 3' NTR (see Fig. 5 B). The IRES elements are one of the two HCV-IRES variants from nucleotides 1-377 or nucleotides 1-389 as well as the IRES of the encephalomyocarditis virus, which controls the translation of the HCV sequence downstream from the genes for NS2 or NS3 up to the authentic 3' end of the genome.

The two HCV-IRES variants mentioned were determined in the following way:

10

Based on deletion analyses of the 3' end of the HCV-IRES (Reynolds et al.1995, *EMBO J.* 14, 6010) various segments of the 5' NTR were fused with the NPT gene and analyzed with regard to the maximum number of colonies formed by means of co-transfection with a plasmid containing the T7 RNA polymerase gene. The best results were achieved with the HCV sequences from 1-377 to 1-389. Due to the fact that the AUG start codon of the HCV polyprotein is located at position 342 and is consequently included in the IRES sequence, 12 or 16 amino acids of the HCV core protein ("core protein") were fused with the neomycin phosphotransferase (see Fig. 1 A).

Accordingly, these modified HCV subgenomic constructs were given the designations I<sub>377</sub>/NS2-3' (or I<sub>377</sub>/NS3-3') and I<sub>389</sub>/NS2-3' (or I<sub>389</sub>/NS3-3'). They are illustrated in a diagram in Fig. 1A.

Different cell lines and primary cell cultures of human hepatocytes were transfected with in vitro transcripts of these modified parental HCV subgenomic constructs I<sub>377</sub>/NS2-3' (or I<sub>377</sub>/NS3-3') and I<sub>389</sub>/NS2-3' (or I<sub>389</sub>/NS3-3').

As a parallel negative control for all transfection experiments, an appropriately modified, but defective subgenome was constructed for every modified parental HCV subgenomic construct, which differs from the parental

construct due to the fact, that it has a deletion of 10 amino acids within the reading frame comprising the active site of the NS5B RNA polymerase (Behrens et al., 1996, *EMBO J.* 15, 12; and Lohmann et al., 1997, *J. Virol.* 71, 8416).

5

(C) Synthesis of selectable HCV genomic constructs

A NS2-3' subgenomic construct, linked at its 5' end to a fragment of the luciferase gene and the complete EMCV-IRES, was restricted with NcoI and SpeI and purified using preparative agarose gel electrophoresis. The vector obtained in this way was ligated with a NcoI/NotI-HCV fragment corresponding to the nucleotide positions 342 to 1968 of the HCV genome and a NotI/SpeI fragment corresponding to the nucleotide positions 1968-9605 in a 3-factor-ligation. The resulting construct, where the complete HCV open reading frame and the 3' NTR lie downstream from the luciferase gene fragment and the EMCV-IRES, was then restricted with PmeI and SpeI and ligated with the analogously restricted I<sub>389</sub>/NS3-3'/wt subgenomic construct vector. This selectable HCV genomic construct is illustrated in Fig. 4.

(D) Production of in vitro transcripts corresponding to the HCV-RNA constructs

The previously described purified plasmid DNAs were linearized with ScaI and used for an in vitro transcription reaction after phenol/chloroform extraction and isopropanol precipitation using the following components: 80 mM HEPES, pH 7.5, 12,5 mM MgCl<sub>2</sub>, 2 mM Spermidine, 40 mM Dithiothreitol, 2 mM of each NTP, 1 unit RNasin/ $\mu$ l, 50  $\mu$ g/ml restricted DNA and approx. 2 units/ $\mu$ l T7 RNA polymerase. After 2 hrs. at 37 °C half of the amount of T7 polymerase was added and the reaction mixture was incubated for two further hours. In order to remove DNA the mixture was extracted with acid phenole (U. Kedzierski, J.C. Porte,

1991, Bio Techniques 10, 210), precipitated with isopropanol, the pellet was dissolved in water and incubated with DNase (2 units per  $\mu\text{g}$  DNA) for 60 min. at 37 °C. After subsequent extraction with acid phenole, acid phenole/chloroform and chloroform as well as isopropanol precipitation the dissolved RNA was quantified using optical density measurement and its integrity was checked using formaldehyde-agarose gel electrophoresis.

10

**Example 2: Transfection experiments with the hepatoma cell line Huh 7**

With all transfection experiments it was carefully ensured that any template DNA had been removed beforehand so as to avoid the possibility of this DNA integrating in transfected cells and conferring a neomycin resistance to them independent from HCV replication. The reaction mixture was therefore treated with 2 units of DNase per  $\mu\text{g}$  DNA for 60 min. at 37 °C after the in vitro transcription and extracted with acid phenole, acid phenole/chloroform and chloroform. Prior to being used for the transfection the precipitated RNA was analyzed using formaldehyde-agarose gel electrophoresis.

25

Three separate transfection experiments were carried out with the highly differentiated human hepatoma cell line Huh 7 (according to Nakabayashi et al. 1982, *Cancer Res.* **42**, 3858). Each time 15  $\mu\text{g}$  RNA were introduced into  $8 \times 10^6$  Huh 7 cells by electroporation and the cells were seeded in culture dishes with a diameter of 10 cm. 24 hours after seeding, neomycin (= G418) was added in a final concentration of 1 mg/ml. The culture medium was changed twice a week. After 3 - 5 weeks small colonies were visible, which were isolated and grown under the same culture conditions.

35

The cell clones obtained during the first experiment were isolated and subpassaged. Most of the clones died during this procedure, and the final yield was only 9 clones from cells, which had been transfected with the parental HCV subgenomic constructs and 1 clone (clone 8-1) from cells, which had been transfected with a defective HCV genomic construct, namely a defective NS2-3' HCV-RNA. Apart from an extended doubling time and the occasional occurrence of irregularly shaped cells, no consistent morphological differences were found between these 9 cell clones and the single cell clone (clone 8-1) or the parental Huh 7 cells.

The main criteria for functioning HCV genomic constructs are the formation of viral RNA with the correct size and the absence of (integrated) plasmid DNA, which could transfer or mediate on a G418 resistance.

To determine the HCV-RNA in the Huh 7 cells, the total RNA was isolated and analyzed by means of the common Northern Blot method using a plus-strand specific ribo probe (RNA probe). For this purpose the total RNA was isolated from the respective cell clones according to the method described by Chomczynski and Sacchi 1987, *Anal. Biochem.* 162, 156 and 10  $\mu\text{g}$  RNA, which is equivalent to the total RNA content of  $0,5 - 1 \times 10^6$  cells, are separated using denatured formaldehyde-agarose gel electrophoresis (lanes 3 to 12 of Fig. 1 B). At the same time  $10^9$  in vitro transcripts (ivtr.), which correspond to the I<sub>389</sub>/NS2-3'/wt or the I<sub>389</sub>/NS3-3'/wt replicon RNAs, are separated as well as size markers with authentic sequence (lane 1 or lane 2). The separated RNA was transferred onto nylon-based membranes and hybridized with a radioactively labeled plus-strand specific RNA probe, which was complementary to the complete NPT gene and the HCV-IRES of nucleotide 377 to nucleotide 1. The positions of the HCV specific RNAs (arrows) and the 28S rRNA are specified to the right of lane 12, the size (amount of nucleotides) of the RNA marker

is specified to left of lane 1. The RNA marker fragments contain HCV sequences and therefore hybridize with the ribo probe (= RNA probe). The results of this analysis are illustrated in Fig. 1 B.

5

With the exception of clone 8-1 transfected with the defective HCV genomic construct, all cell clones produced homogenous HCV-RNAs of correct lengths (approx. 8640 nucleotides in the case of the NS2-3' and approx. 7970  
10 nucleotides in the case of NS3-3' replicon). This result is an indication of the fact, that the functional replicons or the functional HCV genomic constructs transfer the G418 resistance. In order to exclude, that the G418 resistance is caused by a plasmid DNA, which is integrated in the  
15 genome of the Huh 7 host cell and transcribed under control of a cellular promoter, the DNA of each clone was analyzed by means of a NPT gene specific PCR. Consequently the DNA was isolated from the selected Huh 7 cell clones by means of digestion with proteinase K (40µg/ml, 1h, 37 °C) in  
20 10mMTris, pH7,5, 1mM EDTA, 0,5 % SDS and subsequent extraction with phenol, phenol/chloroform and isopropanol precipitation. The DNA precipitate was dissolved in 10 mM Tris (pH 7,5) and 1 mM EDTA and incubated with Rnase A for 1 hour. Following a phenol/chloroform extraction and  
25 ethanol precipitation, 1 µg DNA, equivalent to 4 - 8 x 10<sup>4</sup> cells, was analyzed by means of PCR using NPT gene specific primers (5'-TCAAGACCGACCTG TCCGGTGCCC-3' and 5'-CTTGAGCCTGGCGAACAGTTCGGC-3'), and a DNA fragment consisting of 379 nucleotides was generated. The specificity of the  
30 PCR product was established by means of the Southern Blot method, in which a DNA fragment labeled by digoxigenin was used, which corresponds to the NPT gene. As positive controls (for the detection of possibly contaminating nucleic acids) the PCR method was carried out with 10<sup>7</sup>  
35 plasmid molecules or 1 µg DNA from a BHK cell line, which was stably transfected with a neomycin resistance gene, and

as negative control, the PCR was carried out with the same reagents but without added DNA.

The results of this analysis are illustrated in Fig. 1 C. The lanes 1 and 2 represent the positive controls, lane 13 represents the negative control. The numbers to the left of lane 1 indicate the sizes of the nucleotide marker molecules.

A NPT-DNA could not be detected in any cell clone, apart from clone 7-3 (Fig.1C, lane 3), which was obtained from cells after transfection with a NS2-3' replicon/NS2-3'HCV genomic construct, and clone 8-1 (Fig.1C, lane 12), which was obtained from cells after transfection with a defective HCV genomic construct. This result was another indication of the fact, that the G418 resistance of most clones was passed on by the replicating HCV-RNA. But even regardless of these results, it is unlikely, that integrated plasmid DNA produces HCV-RNAs of correct sizes, because the plasmids used for the in vitro transcription contain neither an eukaryotic promoter nor a polyadenylation signal. In the case of clone 7-3 the resistance is therefore very probably passed on by the HCV-RNA construct or the replicating HCV-RNA as well as by an integrated NPT DNA sequence, whereas the resistance of the cells of clone 8-1 is only caused by the integrated plasmid DNA.

Clone 9-13 (Fig. 1 B, lane 11) was subject to further tests, to confirm that the G418 resistance was passed on by a self-replicating HCV-RNA. Clone 8-1, which carries the integrated copies of the NPT gene, was used throughout as negative control. A PCR was carried out, which allows the detection of < 1000 NPT gene copies in ~ 40.000 cells, with the aim to rigorously exclude the presence of NPT-DNA in clone 9-13. The result of this PCR is illustrated in Fig. 2A. The PCR proceeded in detail as follows:

During the test,  $10^6$  -  $10^2$  plasmid molecules (I<sub>377</sub>/NS3-3'/wt) were used either directly (lanes 7 - 11) or after adding each 1  $\mu$ g 9-13 DNA (lanes 2 - 6). The specificity of the amplified DNA fragment was determined by Southern Blot  
5 using a NPT specific probe. A PCR without DNA probe was carried out as negative control (lane 12).

Even with this sensitive method, no plasmid DNA could be detected in one  $\mu$ g DNA of the cell clone 9-13 (lane 1). To  
10 estimate the amount of HCV plus- and minus-strand RNAs in these cells, a dilution series of total RNA was analyzed with the Northern Blot method using a plus- or minus-strand specific radioactively labeled ribo probe (= RNA probe). For this purpose, 8, 4 or 2  $\mu$ g of total RNA, which have  
15 been isolated from the cell clones 9-13 and 8-1, were analyzed in parallel to known amounts of in vitro transcripts with plus- or minus-strand polarity (control RNAs) in the Northern Blot method and were then subjected to a hybridization. The hybridization was carried out with  
20 a plus-strand specific ribo probe, which covered the complete NPT gene and the HCV-IRES ('plus-strand', top panel), or with a minus-strand specific RNA probe, which was complementary to the NS3 sequence ('minus-strand', bottom panel). The arrows mark the positions of replicon  
25 RNA. The results of this analysis are illustrated in Fig. 2 B.

Approx.  $10^8$  copies/ $\mu$ g total RNA were detected in the case of the plus-strand, which is equivalent to 1000 - 5000 HCV-  
30 RNA molecules per cell, whereas the amount of minus-strand RNA was 5- to 10-times less. This result corresponds to the assumption, that the minus-strand RNA is the replicative intermediate form or intermediate copy, which is used as a template for the synthesis of the plus-strand molecules.

35

Due to the fact that the reaction is mainly catalyzed by the viral RNA-dependent RNA polymerase, the synthesis of

the HCV-RNAs should be resistant to dactinomycin, an antibiotic, which selectively inhibits the RNA synthesis of DNA templates, but not the RNA synthesis from RNA templates. To confirm this assumption, cells were incubated with [<sup>3</sup>H] uridine in the presence of dactinomycin, the radioactively labeled RNAs were extracted, separated by means of denaturing agarose gel electrophoresis and analyzed with the aid of a common Bio-Imager using a [<sup>3</sup>H]-sensitive screen. For this purpose approx. 5 x 10<sup>5</sup> cells of the clones 9-13 and 8-1 were incubated at a time with 100 µCi [<sup>3</sup>H] uridine for 16 hrs. in the absence (-) or presence (+) of 4 µg/ml of dactinomycin (Dact). Following this labeling reaction the total RNA was prepared and analyzed by means of formaldehyde-agarose gel electrophoresis. Only 1/10 of the total RNA is illustrated in the first two lanes. The radioactively labeled RNA was visualized using a BAS-2500 Bio-Imager (Fuji).

The results of this analysis are illustrated in Fig. 2 C. In accordance with the inhibitor profile of the NS5B polymerase (Behrens et al., 1996, *EMBOJ.* 15, 12 and Lohmann et al., 1997, *J. Virol.* 71, 8416), dactinomycin had no influence on the replication of the HCV RNA, whereas the synthesis of cellular RNA was inhibited. A RT-PCR was carried out for recloning the replicating sequences, to confirm the identity of the viral RNA. The sequence analysis of the recloned RNA showed that the RNA in clone 9-13 is HCV specific and corresponds to the transfected transcript of the HCV construct I<sub>377</sub>/NS3-3'/wt.

For the analysis of the viral proteins, at first the respective cells were metabolically radioactively labeled with [<sup>35</sup>S] methionine/cysteine, then lysed and afterwards the HCV specific proteins were isolated from the cell lysates by means of immunoprecipitation. The results of these analyses are illustrated in Fig. 3 A. The detailed procedure was as follows: Cells of the cell clones 9-13

(wt) and 8-1 ( $\Delta$ ) were metabolically radioactively labeled by treating them with a protein labeling mixture familiar to the expert and available on the market (for example. NEN Life Science). The HCV specific proteins were separated  
5 from the cell lysate by immunoprecipitation (IP) under non-denaturing conditions (for example according to Bartenschlager et al., 1995, *J. Virol.* 69, 7519) using three different antisera (3/4, 5A, 5B, according to the labeling on the top end of the lanes 1 to 12). The immune  
10 complexes were analyzed by means of tricine SDS-PAGE and made visible by means of autoradiography. To obtain authentic size markers, the homologous replicon construct I<sub>377</sub>/NS3-3'/wt was subject to a transient expression by the vaccinia virus-T7 hybrid system in the Huh 7 cells. The  
15 resulting products were used as size markers (lanes 7 - 9) parallel to the cells of the clones 9-13 and 8-1. Identified HCV proteins are labeled on the left edge of lane 1, the molecular weights (in Kilodalton) are specified on the right edge of lane 9. It should be noted that the  
20 NS3/4 specific antiserum ('3/4') used preferably reacts with NS4A and NS4B causing an underrepresentation of NS3.

All viral antigens could unambiguously be detected, and their apparent molecular weights did not show any  
25 difference to those being detected after transient expression of the same bicistronic HCV-RNA construct in the original Huh 7 cells. An immunofluorescence detection reaction was carried out using NS3 and NS5A specific antisera, to determine the subcellular distribution of the  
30 viral antigens (for example according to Bartenschlager et al., 1995, *J. Virol.* 69, 7519). For this purpose cells of the clones 9-13 (wt) and 8-1 ( $\Delta$ ) were fixed with methanol/acetone 24 hrs. after incubating on coverslips and incubated with polyclonal NS3 or NS5A specific antisera.  
35 The bound antibodies were made visible with a commercially available FITC conjugated anti-rabbit antiserum. The cells

were counterstained with the 'Evans Blue' stain to suppress  
unspecific fluorescence signals.

The results of this detection test are illustrated in Fig.  
5 3 B. A strong fluorescence in the cytoplasm could be  
detected with both antisera. The NS5A specific antiserum  
also caused a slight nuclear fluorescence, which indicates  
that at least small amounts of this antigen also reach the  
nucleus. But the generally dominating presence of the viral  
10 antigens in the cytoplasm are a strong indication that HCV-  
RNA replication occurs in the cytoplasm, as is the case  
with most RNA viruses.

These results prove clearly that the establishment of a  
15 cell culture system for the HCV could be accomplished with  
the test arrangement described, the efficiency of which  
surpasses everything known up until now by far and for the  
first time allows the detection of viral nucleic acids and  
proteins with conventional and approved biochemical  
20 methods. This efficiency actually allows detailed  
examination of HCV pathogenesis, genetic analyses of  
different HCV functions and a precise study of the  
virus/host cell interaction, through which new starting  
points for the development of a antiviral therapy can be  
25 defined.

### **Example 3: Transfection of Huh 7 cells with HCV total constructs**

30 Huh7 cells are transfected and selected as described in  
example 2, whereby in this case selectable constructs are  
used, which contain the complete virus genome.  
Corresponding to Example 2, the resulting cell clones are  
35 tested for the absence of HCV-DNA by means of PCR and the  
productive replication of HCV-RNA is then established by  
means of Northern Blot, [<sup>3</sup>H]uridine labeling in the

presence of dactinomycin, detection of the viral proteins or antigens preferably with the aid of the Western Blot, the immunoprecipitation or immunofluorescence. In contrast to the arrangements described in Example 2, the construct described here makes it possible to obtain more complete and very likely infectious viruses, which has not been the case in the subgenomic constructs described in Example 2. These viruses existing in the cell and the cell culture supernatant are concentrated for example by means of ultracentrifugation, immunoprecipitation or precipitation with polyethyleneglycol, and all exogenous nucleic acids, which means those that are not incorporated into the virus particle, are digested by incubation with nucleases (RNase, DNase, micrococcal nuclease). In this way all contaminating nucleic acids, which are not included in the protecting virus particle, can be removed. After inactivation of the nucleases, the protected viral RNA is isolated for example by means of incubation with proteinase K in a buffer containing SDS, by extracting with phenol and phenol/chloroform and detected by means of Northern Blot or RT-PCR using HCV specific primers. Also in this test arrangement, the combination of the HCV consensus genome described with a selection marker was crucial for the efficient production of viral RNA, viral protein and therefore HCV particles.

**Example 4: Production and application of a HCV-RNA construct, whereby the resistance gene is linked to the HCV subgenomic sequence via a ribozyme or a recognition site for a ribozyme.**

A HCV-RNA construct is produced according to Example 1 or Example 3, where an antibiotic resistance gene is linked to the HCV-RNA sequence through a ribozyme or a recognition sequence for a ribozyme. These constructs are illustrated in a diagram in Fig. 7. Huh 7 cells are transfected with

this HCV-RNA construct as described in Example 2. A selection with the appropriate antibiotic follows the transfection into the cells. The inserted ribozyme is activated in the cell clones obtained in the procedure or, in the case of a construct, which carries a recognition sequence for a ribozyme, the ribozyme is transfected into the cell (for example by means of transfection of a ribozyme construct or infection with a viral expression vector, into which the respective ribozyme has been inserted). In both cases the resistance gene is separated from the HCV-RNA sequence by the ribozyme-dependent cleavage. The result in the case of the HCV genome is an authentic HCV genome without a resistance gene, which can form authentic infectious virus particles. A HCV replicon without resistance gene is created in the case of the HCV subgenomic constructs.

**Example 5: Co-transfection of a HCV-RNA construct with a separate luciferase transfection construct**

A HCV-RNA construct is produced according to Example 1 (A) or Example 3 or Example 4. At the same time a transfection construct is produced, which comprises the luciferase gene, whereby this luciferase gene is linked to a first nucleotide sequence, which encodes a HCV protease (for example NS3 protease) cleavage site, to a second nucleotide sequence, which encodes for another protein or a part of another protein. HCV-RNA construct and transfection construct are transfected into any host cells, preferably hepatoma cells, most preferably Huh 7 cells. This can be realised as described in Example 2. The product of the modified luciferase gene is a luciferase fusion protein, where the luciferase is inactivated due to the fusion with the foreign part. The fusion protein, which contains a recognition sequence for a HCV protease, is cleaved in transfected cells with high HCV replication, and

consequently the active form of the luciferase, which can be identified through luminometric measurement, is released. If the replication of the HCV-RNA construct is inhibited the fusion protein will not be cleaved and no active luciferase will be released. The quantitative determination of the luciferase is therefore a measure for the replication of the HCV construct. Instead of the luciferase gene, another reporter gene can just as easily be used, which is modified in the same way, so that its expression depends on the viral replication, although this reporter gene is not part of the HCV construct. A cellular protein, which is deactivated or activated by the HCV proteins or nucleic acid, can also be used as a so called surrogate marker. The expression or activity of this surrogate marker is in this case a measure for the replication of the viral DNA.

**Example 6: Production of HCV subgenomic constructs with integrated foreign genes to be used as liver cell specific vector in gene therapy**

These recombinant and selectable HCV subgenomic constructs are transfected in transcomplementing helper cell lines, which means in cell lines, which express the missing functions (for example the structural proteins) in an inducible or constitutive way. Cell clones containing a functional HCV subgenomic construct can be established through appropriate selection. The viral structural proteins expressed from the host cell allow the formation of virus particles, into which the RNA of the HCV subgenomic constructs will be transfected. The results are therefore virus-like particles, which contain a HCV subgenomic construct according to the invention including the inserted foreign gene and which can transmit this to other cells by means of infection. An example for this construct is illustrated in Fig. 8.

It is also possible to use the HCV subgenomic construct with integrated foreign gene directly as an expression vector. This involves the same method as mentioned previously except that cell lines, which do not express transcomplementing factors, are transfected. In this case the HCV construct is only used as an expression vector.

10 **Example 7: Production of cell culture adapted HCV-RNA constructs**

**(A) Method of isolation**

15 The following method was used to determine adaptive mutations and to produce cell culture adapted HCV-RNA constructs: cells were transfected with a HCV-RNA construct as described in Examples 1 and 2 and G418-resistant cell clone produced. For the determination of ability to replicate (understood in this context to be the number of G418 resistant cell clones obtained per microgram of transfected HCV-RNA or HCV-RNA construct), the total RNA from one of the cell clones, [the 9-13 clone (Fig. 1 B, lane 11)], was isolated and the quantity of HCV-RNA  
20 contained within it was determined by Northern Blot as described in Fig. 2 B. Ten micrograms of the total RNA, containing approx.  $10^9$  molecules of HCV-RNA, was then transfected into naive Huh-7 cells using electroporation (Fig. 9). In parallel,  $10^9$  in vitro transcripts of the analogous neo-HCV-RNA, which had been adjusted with isolated total RNA from naive Huh-7 cells to a total RNA quantity of 10  $\mu$ g, were transfected in naive Huh-7 cells. After selection with G418, the number of cell colonies, expressed as 'colony forming units (cfu) per microgram  
25 RNA', was determined in both these cultures. At a concentration of 500  $\mu$ g/ml G418 in the selection medium, the number of colonies obtained with the HCV-RNA contained

in **isolated** total RNA from clone 9-13, was approx. 100,000 cfu per microgram HCV-RNA. In contrast, only 30 - 50 colonies were obtained with the same quantity of in vitro transcribed HCV-RNA. This result confirms that the specific infectivity of the HCV-RNA isolated from the cell clones is approx. 1,000 - 10,000 times higher than the infectivity of the analogous in vitro transcripts. The experimental approach is shown in Fig. 9.

10 With the aid of 'long-distance RT-PCR', the HCV-RNA was amplified from the total RNA of the 9-13 cells, the PCR amplificate was cloned and numerous clones were sequenced. A comparison of the sequences of these recloned RNAs with the sequence of the RNA originally transfected into the naive Huh-7 cells, showed that the recloned RNAs possessed numerous amino acid exchanges distributed over the whole HCV sequence (Fig. 10). *Sfi*I fragments of these recloned mutants were used to replace the analogous *Sfi*I fragment of the original replicon construct, and RNAs of the respective mutants were transfected in naive Huh-7 cells. After selection with G418 the number of colonies created was determined for each HCV-RNA mutant. While only 30 - 50 colonies per microgram RNA were obtained with the parental RNA the number of colonies was noticeably higher for two of the recloned variants (Fig. 10). In the case of the HCV-RNA constructs 9-13I and 9-13C the specific infectivity was increased to 100 - 1,000 cfu per microgram RNA and for 9-13F replicon it was 1,000 - 10,000 cfu per microgram RNA. These results show that the amino acid exchanges in the analysed NS3-5B regions of the mutants 9-13I, 9-13C and particularly of 9-13F, led to a considerable increase in ability to replicate. In contrast all the other HCV-RNA constructs (9-13 A, B, G, H und K) were no longer able to replicate, they thus contained lethal mutations.

35

In order to answer the question which of the amino acid exchanges in the 9-13F construct led to an increase in

replication, the exchanges were introduced separately or in combination into the parental HCV-RNA construct, and the corresponding RNAs transfected in naive Huh-7 cells. The result of the transfection with these RNAs is summarised in Table 1. From this it is evident that in the present example the high ability to replicate is determined by several mutations. The amino acid exchanges in the HCV-RNA regions NS5A and NS4B make the greatest contribution. The single exchanges in the NS3-Region also make a contribution and perhaps they are synergistic.

These results confirm that it was through the G418 selection of the cells transfected with the neo-HCV-RNA construct that there was enrichment of those HCV-RNAs having noticeably higher ability to replicate. HCV-RNA constructs with greatly differing replication efficiencies can be selected using the experimental approach described here. The higher the concentration of the antibiotic in the selection medium, in/on which the HCV-RNA construct containing cells are cultivated for selection, the higher must be the extent of adaptive mutations and hence replication efficiency of the relevant HCV-RNA constructs, to allow the cells to grow under these conditions. If the selections are carried out using lower antibiotic concentrations, cells can survive and multiply, but the HCV-RNA construct shows a comparatively lower replication efficiency and fewer adaptive mutations.

As has been shown, the 9-13F HCV-RNA construct described so far, which contains several adaptive mutations, had a higher replication efficiency than the parental HCV-RNA. In order to obtain HCV-RNAs with even higher replication in cell culture, the HCV-RNA contained in the total RNA of a selected cell clone was passaged several times in naive Huh-7 cells. The selected 5-15 cell clone, was obtained by transfection with the HCV-RNA construct I<sub>389</sub>/NS3-3' (Fig. 1). It largely corresponds to the cell clone 9-13, produced

by transfection with a HCV-RNA construct, having a HCV-IRES shorter by 22 nucleotides (I<sub>377</sub>/NS3-3'; Fig. 1). Ten micrograms of total RNA, isolated from cell clone 5-15, were transfected into naive Huh-7 cells using electroporation and the cells subjected to a selection with 1 mg/ml G418. The total RNA from one of the cell clones thus produced was again isolated, transfected into naive Huh-7 cells and selected in the same way. This process was repeated a total of four times. After the fourth passage the total RNA was isolated from a cell clone and the neo-HCV-RNA amplified with the aid of the 'long-distance RT-PCR'. The amplified DNA fragment was digested with the restriction enzyme *SfiI* and inserted into the *SfiI*-restricted parental construct I<sub>389</sub>/NS3-3'. Over 100 DNA clones were obtained altogether and then analysed by means of restriction digestion. In vitro transcribed RNA of about 80 of these clones was each transfected into naive Huh-7 cells and subjected to a selection with 500mg/ml G418. Of the 80 neo-HCV-RNA variants examined, the great majority proved to be replication defective. However, the specific infectivity, expressed as 'colony forming units' per microgram RNA, was noticeably increased in the case of two mutants, 5.1 and 19 (Table 2). Through several passages of the RNA in cell culture it is clear that HCV-RNAs are produced whose replication efficiency due to mutations („adaptive mutations“) is several orders of magnitude higher than the original RNA cloned from patients.

### **(B) Modified method**

Adaptive mutations produced and identified according to (A) can be transferred into a HCV-RNA construct with low ability to replicate. This leads to a huge increase in the replication of this construct. The increase is so great it can be demonstrated that HCV-RNAs transfected into cell culture can replicate even in the absence of selection pressure Fig. 12 shows a comparison of the replication

efficiency of HCV-RNAs, which corresponded either to the starting sequence or to the adaptive sequences 9-13F or 5.1. For the purposes of simple measurement, the neo-gene was removed and replaced by the gene for luciferase. The negative control used was again a HCV-RNA construct that was replication defective due to a deactivating mutation in the NS5B RNA polymerase. Already 24 hours after transfection a noticeable difference is evident in luciferase activity between the defective RNA and the 9-13F or 5.1 constructs, while hardly any difference could be seen between the defective RNA (318 DN) and the parental RNA construct (wt) that possessed no adaptive mutations. During the whole period of observation, the highest luciferase activity, and thus highest replication, was obtained with the 5.1-RNA. These results not only confirm the high replication efficiency of this RNA, but also show that it is possible to create a cell culture system with adapted HCV-RNA constructs for which the presence of a selectable gene is no longer necessary. A summary of the nucleotide and amino acid differences between the starting construct and the mutants 9-13F, 5.1 and 19 is presented in Table 3.

**Example 8: Production of cell culture adapted HCV-RNA full-length genome**

In the examples 1 to 7 a subgenomic HCV-RNA was used which lacked the whole structural protein region from core up to p7 or even NS2. It will be shown in this example that it is possible to make a HCV full-length genome replicate in cell culture with the aid of an adapted NS3-5B sequence. For this purpose the *SfiI* fragment of the highly adapted HCV-RNA 5.1 produced according to Example 7 is first transferred into a selectable HCV full-length genome (Fig. 12). This HCV genome was transfected into naive Huh-7 cells and subjected to selection with various G418

concentrations. Depending on the strength of selection (the G418 concentration), a varying large number of cell clones was obtained (Fig. 12 B). By contrast no colonies were obtained with the parental HCV full-length genome containing no adaptive mutations, as was the case for the negative control, which was replication defective due to a deactivating mutation in the NS5B RNA polymerase. To confirm that the thus resulting cell clones really contained an autonomously replicating HCV full-length construct, total RNA from several cell clones was isolated and analysed by means of the Northern Blot method. The full-length HCV-RNA was clearly detectable in all cell clones (Fig. 12). It is thus clearly confirmed, that with the aid of cell culture adapted HCV sequences it is possible to produce a HCV full-length genome, which replicates highly efficiently and autonomously in a cell line, i.e. adapted HCV full-length genomes can also be produced with the system of the invention. Furthermore, as this clone possesses the complete HCV sequence, i.e. it also possesses the structural proteins necessary for virus particle formation, it is possible to produce large quantities of infectious virus particles in cell cultures with this system. As a confirmation of these viruses, cell-free cell supernatants carrying a replicating HCV full-length genome, are added to naive Huh-7 cells and the thus infected cells subjected to selection with G418. Each cell clone growing under these conditions originates from an infected cell. The viruses in the cell culture supernatant of cells possessing a replicating HCV full-length genome can be enriched and purified using various known state of the art methods such as ultracentrifugation or microdialysis. They can then be used for the infection of naive cells. Using this method it is clearly demonstrated that cell culture adapted full-length genomes can be produced with the HCV cell culture system of the invention. These genomes replicate with high efficiency in cells and produce infectious viruses. The latter can be detected by

infection of an experimental animal, preferably a chimpanzee.

5 **Example 9: Production of HCV full-length constructs and HCV subgenomic constructs with reporter genes**

A HCV-RNA construct is produced in which a reporter gene is inserted in place of the antibiotic resistance gene (Fig. 10 13). Replication can thereby be determined through the quantity or activity of the reporter gene or reporter gene product. The reporter gene is preferably a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene 15 (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) or the SEAP gene (secreted alkaline phosphatase gene). This reporter gene and its products, namely the relevant reporter proteins, can be detected for example using 20 fluorescence, chemiluminescence, colorimetrically or by means of immunological methods (for example, enzyme-linked immunosorbent assay, ELISA).

The reporter gene can be expressed either from a separate IRES or in the form of a fusion protein, which is active 25 either as such or fused with a HCV protein via a proteolytically cleavable amino acid sequence in such a way that the reporter is separated from the HCV protein by cleavage of a cellular or viral (HCV) protease.

30

**Example 10: Production of HCV full-length constructs with integrated foreign genes used as liver cell specific vectors for gene therapy or as expression vectors.**

35

The construct (Fig. 14) is transfected in cells and leads to the formation of HCV virus particles that can be used

for the infection of further cells. Since the virus particles have encapsidated RNA with a foreign gene, it can be used in the infected cells for the production of the protein coded by this foreign gene. Cells transfected with  
5 the construct also express the foreign gene.

**Example 11: Production of monocistronic HCV-RNA constructs in which the resistance gene product is expressed as a fusion protein with the HCV portion.**  
10

It is an advantage for some tests if the HCV-RNA construct does not possess a heterologous IRES element. Tests of this  
15 type are, for example, the determination of interferon resistance. If a cell possessing a HCV-RNA construct is incubated with interferon alpha or beta, a reduction in replication of the HCV-RNA results. In order to explain the mechanism of this effect it is necessary for the HCV-RNA  
20 construct not to possess any heterologous IRES, as otherwise it is not possible to determine whether the interferon mediated inhibition is via inhibition of the HCV replication or inhibition of the heterologous IRES. For this reason constructs are produced for which the  
25 resistance gene is fused with a HCV protein (Fig. 15). Either the fusion protein is active as such or the resistance gene product is linked to a HCV protein via a proteolytically cleavable amino acid sequence in such a way that it is separated from the HCV protein by a cellular or  
30 viral (HCV) protease.

**Table 1:** Specific infectivities (cfu/ $\mu$ g RNA) of HCV RNA constructs with adaptive mutations found with the 9-13F mutant introduced into the parental construct I<sub>389</sub>/NS3-3'/wt

amino acid exchange <sup>1</sup>	HCV protein	cfu/ $\mu$ g RNA <sup>2</sup>
none		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	1000-5000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5A	30 - 60
all together		5000

5

<sup>1</sup> amino acid change in the polyprotein of the HCV isolate con 1 (EMBL-gene bank No. AJ238799); amino acids are given in single letter code.

<sup>2</sup> Colony forming units (number of cell clones) obtained with a selection of 500  $\mu$ g/ml G418.

10

**Table 2:** Specific infectivities (cfu/ $\mu$ g RNA) of the parental HCV RNA construct I<sub>389</sub>/NS3-3'/wt and the variants 9-13I, 9-13F, 5.1 and 19.

5

Transfected RNA variant	cfu/ $\mu$ g RNA <sup>1</sup>
wild type	30 - 50
9-13 I	100 - 1.000
9-13 F	1.000 - 10.000
5.1	50.000 - 100.000
19	50.000 - 100.000

<sup>1</sup> Colony forming units (number of cell clones) obtained with a selection of 500 $\mu$ g/ml G418.

**Table 3:** Nucleotide and amino acid sequence differences between the parental HCV RNA construct I<sub>389</sub>/NS3-3'/wt and the mutants 9-13I, 9-13F, 5.1 und 19

Clone	nt-position	nt-exchange	aa-exchange	
9-13 I	3685	C > T	P > L	
	4933	C > T	T > M	
	5249	T > C	-	
	8486	C > T	-	
	8821	G > A	W > stop	
	8991	C > G	R > G	
	9203	A > G	-	
	9313	T > C	F > S	
	9346	T > C	V > A	
	9-13 F	3866	C > T	-
4188		A > G	R > G	
4489		A > C	E > A	
4562		G > A	-	
4983		T > C	-	
5071		A > G	K > R	
5166		A > G	K > E	
6147		C > T	P > S	
6829		A > G	E > G	
7329		A > G	K > E	
7664		A > G	I > V	
8486		C > T	-	
8991		C > G	R > G	
NK5.1		4180	C > T	T > I
		4679	C > T	-
	4682	T > C	-	
	5610	C > A	L > I	
	6437	A > G	-	
	6666	A > G	N > D	

	6842	C > T	-
	6926	C > T	-
	6930	T > C	S > P
	7320	C > T	P > S
	7389	A > G	K > E
NK19	3946	A > G	E > G
	4078	C > G	A > G
	4180	C > T	T > I
	4682	T > C	-
	5610	C > A	L > I
	5958	A > T	M > L
	6170	T > A	-
	6596	G > A	-
	6598	C > G	A > G
	6833	C > T	-
	6842	C > T	-
	6930	T > C	S > P
	7141	A > G	E > G
	7320	C > T	P > S
	7389	A > G	K > E
	7735	G > A	S > N

Given are the differences between the nucleotide and amino acid sequences of the parental HCV RNA sequence con 1 (EMBL-gene bank No. AJ238799) and those of the cell culture adapted HCV RNAs. Numbers refer to the nucleotide and amino acid positions of the con 1 isolate. nt, nucleotide; aa, amino acid.

## SEQUENCE LISTING

&lt;110&gt; Bartenschlager, Ralf

&lt;120&gt; Hepatitis C Virus cell culture system

&lt;130&gt; ba-1

&lt;140&gt; 199 15 178.4

&lt;141&gt; 1999-04-03

&lt;160&gt; 11

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 11076

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C Virus

&lt;400&gt; 1

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctct cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcagggtc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcc cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgccag 900
ccgaactgtt cgccaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcacc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctagecggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccggg aacctggccc tgtcttcttg 1380
acgagcattc ctagggtctt tccccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgtg ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740
tcgagggtta aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaactg taacaccaac 1860
cgccgccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcgggtg agtttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggctc 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccagagg tagggcctgg 2040
gctcagcccg ggtaccctct gccctctat ggcaatgagg gcttgggggtg ggcaggatgg 2100
ctcctgtcac cccgtggctc tcggcctagt tggggcccca cggacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggatc cgataaccct acgtgcggct tcgccgatct catggggtac 2220
attccgctcg tcggcgcccc cctagggggc gctgccaggg ccctggcgca tggcgtccgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc cttttctatc 2340
ttccttttgg ctttggctgtc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcgacatga tcatgatac cccgggtgct gtgccctgct tcggggagaa caactcctcc 2520
cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga tttgctcgtt ggggcggctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttctc gtcgccagc tgttcacct ctcgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggtcac 2760

```

cgtatggctt	gggatatgat	gatgaactgg	tcacctacag	cagccctagt	ggtatcgcag	2820
ttactccgga	tcccacaagc	tgtcgtggat	atgggtggcg	gggccattg	gggagtccta	2880
gcgggccttg	cctactattc	catgggtggg	aactgggcta	aggttctgat	tgtgatgcta	2940
ctctttgccg	gcgttgacgg	gggaacctat	gtgacagggg	ggacgatggc	caaaaacacc	3000
ctcgggatta	cgccctctt	ttcaccggg	tcateccaga	aatccagct	tgtaaacacc	3060
aacggcagct	ggcacatcaa	caggactgcc	ctgaactgca	atgactccct	caacactggg	3120
ttccttgctg	cgctgttcta	cgtgcacaag	ttcaactcat	ctggatgccc	agagcgcag	3180
gccagctgca	gccccatcga	cgcgttcgct	caggggtggg	ggcccatcac	ttacaatgag	3240
tcacacagct	cggaccagag	gccttattgt	tggcactacg	caccccgccc	gtgcggtatc	3300
gtaccgcggg	cgcaggtgtg	tggctccagt	tactgcttca	ccccaaagcc	tgtcgtgggtg	3360
gggacgaccg	accggttcgg	cgtccctacg	tacagttggg	gggagaatga	gacggacgtg	3420
ctgcttctta	acaacacgcg	gccgccgcaa	ggcaactggg	ttggctgtac	atggatgaat	3480
agcactgggt	tcaccaagac	gtgcgggggc	cccccggtga	acatcggggg	gatcggcaat	3540
aaaaccttga	cctgccccac	ggactgcttc	cggaaagcacc	ccgaggccac	ttacaccaag	3600
tgtggttcgg	ggccttggtt	gacacccaga	tgcttggtcc	actaccata	caggctttgg	3660
cactaccctt	gcactgtcaa	ctttaccatc	ttcaaggtta	ggatgtacgt	ggggggagtg	3720
gagcacaggc	tcgaagccgc	atgcaattgg	actcgaggag	agcgttgtaa	cctggaggac	3780
agggacagat	cagagcttag	cccgtgctg	ctgtctacaa	cggagtggca	ggtattgcc	3840
tgcttcttca	ccaccctacc	ggctctgtcc	actggtttga	tccatctcca	tcagaacgtc	3900
gtggacgtac	aatacctgta	cggtataggg	tcggcggttg	tctcctttgc	aatcaaattg	3960
gagtatgtcc	tgttgcctct	ccttcttctg	gcggacgcgc	gcgtctgtgc	ctgcttgtgg	4020
atgatgctgc	tgatagctca	agctgaggcc	gccctagaga	acctggtggt	cctcaacgcg	4080
gcatccgtgg	ccggggcgca	tggcattctc	tccttctctg	tgcttctctg	tgctgcctgg	4140
tacatcaagg	gcaggctggt	cctgggggcg	gcataatgcc	tctacggcgt	atggccgcta	4200
ctcctgctcc	tgctggcggt	accaccacga	gcatacggca	tggaccggga	gatggcagca	4260
tcgtgcggag	gcgcgggttt	cgtaggtctg	atactcttga	ccttgtcacc	gcactataag	4320
ctgttcctcg	ctaggctcat	atgggtggtta	caatatttta	tcaccagggc	cgaggcacac	4380
ttgcaagtgt	ggatcccccc	cctcaacggt	cggggggggc	gcgatgccgt	catcctctc	4440
acgtgcgcga	tcacccaga	gctaattctt	accatcacca	aaatcttgct	cgccatactc	4500
ggteccactca	tgggtgctcca	ggctggtata	accaaagtgc	cgtacttctg	gcgcgcacac	4560
gggctcattc	gtgcatgcat	gctggtgcgg	aagggtgctg	ggggtcatta	tgtccaaatg	4620
gctctcatga	agttggccgc	actgacaggt	acgtacgttt	atgaccatct	cacccctctg	4680
cgggactggg	cccacgcggg	cctacgagac	cttgcgggtg	cagttgagcc	cgctcgtctc	4740
tctgatatgg	agaccaaggt	tatcacctgg	ggggcagaca	ccgcggcgtg	tggggacatc	4800
atcttggggc	tgcccgtctc	cgcccgcagg	gggagggaga	tacatctggg	accggcagac	4860
agccttgaag	ggcaggggtg	gcgactcctc	gcgcctatta	cggcctactc	ccaacagacg	4920
cgaggcctac	ttggctgcat	catcactagc	ctcacaggcc	gggacaggaa	ccaggctcgag	4980
ggggaggtcc	aagtggctctc	caccgcaaca	caatctttcc	tggcgacctg	cgtcaatggc	5040
gtgtgttggg	ctgtctatca	tggtgccggc	tcaaagacc	ttgcccggcc	aaagggccca	5100
atcacccaaa	tgtacaccaa	tgtggaccag	gacctcgtcg	gctggcaagc	gcccccggg	5160
gcgcgttctt	tgacaccatg	cacctgcggc	agctcggacc	tttacttggg	cacgaggcat	5220
gccgatgtca	ttccggtgcg	ccggcggggc	gacagcaggg	ggagcctact	ctccccaggg	5280
cccgtctcct	acttgaaggg	ctcttcgggg	ggtccactgc	tctgcccctc	ggggcacgct	5340
gtgggcatct	ttcgggctgc	cgtgtgcacc	cgaggggttg	cgaaggcggg	ggactttgta	5400
cccgtcgagt	ctatggaaac	cactatgcgg	tccccggtct	tcacggacaa	ctcgtcccct	5460
ccggccgtac	cgcagacatt	ccagggtggc	catctacacg	cccctactgg	tagcggcaag	5520
agcactaagg	tgccggctgc	gtatgcagcc	caaggggtata	agggtgcttg	cctgaaccgg	5580
tccgtcgccg	ccaccctagg	tttcggggcg	tatatgtcta	aggcacatgg	tatcgaccct	5640
aacatcagaa	ccggggtaag	gaccatcacc	acgggtgccc	ccatcacgta	ctccacctat	5700
ggcaagtttc	ttgcccagcg	tggttgctct	gggggcgcct	atgacatcat	aatatgtgat	5760
gagtgccact	caactgactc	gaccactatc	ctgggcatcg	gcacagtcc	ggaccaagcg	5820
gagacggctg	gagcgcgact	cgtcgtgctc	gccaccgcta	cgctccggg	atcggtcacc	5880
gtgccacatc	caaacatcga	ggaggtggct	ctgtccagca	ctggagaaat	ccccttttat	5940
ggcaaagcca	tccccatcga	gaccatcaag	ggggggaggc	acctcatttt	ctgccattcc	6000
aagaagaaat	gtgatgagct	cgccgcgaag	ctgtccggcc	tcggactcaa	tgtgttagca	6060
tattaccggg	gccttgatgt	atccgtcata	ccaactagcg	gagacgtcat	tgtcgtagca	6120
acggacgctc	taatgacggg	ctttaccggc	gatttcgact	cagtgatcga	ctgcaatata	6180
tgtgtcacc	agacagtcca	cttcagcctg	gacccgacct	tcaccattga	gacgacgacc	6240
gtgccacaag	acgcgggtgc	acgctcgcag	cggcagggca	ggactggtag	gggcaggatg	6300
ggcatttaca	ggtttgtgac	tccaggagaa	cggccctcgg	gcatgttcga	ttcctcggtt	6360
ctgtgcgagt	gctatgacgc	gggctgtgct	tggtagcagc	tcacgcccgc	cgagacctca	6420
gtaggttgc	gggcttacct	aaacacacca	gggttgcccg	tctgccagga	ccatctggag	6480
ttctgggaga	gcgtctttac	aggcctcacc	cacatagacg	cccatttctt	gtcccagact	6540
aagcaggcag	gagacaactt	cccctacctg	gtagcatacc	aggctacggg	gtgcgccagg	6600
gctcaggctc	cacctccatc	gtgggaccaa	atgtggaagt	gtctcatacg	gctaaagcct	6660
acgctgcacg	ggccaacgcc	cctgctgtat	aggctgggag	ccgttcaaaa	cgaggttact	6720
accacacacc	ccataaccaa	atacatcatg	gcattgcatg	cggctgacct	ggaggctcgtc	6780

acgagcacct	gggtgctggt	aggcggagtc	ctagcagctc	tggccgcgta	ttgcctgaca	6840
acaggcagcg	tggtcattgt	gggcaggatc	atcttgtccg	gaaagccggc	catcattccc	6900
gacagggaa	tcctttaccg	ggagttcgat	gagatggaag	agtgcgcctc	acacctccct	6960
tacatcgaac	agggaaatgca	gctcgcgcaa	caattcaaac	agaaggcaat	cgggttgctg	7020
caaacagcca	ccaagcaagc	ggaggctgct	gctcccgtgg	tggaatccaa	gtggcggacc	7080
ctcgaagcct	tctggggcgaa	gcataatgtg	aatttcatca	gccggataca	atatttagca	7140
ggcttgtcca	ctctgcctgg	caaccccgcg	atagcatcac	tgatggcatt	cacagcctct	7200
atcaccagcc	cgctcaccac	ccaacatacc	ctcctgttta	acatcctggg	gggatgggtg	7260
gccgcccaac	ttgctcctcc	cagcgtgctg	tctgctttcg	taggcgccgg	catcgctgga	7320
gcggctggtg	gcagcatagg	ccttgggaag	gtgcttgtgg	atattttggc	aggttatgga	7380
gcaggggtgg	caggcgcgct	cgtggccttt	aaggatcatga	gccggcgagat	gccctccacc	7440
gaggacctgg	ttaacctact	ccctgctatc	ctctcccctg	gcgccctagt	cgtcggggtc	7500
gtgtgcgag	cgatactgag	tccggcacgtg	ggcccagggg	agggggctgt	gcagtggatg	7560
aaccggctga	tagcgttcgc	ttcgcggggg	aaccacgtct	ccccacgca	ctatgtgctt	7620
gagagcgag	ctgcagcag	tgtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
ctgaagaggg	ttcaccagtg	gatcaacgag	gactgctcca	cgccatgctc	cggctcgtgg	7740
ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
tccaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
catggaacat	tccccattaa	cgcgtacacc	acgggcccct	gcacgccctc	cccggcgcca	8040
aattattcta	gggcgctgtg	gcccgtggct	gctgaggagt	acgtggaggt	tacgcgggtg	8100
ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaagtgcc	gtgtcagggt	8160
ccggcccccg	aattcttcac	agaagtggat	ggggtgcggt	tgcacaggta	cgctccagcg	8220
tgcaaaccce	tctacggga	ggaggtcaca	ttcctggtcg	ggctcaatca	atacctgggt	8280
gggtcacagc	tcccatgcga	gcccgaaacc	gacgtagcag	tgctcacttc	catgctcacc	8340
gaccctccc	acattacggc	ggagacggct	aagcgtaggc	tggccagggg	atctcccccc	8400
tccttggcca	gctcatcagc	tagccagctg	tctgcccctt	ccttgaaggc	aacatgcact	8460
accgctcatg	actccccgga	cgctgacctc	atcgaggcca	acctcctgtg	gcccagggag	8520
atgggcggga	acatcaccgg	cgtggagtca	gaaaataagg	tagtaatttt	ggactctttc	8580
gagccgctcc	aagcggagga	ggatgagagg	gaagtatccg	ttccggcgga	gatcctgcgg	8640
aggtccagga	aattccctcg	agcgtatgcc	atatgggcac	gcccggatta	caaccctcca	8700
ctgttagagt	cctggaagga	cccggactac	gtccctccag	tggtagacgg	gtgtccattg	8760
ccgctgccca	aggcccctcc	gataccacct	ccacggagga	agaggacggg	tgctcctgtca	8820
gaatctaccg	tgtcttctgc	cttggcggag	ctcgccacaa	agaccttcgg	cagctccgaa	8880
tcgtcggccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
gacgcgggat	ccgacgttga	gtcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
gatcccgatc	tcagcagcgg	gtcttggctt	accgtaagcg	aggaggctag	tgaggacgct	9060
gtctgctgct	cgatgtccta	cacatggaca	ggcgcccctga	tcacgccatg	cgctgcggag	9120
gaaaccaagc	tgcccataca	tgcactgagc	aactctttgc	tccgtcacca	caacttggct	9180
tatgctacaa	catctcgcag	cgcaagcctg	cggcagaaga	aggtcacctt	tgacagactg	9240
caggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaagc	tgacgcccc	acattcggcc	9360
agatctaata	ttggctatgg	ggcaaaggac	gtccggaaacc	tatccagcaa	ggccgttaac	9420
cacatccgct	ccgtgtggaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccacc	9480
atcatggcaa	aaaatgaggt	tttctgcgct	caaccagaga	aggggggccc	caagccagct	9540
cgcttatcgc	tattcccaga	tttggggggt	cgtgtgtgcg	agaaaatggc	cctttacgat	9600
gtggtctcca	ccctccctca	ggccgtgatg	ggctcttcat	acggattcca	atactctcct	9660
ggacagcggg	tcgagttcct	ggtgaatgcc	tggaaagcga	agaaatgcc	tatgggcttc	9720
gcataatgaca	cccgtgtttt	tgactcaacg	gtcactgaga	atgacatccg	tggtgaggag	9780
tcaatctacc	aatgttgtga	cttggccccc	gaagccagac	aggccataag	gtcgtcaca	9840
gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
cggtgccgcg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttacttg	9960
aaggccgctg	cggcctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgtatgcgga	10020
gacgaccttg	tcgttatctg	tgaagcgcg	gggacccaag	aggacgaggc	gagcctacgg	10080
gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccgcc	caaaccagaa	10140
tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcatct	10200
ggcaaaaggg	tgtactatct	caccctgac	cccaccacc	cccttgcgcg	ggctgcgtgg	10260
gagacagcta	gacacactcc	agtcaattcc	tggctaggca	acatcatcat	gtatgcgcc	10320
accttgtggg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
gacctacctc	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagttac	10500
tctccagggtg	agatcaatag	ggtggcttca	tgcctcagga	aacttggggg	accgcccttg	10560
cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtcca	gggggggagg	10620
gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
ccaatcccgg	ctgcgtccca	gttggattta	tccagctggt	tcgttgctgg	ttacagcggg	10740
ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgtc	ggttcatgtg	gtgcctactc	10800

ctactttctg	taggggtagg	catctatcta	ctccccaacc	gatgaacggg	gagctaaaca	10860
ctccaggcca	ataggccatc	ctgttttttt	ccotTTTTTT	ttttcttttt	tttttttttt	10920
tttttttttt	tttttttttc	tcctTTTTTT	ttcctctttt	tttctttttc	tttcttttgg	10980
tggtccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
agagagtgct	gatactggcc	tctctgcaga	tcaagt			11076

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

<400> 2

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtggtgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgtg gaccgtgcac catgagcacg aatcctaac 360
ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480
ctgatgccgc cgtgttcggy ctgtcagcgc agggggcgccc ggttcttttt gtcaagaccg 540
acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600
cgacgggctg tccttgccga gctgtgctcg acgttgtcac tgaagcggga agggactggc 660
tgctattggg cgaagtgccg gggcaggatc tcctgtcac tcccttgcct cctgccgaga 720
aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780
cattcgacca ccaagcgaaa catcgcacgc agcagcacg tactcggatg gaagccggtc 840
ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900
ccaggctcaa ggcgcgcgat cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 960
gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg atccatcgac tgtggccggc 1020
tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080
ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgcgct cccgattcgc 1140
agcgcacgc cttctatcgc cttcttgacg agttcttctg agtttaaca gaccacaacg 1200
gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260
cgaagccgct tgggaataagg ccggtgtgcy tttgtctata tgttattttc caccatattg 1320
ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct 1380
aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440
gttccctctg aagcttcttg aagacaaaca acgtctgtag cgacccttg caggcagcgg 1500
aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560
gcaaaggcgg cacaaccca gtgccacggt gtgagttgga tagttgtgga aagagtcaa 1620
tggctctcct caagcgtatt caacaagggg ctgaaggatg ccagaaggc acccattgt 1680
atgggatctg atctggggcc tcgggtgcaca tgctttacat gtgttttagt gaggttaaaa 1740
aacgtctagg cccccgaac cacggggagc tggttttcct ttgaaaaaca cgataatacc 1800
atggaccggg agatggcagc atcgtgcgga ggcgcggttt tcgtaggtct gatactcttg 1860
accttgtcac cgcactataa gctgttctc gctaggctca tatggtggtt acaatatttt 1920
atcaccaggg ccgaggcaca cttgcaagtg tggatcccc ccctcaacgt tcgggggggc 1980
cgcgatgccg tcatcctcct cacgtgcgcy atccaccag agctaactct taccatcacc 2040
aaaatcttgc tcgccatact cggtcactc atggtgctcc aggttggtat aaccaaagtg 2100
ccgtacttcg tgcgcgcaca cgggtcatt cgtgcatgca tgctgggtgc gaaggttgct 2160
gggggtcatt atgtccaaat ggctctcatg aagttggcgg cactgacagg tacgtacgtt 2220
tatgaccatc tcacccact gcgggactgg gccacgcgg gcctacgaga ccttgccgtg 2280
gcagttgagc ccgtcgtctt ctctgatatg gagaccaagg ttatcacctg gggggcagac 2340
accgcccgt gtggggacat catcttgggc ctgcccgtct ccgcccgcag ggggaggag 2400
atacatctgg gaccggcaga cagccttga gggcaggggt ggcgactcct cgcgcctatt 2460
acggcctact cccaacagac gcgaggccta cttgggtgca tcatcactag cctcacaggc 2520
cgggacagga accaggtcga gggggaggtc caagtgtct ccaccgcaac acaatctttc 2580
ctggcgacct gcgtcaatgg cgtgtgttgg actgtctatc atggtgcccg ctcaaagacc 2640
cttgccggcc caaagggcc aatcacccaa atgtacacca atgtggacca ggacctcgtc 2700
ggctggcaag cgcccccgg ggcgcgttcc ttgacaccat gcacctgcgg cagctcggac 2760
ctttacttgg tcacgaggca tgccgatgtc attccgggtc gccggcgggg cgacagcagg 2820
gggagcctac tctccccag gcccgctctc tacttgaagg gctcttcggg cgggtccactg 2880
ctctgccct cggggcacgc tgtgggcatc tttcgggctg ccgtgtgcac ccgaggggtt 2940
gcgaaggcgg tggactttgt acccgtcgag tctatggaaa cactatgcy gtecccggtc 3000
ttcacggaca actcgtcccc tccggccgta ccgcagacat tccagggtgg ccatctacac 3060
gccctactg gtagecggca gagcactaag gtgccggctg cgtatgcagc ccaagggtat 3120
aaggtgcttg tcctgaacc gtccgtcgcc gccaccctag gtttcggggc gtatatgtct 3180
aaggcacatg gtatcgacc taacatcaga accggggtaa ggaccatcac cacgggtgcc 3240
cccatcacgt actccaccta tggcaagttt cttgcgcagc gtggttgctc tgggggcgcc 3300
tatgacatca taatatgtga tgagtgccac tcaactgact cgaccactat cctgggcate 3360
ggcacagtcc tggaccaagc ggagacggct ggagcgcgac tcgtcgtgct cgcaccgct 3420
acgcctccgg gatcgggtcac cgtgccacat ccaaacatcg aggaggtggc tctgtccagc 3480
actggagaaa tcccctttta tggcaagcc atccccatcg agaccatcaa gggggggagg 3540
cacctcattt tctgcoattc caagaagaaa tgtgatgagc tcgcccgcaa gctgtccggc 3600

```

ctcggactca	atgctgtagc	atattaccgg	ggccttgatg	tatccgtcat	accaactagc	3660
ggagacgtca	ttgtcgtagc	aacggacgct	ctaatagacg	gctttaccgg	cgatttcgac	3720
tcagtgatcg	actgcaatac	atgtgtcacc	cagacagtcg	acttcagcct	ggacccgacc	3780
ttcaccattg	agacgacgac	cgtgccacaa	gacgcggtgt	cacgctcgca	gcggcgaggc	3840
aggactggta	ggggcaggat	gggcatttac	aggtttgtga	ctccaggaga	acggccctcg	3900
ggcatgttcg	attcctcggt	tctgtgcgag	tgctatgacg	cgggctgtgc	ttggtacgag	3960
ctcacgcccg	ccgagacctc	agttagggtg	cgggcttacc	taaacacacc	agggttgccc	4020
gtctgccagg	accatctgga	gttctgggag	agcgtcttta	caggcctcac	ccacatagac	4080
gcccatttct	tgtcccagac	taagcaggca	ggagacaact	tcccctacct	ggtagcatac	4140
caggctacgg	tgtgcgcccag	ggctcaggct	ccacctccat	cgtgggacca	aatgtggaag	4200
tgtctcatac	ggctaaagcc	tacgctgcac	gggccaacgc	ccctgctgta	taggctggga	4260
gccgttcaaa	acgagggtac	taccacacac	cccataacca	aatacatcat	ggcatgcatg	4320
tcggctgacc	tggaggtcgt	cacgagcacc	tgggtgctgg	taggcggagt	cctagcagct	4380
ctggcccgct	attgcctgac	aacaggcagc	gtggtcattg	tgggcaggat	catcttgtcc	4440
ggaaagccgg	ccatcattcc	cgacagggaa	gtcctttacc	gggagttcga	tgagatggaa	4500
gagtgcgcct	cacacctccc	ttacatcgaa	cagggaatgc	agctcgccga	acaattcaaa	4560
cagaaggcaa	tcgggttgct	gcaaacagcc	accaagcaag	cggaggctgc	tgctcccgtg	4620
gtggaatcca	agtggcggac	cctcgaagcc	ttctgggcga	agcatatgtg	gaatttcac	4680
agcgggatac	aatatthtagc	aggcttgtcc	actctgcctg	gcaaccccgc	gatagcatca	4740
ctgatggcat	tcacagcctc	tatcaccagc	ccgctcacca	cccaacatac	cctcctgttt	4800
aacatcctgg	ggggatgggt	ggccgcccac	cttgcctcct	ccagcgtgc	ttctgctttc	4860
gtaggcgccc	gcatcgctgg	agcggctggt	ggcagcatag	gccttgggaa	ggtgcttgtg	4920
gatattttgg	caggttatgg	agcaggggtg	gcaggcgcgc	tcgtggcctt	taaggctcatg	4980
agcggcgaga	tgccctccac	cgaggacctg	gttaacctac	tcctgctat	cctctcccct	5040
ggcgccctag	tcgtcggggt	cgtgtgcgca	gcgatactgc	gtcggcacgt	gggcccaggg	5100
gagggggctg	tgcaagtggat	gaaccggctg	atagcgttcg	cttcgcgggg	taaccacgtc	5160
tccccacgc	actatgtgcc	tgagagcgcac	gctgcagcac	gtgtcactca	gatcctctct	5220
agtcttacca	tcactcagct	gctgaagagg	cttcaccagt	ggatcaacga	ggactgctcc	5280
acgccatgct	ccggctcgtg	gctaagagat	gtttgggatt	ggatatgcac	ggtgttgact	5340
gatttcaaga	cctggctcca	gtccaagctc	ctgcgcgat	tgccgggagt	ccccttcttc	5400
tcatgtcaac	gtgggtacaa	gggagtctgg	cggggcgacg	gcatcatgca	aaccacctgc	5460
ccatgtggag	cacagatcac	cggacatgtg	aaaaacgggt	ccatgaggat	cgtggggcct	5520
aggacctgta	gtaaacagtg	gcatggaaca	ttccccatta	acgcgtacac	cacgggcccc	5580
tgcacgccc	ccccggcgcc	aaattattct	agggcgctgt	ggcgggtggc	tgctgaggag	5640
tacgtggagg	ttacgcgggt	gggggatttc	cactacgtga	cgggcatgac	cactgacaac	5700
gtaaagtgcc	cgtgtcaggt	tccggcccc	gaattcttca	cagaagtgga	tgggggtgcg	5760
ttgcacaggt	acgctccagc	gtgcaaacc	ctcctacggg	aggaggtcac	attcctggtc	5820
gggctcaate	aatacctggt	tgggtcacag	ctcccattgc	agcccgaacc	ggacgtagca	5880
gtgctcactt	ccatgctcac	cgacccctcc	cacattacgg	cggagacggc	taagcgtagg	5940
ctggccaggg	gatctcccc	ctccttggcc	agctcatcag	ctagccagct	gtctgcgcct	6000
tccttgaagg	caacatgcac	taccgctcat	gactccccgg	acgctgacct	catcgaggcc	6060
aacctcctgt	ggcggcagga	gatgggcggg	aacatcacc	gcgtggagtc	agaaaataag	6120
gtagtaattt	tggactcttt	cgagccgctc	caagcggagg	aggatgagag	ggaagtatcc	6180
gttccggcgg	agatcctgcg	gaggtccagg	aaattccctc	gagcgtatgc	catatgggca	6240
cgcccgatt	acaacctcc	actgttagag	tcctggaagg	accgggacta	cgccctcca	6300
gtggtacacg	ggtgtccatt	gccgcctgcc	aaggccctc	cgataccacc	tccacggagg	6360
aagaggacgg	ttgtcctgtc	agaatctacc	gtgtctctct	ccttggcgga	gctcgccaca	6420
aagaccttcg	gcagctccga	atcgtcggcc	gtcgacagcg	gcacggcaac	ggcctctcct	6480
gaccagccct	ccgacgacgg	cgacgcggga	tcgacgctt	agtcgtactc	ctccatgccc	6540
ccccttgagg	gggagccggg	ggatcccgat	ctcagcgacg	ggtcttggtc	taccgtaagc	6600
gaggaggcta	gtgaggacgt	cgtctgctgc	tcgatgtcct	acacatggac	aggcgccctg	6660
atcacgccat	gcgctgcgga	ggaaaccaag	ctgcccatac	atgcaactgag	caactctttg	6720
ctccgtcacc	acaacttgg	ctatgctaca	acatctcgca	gcgcaagcct	gcggcagaag	6780
aaggtcacct	ttgacagact	gcaggtcctg	gacgaccact	accgggacgt	gctcaaggag	6840
atgaaggcga	aggcgtccac	agttaaggct	aaacttctat	ccgtggagga	agcctgtaag	6900
ctgacgcccc	cacattcggc	cagatctaaa	tttggtatg	gggcaaagga	cgcccggaac	6960
ctatccagca	aggccgttaa	ccacatccgc	tccgtgtgga	aggacttgc	ggaagacact	7020
gagacaccaa	ttgacaccac	catcatggca	aaaaatgagg	ttttctgcgt	ccaaccagag	7080
aaggggggccc	gcaagccagc	tcgccttacc	gtattcccag	atthgggggt	tcgtgtgtgc	7140
gagaaaatgg	ccctttacga	tgtggtctcc	accctccctc	aggccgtgat	gggctcttca	7200
tacggattcc	aatactctcc	tggacagcgg	gtcgagttcc	tgggtgaatgc	ctggaaagcg	7260
aagaaatgcc	ctatgggctt	cgcatatgac	accgctggt	ttgactcaac	ggtcactgag	7320
aatgacatcc	gtggtgagga	gtcaatctac	caatgttgtg	acttggcccc	cgaagccaga	7380
caggccataa	ggtcgctcac	agagcggctt	tacatcgggg	gccccctgac	taattctaaa	7440
gggcagaact	gcggctatcg	ccggtgccgc	gcgagcgggt	tactgacgac	cagctgcggt	7500
aataccctca	catgttactt	gaaggccgct	gcggcctgtc	gagctgcgaa	gctccaggac	7560
tgcacgatgc	tcgtatgcgg	agacgacctt	gtcgttatct	gtgaaagcgc	ggggacccaa	7620

gaggacgagg	cgagcctacg	ggccttcacg	gaggctatga	ctagatactc	tgccccccct	7680
ggggacccgc	ccaaaccaga	atacgacttg	gagttgataa	catcatgctc	ctccaatgtg	7740
tcagtcgcg	acgatgcac	tggcaaaagg	gtgtactatc	tcaccctgga	ccccaccacc	7800
ccccttgcg	gggctgctg	ggagacagct	agacacactc	cagtcaattc	ctggctaggc	7860
aacatcatca	tgtatgcgcc	caccttgtgg	gcaaggatga	tcctgatgac	tcatttcttc	7920
tccatccttc	tagctcagga	acaacttgaa	aaagccctag	attgtcagat	ctacggggcc	7980
tgttactcca	ttgagccact	tgacctacct	cagatcattc	aacgactcca	tggccttagc	8040
gcattttcac	tccatagtta	ctctccaggt	gagatcaata	gggtggcttc	atgcctcagg	8100
aaacttgggg	taccgccctt	gcgagtctgg	agacatcggg	ccagaagtgt	ccgcgctagg	8160
ctactgtccc	agggggggag	ggctgccact	tgtggcaagt	acctcttcaa	ctgggcagta	8220
aggaccaagc	tcaaactcac	tccaatcccg	gctgcgtccc	agttggattt	atccagctgg	8280
ttcgttgctg	gttacagcgg	gggagacata	tatcacagcc	tgtctcgtgc	ccgaccccg	8340
tggttcatgt	ggtgcctact	cctactttct	gtaggggtag	gcatctatct	actccccaac	8400
cgatgaacgg	ggagctaaac	actccaggcc	aataggccat	cctgtttttt	tccctttttt	8460
ttttctttt	ttttttttt	ttttttttt	ttttttttt	ctcctttttt	tttctcttt	8520
ttttccttt	ctttcctttg	gtggctccat	cttagcccta	gtcacggcta	gctgtgaaag	8580
gtccgtgagc	cgcttgactg	cagagagtgc	tgatactggc	ctctctgcag	atcaagt	8637

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

<400> 3

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctct cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggtat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggccaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggetg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgccag 900
ccgaactgtt cggcaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgatc gcagcgcac gccttctatc gccttcttga cgagttcttc tgagtttaa 1200
cagaccacia cggtttccct ctagcgggat caattccgcc cctctcctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctaggggtct tccccctctc gccaaaggaa tgcaaggtct gttgaatgct 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcagccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgtg ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta gggccccgga accacgggga cgtggttttc ctttgaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
tgcacatca ctagcctcac aggccgggac aggaaccagg tcgaggggga ggtccaagtg 1920
gtctccaccg caacacaatc tttcctggcg acctgcgtca atggcgtgtg ttggactgtc 1980
tatcatggtg ccggtcaaaa gacccttgcc ggcccaaagg gcccaatcac ccaaatgtac 2040
accaatgtgg accaggacct cgtcggctgg caagcgcgcc ccggggcgcg ttccttgaca 2100
ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160
gtgcgcgggc ggggcgacag cagggggagc ctactctccc ccaggcccgt ctctacttg 2220
aagggctctt cgggcggctc actgctctgc cctcggggc acgctgtggg catctttcgg 2280
gctgccgtgt gcaccgaggg ggttgcgaag gcggtggact ttgtaccctg cgagtctatg 2340
gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt ccctccggc cgtaccgcag 2400
acattccagg tggcccatct acacgccct actggtagcg gcaagagcac taaggtgccg 2460
gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgtccgt cgccgccacc 2520
ctaggtttcg gggcgtatat gtctaaggca catggtatcg accctaacad cagaaccggg 2580
gtaaggacca tcaccacggg tgccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacgggtggtt gctctggggg cgcctatgac atcataatat gtgatgagtg ccaactcaact 2700
gactcgacca ctatcctggg catcggcaca gtcctggacc aagcggagac ggctggagcg 2760
cgactcgtcg tgctcgccac cgctacgct ccgggatcgg tcaccgtgcc acatccaaac 2820
atcgaggagg tggctctgtc cagcaactgga gaaatcccc tttatggcaa agccatcccc 2880
atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
gagctcgccg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaatg 3060
acgggcttta ccggcgattt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
gtcgacttca gcctggacc gacctcacc attgagacga cgaccgtgcc acaagacgcg 3180
gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgtat 3300
gacgcgggct gtgcttggtg cgagctcacg cccgccgaga cctcagttag gttgcgggct 3360
tacctaaaca caccagggtt gcccgctctgc caggaccatc tggagttctg ggagagcgtc 3420
tttacaggcc tcaccacat agacgccat ttcttgtccc agactaagca ggcaggagac 3480
aacttccct acctggtagc ataccaggct acggtgtgcg ccagggtca ggctccacct 3540

```

ccatcgtggg	accaaagtgt	gaagtgtctc	atacggctaa	agcctacgct	gcacggggcca	3600
acgcccctgc	tgtataggct	gggagccggt	caaaacgagg	ttactaccac	acaccccata	3660
accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
attgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccacag	ggaagtcctt	3840
taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcccttacat	cgaacaggga	3900
atgcagctcg	ccgaacaatt	caaacagaag	gcaatcgggt	tgctgcaaac	agccaccaag	3960
caagcggagg	ctgctgctcc	cgtggtggaa	tccaagtggc	ggacctcga	agccttctgg	4020
gcgaagcata	tgtggaattt	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgcctc	4140
accaccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
cctcccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tgttggcagc	4260
ataggccttg	ggaaggtgct	tgtggatatt	ttggcagggt	atggagcagg	ggtggcaggc	4320
gcgctcgtgg	cctttaaggt	catgagcggc	gagatgccct	ccaccgagga	cctggttaac	4380
ctactccctg	ctatcctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgcagcgata	4440
ctgctcggc	acgtgggccc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
ttcgtctcgc	gggtaacca	cgtctccccc	acgcactatg	tgctgagag	cgacgctgca	4560
gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggcttcac	4620
cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggctaag	agatgtttgg	4680
gattggatat	gcacggtggt	gactgatttc	aagacctggc	tccagtccaa	gctcctgccg	4740
cgattgccgg	gagtcccctt	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
gacggcatca	tgcaaaccac	ctgcccattg	ggagcacaga	tcaccggaca	tgtgaaaaac	4860
ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
attaacgcgt	acaccacggg	cccctgcacg	ccctccccgg	cgccaaatta	ttctagggcg	4980
ctgtggcggg	tggctgctga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
gtgacgggca	tgaccactga	caacgtaaag	tgcccgtgtc	aggttccggc	ccccgaattc	5100
ttcacagaag	tggatggggg	gcggttgca	aggtacgctc	cagcgtgcaa	accctccta	5160
cgggaggagg	tcacattcct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctccca	5220
tgcgagcccg	aaccggacgt	agcagtgtct	acttccatgc	tcaccgacc	ctcccacatt	5280
acggcggaga	cggctaagcg	taggctggcc	aggggatctc	ccccctcctt	ggccagctca	5340
tcagctagcc	agctgtctgc	gccttccttg	aaggcaacat	gcactacccg	tcattgactcc	5400
ccggacgctg	acctcatcga	ggccaacctc	ctgtggcggc	aggagatggg	cgggaacatc	5460
acccgcgtgg	agtcagaaaa	taaggtagta	atthttggact	ctttcgagcc	gctccaagcg	5520
gaggaggatg	agagggaagt	atccgttccg	gcggagatcc	tgccggaggtc	caggaaattc	5580
cctcgagcga	tgcccatatg	ggcacgcccg	gattacaacc	ctccactggt	agagtcctgg	5640
aaggacccgg	actacgtccc	tccagtggta	cacgggtgtc	cattgcccgc	tgccaaggcc	5700
cctccgatac	cacctccacg	gaggaagagg	acggttgtcc	tgtcagaatc	taccgtgtct	5760
tctgccttgg	cggagctcgc	cacaaagacc	ttcggcagct	ccgaatcgtc	ggccgctcgac	5820
agcggcacgg	caacggcctc	tctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
gttgagtcgt	actcctccat	gccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
gacgggtctt	ggtctaccgt	aagcgaggag	gctagtgagg	acgtcgtctg	ctgctcgatg	6000
tcctacacat	ggacaggcgc	cctgatcacg	ccatgcgctg	cggaggaaac	caagctgccc	6060
atcaatgcac	tgagcaactc	tttgctccgt	caccacaact	tggtctatgc	tacaacatct	6120
cgacgcgcaa	gcctgcggca	gaagaaggtc	acctttgaca	gactgcagg	cctggacgac	6180
cactaccggg	acgtgctcaa	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
ctatccgtgg	aggaagcctg	taagctgacg	ccccacatt	cggccagatc	taaatttggc	6300
tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgctccgtg	6360
tgggaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
gaggttttct	gcgtccaacc	agagaagggg	ggccgcaagc	cagctcgctt	tatcgtatct	6480
ccagatttgg	gggttcgtgt	gtgagagaaa	atggcccttt	acgatgtgg	ctccaccctc	6540
cctcaggccg	tgatgggctc	ttcatacgg	ttccaatact	ctcctggaca	gcgggtcgag	6600
ttcctggtga	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacaccgcg	6660
tgttttgact	caacggtcac	tgagaatgac	atccgtggtg	aggagtcaat	ctaccaatgt	6720
tgtgacttgg	ccccgaagc	cagacaggcc	ataaggctgc	tcacagagcg	gctttacatc	6780
gggggcccc	tgactaatc	taaagggcag	aactgcggct	atcgccgggtg	ccgcgcgagc	6840
ggtgtactga	cgaccagctg	cggtaatacc	ctcacatggt	acttgaaggc	cgctgcggcc	6900
tgctgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
atctgtgaaa	gcgcggggac	ccaagaggac	gagggcagcc	tacgggcctt	cacggaggct	7020
atgactagat	actctgcccc	ccctggggac	ccgccaaac	cagaatacga	cttgaggttg	7080
ataacatcat	gctcctccaa	tgtgtcagtc	gcgcacgatg	catctggcaa	aaggggtgtac	7140
tatctcacc	gtgacccac	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
actccagtca	attcctggct	aggcaacatc	atcatgtatg	cgcccacctt	gtgggcaagg	7260
atgatcctga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
ctagattgtc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
attcaacgac	tccatggcct	tagcgcattt	tactccata	gttactctcc	aggtgagatc	7440
aataggggtg	cttcatgcct	caggaaactt	ggggtaccgc	ccttgcgagt	ctggagacat	7500
cgggccagaa	gtgtccgcgc	taggctactg	tcccaggggg	ggagggctgc	cacttgtggc	7560

aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggctgcg	7620
tcccagttgg	atztatccag	ctggttcgtt	gctggttaca	gcgggggaga	catatatcac	7680
agcctgtctc	gtgcccgacc	ccgctggttc	atgtgggtgcc	tactcctact	ttctgtaggg	7740
gtaggcacct	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
ccatcctggt	tttttcctt	tttttttttc	tttttttttt	tttttttttt	tttttttttt	7860
ttttctcctt	ttttttcct	ctttttttcc	ttttctttcc	tttgggtggct	ccatcttagc	7920
cctagtcacg	gctagctgtg	aaaggtccgt	gagccgcttg	actgcagaga	gtgctgatac	7980
tggcctctct	gcagatcaag	t				8001

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

<400> 4

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgtc gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480
ctgatgccgc cgtgttcggg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540
acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600
cgacgggctg tccttgccga gctgtgctcg acgttgtcac tgaagcggga agggactggc 660
tgctattggg cgaagtgcg gggcaggatc tcctgtcacc tcacctgct cctgccgaga 720
aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780
cattcgacca ccaagcgaaa catcgcacgc agcagcacg tactcggatg gaagccggtc 840
ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900
ccaggctcaa ggcgcgcacg cccgacggcg aggatctcgt cgtgacccat ggcgatgct 960
gcttgccgaa tatcatgggt gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1020
tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080
ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgcgct cccgattcgc 1140
agcgcacgc cttctatcgc cttcttgacg agttcttctg agtttaaaca gaccacaacg 1200
gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260
cgaagccgct tggaataagg ccgggtgtgcg tttgtctata tgttattttc caccatattg 1320
ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct 1380
aggggtcttt cccctctcgc caaaggaatg caaggctctg tgaatgtcgt gaaggaagca 1440
gttccctctg aagcttcttg aagacaaaca acgtctgtag cgacccttg caggcagcgg 1500
aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacact 1560
gcaaaggcgg cacaaccoca gtgccacggt gtgagttgga tagttgtgga aagagtcaaa 1620
tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggc accccattgt 1680
atgggatctg atctggggcc tcgggtgcaca tgctttacat gtgttttagtc gaggttaaaa 1740
aacgtctagg cccccgaac cacggggacg tggttttcct ttgaaaaaca cgataatacc 1800
atggcgccca ttacggccta ctcccaacag acgcgagggc tacttggtg catcatcact 1860
agcctcacag gccgggacag gaaccaggtc gagggggagg tccaagtggc ctccaccgca 1920
acacaatctt tcctggcgac ctgcgtcaat ggcgtgtgtt ggactgtcta tcatggtgcc 1980
ggctcaaaga cccttgccgg cccaaagggc ccaatcacc aaatgtacac caatgtggac 2040
caggacctcg tcggctggca agcgccccc ggggcgcggt ccttgacacc atgcacctgc 2100
ggcagctcgg accttactt ggtcacgagg catgccgatg tcattccggg gcgccggcgg 2160
ggcgacagca gggggagcct actctcccc aggccgctct cctacttgaa gggctcttcg 2220
ggcggctccac tgctctgccc ctcggggcac gctgtgggca tcttctgggc tgccgtgtgc 2280
acctgagggg ttgcgaaggc ggtggacttt gtaccgctcg agtctatgga aaccactatg 2340
cggctccccg tcttcacgga caactcgtcc cctccggccg taccgcagac attccagggtg 2400
gcccattctac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 2460
gcccagggtg ataagggtgct tgtcctgaac ccgtccgtcg ccgccaccct aggtttcggg 2520
gcgtatatgt ctaaggcaca tggtatcgac cctaaccatca gaaccggggg aaggaccatc 2580
accacgggtg ccccatcac gtactccacc tatggcaagt ttcttgccga cgggtggttg 2640
tctggggcgc cctatgacat cataatatgt gatgagtgcc actcaactga ctcgaccact 2700
atcctgggca tcggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtcgtg 2760
ctcgccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggagggtg 2820
gctctgtcca gactggaga aatccccttt tatggcaaag ccatccccat cgagaccatc 2880
aaggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccgcg 2940
aagctgtccg gcctcggact caatgctgta gcatattacc ggggccttga tgtatccgtc 3000
ataccaacta gcggagacgt cattgtcgtc gcaacggacg ctctaatac gggctttacc 3060
ggcgatttgc actcagtgat cgactgcaat acatgtgtca cccagacagt cgacttcagc 3120
ctggaccgca ccttcacat tgagacgacg accgtgccac aagacgcggg gtcacgctcg 3180
cagcggcgag gcaggactgg taggggcagg atgggcattt acaggtttgt gactccagga 3240
gaacggccct cgggcatggt cgattcctcg gttctgtgcg agtgctatga cgcgggctgt 3300
gcttggtagc agctcacgcc cgcgagacc tcagttaggt tgcgggctta cctaaacaca 3360
ccagggttgc ccgtctgcca ggaccatctg gagttctggg agagcgtctt tacaggcctc 3420
acctacatag acgcccattt cttgtcccag actaagcagg caggagacaa ctccccctac 3480
ctggtagcat accaggctac ggtgtgcgcc agggctcagg ctccacctcc atcgtgggac 3540
caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gccctgctg 3600

```

tataggctgg	gagccgttca	aaacgaggtt	actaccacac	acccataac	caaatacatc	3660
atggcatgca	tgtcggctga	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	3720
gtcctagcag	ctctggccgc	gtattgctg	acaacaggca	gcgtggctcat	tgtgggcagg	3780
atcatcttgt	ccggaagcc	ggccatcatt	cccagacagg	aagtccttta	ccgggagttc	3840
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacagggaat	gcagctcgcc	3900
gaacaattca	aacagaaggc	aatcgggttg	ctgcaaacag	ccaccaagca	agcggaggct	3960
gctgctcccg	tgggtggaatc	caagtggcgg	accctcgaag	ccttctgggc	gaagcatatg	4020
tggaatttca	tcagcgggat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
gcatagcat	cactgatggc	attcacagcc	tctatcacca	gcccgtcac	cacccaacat	4140
accctcctgt	ttaacatcct	ggggggatgg	gtggccgccc	aacttgctcc	tcccagcgt	4200
gcttctgctt	tcgtaggcgc	cggcatcgct	ggagcggctg	ttggcagcat	aggccttggg	4260
aagggtgctt	tggatatttt	ggcaggttat	ggagcagggg	tggcaggcgc	gctcgtggcc	4320
tttaaggtea	tgagcggcga	gatgccctcc	accgaggacc	tggttaacct	actccctgct	4380
atcctctccc	ctggcgcctt	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4440
gtgggcccag	gggagggggc	tgtgcagtgg	atgaaccggc	tgatagcgtt	cgcttcgcgg	4500
ggtaaccacg	tctccccac	gcactatgtg	cctgagagcg	acgctgcagc	acgtgtcact	4560
cagatcctct	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gtggatcaac	4620
gaggactgct	ccacgccatg	ctccggctcg	tggctaagag	atgtttggga	ttggatatgc	4680
acggtgttga	ctgatttcaa	gacctggctc	cagtccaagc	tctgcccgcg	attgccggga	4740
gtccccttct	tctcatgtca	acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	4800
caaaccacct	gcccattgtg	agcacagatc	accggacatg	tgaaaaacgg	ttccatgagg	4860
atcgtggggc	ctaggacctg	tagtaacacg	tggcatggaa	cattccccat	taacgcgtac	4920
accacggggc	cctgcacgcc	ctccccggcg	ccaaattatt	ctagggcgct	gtggcgggtg	4980
gctgctgagg	agtacgtgga	ggttacgcgg	gtgggggatt	tccactacgt	gacgggcatg	5040
accactgaca	acgtaaagtg	cccgtgtcag	gttccggccc	ccgaattctt	cacagaagtg	5100
gatgggggtg	ggttgcacag	gtacgctcca	gcgtgcaaac	ccctcctacg	ggaggaggtc	5160
acattcctgg	tcgggctcaa	tcaataacctg	gttgggtcac	agctcccatg	cgagcccga	5220
ccggacgtag	cagtgtcac	ttccatgctc	accgacctt	cccacattac	ggcggagacg	5280
gctaagcgta	ggctggccag	gggatctccc	ccctccttgg	ccagctcctc	agctagccag	5340
ctgtctgcgc	cttccctgaa	ggcaacatgc	actaccgctc	atgactcccc	ggacgctgac	5400
ctcatcgagg	ccaacctcct	gtggcggcag	gagatgggcg	ggaacatcac	ccgcgtggag	5460
tcagaaaata	aggtagtaat	tttggactct	ttcgagccgc	tccaagcggga	ggaggatgag	5520
agggaaagtat	ccgttcgggc	ggagatcctg	cggaggcca	ggaaattccc	tcgagcgatg	5580
cccatatggg	cacgcccgga	ttacaacctt	ccactgttag	agtcctggaa	ggaccgggac	5640
tacgtccctc	cagtggta	cgggtgtcca	ttgccgcctg	ccaaggcccc	tccgatacca	5700
cctccacgga	ggaagaggac	ggttgtcctg	tcagaatcta	ccgtgtcttc	tgccttggcg	5760
gagctcgcca	caaagacctt	cggcagctcc	gaatcgtcgg	ccgtcgacag	cggcacggca	5820
acggcctctc	ctgaccagcc	ctccgacgac	ggcgcgcggg	gatccgacgt	tgagtcgtac	5880
tcctccatgc	cccccttga	gggggagccg	ggggatcccg	atctcagcga	cgggtcttgg	5940
tctaccgtaa	gcgaggaggc	tagtgaggac	gtcgtctgct	gctcgtatgc	ctacacatgg	6000
acaggcgccc	tgatcacgcc	atgcgctgcg	gaggaaacca	agctgcccat	caatgcaactg	6060
agcaactctt	tgctccgtca	ccacaacttg	gtctatgcta	caacatctcg	cagcgcgaagc	6120
ctgcggcaga	agaaggtcac	ctttgacaga	ctgcaggctc	tggacgacca	ctaccgggac	6180
gtgctcaagg	agatgaaggc	gaaggcgtcc	acagttaagg	ctaaacttct	atccgtggag	6240
gaagcctgta	agctgacgcc	cccacattcg	gccagatcta	aatttggcta	tggggcaaag	6300
gacgtccgga	acctatccag	caaggccggt	aaccacatcc	gctccgtgtg	gaaggacttg	6360
ctggaagaca	ctgagacacc	aattgacacc	accatcatgg	caaaaaatga	ggttttctgc	6420
gtccaaccag	agaagggggg	ccgcaagcca	gctcgcctta	tcgtattccc	agatttgggg	6480
gttcgtgtgt	gcgagaaaat	ggccctttac	gatgtggtct	ccacctccc	tcaggccgtg	6540
atgggctctt	catacggatt	ccaatactct	cctggacagc	gggtcgagtt	cctggtgaat	6600
gcctggaaag	cgaagaaatg	ccctatgggc	ttcgcatatg	acaccgctg	ttttgactca	6660
acggtcactg	agaatgacat	ccgtgttgag	gagtcaatct	accaatggtg	tgacttggcc	6720
cccgaagcca	gacaggccat	aaggtcgctc	acagagcggc	tttacatcgg	gggccccctg	6780
actaattcta	aagggcagaa	ctgcggctat	cgccgggtgc	gcgcgagcgg	tgtactgacg	6840
accagctgcg	gtaataacct	cacatgttac	ttgaaggccg	ctgcggcctg	tcgagctgcg	6900
aagctccagg	actgcacgat	gctcgtatgc	ggagacgacc	ttgtcgttat	ctgtgaaagc	6960
gcggggacc	aagaggacga	ggcgcgccta	cgggccttca	cggaggctat	gactagatac	7020
tctgcccccc	ctggggacc	gcccacaacca	gaatacgact	tggagttgat	aacatcatgc	7080
tcctccaatg	tgctcagtcgc	gcacgatgca	tctggcaaaa	gggtgtacta	tctcaccctg	7140
gacccacca	cccccttgc	gcgggctgcg	tgggagacag	ctagacacac	tccagtcaat	7200
tcctggctag	gcaacatcat	catgtatgcg	cccaccttgt	gggcaaggat	gatcctgatg	7260
actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaagccct	agattgtcag	7320
atctacgggg	cctgttactc	cattgagcca	cttgacctac	ctcagatcat	tcaacgactc	7380
catggcctta	gcgcattttc	actccatagt	tactctccag	gtgagatcaa	taggggtggct	7440
tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	7500
gtccgcgcta	ggctactgtc	ccaggggggg	agggctgcca	cttgtggcaa	gtacctcttc	7560
aactgggcag	taaggaccaa	gctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	7620

ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	7680
gcccgacccc	gctggttcat	gtggtgccta	ctcctacttt	ctgtaggggt	aggcatctat	7740
ctactcccca	accgatgaac	ggggagctaa	acactccagg	ccaataggcc	atcctgtttt	7800
ttccctttt	ttttttctt	ttttttttt	ttttttttt	ttttttttt	ttctcctttt	7860
tttttctct	tttttctt	ttcttctt	tggtggctcc	atcttagccc	tagtcacggc	7920
tagctgtgaa	aggtccgtga	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	7980
agatcaagt						7989

<210> 5  
 <211> 8649  
 <212> DNA  
 <213> Hepatitis C Virus

<400> 5

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctct cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tctttctctg gatcaaccog ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctctga gaccgtgcac catgagcacg aatcctaacc 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtc ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcc cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgccag 900
ccgaactgtt cgccaggctc aaggcgcgca tgcccgacgg cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctt cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcate gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtctt tccccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgtt tcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggcccccgga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatggaccg ggagatggca gcatcgtgcg gaggcgcggt tttcgtaggt 1860
ctgatactct tgacctgtc accgcactat aagctgttcc tcgctaggct catatgggtg 1920
ttacaatatt ttatcaccag ggccgaggca cacttgcaag tgtggatccc cccctcaac 1980
gttcgggggg gcgcgatgc cgtcatctc ctacagtgcg cgatccacc agagctaate 2040
tttaccatca caaaatctt gctcgcata ctcggtccac tcatgggtgct ccaggctgg 2100
ataaccaaag tgccgtactt cgtgcgcgca cacgggctca ttcgtgcatg catgctgggt 2160
cggaagggtg ctgggggtca ttatgtccaa atggctctca tgaagttggc cgactgaca 2220
ggtagctacg tttatgacca tctcacccca ctgcgggact gggcccacgc gggcctacga 2280
gaccttgccg tggcagttga gcccgctcgc ttctctgata tggagaccaa ggttatcacc 2340
tggggggcag acaccgcggc gtgtggggac atcatcttgg gcctgcccgt ctccgcccgc 2400
agggggaggg agatacatct gggaccggca gacagcctg aagggcaggg gtggcgactc 2460
ctcgcgccta ttacggccta ctcccaacag acgcgaggcc tacttggctg catcatcact 2520
agcctcacag gccgggacag gaaccaggct gagggggagg tccaagtggc ctccaccgca 2580
acacaatctt tcttggcgac ctgctgcaat ggcgtgtgtt ggactgtcta tcatgggtgcc 2640
ggctcaaaga cccttgccgg cccaaagggc ccaatcacc aatgtacac caatgtggac 2700
caggacctcg tcggctggca agcgccccc ggggcgcggt ccttgacacc atgcacctgc 2760
ggcagctcgg accttactt ggtcacgagg catgcccgat tcattccggt gcgcggcg 2820
ggcgacagca gggggagcct actctcccc aggccgctc cctacttgaa gggctcttcg 2880
ggcggctccac tgctctgccc ctccgggca gctgtgggca tcttccgggc tgccgtgtgc 2940
accgagggg ttgcgaaggc ggtggacttt gtaccctcgc agtctatgga aaccactatg 3000
cggctcccgg tcttcacgga caactcgtcc cctccggccg taccgcagac attccagggt 3060
gcccattctac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 3120
gccaagggt ataaggtgct tgtcctgaac ccgtccgtcg ccgccaccct aggtttcggg 3180
gcgtatatgt ctaaggcaca tggatcgcac cctaacaatca gaaccggggg aaggaccatc 3240
accacgggtg ccccatcac gtactccacc tatggcaagt ttcttgccga cgggtggttc 3300
tctggggggc cctatgacat cataatatgt gatgagtgc actcaactga ctcgaccact 3360
atcctgggca tcggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtcgtg 3420
ctcgcaccgg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggaggtg 3480
gctctgtcca gcaactggaga aatccccctt tatggcaag ccatcccat cgagaccatc 3540
aaggggggga ggcacctcat tttctgccc tccaagaaga aatgtgatga gctcgcgcg 3600

```

aagctgtccg	gcctcggact	caatgctgta	gcatattacc	ggggccttga	tgtatccgtc	3660
ataccaacta	gcgagacgt	cattgtcgta	gcaacggacg	ctctaagac	gggctttacc	3720
ggcgatttcg	actcagtgat	cgactgcaat	acatgtgtca	cccagacagt	cgacttcagc	3780
ctggaccoga	ccttcacat	tgagacgacg	accgtgccac	aagacgcggt	gtcacgctcg	3840
cagcggcgag	gcaggactgg	taggggcagg	atgggcattt	acaggtttgt	gactccagga	3900
gaacggccct	cgggcatggt	cgattcctcg	gttctgtgcg	agtgctatga	cgcgggctgt	3960
gcttggtagc	agctcacgcc	cgccgagacc	tcagttaggt	tgcgggctta	cctaaacaca	4020
ccagggttgc	ccgtctgcca	ggaccatctg	gagttctggg	agagcgtctt	tacaggcctc	4080
accacatag	acgcccattt	cttgtcccag	actaagcagg	caggagacaa	cttcccctac	4140
ctggtagcat	accaggctac	ggtgtgcgcc	agggctcagg	ctccacctcc	atcgtgggac	4200
caaatgtgga	agtgtctcat	acggctaaag	cctacgctgc	acgggccaac	gcccctgctg	4260
tataggctgg	gagccgttca	aaacgagggt	actaccacac	accccataac	caaatacatc	4320
atggcatgca	tgtcggctga	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcggg	4380
gtcctagcag	ctctggccgc	gtattgcctg	acaacaggca	gcgtggctcat	tgtgggcagg	4440
atcatcttgt	ccggaagcc	ggccatcatt	cccagacagg	aagtccttta	ccgggagttc	4500
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacagggaat	gcagctcgcc	4560
gaacaattca	aacagaaggc	aatcgggttg	ctgcaaacag	ccaccaagca	agcggaggct	4620
gctgctcccg	tgggtggaatc	caagtggcgg	accctcgaag	ccttctgggc	gaagcatatg	4680
tggaatttca	tcagcgggat	acaatattta	gcaggcttgt	ccactctgce	tggcaacccc	4740
gcatagcat	cactgatggc	attcacagcc	tctatcacca	gcccgtcac	caccaacat	4800
accctcctgt	ttaacatcct	ggggggatgg	gtggccgcc	aacttgcctc	tcccagcgt	4860
gcttctgctt	tcgtaggcgc	cggcatcgct	ggagcggctg	ttggcagcat	aggccttggg	4920
aaggtgcttg	tggatatttt	ggcaggttat	ggagcagggg	tggcagggcg	gctcgtggcc	4980
tttaaggcca	tgagcggcga	gatgccctcc	accgaggacc	tggttaacct	actccctgct	5040
atcctctccc	ctggcgccct	agtctgctgg	gtcgtgtgcg	cagcgatact	gcgtcggcac	5100
gtgggcccag	gggagggggc	tgtgcagtgg	atgaaccggc	tgatagcgtt	cgcttcgctg	5160
ggtaaccacg	tctccccac	gcactatgtg	cctgagagcg	acgctgcagc	acgtgtcact	5220
cagatcctct	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gtggatcaac	5280
gaggactgct	ccacgccatg	ctccggctcg	tggctaagag	atgtttggga	ttggatatgc	5340
acggtgttga	ctgatttcaa	gacctggctc	cagtccaagc	tcttgcgctg	attgccggga	5400
gtccccttct	tctcatgtca	acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	5460
caaaccacct	gcccattgtg	agcacagatc	accggacatg	tgaaaaacgg	ttccatgagg	5520
atcgtggggc	ctaggacctg	tagtaacacg	tggcatggaa	cattccccat	taacgcgtac	5580
accacggggc	cctgcacgcc	ctccccggcg	ccaaattatt	ctagggcgct	gtggcgggtg	5640
gctgctgagg	agtacgtgga	ggttacgcgg	gtgggggatt	tccactacgt	gacgggcatg	5700
accactgaca	acgtaaagtg	cccgtgtcag	gttccggccc	ccgaattctt	cacagaagtg	5760
gatgggggtg	ggttgcaacg	gtacgctcca	gcgtgcaaac	ccctcctacg	ggaggaggtc	5820
acattcctgg	tcgggctcaa	tcaataacctg	gttgggtcac	agctcccatg	cgagcccga	5880
ccggacgtag	cagtgtctac	tccatgctc	accgacctc	cccacattac	ggcggagacg	5940
gctaagcgta	ggctggccag	gggatctccc	ccctccttgg	ccagctcatc	agctagccag	6000
ctgtctgctg	cttccttgaa	ggcaacatgc	actaccgctc	atgactcccc	ggacgctgac	6060
ctcatcgagg	ccaacctcct	gtggcggcag	gagatgggcg	ggaacatcac	ccgctgggag	6120
tcagaaaata	aggtagtaat	tttggactct	ttcgagccgc	tccaagcggg	ggaggatgag	6180
agggaaagtat	ccgttccggc	ggagatcctg	cggagggtcca	ggaaattccc	tcgagcgatg	6240
cccatatggg	cacgcccggg	ttacaacctt	ccactgttag	agtcctggaa	ggacccggac	6300
tacgtccctc	cagtgggtaca	cgggtgtcca	ttgcccctg	ccaaggcccc	tccgatacca	6360
cctccacgga	ggaagaggac	ggttgtcctg	tcagaatcta	ccgtgtcttc	tgccttggcg	6420
gagctcgcca	caaagacctt	cggcagctcc	gaatcgtcgg	ccgtcgacag	cggcacggca	6480
acggcctctc	ctgaccagcc	ctccgacgac	ggcgcgcggg	gatccgacgt	tgagtcgtac	6540
tcttccatgc	cccccttga	gggggagccg	ggggatcccg	atctcagcga	cgggtcttgg	6600
tctaccgtaa	gcgaggaggc	tagtgaggac	gtcgtctgct	gctcgatgtc	ctacacatgg	6660
acaggcggcc	tgatcacgcc	atgcgctgcg	gaggaaacca	agctgcccac	caatgcactg	6720
agcaactctt	tgctccgtca	ccacaacttg	gtctatgcta	caacatctcg	cagcgcaagc	6780
ctgcggcaga	agaagggtcac	ctttgacaga	ctgcaggctc	tggacgacca	ctaccgggac	6840
gtgctcaagg	agatgaaggc	gaaggcgtcc	acagttaagg	ctaaactctt	atccgtggag	6900
gaagcctgta	agctgacgcc	cccacattcg	gccagatcta	aatttggcta	tggggcaaag	6960
gacgtccgga	acctatccag	caaggccgtt	aaccacatcc	gctccgtgtg	gaaggacttg	7020
ctggaagaca	ctgagacacc	aattgacacc	accatcatgg	caaaaaatga	ggttttctgc	7080
gtccaaccag	agaagggggg	ccgcaagcca	gctcgcctta	tcgtattccc	agatttgggg	7140
gttcgtgtgt	gcgagaaaat	ggccctttac	gatgtggtct	ccaccctccc	tcaggccgtg	7200
atgggctctt	catacggatt	ccaatactct	cctggacagc	gggtcgagtt	cctgggtgaat	7260
gcctggaaag	cgaagaaatg	ccctatgggc	ttcgcatatg	acaccgctg	ttttgactca	7320
acggtcactg	agaatgacat	ccgtgttgag	gagtcaatct	accaatgttg	tgacttggcc	7380
cccgaagcca	gacaggccat	aaggtcgtct	acagagcggc	tttacatcgg	gggccccctg	7440
actaattcta	aagggcagaa	ctgcggctat	cgccgggtgc	gcgcgagcgg	tgtactgacg	7500
accagctgcg	gtaataacct	cacatgttac	ttgaaggccg	ctgcggcctg	tcgagctgcg	7560
aagctccagg	actgcacgat	gctcgtatgc	ggagacgacc	ttgtcgttat	ctgtgaaagc	7620

gcggggaccc	aagaggacga	ggcgagccta	cgggccttca	cggaggctat	gactagatac	7680
tctgcccccc	ctggggaccc	gcccacacca	gaatacgact	tggagttgat	aacatcatgc	7740
tctccaatg	tgteagtcgc	gcacgatgca	tctggcaaaa	gggtgtacta	tctcaccctg	7800
gacccaccca	cccccttgc	gcgggctgcg	tgggagacag	ctagacacac	tccagtcaat	7860
tctgggctag	gcaacatcat	catgtatgcg	cccaccttgt	gggcaaggat	gatcctgatg	7920
actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaaagccct	agattgtcag	7980
atctacgggg	cctgttactc	cattgagcca	cttgacctac	ctcagatcat	tcaacgactc	8040
catggcctta	gcgcattttc	actccatagt	tactctccag	gtgagatcaa	taggggtggct	8100
tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	8160
gtccgcgcta	ggctactgtc	ccaggggggg	agggctgccca	cttgtggcaa	gtacctcttc	8220
aactgggcag	taaggaccaa	gctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	8280
ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	8340
gcccgacccc	gctggttcat	gtgggtgccta	ctcctacttt	ctgtaggggt	aggcatctat	8400
ctactcccca	accgatgaac	ggggagctaa	acactccagg	ccaataggcc	atcctgtttt	8460
tttccctttt	ttttttctt	ttttttttt	ttttttttt	ttttttttt	ttctcctttt	8520
tttttctct	tttttctct	ttcttctct	tggtggctcc	atcttagccc	tagtcacggc	8580
tagctgtgaa	aggcccgatg	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	8640
agatcaagt						8649

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

<400> 6

```

gccagcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctct cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tcccttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggtat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacctgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcacctg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggtgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaaagccg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 900
ccgaactggt cgcaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcac gccttctatc gccttcttga cgagttcttc tgagtttaa 1200
cagaccacaa cggtttccct ctagcgggat caattccgcc cctctcctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctaggggtct tccccctct gccaaaggaa tgcaaggctc gttgaatgct 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaa gccacgtgta 1560
taagatacac ctgcaaagge ggcacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgt tcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740
tcgaggtaa aaaacgtcta ggcccccgga accacgggga cgtggtttct ctttgaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
tgcatcatca ctagcctcac aggcggggac aggaaccagg tcgaggggga ggtccaagtg 1920
gtctccaccg caacacaatc tttcctggcg acctgcgtca atggcgtgtg ttggactgtc 1980
tatcatgggt ccggtcaaa gacccttgcc ggcccaaagg gcccaatcac ccaaatgtac 2040
accaatgtgg accaggacct cgtcggctgg caagcgcgcc ccggggcgcg ttccttgaca 2100
ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160
gtgcgccggc gggcgacag cagggggagc ctactctccc ccaggcccgt ctactacttg 2220
aagggtctt cggcggtcc actgctctgc cctcggggc atgctgtggg catctttcgg 2280
gctgccgtgt gcacctgagg ggttgcgaag gcggtggact ttgtaccctg cgagtctatg 2340
gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
acattccagg tggccatct acacgccct actggtagcg gcaagagcac taaggtgccg 2460
gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgtccgt cgccgccacc 2520
ctaggttctg gggcgtatat gtctaaggca catggtatcg accctaacat cagaaccggg 2580
gtagggacca tcaccacggg tgccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacggtggtt gctctggggg cgctatgac atcataatat gtgatgagtg ccaactcaact 2700
gactcgacca ctatcctggg catcggcaca gtcctggacc aagcggagac ggctggagcg 2760
cgactcgtcg tgctcgccac cgctacgcct ccgggatcgg tcaccgtgcc acatccaaac 2820
atcgaggagg tggctctgtc cagcactgga gaaatcccct tttatggcaa agccatcccc 2880
atcgcgacca tcaagggggg gaggcacctc atttctgccc attccaagaa gaaatgtgat 2940
gagctcggcg cgaagctatc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaatg 3060
acgggcttta ccggcgattt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
gtcgacttca gcctggacct gacctcacc attgagacga cgaccgtgcc acaagacgcg 3180
gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgtat 3300
gacgcgggct gtgcttggtg cgagctcagc ccgcccgaga cctcagttag gttgcgggct 3360
tacctaaaca caccagggt gcccgctgc caggaccatc tggagttctg ggagagcgtc 3420
tttacaggcc tcaccacat agacgccat ttcttgtccc agactaggca ggcaggagac 3480
aacttcccct acctggtagc ataccaggct acggtgtgcg ccagggtca ggctccacct 3540
ccatcgtggg accaaatgtg ggagtgtctc atacggctaa agcctacgct gcacgggcca 3600

```

acgcccctgc	tgtataggct	gggagccggt	caaaacgagg	ttactaccac	acaccccata	3660
accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
attgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccagacg	ggaagtcctt	3840
taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcctttacat	cgaacaggga	3900
atgcagctcg	ccgaacaatt	caaacagaag	gcaatcgggt	tgctgcaaac	agccaccaag	3960
caagcggagg	ctgctgctcc	cgtggtggaa	tccaagtggc	ggaccctcga	agccttctgg	4020
gcgaagcata	tgtggaattt	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgtct	4140
accaccaaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
cctcccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tgttggcagc	4260
ataggccttg	ggaaggtgct	tgtggatatt	ttggcagggt	atggagcagg	ggtggcaggc	4320
gcgctcgtgg	cctttaaggt	catgagcggc	gagatgccct	ccaccgagga	cctggttaac	4380
ctactccctg	ctatcctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgcagcgata	4440
ctgcgtcggc	acgtgggccc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
ttcgcttcgc	ggggtaacca	cgtctcccc	acgcactatg	tgtctgagag	cgacgctgca	4560
gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggcttcac	4620
cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggctaag	agatgtttgg	4680
gattggatat	gcacgggtgtt	gactgatttc	aagacctggc	tccagtccaa	gctcctgccg	4740
cgattgccgg	gagtcacctt	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
gacggcatca	tgcaaaccac	ctgccatgt	ggagcacaga	tcaccggaca	tgtgaaaaac	4860
ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
attaacgcgt	acaccacggg	ccctgcacg	ccctccccgg	cgccaaatta	ttctagggcg	4980
ctgtggcggg	tggctgctga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
gtgacgggca	tgaccactga	caacgtaaag	tgcccgtgtc	aggttccggc	ccccgaattc	5100
ttcacagaag	tggatggggg	gcggttgca	aggtagcctc	cagcgtgcaa	accctccta	5160
cgggaggagg	tcacattcct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctcca	5220
tgcgggcccg	aaccggacgt	agcagtgtc	acttccatgc	tcaccgacc	ctcccacatt	5280
acggcggaga	cggctaagcg	taggctggcc	aggggatctc	ccccctcctt	ggccagctca	5340
tcagctagcc	agctgtctgc	gccttccttg	aaggcaacat	gcactaccgg	tcatgactcc	5400
ccggacgctg	acctcatcga	ggccaacctc	ctgtggcggc	aggagatggg	cgggaacatc	5460
accgcgctgg	agtcagaaaa	taaggtagta	atthtggact	ctttcgagcc	gctccaagcg	5520
gaggaggatg	agagggaaag	atccgttccg	gcggagatcc	tgcggagggtc	caggaaattc	5580
cctcgagcga	tgcccatatg	ggcacgcccg	gattacaacc	ctccactggt	agagtcctgg	5640
aaggaccggg	actacgtccc	tccagtggta	cacgggtgtc	cattgccgcc	tgccaaggcc	5700
cctccgatac	cacctccacg	gagggagagg	acggttgctc	tgctcagaatc	taccgtgtct	5760
tctgccttgg	cggagctcgc	cacaaagacc	ttcggcagct	ccgaatcgtc	ggccgtcgac	5820
agcggcacgg	caacggcctc	tctgaccag	ccctccgacg	acggcgcgcg	gggatccgac	5880
gttgagtcgt	actcctccat	gccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
gacgggtctt	ggtctaccgt	aagcggagg	gctagtgagg	acgtcgtctg	ctgctcgatg	6000
tcctacacat	ggacagggcg	cctgatcagc	ccatgcgctg	cggaggaaac	caagctgccc	6060
gtcaatgcac	tgagcaactc	tttgcctcgt	caccacaact	tggtctatgc	tacaacatct	6120
cgcagcgcaa	gcctgcggca	gaagaaggtc	acctttgaca	gactgcaggt	cctggacgac	6180
cactaccggg	acgtgctcaa	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
ctatccgtgg	aggaagcctg	taagctgacg	ccccacatt	cggccagatc	taaatttggc	6300
tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgctccgtg	6360
tggaaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
gaggttttct	gcgtccaacc	agagaagggg	ggccgcaagc	cagctcgcct	tatcgtatct	6480
ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggt	ctccaccctc	6540
cctcaggccg	tgatgggctc	ttcatacgga	ttccaatact	ctcctggaca	gcgggtcgag	6600
ttcctgggtg	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacaccgcg	6660
tgttttgact	caacggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
tgtgacttgg	ccccgaagc	cagacaggcc	ataaggctcg	tcacagagcg	gctttacatc	6780
gggggcccc	tgactaattc	taaagggcag	aactgcggct	atcgcgggtg	ccgcgcgagc	6840
ggtgtactga	cgaccagctg	cggtaatacc	ctcacatggt	atthgaaggc	cgctgcggcc	6900
tgtcgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
atctgtgaaa	gcgcggggac	ccaagaggac	gagggcagcc	tacgggcctt	cacggaggct	7020
atgactagat	actctgcccc	ccctggggac	ccgcccaaac	cagaatacga	cttgaggttg	7080
ataacatcat	gctcctccaa	tgtgtcagtc	gcgcacgatg	catctggcaa	aaggggtgtac	7140
tatctcacc	gtgacccac	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
actccagtea	attcctggct	aggcaacatc	atcatgtatg	cgccacctt	gtgggcaagg	7260
atgatcctga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
ctagattgtc	agatctacgg	ggcctgttac	tcattgagc	cacttgacct	acctcagatc	7380
attcaacgac	tccatggcct	tagcgcattt	tcactccata	gttactctcc	aggtgagatc	7440
aataggggtg	cttcatgcct	caggaaactt	ggggtaccgc	ccttgcgagt	ctggagacat	7500
cgggccagaa	gtgtccgcgc	taggctactg	tcccaggggg	ggagggctgc	cacttgtggc	7560
aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggctgcg	7620

tcccagttgg	atztatccag	ctggttcggt	gctggttaca	gcgggggaga	catatatcac	7680
agcctgtctc	gtgcccgacc	ccgctgggtc	atgtgggtgcc	tactcctact	ttctgtaggg	7740
gtaggcatct	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
ccatcctggt	ttttccctt	ttttttttc	ttttttttt	ttttttttt	ttttttttt	7860
ttttctcct	ttttttcct	ctttttttc	ttttctttc	tttgggtggt	ccatcttagc	7920
cctagtcacg	gctagctgtg	aaagggtccg	gagccgcttg	actgcagaga	gtgctgatac	7980
tggcctctct	gcagatcaag	t				8001

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

<400> 7

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctct cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaac 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtct ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgccag 900
ccgaactgtt cggcaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctt cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcac gccttctatc gccttcttga cgagttcttc tgagtttaa 1200
cagaccacaa cggtttcctt ctagcgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctaggggtct tcccctctc gccaaaggaa tgcaaggctc gttgaatgct 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgt tcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta gccccccga accacgggga cgtggttttc ctttgaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaact taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcgggtg agtttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggctc 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcggc agcccgaggg tagggcctgg 2040
gctcagcccg ggtaccctct gccctctat ggcaatgagg gcttgggggt ggcaggatgg 2100
ctcctgtcac cccgtggctc tcggcctagt tggggcccca cggacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggteat cgataccctc acgtgcggct tcgccgatct catgggggtac 2220
attccgctcg tcggcgcccc cctagggggc gctgccaggg ccctggcgca tggcgtccgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttctc cttttctatc 2340
ttccttttgg ctttgcgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcgacatga tcatgcatc cccgggtgc gtgccctgcg ttccggagaa caactcctcc 2520
cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga tttgctcgtt ggggcggctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttctc gtcgccagc tgttcacctt ctgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatac ccggccactg gacaggtcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagc tgcgtggat atggtggcgg gggccattg gggagtccta 2880
gcgggccttg cctactattc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttgccg gcgttgacgg gggaaacctat gtgacagggg ggacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt tccaccggg tcatcccaga aatccagct tgtaaacacc 3060
aacggcagct ggcaatcaa caggactgcc ctgaactgca atgactcct caaactggg 3120
ttccttgcgt cgctgttcta cgtgcacaag ttcaactcat ctggatgcc agagcgcatg 3180
gccagctgca gcccatcga cgcgttcgct caggggtggg ggccatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcactacg cccccggcc gtgcggtatc 3300
gtaccccgcg cgcaggtgtg tggccagtgc tactgcttca cccaagccc tgcgtgggtg 3360
gggacgaccg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgcg gccgcgcaa ggcaactggt ttggctgtac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc ccccgtgta acatcggggg gatcggcaat 3540
aaaaccttga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600

```

tgtgggtcgg	ggccttggtt	gacacccaga	tgcttggtec	actaccata	caggctttgg	3660
cactaccct	gcactgtcaa	ctttaccatc	ttcaaggta	ggatgtacgt	ggggggagt	3720
gagcacaggc	tcgaagccgc	atgcaattgg	actcgaggag	agcgttgtaa	cctggaggac	3780
agggacagat	cagagcttag	cccgtctctg	ctgtctacaa	cggagtggca	ggtattgcc	3840
tgttccttca	ccaccctacc	ggctctgtcc	actggtttga	tccatctcca	tcagaacgtc	3900
gtggacgtac	aatacctgta	cggtataggg	tcggcggttg	tctcctttgc	aatcaaatgg	3960
gagtatgtcc	tgttgctctt	ccttctctctg	gcgacgcgc	gcgtctgtgc	ctgcttgtgg	4020
atgatgctgc	tgatagctca	agctgaggcc	gccctagaga	acctggtggt	cctcaacgcg	4080
gcatccgtgg	ccggggcgca	tggcattctc	tcttctctcg	tgttctctctg	tgctgcctgg	4140
tacatcaagg	gcaggctggt	ccctggggcg	gcatatgcc	tctacggcgt	atggccgcta	4200
ctcctgctcc	tgctggcggt	accaccacga	gcatacgcca	tggaccggga	gatggcagca	4260
tcgtgcggag	gcgcgggttt	cgtaggtctg	atactcttga	ccttgtcacc	gcactataag	4320
ctgctcctcg	ctaggctcat	atggtggtta	caatatthta	tcaccagggc	cgaggcacac	4380
ttgcaagtgt	ggatccccc	cctcaacggt	cgggggggcc	gcgatgccgt	catcctcctc	4440
acgtgcgcga	tccaccacga	gctaactctt	accatcacca	aatcttgc	cgccatactc	4500
ggccactca	tggtgctcca	ggctgggata	accaaagtgc	cgtacttctg	gcgcgcacac	4560
gggctcattc	gtgcatgcat	gctgggtcgg	aagggtgctg	ggggtcatta	tgtccaaatg	4620
gctctcatga	agttggccgc	actgacaggt	acgtacgttt	atgaccatct	cacccactg	4680
cgggactggg	cccacgcggg	cctacgagac	cttgcgggtg	cagttgagcc	cgctcgtctc	4740
tctgatatgg	agaccaaggt	tatcacctgg	ggggcagaca	ccgcggcgtg	tggggacatc	4800
atcttggggc	tgcccgtctc	cgcccgcagg	gggagggaga	tacatctggg	accggcagac	4860
agccttgaag	ggcaggggtg	gcgactcctc	gcgcctatta	cggcctactc	ccaacagacg	4920
cgaggcctac	ttggctgcat	catcactagc	ctcacaggcc	gggacaggaa	ccaggctcgag	4980
ggggaggtcc	aagtggctctc	caccgcaaca	caatctttcc	tggcgacctg	cgtcaatggc	5040
gtgtgttggg	ctgtctatca	tggtgccggc	tcaaagacc	ttgccggccc	aaagggccca	5100
atcacccaaa	tgtacaccaa	tgtggaccag	gacctcgtcg	gctggcaagc	gccccccggg	5160
gcgcgttctc	tgacaccatg	cacctgcggc	agctcggacc	tttacttggg	cacgaggcat	5220
gccgatgtca	ttccgggtgcg	ccggcggggc	gacagcaggg	ggagcctact	ctccccagg	5280
cccgtctcct	acttgaaggg	ctcttcgggc	ggtccactgc	tctgcccctc	ggggcatgct	5340
gtgggcatct	ttcgggctgc	cgtgtgcacc	cgaggggttg	cgaaggcggg	ggactttgta	5400
cccgtcgagt	ctatggaaac	cactatgcgg	tcccgggtct	tcacggacaa	ctcgtcccct	5460
ccggccgtac	cgcagacatt	ccagggtggc	catctacacg	cccctactgg	tagcggcaag	5520
agcactaagg	tgccggctgc	gtatgcagcc	caagggtata	aggtgcttgt	cctgaaccgg	5580
tccgtcgccg	ccaccctagg	tttcggggcg	tatatgtcta	aggcacatgg	tatcgaccct	5640
aacatcagaa	ccggggtagg	gaccatcacc	acgggtgcc	ccatcacgta	ctccacctat	5700
ggcaagtthc	ttgccgacgg	tggttgctct	gggggcgcct	atgacatcat	aatatgtgat	5760
gagtgccact	caactgactc	gaccactatc	ctgggcatcg	gcacagtcct	ggaccaagcg	5820
gagacggctg	gagcgcgact	cgctcgtgctc	gccaccgcta	cgctccggg	atcggtcacc	5880
gtgccacatc	caaacatcga	ggagggtggct	ctgtccagca	ctggagaaat	ccccttttat	5940
ggcaaagcca	tcccacatcg	gaccatcaag	ggggggaggc	acctcatttt	ctgccattcc	6000
aagaagaaat	gtgatgagct	cgccgcgaag	ctatccggcc	tcggactcaa	tgctgtagca	6060
tattaccggg	gccttgatgt	atccgtcata	ccaactagcg	gagacgtcat	tgctcgtagca	6120
acggacgctc	taatgacggg	ctttaccggc	gatttcgact	cagtgatcga	ctgcaatata	6180
tgtgtcacc	agacagtcga	cttcagcctg	gaccgcacct	tcaccattga	gacgacgacc	6240
gtgccacaag	acgcgggtgc	acgctcgcag	cggcagggca	ggactggtag	gggcaggatg	6300
ggcatttaca	ggtttgtgac	tccaggagaa	cggccctcgg	gcatgttcga	ttcctcgggt	6360
ctgtgcgagt	gctatgacgc	gggctgtgct	tggtacgagc	tcacgcccgc	cgagacctca	6420
gttaggttgc	gggcttacct	aaacacacca	gggctgcccg	tctgccagga	ccatctggag	6480
ttctggggaga	gcgtctttac	aggcctcacc	cacatagacg	cccatttctt	gtcccagact	6540
aggcaggcag	gagacaactt	cccctacctg	gtagcatacc	aggctacggg	gtgcgccagg	6600
gctcaggctc	cacctccatc	gtgggaccaa	atgtgggagt	gtctcatacg	gctaaagcct	6660
acgctgcacg	ggccaacgcc	cctgctgtat	aggctgggag	ccgttcaaaa	cgaggttact	6720
accacacacc	ccataaccaa	atacatcatg	gcatgcatgt	cggctgacct	ggaggctcgtc	6780
acgagcacct	gggtgctggt	aggcggagtc	ctagcagctc	tggccgcgta	ttgcctgaca	6840
acaggcagcg	tggtcattgt	gggcaggatc	atcttgtccg	gaaagccggc	catcattccc	6900
gacaggggag	tcctttaccg	ggagtctgat	gagatggaag	agtgcgcctc	acacctccct	6960
tacatcgaac	agggaatgca	gctcgcggaa	caattcaaac	agaaggcaat	cgggttgctg	7020
caaacagcca	ccaagcaagc	ggaggctgct	gctcccgtgg	tggaaatcaa	gtggcggacc	7080
ctcgaagcct	tctgggcgaa	gcatatgtgg	aatttcatca	gcgggataca	atatttagca	7140
ggcttgtcca	ctctgcctgg	caaccccgcg	atagcatcac	tgatggcatt	cacagcctct	7200
atcaccagcc	cgctcaccac	ccaacatacc	ctcctgttta	acatcctggg	gggatgggtg	7260
gccgcccac	ttgctcctcc	cagcgtgct	tctgctttcg	taggcgccgg	catcgctgga	7320
gcggctgttg	gcagcatagg	ccttgggaag	gtgcttgtgg	atattttggc	aggttatgga	7380
gcaggggttg	caggcgcgct	cgtggccttt	aaggctcatga	gcggcgagat	gccctccacc	7440
gaggacctgg	ttaacctact	ccctgctatc	ctctcccctg	gcgccttagt	cgctcggggtc	7500
gtgtgcgcag	cgatactgcg	tcggcacgtg	ggcccagggg	agggggctgt	gcagtggatg	7560
aaccggctga	tagcgttcgc	ttcgcggggg	aaccacgtct	ccccacgca	ctatgtgtct	7620

gagagcgacg	ctgcagcacg	tgtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
ctgaagaggg	ttcaccagtg	gatcaacgag	gactgctcca	cgccatgctc	cggctcgtgg	7740
ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atctcaagac	ctggctccag	7800
tccaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
catggaacat	tccccattaa	cgcgtacacc	acgggccctc	gcacgccctc	cccggcgcca	8040
aattattcta	gggcgctgtg	gcgggtggct	gctgaggagt	acgtggaggt	tacgcgggtg	8100
ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaagtgcc	gtgtcagggt	8160
ccggcccccg	aattcttcac	agaagtggat	ggggtgcggg	tgcacaggta	cgctccagcg	8220
tgcaaaccct	tcctacggga	ggaggtcaca	ttcctggctg	ggctcaatca	atacctgggt	8280
gggtcacagc	tcccatgcgg	gcccgaaccg	gacgtagcag	tgctcacttc	catgctcacc	8340
gaccctccc	acattacggc	ggagacggct	aagcgtaggc	tggccagggg	atctcccccc	8400
tccttgcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
accgctcatg	actccccgga	cgctgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
atgggcggga	acatcacccg	cgtggagtca	gaaaataagg	tagtaatttt	ggactctttc	8580
gagccgctcc	aagcggagga	ggatgagagg	gaagtatccg	ttccggcgga	gatcctgcgg	8640
aggtccagga	aattccctcg	agcgatgcc	atatgggcac	gcccggatta	caaccctcca	8700
ctgttagagt	cctggaagga	cccggactac	gtccctccag	tggtagacag	gtgtccattg	8760
ccgctgcca	aggccccctc	gataccacct	ccacggaggg	agaggacggg	tgtcctgtca	8820
gaatctaccg	tgtcttctgc	cttggcggag	ctcgccacaa	agaccttcgg	cagctccgaa	8880
tcgtcggccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
gacgcgggat	ccgacgttga	gtcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
gatcccgatc	tcagcgacgg	gtcttggctc	accgtaagcg	aggaggctag	tgaggacgtc	9060
gtctgctgct	cgatgtccta	cacatggaca	ggcgccctga	tcacgccatg	cgctgcggag	9120
gaaaccaagc	tgcccgtcaa	tgactgagc	aactctttgc	tccgtcacca	caacttggtc	9180
tatgctacaa	catctcgcag	cgcaagcctg	cggcagaaga	aggtcacctt	tgacagactg	9240
caggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaagc	tgacgcccc	acattcggcc	9360
agatctaaat	ttggctatgg	ggcaaaggac	gtccggaacc	tatccagcaa	ggccgttaac	9420
cacatccgct	ccgtgtggaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccacc	9480
atcatggcaa	aaaatgaggt	tttctgcgtc	caaccagaga	aggggggccc	caagccagct	9540
cgccttatcg	tattcccaga	tttggggggt	cgtgtgtgcg	agaaaatggc	cctttacgat	9600
gtggtctcca	ccctccctca	ggccgtgatg	ggctcttcat	acggattcca	atactctcct	9660
ggacagcggg	tcgagttcct	ggtgaatgcc	tggaaagcga	agaaatgcc	tatgggcttc	9720
gcatatgaca	cccgtgtttt	tgactcaacg	gtcactgaga	atgacatccg	tgttgaggag	9780
tcaatctacc	aatgttgtga	cttgcccccc	gaagccagac	aggccataag	gtcgctcaca	9840
gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
cggtgcccg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttatttg	9960
aaggccgctg	cggcctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgtatgcgga	10020
gacgaccttg	tcgttatctg	tgaaagcgcg	gggacccaag	aggacgaggc	gagcctacgg	10080
gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccgcc	caaaccagaa	10140
tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcatct	10200
ggcaaaagg	tgtactatct	caccctgtgac	cccaccacc	cccttgcgcg	ggctgcgtgg	10260
gagacagcta	gacacactcc	agtcaattcc	tggctaggca	acatcatcat	gtatgcgccc	10320
accttggtgg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
gacctacctc	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagttac	10500
tctccagggtg	agatcaatag	ggtggcttca	tgctcagga	aacttggggg	accgcccttg	10560
cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtccca	gggggggagg	10620
gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
ccaatcccgg	ctgcgtccca	gttggattta	tccagctggg	tcgttgctgg	ttacagcggg	10740
ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgct	ggttcatgtg	gtgcctactc	10800
ctactttctg	taggggtagg	catctatcta	ctccccacc	gatgaacggg	gagctaaaca	10860
ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
tttttttttt	tttttttttc	tccttttttt	ttctcttttt	tttctttttc	tttcttttgg	10980
tggctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
agagagtgct	gatactggcc	tctctgcaga	tcaagt			11076

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

<400> 8

```

gccagcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctct cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc 240
gogagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaac 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctccgcgccag 900
ccgaactgtt cgcaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcac gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctaggggtct tcccctctc gccaaaggaa tgcaaggctt gttgaatgtc 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaagge ggcacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgtg ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta gccccccga accacgggga cgtggtttcc ctttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
tgcatcatca ctagectcac aggcggggac aggaaccagg tcgaggggga ggtccaagtg 1920
gtctccaccg caacacaatc tttcctggcg acctgcgtca atggcgtgtg ttggactgtc 1980
tatcatggtg ccggctcaaa gacccttgcc ggcccaaaagg gcccaatcac ccaaagtac 2040
accaatgtgg accaggacct cgtcggctgg caagcgcgcc ccgggggcgc ttccttgaca 2100
ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160
gtgcgccggc ggggcgacag cagggggagc ctactctccc ccaggcccgt ctactacttg 2220
aagggtctt cgggcggctc actgctctgc cctcggggc acgctgtggg catcttctcg 2280
gctgccgtgt gcacccgagg ggttgcgaag gcggtggact ttgtacctgt cgagtctatg 2340
gaaaccacta tgccgtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
acattccagg tggccatct acacgcccct actggtagcg gcaagagcac taaggtgccg 2460
gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgtccgt cgccgccacc 2520
ctaggtttcg gggcgatat gtctaaggca catggtatcg accctaacat cagaatcggg 2580
gtaaggacca tcaccacggg tgccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacggtggtt gctctggggg cgcctatgac atcataatat gtgatgagtg ccaactcaact 2700
gactcgacca ctatcctggg catcggcaca gtccctggacc aagcggagac ggctggagcg 2760
cgactcgtcg tgctcgcacc cgctacgcct ccgggatcgg tcaccgtgcc acatccaaac 2820
atcgaggagg tggctctgtc cagcactgga gaaatcccct tttatggcaa agccatcccc 2880
atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
gagctcggcg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaattg 3060
acgggcttta ccggtgactt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
gtcgacttca gcctggacct gaccttacc attgagacga cgaccgtgcc acaagacgcg 3180
gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgtat 3300
gacgcgggct gtgcttggtg cgagctcacg cccgccgaga cctcagttag gttgcgggct 3360
tacctaaaca caccagggtt gcccgctctg caggaccatc tggagttctg ggagagcgtc 3420
tttacaggcc tcaccacat agacgcccct ttcttgtccc agactaagca ggcaggagac 3480
aacttcccct acctggtagc ataccaggct acggtgtgcg ccagggtca ggctccacct 3540
ccatcgtggg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacgggcca 3600

```

acgcccctgc	tgtataggct	gggagccggt	caaaacgagg	ttactaccac	acaccccata	3660
accaaataca	tcatggcatg	catgtcggct	gacctggagg	togtcacgag	cacctgggtg	3720
ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
attgtgggca	ggatcatcct	gtccggaaag	ccggccatca	ttcccagacag	ggaagtcctt	3840
taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcccttacat	cgaacaggga	3900
atgcagctcg	ccgaacaatt	caaacagaag	gcaatcgggt	tgctgcaaac	agccaccaag	3960
caagcggagg	ctgctgctcc	cgtgggtgga	tccaagtggc	ggaccatcga	agccttctgg	4020
gcgaagcata	tgtggaatth	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgcctc	4140
accacccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
cctcccagecg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tgttggcagc	4260
ataggccttg	ggaaggtgct	tgtggatatt	ttggcaggtt	atggagcagg	ggtggcaggc	4320
gcgctcgtgg	cctttaaggt	catgagcggc	gagatgccct	ccaccgagga	cctggttaac	4380
ctactccctg	ctatcctctc	ccctggcgcc	ctagtcgtcg	gggtcgtgtg	cgcagcgata	4440
ctgcgtcggc	acgtgggccc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
ttcgcttcgc	ggggtaacca	cgtctcccc	acgcactatg	tgctgagag	cgacgctgca	4560
gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggcttcac	4620
cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggctaag	agatgtttgg	4680
gattggatat	gcacgggtgtt	gactgatttc	aagacctggc	tccagtccaa	gctcctgccg	4740
cgattgccgg	gagtcccctt	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
gacggcatca	tgcaaaccac	ctgcccattg	ggggcacaga	tcaccggaca	tgtgaaaaac	4860
ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
attaacgcgt	acaccacggg	cccctgcacg	ccctccccgg	cgccaaatta	ttctagggcg	4980
ctgtggcggg	tggctgctga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
gtgacgggca	tgaccactga	cgacgtaaag	tgcccgtgtc	aggttccggc	ccccgaattc	5100
ttcacagaag	tggatggggg	gcggttgca	aggtacgctc	cagcgtgcaa	accctccta	5160
cgggaggagg	tcacattcct	ggtcgggctc	aatcaatacc	tgggtgggtc	acagctccca	5220
tgogagcccg	aaccggatgt	agcagtgtct	acttccatgc	tcaccgacc	ctcccacatt	5280
acggcggaga	cggctaagcg	taggctggcc	aggggatctc	ctccccctt	ggccagctca	5340
tcagctagcc	agctgtctgc	gccttccttg	aaggcaacat	gcactacccg	tcatgactcc	5400
ccggacgctg	acctcatcga	ggccaacctc	ctgtggcggc	aggagatggg	cgggaacatc	5460
acccgcgtgg	agtcagaaaa	taaggtagta	atthtggact	ctttcgagcc	gctccaagcg	5520
gaggaggatg	agagggaaag	atccgttccg	gcggagatcc	tgccggaggtc	caggaaattc	5580
cctcgagcga	tgcccatatg	ggcacgcccg	gattacaacc	ctccactgtt	agagtctctg	5640
aaggacccgg	actacgtccc	tccagtggta	cacgggtgtc	cattgccgcc	tgccaaggcc	5700
cctccgatac	caccttcacg	gaggaagagg	acggttgctc	tgtcagaatc	taccgtgtct	5760
tctgccttgg	cggagctcgc	cacagagacc	ttcggcagct	ccgaatcgtc	ggccgtcgac	5820
agcggcacgg	caacggcctc	tctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
gttgagtcgt	actcctccat	gccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
gacgggtctt	ggtctaccgt	aagcgaggag	gctagtgagg	acgtcgtctg	ctgctcgatg	6000
tcctacacat	ggacaggcgc	cctgatcacg	ccatgcgctg	cggaggaaac	caagctgccc	6060
atcaatgcac	tgagcaactc	tttgctccgt	caccacaact	tgggtctatgc	tacaacatct	6120
cgcagcgcaa	gcctgcggca	gaagaaggtc	acctttgaca	gactgcaggt	cctggacgac	6180
cactaccggg	acgtgctcaa	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
ctatccgtgg	aggaagcctg	taagctgacg	ccccacatt	cggccagatc	taaatttggc	6300
tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgctccgtg	6360
tggaaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
gaggttttct	gcgtccaacc	agagaagggg	ggccgcaagc	cagctcgcct	tatcgtatct	6480
ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggg	ctccaccctc	6540
cctcaggccg	tgatgggctc	ttcatacggg	ttccaatact	ctcctggaca	gcgggtcgag	6600
ttcctgggtg	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacacccgc	6660
tgttttgact	caacgggtcac	tgagaatgac	atccgtggtg	aggagtcaat	ctaccaatgt	6720
tgtgacttgg	ccccgaagc	cagacaggcc	ataaggctcg	tcacagagcg	gctttacatc	6780
gggggcccc	tgactaatte	taaagggcag	aactgcggct	atcgccgggtg	ccgcgcgagc	6840
ggtgtactga	cgaccagctg	cggtaatacc	ctcacatggt	acttgaaggc	cgctgcggcc	6900
tgtcgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
atctgtgaaa	gcgcggggac	ccaagaggac	gagggcagcc	tacgggcctt	cacggaggct	7020
atgactagat	actctgcccc	ccctggggac	ccgcccacac	cagaatacga	cttggagttg	7080
ataacatcat	gctcctccaa	tgtgtcagtc	gcgcacgatg	catctggcaa	aagggtgtac	7140
tatctcacc	gtgacccac	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
actccagctc	attcctggct	aggcaacatc	atcatgtatg	cgcccacctt	gtgggcaagg	7260
atgatcctga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
ctagattgtc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
attcaacgac	tccatggcct	tagcgcattt	tactccata	gttactctcc	aggtgagatc	7440
aataggggtg	cttcatgcct	caggaaactt	ggggtaaccg	ccttgogagt	ctggagacat	7500
cgggccagaa	gtgtccgcgc	taggctactg	tcccaggggg	ggagggctgc	cacttgtggc	7560
aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggctgcg	7620

tcccagttgg	atztatccag	ctggttcgtt	gctggttaca	gcgggggaga	catatatcac	7680
agcctgtctc	gtgcccgacc	ccgctggttc	atgtgggtgcc	tactcctact	ttctgtaggg	7740
gtaggcatct	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
ccatcctggt	ttttccctt	tttttttttc	tttttttttt	tttttttttt	tttttttttt	7860
ttttctcctt	ttttttcct	ctttttttcc	ttttctttcc	tttgggtggct	ccatcttagc	7920
cctagtcacg	gctagctgtg	aaaggteccg	gagccgcttg	actgcagaga	gtgctgatac	7980
tggcctctct	gcagatcaag	t				8001

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

<400> 9

```

gccagcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctct cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tcccttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggccc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacctgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaaagcgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgcgccg 900
ccgaactgtt cggcaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctt cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcacc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttcctt ctagecggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtctt tcccctctc gccaaaggaa tgcaaggctt gttgaatgtc 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgtt tcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcggtag agtttacctg 1920
ttgcccgcga ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggctc 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccc ggtaccctg gccctctat ggcaatgagg gcttgggggtg ggcaggatgg 2100
ctcctgtcac cccgtggctc tcggcctagt tggggcccca cggacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggatc cgataacctc acgtgcggct tcgccgatct catggggtac 2220
attccgctcg tcggcgcccc cctagggggc gctgccaggg ccctggcgca tggcgtccgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc cttttctatc 2340
ttccttttgg ctttgcgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcgacatga tcatgcatac cccgggtgc gtgccctgcg ttcgggagaa caactcctcc 2520
cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtoga tttgctcggt ggggcggctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttctc gtcgcccagc tgttcacctt ctgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagc tgtcgtggat atgggtggcg gggccattg gggagtcta 2880
gcgggccttg cctactattc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttgccg gcgttgacgg gggaaacctat gtgacagggg ggacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt tcaccgggg tcatcccaga aatccagct tgtaaacc 3060
aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactcctt caactctgg 3120
ttccttgctg cgctgtteta cgtgcacaag ttcaactcat ctggatgccc agagcgcag 3180
gccagctgca gcccacatga cgcgttcgct caggggtggg ggccatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcaactac cccccggcc gtgcggtatc 3300
gtaccccgcg cgcagggtgt tggccagtg tactgcttca cccaagccc tgtcgtggtg 3360
gggacgaccg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgcg gccgcgcaa ggcaactggt ttggctgtac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc ccccgtgta acatcggggg gatcggcaat 3540
aaaacctga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600

```

tgtggttcgg	ggcettgggt	gacacccaga	tgcttggtec	actaccata	caggctttgg	3660
cactaccct	gcactgtcaa	ctttaccatc	ttcaaggta	ggatgtacgt	ggggggagt	3720
gagcacaggc	tcgaagccgc	atgcaattgg	actcgaggag	agcgttgtaa	cctggaggac	3780
agggacagat	cagagcttag	cccgtgctg	ctgtctacia	cggagtggca	ggtattgcc	3840
tgttccttca	ccaccctacc	ggctctgtcc	actggtttga	tccatctcca	tcagaacgtc	3900
gtggacgtac	aatacctgta	cggtataggg	tcggcggttg	tctcctttgc	aatcaaatgg	3960
gagtatgtcc	tgttgctctt	ccttcttctg	gcggacgcgc	gcgtctgtgc	ctgcttgtgg	4020
atgatgctgc	tgatagctca	agctgaggcc	gccctagaga	acctgggtgg	cctcaacgcg	4080
gcacccgtgg	ccggggcgca	tgccattctc	tccttctctg	tgttcttctg	tgctgcctgg	4140
tacatcaagg	gcaggctggg	ccctggggcg	gcataatgcc	tctacggcgt	atggccgcta	4200
ctcctgctcc	tgctggcggt	accaccacga	gcatacgcga	tggaccggga	gatggcagca	4260
tcgtgcggag	gcgcgggttt	cgtaggtctg	atactcttga	ccttgtcacc	gcactataag	4320
ctgttctctg	ctaggctcat	atgggtggta	caatatttta	tcaccagggc	cgaggcacac	4380
ttgcaagtgt	ggatcccccc	cctcaacggt	cgggggggcc	gcgatgccgt	catcctcctc	4440
acgtgcgcga	tcacccaga	gctaactctt	accatcacca	aatcttgc	cgccatactc	4500
ggtccactca	tggtgctcca	ggctggata	accaaagtgc	cgtacttctg	gcgcgcacac	4560
gggctcattc	gtgcatgcat	gctgggtcgg	aaggttgc	ggggtcatta	tgtccaaatg	4620
gctctcatga	agttggcgc	actgacaggt	acgtacgttt	atgaccatct	cacccactg	4680
cgggactggg	cccacgcggg	cctacgagac	cttgcgggtg	cagttgagcc	cgctccttc	4740
tctgatatgg	agaccaaggt	tatcacctgg	ggggcagaca	ccgcggcgtg	tggggacatc	4800
atcttggggc	tgcccgctc	cgcccgcagg	gggagggaga	tacatctggg	accggcagac	4860
agccttgaag	ggcaggggtg	gcgactcctc	gcgcctatta	cgccctactc	ccaacagacg	4920
cgaggcctac	ttggctgcat	catcactagc	ctcacaggcc	gggacaggaa	ccaggctcag	4980
ggggaggctc	aagtggctc	caccgcaaca	caatctttcc	tggcgacctg	cgtcaatggc	5040
gtgtgttggg	ctgtctatca	tggtgcccgc	tcaaagacc	ttgccggccc	aaagggccca	5100
atcacccaaa	tgtacaccaa	tgtggaccag	gacctcgtcg	gctggcaagc	gccccccggg	5160
gcgcgttctc	tgacaccatg	cacctgcggc	agctcggacc	tttacttggg	cacgaggcat	5220
gccgatgtca	ttccgggtgc	ccggcggggc	gacagcaggg	ggagcctact	ctccccagg	5280
cccgtctcct	acttgaagg	ctcttcgggc	ggctccactgc	tctgcccctc	ggggcacgct	5340
gtgggcatct	ttcgggctgc	cgtgtgcacc	cgaggggttg	cgaaggcggg	ggactttgta	5400
cccgtcagat	ctatggaaac	cactatgcgg	tccccggctc	tcacggacaa	ctcgtcccct	5460
ccggccgtac	cgcagacatt	ccagggtggc	catctacacg	cccctactgg	tagcggcaag	5520
agcactaagg	tgccggctgc	gtatgcagcc	caagggtata	aggtgcttgt	cctgaacccg	5580
tccgtcggcc	ccaccctagg	tttcggggcg	tatatgtcta	aggcacatgg	tatcgaccct	5640
aacatcagaa	tcggggtaag	gaccatcacc	acgggtgccc	ccatcacgta	ctccacctat	5700
ggcaagtttc	ttgccgacgg	tggttgcctc	gggggcgcct	atgacatcat	aatatgtgat	5760
gagtgccact	caactgactc	gaccactatc	ctgggcatcg	gcacagtcct	ggaccaagcg	5820
gagacggctg	gagcgcgact	cgctcgtgctc	gccaccgcta	cgctccggg	atcggtcacc	5880
gtgccacatc	caaacatcga	ggagggtggc	ctgtccagca	ctggagaaat	ccccttttat	5940
ggcaaagcca	tccccatcga	gaccatcaag	ggggggaggc	acctcatttt	ctgccattcc	6000
aagaagaaat	gtgatgagct	cgccgcgaag	ctgtccggcc	tcggactcaa	tgctgtagca	6060
tattaccggg	gccttgatgt	atccgtcata	ccaactagcg	gagacgtcat	tgctgtagca	6120
acggacgctc	taatgacggg	ctttaccggt	gacttcgact	cagtgatcga	ctgcaataca	6180
tgtgtcacc	agacagtcga	cttcagcctg	gacccgacct	tcaccattga	gacgacgacc	6240
gtgccacaag	acgcgggtgc	acgctcgcag	cgccgaggca	ggactggtag	gggcaggatg	6300
ggcatttaca	ggtttgtgac	tccaggagaa	cgccctcgg	gcattgtcga	ttcctcgggt	6360
ctgtgcgagt	gctatgacgc	gggtgtgct	tggtacgagc	tcacgcccgc	cgagacctca	6420
gtaggttgc	gggcttacct	aaacacacca	gggttgcccg	tctgccagga	ccatctggag	6480
ttctgggaga	gcgtctttac	aggcctcacc	cacatagacg	cccatttctt	gtcccagact	6540
aagcaggcag	gagacaactt	cccctacctg	gtagcatacc	aggctacggg	gtgcgccagg	6600
gctcaggctc	cacctccatc	gtgggaccaa	atgtggaagt	gtctcatacg	gctaaagcct	6660
acgctgcacg	ggccaacgcc	cctgctgtat	aggctgggag	ccgttcaaaa	cgaggttact	6720
accacacacc	ccataaccaa	atacatcatg	gcattgcatt	cggtgacct	ggaggctcgtc	6780
acgagcacct	gggtgctggg	aggcggagtc	ctagcagctc	tggccgcgta	ttgcctgaca	6840
acaggcagcg	tggtcattgt	gggcaggatc	atcttgtccg	gaaagccggc	catcattccc	6900
gacagggaag	tcctttaccg	ggagttcgat	gagatggaag	agtgcgcctc	acacctcct	6960
tacatcgaac	agggaatgca	gctcgcgcaa	caattcaaac	agaaggcaat	cggtttgctg	7020
caaacagcca	ccaagcaagc	ggaggctgct	gctcccgtgg	tggaatccaa	gtggcggacc	7080
atcgaagcct	tctgggcgaa	gcataatgtg	aatttcatca	gcgggataca	atatttagca	7140
ggcttgtcca	ctctgcctgg	caaccccgcg	atagcatcac	tgatggcatt	cacagcctct	7200
atcaccagcc	cgctcaccac	ccaacatacc	ctcctgttta	acatcctggg	gggatgggtg	7260
gccgcccaac	ttgctcctcc	cagcgtgct	tctgctttcg	taggcgccgg	catcgtgga	7320
gcggctgttg	gcagcatagg	ccttgggaag	gtgcttgtgg	atattttggc	aggttatgga	7380
gcaggggtgg	caggcgcgct	cgtggccttt	aaggtcatga	gcggcgagat	gccctccacc	7440
gaggacctgg	ttaacctact	ccctgctatc	ctctcccctg	gcgcctagt	cgctgggggtc	7500
gtgtgcgcag	cgatactgcg	tcggcacgtg	ggcccagggg	aggggctgt	gcagtggatg	7560
aaccggctga	tagcgttcgc	ttcgcggggg	aaccacgtct	ccccacgca	ctatgtgcct	7620

gagagcgacg	ctgcagcacg	tgtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
ctgaagaggc	ttcaccagtg	gatcaacgag	gactgctcca	cgccatgctc	cggctcgtgg	7740
ctaagagatg	tttgggattg	gatatgcacg	gtggtgactg	atttcaagac	ctggctccag	7800
tccaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggggc	acagatcacc	7920
ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
catggaacat	tccccattaa	cgcgtaacac	acgggcccct	gcacgccctc	cccggcgcca	8040
aattattcta	gggcgctgtg	gcgggtggct	gctgaggagt	acgtggagg	tacgcgggtg	8100
ggggatttcc	actacgtgac	gggcatgacc	actgacgacg	taaagtgcc	gtgtcagggt	8160
ccggcccccg	aattcttcac	agaagtggat	ggggtgcggt	tgcacaggta	cgctccagcg	8220
tgcaaacc	tcctacggga	ggaggtcaca	ttcctggtcg	ggctcaatca	atacctgggt	8280
gggtcacagc	tcccatgcga	gcccgaaccg	gatgtagcag	tgctcacttc	catgctcacc	8340
gaccctccc	acattacggc	ggagacggct	aagcgtaggc	tggccagggg	atctcctccc	8400
cccttgcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcaact	8460
accgctcatg	actccccgga	cgctgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
atgggcggga	acatcacccg	cgtggagtca	gaaaataagg	tagtaatttt	ggactctttc	8580
gagccgctcc	aagcggagga	ggatgagagg	gaagtatccg	ttccggcgga	gatcctgcgg	8640
aggtccagga	aattccctcg	agcgatgcc	atatgggcac	gcccggatta	caaccctcca	8700
ctgttagagt	cctggaagga	cccggactac	gtccctccag	tggtagacag	gtgtccattg	8760
ccgcctgcca	aggcccctcc	gataccacct	tcacggagga	agaggacgg	tgtcctgtca	8820
gaatctaccg	tgtcttctgc	cttggcggag	ctcgccacag	agaccttcgg	cagctccgaa	8880
tcgtcggccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
gacgcgggat	ccgacgttga	gtcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
gatcccgatc	tcagcgacgg	gtcttggtct	accgtaagcg	aggaggctag	tgaggacgtc	9060
gtctgctgct	cgatgtccta	cacatggaca	ggcgccctga	tcacgccatg	cgctgcggag	9120
gaaaccaagc	tgccatcaa	tgcactgagc	aactctttgc	tccgtcacca	caacttggtc	9180
tatgctacaa	catctcgcag	cgcaagcctg	cggcagaaga	aggtcacctt	tgacagactg	9240
caggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaagc	tgacgcccc	acattcggcc	9360
agatctaaat	ttggctatgg	ggcaaaggac	gtccggaacc	tatccagcaa	ggccgttaac	9420
cacatccgct	ccgtgtggaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccacc	9480
atcatggcaa	aaaatgaggt	tttctgcgtc	caaccagaga	aggggggccc	caagccagct	9540
cgcttatcgc	tattcccaga	tttggggggt	cgtgtgtgcg	agaaaatggc	cctttacgat	9600
gtggtctcca	ccctccctca	ggccgtgatg	ggctcttcat	acggattcca	atactctcct	9660
ggacagcggg	tcgagttcct	ggtgaatgcc	tggaaagcga	agaaatgcc	tatgggcttc	9720
gcataatgaca	cccgtgtgtt	tgactcaacg	gtcactgaga	atgacatccg	tgttgaggag	9780
tcaatctacc	aatggttga	cttggcccc	gaagccagac	aggccataag	gtcgtcaca	9840
gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
cgggtgccgcg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttacttg	9960
aaggccgctg	cgccctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgtatgcgga	10020
gacgaccttg	tcgttatctg	tgaagcgcg	gggacccaag	aggacgaggc	gagcctacgg	10080
gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccgcc	caaaccagaa	10140
tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcatct	10200
ggcaaaaggg	tgtactatct	caccctgtgac	cccaccacc	cccttgcgcg	ggctgcgtgg	10260
gagacagcta	gacacactcc	agtcaattcc	tggctaggca	acatcatcat	gtatgcgccc	10320
accttgtggg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
gacctacctc	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagttac	10500
tctccagggtg	agatcaatag	ggtggcttca	tgcctcagga	aacttgggg	accgcccttg	10560
cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtccca	gggggggagg	10620
gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
ccaatcccgg	ctgcgtccca	gttggattta	tccagctggt	tcggtgctgg	ttacagcggg	10740
ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgtc	ggttcatgtg	gtgcctactc	10800
ctactttctg	taggggtagg	catctatcta	ctcccccaacc	gatgaacggg	gagctaaaca	10860
ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
tttttttttt	tttttttttc	tccttttttt	ttcctctttt	tttctttttc	tttcttttgg	10980
tggctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
agagagtgtc	gatactggcc	tctctgcaga	tcaagt			11076

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

<400> 10

```

gccagcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctct cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaac 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggccc cgggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcc cagctgtgct cgacctgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagega aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgctgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 900
ccgaactggt cggcaggctc aaggcgcgca tgcccgcagc cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccc aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctt cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcacc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctaggggtct tcccctctc gccaaaggaa tgcaaggctc gttgaatgct 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaacce cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgtg ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
tgcacatca ctagcctcac aggcggggac aggaaccagg tcgaggggga ggtccaagtg 1920
gtctccaccg caacacaatc tttcctggcg acctgcgtca atggcgtgtg ttggactgtc 1980
tatcatggtg ccggtcaaaa gacccttgcc ggcccaaaag gcccaatcac ccaaatgtac 2040
accaatgtgg accaggacct cgtcggctgg caagcgcgcc ccggggcgcg ttccttgaca 2100
ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcattgccg tgctattccg 2160
gtgcgcgggc ggggcgacag cagggggagc ctactctccc ccaggcccgt ctctacttg 2220
aagggtctt cgggcggctc actgctctgc ccctcggggc acgctgtggg catctttcgg 2280
gctgccgtgt gcaccgagg ggttgcgaag gcggtggact ttgtaccctg cgagtctatg 2340
ggaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
acattccagg tggcccatct acacgccctt actggtagcg gcaagagcac taagggtgcc 2460
gctgcgtatg caggccaagg gtataagggt cttgtcctga acccgtccgt cgccgccacc 2520
ctaggtttcg gggcgtatat gtctaaggca catggtatcg accctaacad cagaatcggg 2580
gtaaggacca tcaccacggg tgccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacggtggtt gctctggggg cgcctatgac atcataatat gtgatgagtg ccaactcaact 2700
gactcgacca ctatcctggg catcggcaca gtccctggacc aagcggagac ggctggagcg 2760
cgactcgtcg tgctcggcac cgctacgct ccgggatcgg tcaccgtgcc acatccaaac 2820
atcgaggagg tggctctgtc cagcactgga gaaatcccct tttatggcaa agccatcccc 2880
atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
gagctcgccg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaata 3060
acgggcttta ccggcgactt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
gtcgacttca gcctggacct gacctcacc attgagacga cgaccgtgcc acaagacgcg 3180
gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgtat 3300
gacgcgggct gtgcttggtg cgagctcacg cccgcggaga cctcagttag gttgcgggct 3360
tacctaaca caccagggtt gcccgctctg caggaccatc tggagttctg ggagagcgtc 3420
tttacaggcc tcaccacat agacgcccat ttcttgtccc agactaagca ggcaggagac 3480
aacttccctt acctggtagc ataccaggct acggtgtgcg ccagggtca ggctccacct 3540
ccatcgtggg accaaatgtg gaagtgtctc ataccgctaa agcctacgct gcacgggcca 3600

```

acgccctgc	tgtatagget	gggagccgtt	caaaacgagg	ttactaccac	acaccccata	3660
accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
attgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccagacag	ggaagtcctt	3840
taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tccttacat	cgaacaggga	3900
atgcagctcg	ccgaacaatt	caaacagaag	gcaatcgggt	tgctgcaaac	agccaccaag	3960
caagcggagg	ctgctgctcc	cgtgggtgaa	tccaagtggc	ggaccatcga	agccttctgg	4020
gcgaagcata	tgtggaattt	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgtc	4140
accacccaac	atacctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
cctcccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tgttggcagc	4260
ataggccttg	ggaaggtgct	tgtggatatt	ttggcagggt	atggagcagg	ggtggcaggc	4320
gcgctcgtgg	cctttaaggt	catgagcggc	gagttgcct	ccaccgagga	cctggttaac	4380
ctactccctg	ctatcctctc	ccctggcgcc	ctagtcgtcg	gggtcgtgtg	cgcagcgata	4440
ctgcgtcggc	acgtgggccc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
ttcgcttcgc	ggggtaacca	cgtctcccc	acgcactatg	tgccctgagag	cgacgctgca	4560
gcacgagtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggcttcac	4620
cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggctaag	agatgtttgg	4680
gattggatat	gcacgggtgtt	gactgatttc	aagacctggc	tccagtccaa	gctcctgccg	4740
cgattgccgg	gagtccectt	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
gacggcatca	tgcaaaccac	ctgcccattg	ggagcacaga	tcaccggaca	tgtgaaaaac	4860
ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
attaacgcgt	acaccacggg	cccctgcacg	ccctccccgg	cgccaaatta	ttctagggcg	4980
ctgtggcggg	taggtgctga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
gtgacgggca	tgaccactga	caacgtaaag	tgccogtgc	aggttcgggc	ccccgaattc	5100
ttcacagaag	tggatggggg	gcggttgca	aggtacgctc	cagcgtgcaa	acccctccta	5160
cgggaggagg	tcacattcct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctcca	5220
tgcgagcctg	aaccggatgt	agcagtgctc	acttccatgc	tcaccgacce	ctcccacatt	5280
acggcggaga	cggctaagcg	taggctggcc	aggggatctc	ccccccctt	ggccagctca	5340
tcagctagcc	agctgtctgc	gccttccttg	aaggcaacat	gcactacccg	tcatgactcc	5400
ccggacgctg	acctcatcga	ggccaacctc	ctgtggcggc	aggagatggg	cgggaacatc	5460
acccgcgtgg	agtcagaaaa	taaggtagta	atthtggact	ctttcgagcc	gctccaagcg	5520
gaggaggatg	agaggggagt	atccgttccg	gcggagatcc	tgccggaggtc	caggaaattc	5580
cctcgagcga	tgcccatatg	ggcacgcccg	gattacaacc	ctccactggt	agagtcctgg	5640
aaggacccgg	actacgtccc	tccagtggta	cacgggtgtc	cattgccgcc	tgccaaggcc	5700
cctccgatac	caccttcacg	gaggaagagg	acggttgtcc	tgtcagaate	taccgtgtct	5760
tctgccttgg	cggagctcgc	cacagagacc	ttcggcagct	ccgaatcgtc	ggccgtcgac	5820
agcggcacgg	caacggcctc	tctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
gttgagtcgt	actcctccat	gccccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
gacgggtctt	ggtctaccgt	aagcgaggag	gctagtgagg	acgtcgtctg	ctgctcgatg	6000
tcctacacat	ggacaggcgc	cctgatcacg	ccatgcgctg	cggaggaaac	caagctgccc	6060
atcaatgcac	tgagcaactc	tttgctccgt	caccacaact	tggtctatgc	tacaacatct	6120
cgcagcgcga	acctgcggca	gaagaaggtc	acctttgaca	gactgcagggt	cctggacgac	6180
cactaccggg	acgtgctcaa	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
ctatccgtgg	aggaagcctg	taagctgacg	ccccacatt	cggccagatc	taaatttggc	6300
tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgctccgtg	6360
tgaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
gaggttttct	gcgtccaacc	agagaagggg	ggccgcaage	cagctcgcct	tatcgtattc	6480
ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggg	ctccaccctc	6540
cctcaggccg	tgatgggctc	ttcatacgga	ttccaatact	ctcctggaca	gcgggtcgag	6600
ttcctgggtga	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacacccgc	6660
tgttttgact	caacggtcac	tgagaatgac	atccgtggtg	aggagtcaat	ctaccaatgt	6720
tgtgacttgg	ccccgaagc	cagacaggcc	ataaggctgc	tcacagagcg	gctttacatc	6780
gggggcccc	tgactaatte	taaagggcag	aactgcggct	atcgccgggtg	ccgcgcgagc	6840
ggtgtactga	cgaccagctg	cggtaatacc	ctcacatggt	acttgaaggc	cgctgcggcc	6900
tgctcgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
atctgtgaaa	gcgcggggac	ccaagaggac	gagggcagcc	tacgggcctt	cacggaggct	7020
atgactagat	actctgcccc	ccctggggac	ccgcccaaac	cagaatacga	cttggagttg	7080
ataacatcat	gctcctccaa	tgtgtcagtc	gcgcacgatg	catctggcaa	aagggtgtac	7140
tatctcacc	gtgacccac	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
actccagtca	attcctggct	aggcaacatc	atcatgtatg	cgcccacctt	gtgggcaagg	7260
atgatcctga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
ctagattgtc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
attcaacgac	tccatggcct	tagcgcattt	tactccata	gttactctcc	aggtgagatc	7440
aatagggtgg	cttcatgcct	caggaaactt	ggggtaccgc	ccttgcgagt	ctggagacat	7500
cgggccagaa	gtgtccgcgc	taggctactg	tcccaggggg	ggagggtgc	cacttgtggc	7560
aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tactccaat	cccggctgcg	7620

tcccagttgg	atztatccag	ctggttcggt	gctggttaca	gcgggggaga	catatatcac	7680
agcctgtctc	gtgcccgacc	ccgctgggtc	atgtgggtgcc	tactcctact	ttctgtaggg	7740
gtaggcatct	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
ccatcctggt	ttttccctt	tttttttttc	tttttttttt	tttttttttt	tttttttttt	7860
ttttctcctt	tttttttct	ctttttttcc	ttttctttcc	tttgggtggct	ccatcttagc	7920
cctagtcacg	gctagctgtg	aaaggtccgt	gagccgcttg	actgcagaga	gtgctgatac	7980
tggcctctct	gcagatcaag	t				8001

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

<400> 11

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctcg cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgtg gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg cgggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 900
ccgaactggt cgccaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgacct 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctt cgtgctttac ggtatcgccg 1140
ctcccgatc gcagcgcac gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtctt tccccctctc gccaaaggaa tgcaaggtct gttgaatgct 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc acctggcgac aggtgectct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaagge ggcacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgtg ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggatc tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggcccccgga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcgggtg agtttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggctg 1980
caacctcgtg gaaggcgaca acctatcccc aaggetcgcc agcccagagg tagggcctgg 2040
gctcagcccg ggtacccttg gccctctat ggcaatgagg gcttgggggtg ggcaggatgg 2100
ctcctgtcac cccgtggctc tcggcctagt tggggcccca cggacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggtcat cgataccctc acgtgcggtc tcgccgatct catggggtac 2220
attccgctcg tcggcgcccc cctagggggc gctgccaggg ccctggcgca tggcgtccgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc cttttctatc 2340
ttccttttgg ctttgctgtc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcggacatga tcatgcatac ccccggtgct gtgccctgcg ttcgggagaa caactcctcc 2520
cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga tttgctcgtt ggggcggctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgoggatc tgttttctc gtgcgccagc tgttcacctt ctgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatac ccggccacgt gacaggtcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagc tgtcgtggat atggtggcgg gggcccattg gggagtcta 2880
gcgggccttg cctactattc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttgccg gcgttgacgg gggaaacctat gtgacagggg ggacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt ttcaccggg tcatcccaga aaatccagct tgtaaacacc 3060
aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caaactggg 3120
ttccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgcc agagcgcag 3180
gccagctgca gcccacgca cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcactacg cccccggcc gtgcggtatc 3300
gtaccccgcg cgcaggtgtg tggccagtg tactgcttca cccaagccc tgcgtgggtg 3360
gggacgaccg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgcg gccgcccga ggcaactggt ttggctgtac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc cccccgtgta acatcggggg gatcggcaat 3540
aaaaccttga cctgcccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600

```

tgtggttcgg	ggccttggtt	gacaccacaga	tgcttggtcc	actaccacata	caggctttgg	3660
cactaccctt	gcactgtcaa	ctttaccatc	ttcaagggtta	ggatgtacgt	ggggggagtg	3720
gagcacaggc	tcgaagccgc	atgcaattgg	actcgaggag	agcgttgtaa	cctggaggac	3780
agggacagat	cagagcttag	cccgtctgtg	ctgtctacaa	cggagtggca	ggtattgccc	3840
tgttccttca	ccaccctacc	ggctctgtcc	actggtttga	tccatctcca	tcagaacgtc	3900
gtggacgtac	aatacctgta	cggtataggg	tcggcgggtg	tctcctttgc	aatcaaatgg	3960
gagtatgtcc	tgttgctctt	ccttcttctg	gcgacgcgc	gcgtctgtgc	ctgcttgtgg	4020
atgatgctgc	tgatagctca	agctgaggcc	gccctagaga	acctgggtgg	cctcaacgcg	4080
gcacccgtgg	ccggggcgca	tggcattctc	tccttcctcg	tgttcttctg	tgctgcctgg	4140
tacatcaagg	gcaggctggg	ccctggggcg	gcataatgcc	tctacggcgt	atggccgcta	4200
ctcctgctcc	tgctggcggt	accaccacga	gcatacggca	tggaaccgga	gatggcagca	4260
tcgtgaggag	gcgcgggttt	cgtaggctctg	atactcttga	ccttgctacc	gcactataag	4320
ctgttcctcg	ctaggctcat	atgggtggta	caatatatta	tcaccagggc	cgaggcacac	4380
ttgcaagtgt	ggatccccc	cctcaacggt	cggggggggc	gcgatgccgt	catcctctc	4440
acgtgcccga	tcacccacga	gctaactctt	accatcacca	aaatcttgc	cgccatactc	4500
ggccactca	tggtgctcca	ggctgggtata	accaaagtgc	cgtacttctg	gcgcgcacac	4560
gggctcattc	gtgcatgcat	gctgggtgcg	aaggttgcgt	ggggtcatta	tgtccaaatg	4620
gctctcatga	agttggccgc	actgacaggt	acgtacggtt	atgacctct	cacccactg	4680
cgggactggg	cccacgcggg	cctacgagac	cttgccgttg	cagttgagcc	cgctcgtctt	4740
tctgatatgg	agaccaaggt	tatcacctgg	ggggcagaca	ccgcggcgtg	tggggacatc	4800
atcttgggcc	tgcccgtctc	cgcccgcagg	gggagggaga	tacatctggg	accggcagac	4860
agccttgaag	ggcaggggtg	gcgactcctc	gcgcctatta	cgccctactc	ccaacagacg	4920
cgaggcctac	ttggctgcat	catcactagc	ctcacaggcc	gggacaggaa	ccaggctcgag	4980
ggggaggctc	aagtggctctc	caccgcaaca	caatctttcc	tgccgacctg	cgtcaatggc	5040
gtgtgttggg	ctgtctatca	tggtgccggc	tcaaagacc	ttgccggccc	aaagggccca	5100
atcacccaaa	tgtacaccaa	tgtggaccag	gacctcgtcg	gctggcaagc	gcccccggg	5160
gcgcgttctt	tgacaccatg	cacctgcggc	agctcggacc	tttacttgg	cacgaggcat	5220
gccgatgtca	ttccgggtgcg	ccggcggggc	gacagcaggg	ggagcctact	ctccccag	5280
cccgtctcct	acttgaagg	ctcttcgggc	ggtccactgc	tctgccctc	ggggcacgct	5340
gtgggcatct	ttcgggctgc	cgtgtgcacc	cgaggggttg	cgaaggcgg	ggactttgta	5400
cccgtcaggt	ctatgggaac	cactatgcgg	tccccggtct	tcacggacaa	ctcgtcccct	5460
ccggccgtac	cgacagacatt	ccagggtggc	catctacacg	cccctactgg	tagcggcaag	5520
agcactaagg	tgccggctgc	gtatgcaggc	caagggata	aggtgcttgt	cctgaaccg	5580
tccgtcgcgc	ccaccctagg	tttcggggcg	tatatgtcta	aggcacatgg	tatcgacct	5640
aacatcagaa	tcggggtaag	gaccatcacc	acgggtgccc	ccatcacgta	ctccacctat	5700
ggcaagtttc	ttgccgacgg	tggttgcctc	gggggcgct	atgacatcat	aatatgtgat	5760
gagtgccact	caactgactc	gaccactatc	ctgggcatcg	gcacagctct	ggaccaagcg	5820
gagacggctg	gagcgcgact	cgctcgtgctc	gccaccgcta	cgccctcggg	atcggctacc	5880
gtgccacatc	caaacatcga	ggaggtggct	ctgtccagca	ctggagaaat	ccccttttat	5940
ggcaaagcca	tccccatcga	gaccatcaag	ggggggaggg	acctcatttt	ctgccattcc	6000
aagaagaaat	gtgatgagct	cgccgcgaag	ctgtccggcc	tcggactcaa	tgctgtagca	6060
tattaccggg	gccttgatgt	atccgtcata	ccaactagcg	gagacgtcat	tgctgtagca	6120
acggacgctc	taatgacggg	ctttaccggc	gacttcgact	cagtgatcga	ctgcaataca	6180
tgtgtcacc	agacagtoga	cttcagcctg	gaccgcacct	tcaccattga	gacgacgacc	6240
gtgccacaag	acgcgggtgc	acgctcgcag	cggcgaggca	ggactggtag	gggcaggatg	6300
ggcatttaca	ggtttgtgac	tccaggagaa	cggccctcgg	gcattgtcga	ttcctcgggt	6360
ctgtgcgagt	gctatgacgc	gggctgtgct	tggtacgagc	tcacgcccgc	cgagacctca	6420
gtagggttgc	gggcttacct	aaacacacca	gggttgccc	tctgccagga	ccatctggag	6480
ttctgggaga	gcgtctttac	aggcctcacc	cacatagacg	cccatttctt	gtcccagact	6540
aagcaggcag	gagacaactt	cccctacctg	gtagcatacc	aggctacggt	gtgcgccagg	6600
gctcaggctc	cacctccatc	gtgggaccaa	atgtggaagt	gtctcatacg	gctaaagcct	6660
acgctgcacg	ggccaacgcc	cctgctgtat	aggctgggag	ccgttcaaaa	cgaggttact	6720
accacacacc	ccataaccaa	atacatcatg	gcattgcatt	cggctgacct	ggaggctcgtc	6780
acgagcacct	gggtgctggg	aggcggagtc	ctagcagctc	tggccgcgta	ttgcctgaca	6840
acaggcagcg	tggtcattgt	gggcaggatc	atcttgtccg	gaaagccggc	catcattccc	6900
gacagggaag	tcctttaccg	ggagttcgat	gagatggaag	agtgcgcctc	acacctccct	6960
tacatcgaac	agggaatgca	gctcgcggaa	caattcaaac	agaaggcaat	cgggttgcgtg	7020
caaacagcca	ccaagcaagc	ggaggctgct	gctcccgtgg	tggaaatcaa	gtggcggacc	7080
atcgaagcct	tctgggcgaa	gcataatgtg	aatttcatca	gcgggataca	atatttagca	7140
ggcttgtcca	ctctgcctgg	caaccccgcg	atagcatcac	tgatggcatt	cacagcctct	7200
atcaccagcc	cgctcaccac	ccaacatacc	ctcctgttta	acatcctggg	gggatgggtg	7260
gccgcccac	ttgctcctcc	cagcgtgct	tctgctttcg	taggcgccgg	catcgtgga	7320
gcggctgttg	gcagcatagg	ccttgggaag	gtgcttgtgg	atattttggc	aggttatgga	7380
gcaggggtgg	caggcgcgct	cgtggccttt	aaggtcatga	gcggcgagtt	gccctccacc	7440
gaggacctgg	ttaacctact	ccctgctatc	ctctcccctg	gcgccttagt	cgctcggggtc	7500
gtgtgcgcag	cgatactgcg	tcggcacgtg	ggcccagggg	aggggctgt	gcagtggatg	7560
aaccggctga	tagcgttcgc	ttcgcggggg	aaccacgtct	ccccacgca	ctatgtgcct	7620

gagagcgacg	ctgcagcacg	agtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
ctgaagaggg	ttcaccagtg	gatcaacgag	gactgctcca	cgccatgctc	cggtcgtgg	7740
ctaagagatg	tttgggattg	gatatgcacg	gtggtgactg	atttcaagac	ctggctccag	7800
tccaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
catggaacat	tccccattaa	cgcgtacacc	acgggccctc	gcaogccctc	cccggcgcca	8040
aattattcta	gggcgctgtg	gcgggtaggt	gctgaggagt	acgtggaggt	tacgcgggtg	8100
ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaagtgcc	gtgtcaggtt	8160
ccggcccccg	aattcttcac	agaagtggat	ggggtgcggt	tgcacaggta	cgctccagcg	8220
tgcaaaccce	tctacggga	ggaggtcaca	ttcctggtcg	ggctcaatca	atacctggtt	8280
gggtcacagc	tcccatgcga	gcctgaaccg	gatgtagcag	tgctcacttc	catgctcacc	8340
gaccctccc	acattacggc	ggagacggct	aagcgtaggc	tggccagggg	atctcccccc	8400
cccttgcca	gctcatcagc	tagccagctg	tctgcgctt	ccttgaaggc	aacatgact	8460
accgctcatg	actccccgga	cgctgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
atgggaggga	acatcacccg	cgtggagtca	gaaaataagg	tagtaatttt	ggactcttct	8580
gagcgcctcc	aagcggagga	ggatgagagg	ggagtatccg	ttccggcgga	gatcctgcg	8640
aggtccagga	aattccctcg	agcgtatgcc	atatgggcac	gcccggatta	caaccctcca	8700
ctgttagagt	cctggaagga	cccggactac	gtccctccag	tggtagacgg	gtgtccattg	8760
ccgcctgcca	aggccccctc	gataccacct	tcacggagga	agaggacggg	tgctcctgtca	8820
gaatctaccg	tgtcttctgc	cttggcggag	ctcgccacag	agaccttcgg	cagctccgaa	8880
togteggccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
gacgcgggat	ccgacgttga	gtcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
gatcccgatc	tcagcgacgg	gtcttggtct	accgtaagcg	aggaggctag	tgaggacgtc	9060
gtctgctgct	cgatgtccta	cacatggaca	ggcgccctga	tcacgccatg	cgctgaggag	9120
gaaaccaagc	tgcccatcaa	tgcactgagc	aactctttgc	tcogtcacca	caacttggtc	9180
tatgctacaa	catctcgcag	cgcaaacctg	cggcagaaga	aggctcacctt	tgacagactg	9240
caggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaagc	tgacgcccc	acattcggcc	9360
agatctaaat	ttggctatgg	ggcaaaggac	gtccggaacc	tatccagcaa	ggccgttaac	9420
cacatccgct	ccgtgtggaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccacc	9480
atcatggcaa	aaaatgaggt	tttctgcgct	caaccagaga	aggggggccc	caagccagct	9540
cgcttatcgc	tattcccaga	tttggggggt	cgtgtgtgcg	agaaaatggc	cctttacgat	9600
gtggtctcca	ccctccctca	ggcgtgatg	ggctcttcat	acggattcca	atactctcct	9660
ggacagcggg	tcgagttcct	ggtgaatgcc	tggaaagcga	agaaatgcc	tatgggcttc	9720
gcatatgaca	cccgcctgtt	tgactcaacg	gtcactgaga	atgacatccg	tgttgaggag	9780
tcaatctacc	aatgtttgtg	cttggcccc	gaagccagac	aggccataag	gtcgcctcaca	9840
gagcggcttt	acatcggggg	cccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
cgggtgcgcg	cgagcgggtg	actgacgacc	agctgcggtg	ataccctcac	atgttacttg	9960
aaggccgctg	cggcctgtcg	agctgccaag	ctccaggact	gcacgatgct	cgtatgcgga	10020
gacgaccttg	tcgttatctg	tgaaagcgcg	gggacccaag	aggacgaggc	gagcctacgg	10080
gccttcacgg	aggctatgac	tagatactct	gccccctctg	gggacccgcc	caaaccagaa	10140
tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcatct	10200
ggcaaaaggg	tgtactatct	caccctgac	cccaccacc	cccttgccg	ggctgcgtgg	10260
gagacagcta	gacacactcc	agtcaattcc	tggttaggca	acatcatcat	gtatgcgccc	10320
accttggtgg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
gacctacctc	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagttac	10500
tctccagggtg	agatcaatag	ggtggcttca	tgcctcagga	aacttggggg	accgcccttg	10560
cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtccca	gggggggagg	10620
gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
ccaatcccgg	ctgcgtccca	gttggattta	tccagctggt	tcgttgctgg	ttacagcggg	10740
ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgct	ggttcatgtg	gtgcctactc	10800
ctactttctg	taggggtagg	catctatcta	ctcccccaacc	gatgaacggg	gagctaaaca	10860
ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
tttttttttt	tttttttttc	tccttttttt	ttctcttttt	tttctttttc	tttcttttgg	10980
tggctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
agagagtgct	gatactggcc	tctctgcaga	tcaagt			11076

**CLAIMS**

1. A hepatitis C virus (HCV) cell culture system, which  
comprises mainly eukaryotic cells containing  
5 transfected HCV specific genetic material,  
characterized in that,  
the eukaryotic cells are human hepatoma cells and the  
transfected HCV specific genetic material is a HCV-RNA  
construct, which comprises the HCV specific RNA  
10 segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and  
3' NTR as well as an additional marker gene for  
selection (selection gene).
2. A cell culture system according to Claim 1,  
15 characterized in that,  
the hepatoma cells are derived from a usual hepatoma  
cell line.
3. A cell culture system according to Claim 1,  
20 characterized in that,  
the hepatoma cells are extracted from a primary  
hepatoma cell culture.
4. A HCV-RNA construct, characterized in that,  
25 it comprises the HCV specific RNA segments 5' NTR,  
NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as an  
additional marker gene for selection (selection gene).
5. A HCV-RNA construct according to Claim 4,  
30 characterized in that,  
it comprises one of the nucleotide sequences  
illustrated in the sequence protocols SEQ ID NO: 1 to  
SEQ ID NO: 11.
- 35 6. A HCV-RNA construct according to Claim 4,  
characterized in that,

the 3' NTR has a nucleotide sequence, which has been selected from the group of nucleotide sequences (a) to (i) listed in the following:

- 5 (a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTA  
GCTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC  
CAGACTGGAGT
- 10 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTA  
GTCTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCCCA  
GACTGGAGC
- 15 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTAAT  
CTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCCCA  
GACTGCAGC
- (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTTT  
AGTCTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGC  
CCAGACTGGAGT
- 20 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTA  
GTCTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGC  
CCAGACTGGAGT
- 25 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTA  
GTCTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTG  
CCCAGACTGGAGT
- 30 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTA  
GTCTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC  
CAGACTGGAGT
- 35 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTTT  
TAATCTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGT  
TGCCCAGACTGGAGT

(i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTTT  
AATCTTTTTTTTTTTTTTTCTTTTTTTTTTTGAGAGAGAGAGTCTCACTCTG  
TTGCCCAGACTGGAGT

- 5 7. A HCV-RNA construct according to one of the Claims 4 to 6, characterized in that, the marker gene for selection is a resistance gene and preferably an antibiotic resistance gene.
- 10 8. A HCV-RNA construct according to one of the Claims 4 to 6, characterized in that, the marker gene for selection is a neomycin phosphotransferase gene.
- 15 9. A HCV-RNA construct according to one of the Claims 4 to 8, characterized in that, the marker gene for selection is integrated in the HCV-RNA downstream from the 5' NTR.
- 20 10. A HCV-RNA construct according to one of the Claims 4 to 83, characterized in that, the marker gene for selection is linked to the HCV-RNA via a ribozyme or a recognition sequence for a ribozyme.
- 25 11. A HCV-RNA construct according to one of the Claims 4 to 10, characterized in that, it has an integrated reporter gene.
- 30 12. A HCV-RNA construct according to Claim 11, characterized in that, the reporter gene is a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene),  
35 the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) and the SEAP gene (secreted alkaline phosphatase gene).

13. A HCV-RNA construct according to one of the Claims 4  
to 11, characterized in that,  
their replication has an influence on the expression  
5 of a (cellular) surrogate marker gene.
14. A HCV-RNA construct according to one of the Claims 11  
to 13, characterized in that,  
the resistance gene is cloned into the open reading  
10 frame of the HCV-RNA in such a way that it will only  
be transferred to an active form after proteolytic  
processing.
15. A HCV-RNA construct according to one of the Claims 11  
15 to 14, characterized in that,  
the reporter gene and the marker gene for selection  
are arranged in the construct in such a way, that they  
are expressed as a fusion protein.
- 20 16. A cell culture system according to one of the Claims 1  
to 3, characterized in that,  
the HCV-RNA construct is a construct according to at  
least one of the Claims 4 to 15.
- 25 17. A cell culture system according to Claim 1,  
characterized in that,  
the cells containing the HCV-RNA construct are lodged  
at the DSMZ, Braunschweig, Germany under the deposit  
number DSM ACC2394 (laboratory name HuBl 9-13).
- 30
18. A use of a cell culture system according to one of the  
Claims 1 to 3 or 16 to 17 and/or a HCV-RNA construct  
according to one of the Claims 4 to 15 for the  
production and/or evaluation and/or testing of  
35 treatments and/or diagnostics for the treatment of HCV  
infections in particular.

19. A use of a cell culture system according to one of the Claims 1 to 3 or 16 to 17 and/or a HCV-RNA construct according to one of the Claims 4 to 15 for the production of a vaccine against HCV infections.
- 5
20. A use a HCV-RNA construct according to one of the Claims 4 to 15 for the production of a liver cell specific vector in gene therapy.
- 10
21. A HCV-RNA construct according to one of the Claims 4 to 15, characterized in that, it has an integrated foreign gene and can be used to transfect this foreign gene in a target cell, which can be used to express this foreign gene.
- 15
22. A process for the production of cell culture adapted mutants of a HCV-RNA construct according to one of the claims 4 to 15, in which the mutants have increased replication efficiency compared to the HCV-RNA construct,
- 20
- characterized in that, a cell culture system according to claim 1, in which the transfected HCV specific genetic material is a HCV-RNA construct with a selection gene according to one of the claims 4 to 15, is cultivated on/in the selection medium corresponding to the selection gene, that the cell clones created are collected, and that the HCV-RNA constructs or parts thereof are isolated from these cell clones.
- 25
- 30
23. A process according to claim 22, characterized in that, the isolated HCV-RNA constructs are passaged at least one more time. That is, they are transfected in cells of a cell culture system according to claim 1. The cell culture system of claim 1 thereby obtained, in which the transfected HCV specific genetic material is
- 35

the isolated HCV-RNA construct with a selection gene, is cultivated on/in the selection medium corresponding to the selection gene, the cell clones created are collected, and the HCV-RNA constructs are isolated from these cell clones.

- 5
24. A process for the production of mutants of a HCV full-length genome or of a HCV subgenome or any HCV construct with increased replication efficiency in comparison to the original HCV full-length genome or subgenome or HCV-RNA construct, characterized in that, a cell culture adapted mutant of a HCV-RNA construct is produced and isolated by a process according to claim 22 or 23, the nucleotide and amino acid sequence of these mutants is determined and the type, number and positions of the nucleotide and amino acid mutations is determined by comparison with the nucleotide and amino acid sequence of the original HCV-RNA construct, and these mutations are introduced in an (isolated) HCV full-length genome or a HCV subgenome or any HCV-RNA construct, either by site-directed mutagenesis or by exchange of fragments containing the relevant mutations
- 10
- 15
- 20
- 25
25. A cell culture adapted HCV-RNA construct with high replication efficiency, characterized in that, it is derivable from a HCV-RNA construct according to one of the claims 4 to 15 through nucleotide and/or amino acid mutations and that it is obtainable with a process according to one of the claims 22 to 24.
- 30
- 35 26. A cell culture adapted HCV-RNA construct according to claim 25, characterized in that,

- 5 it carries one or several of the following amino acid exchanges: 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.
- 10 27. A cell culture adapted HCV-RNA construct according to claim 25 or 26, characterized in that, it displays one or several of the nucleotide and/or amino acid exchanges shown in Table 3, where Table 3 is an integral part of this claim.
- 15 28. Cell culture adapted mutant of a HCV-RNA construct or of a HCV full-length genome with increased replication efficiency in comparison to the original HCV-RNA construct or original HCV full-length genome, characterized in that, it is obtainable with a process in which the number and type of mutations in a cell culture adapted HCV-RNA construct according to claim 24 are determined through sequence analysis and sequence comparison, and these mutations are introduced in a HCV-RNA construct, particularly in a HCV-RNA construct according to one of the claims 4 to 15, or in an (isolated) HCV-RNA full-length genome, either by targeted mutagenesis or by exchange of sequence sections containing the relevant mutations.
- 20 29. Hepatitis C virus particles or virus-like particles, characterized in that, they are obtainable by a process according to one of the claims 22 - 24.
- 25 30. Cells infected with Hepatitis C virus particles or virus-like particles, according to claim 29.
- 30
- 35

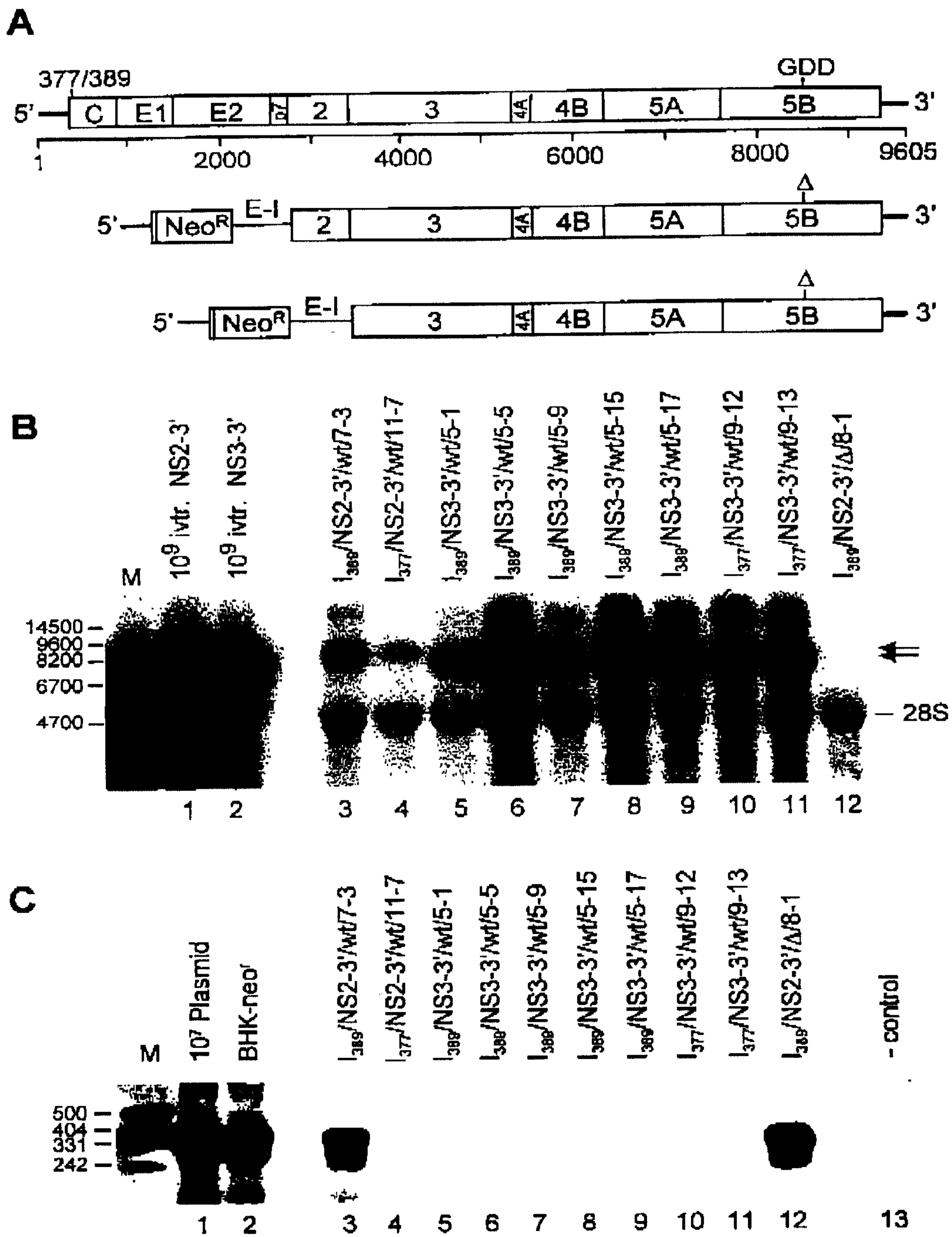


Fig. 1

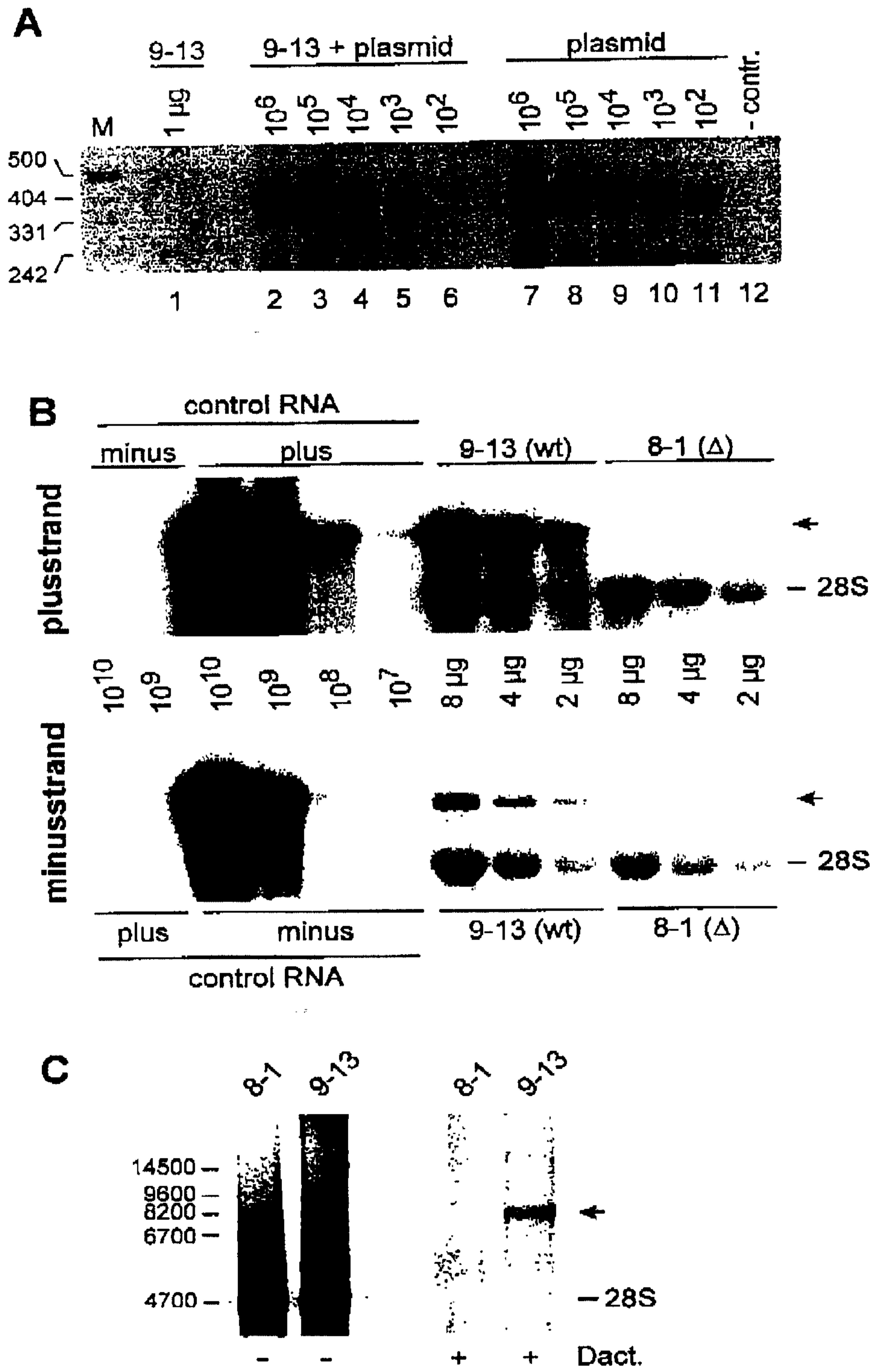


Fig. 2



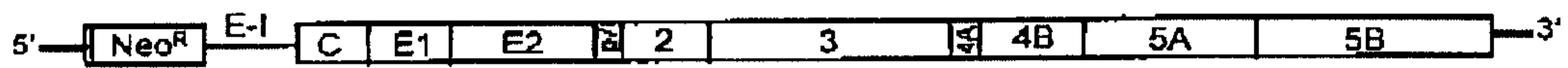


Fig. 4

A



B

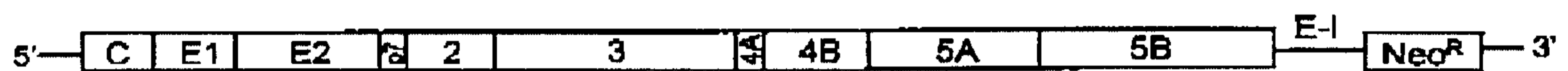
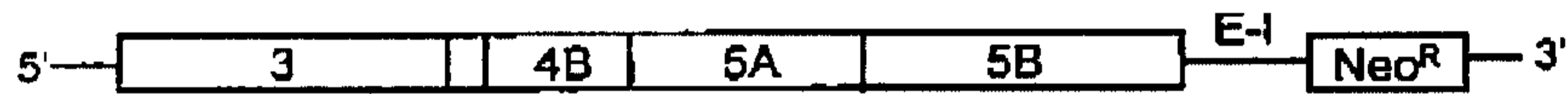


Fig. 5



**Fig. 6**

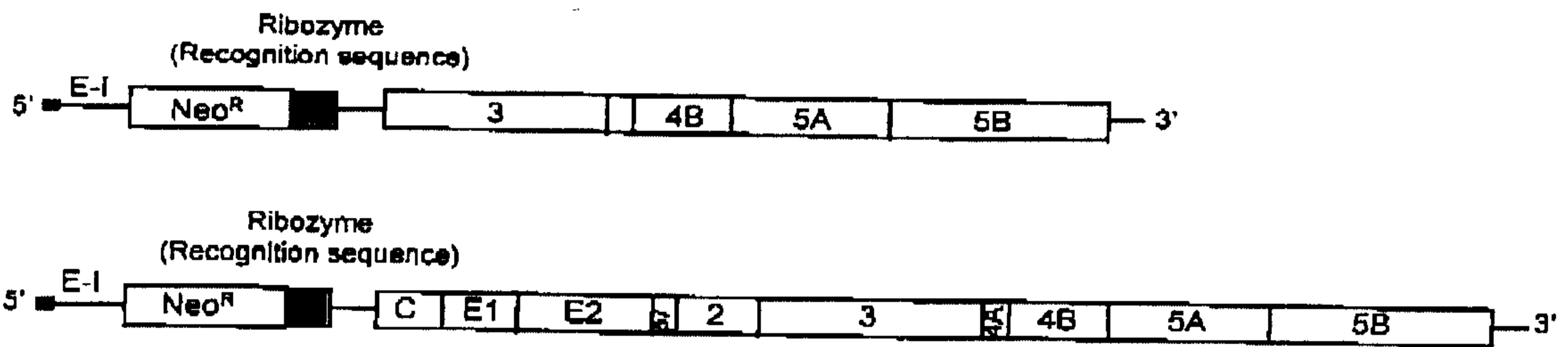


Fig. 7

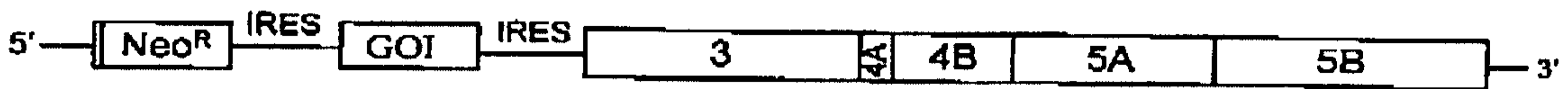


Fig. 8

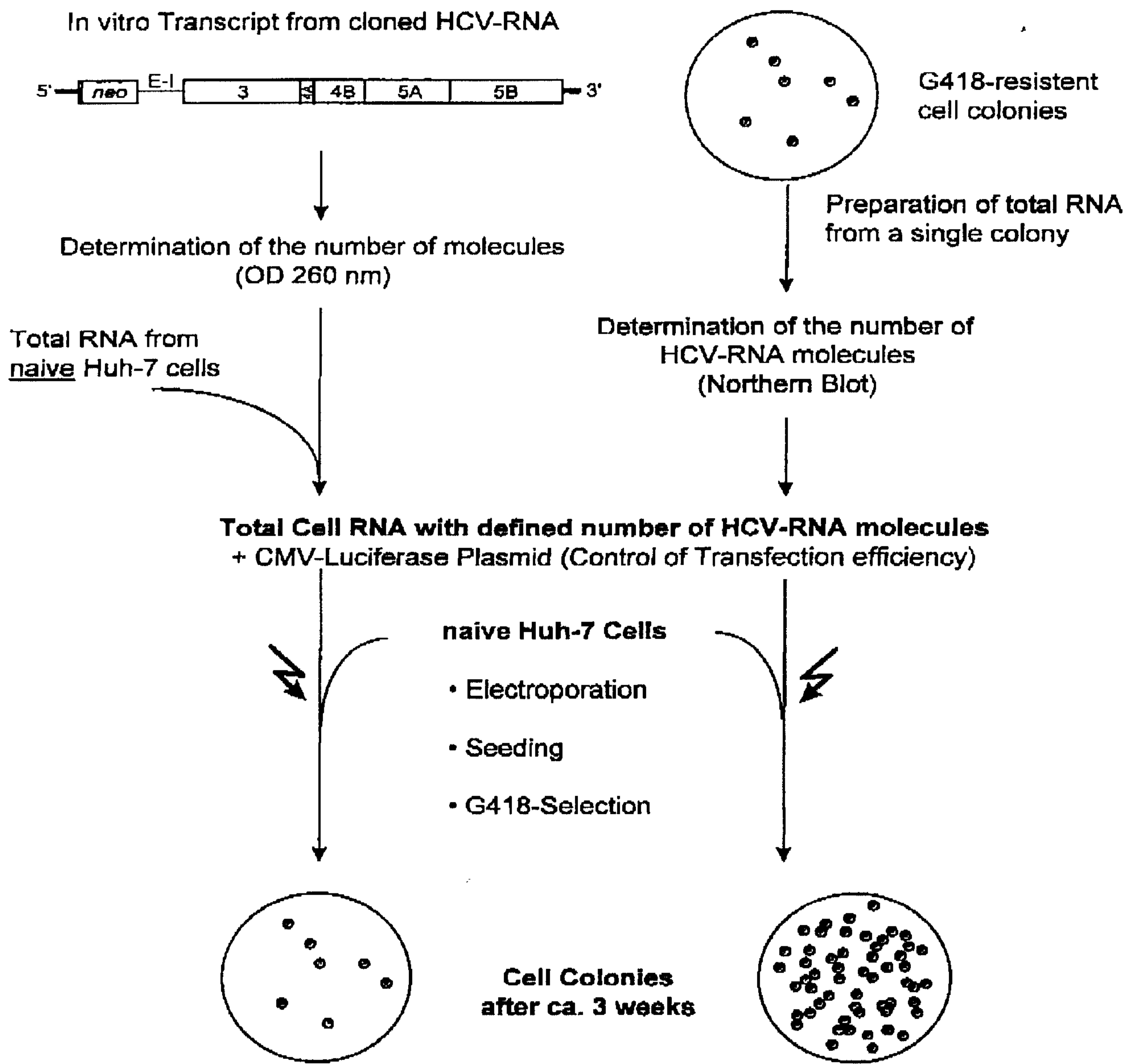


Fig. 9

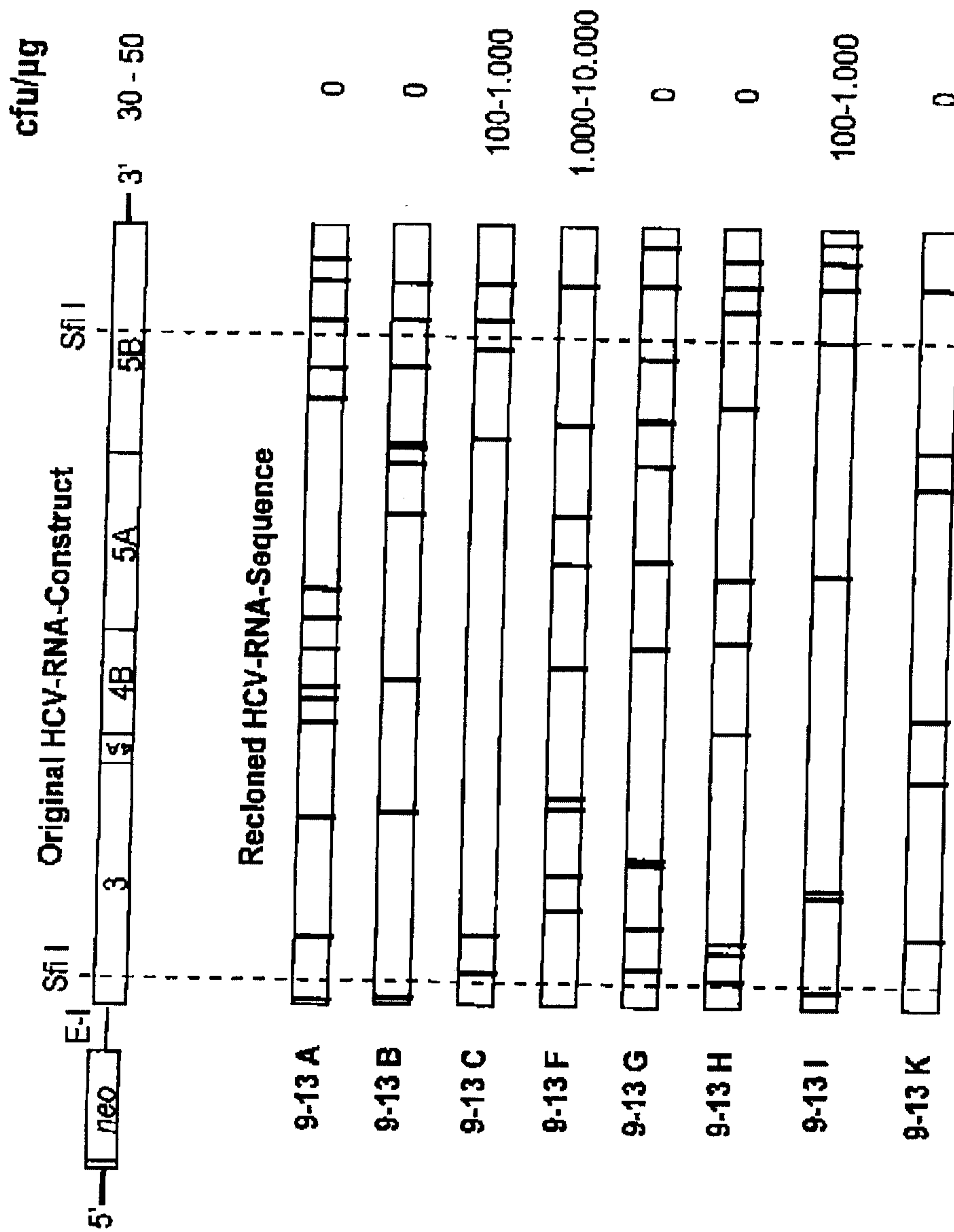
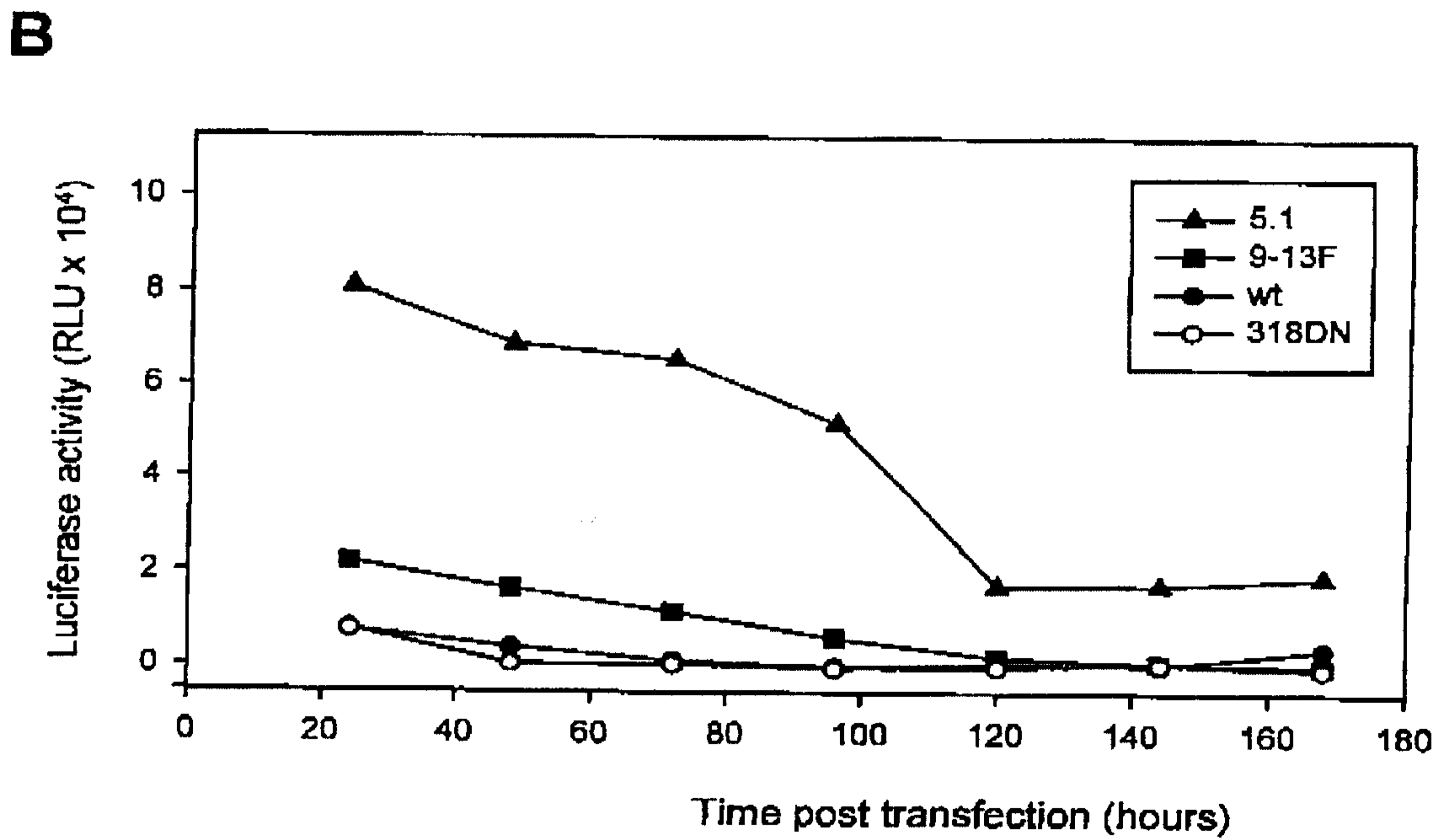
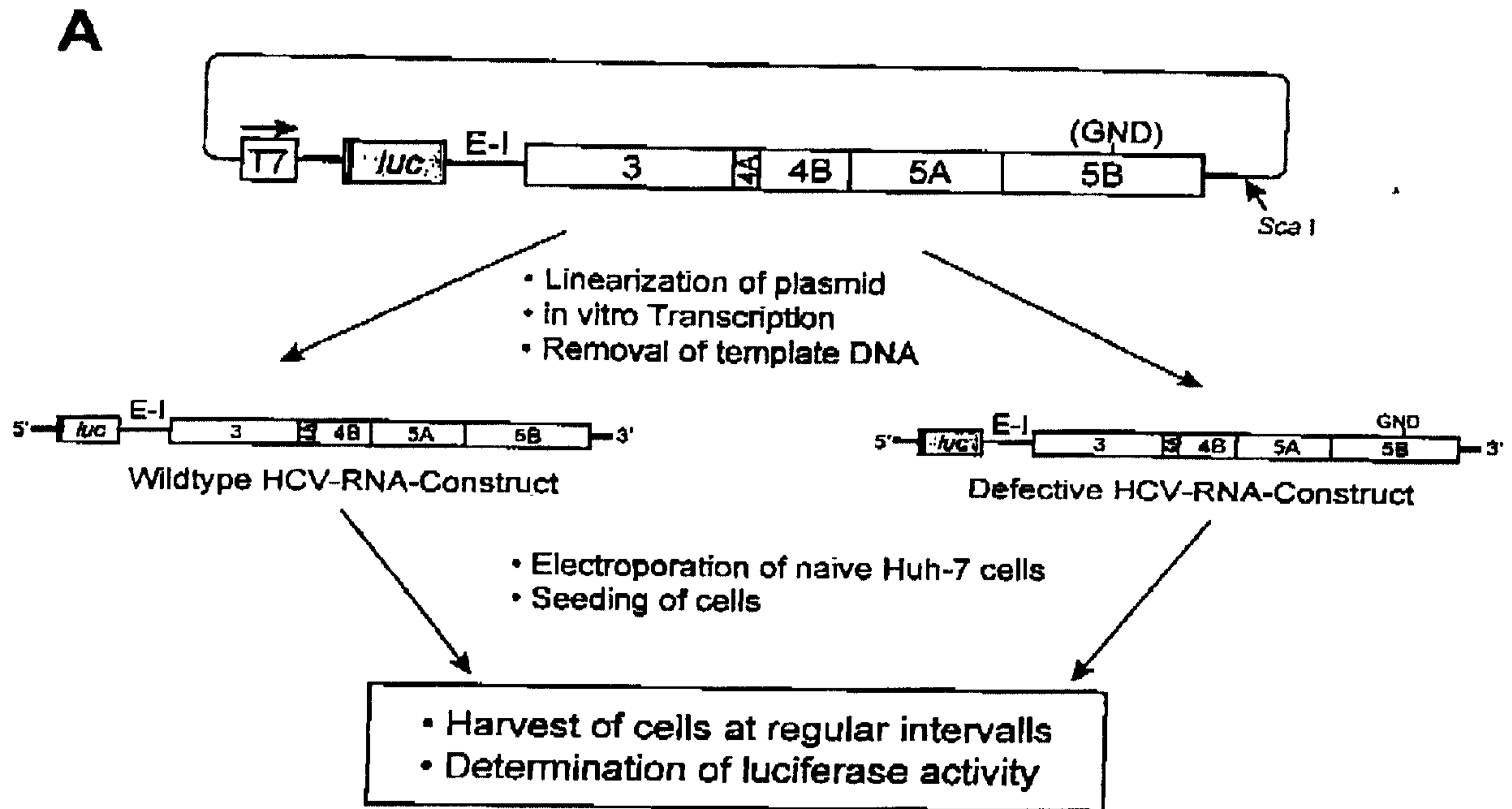
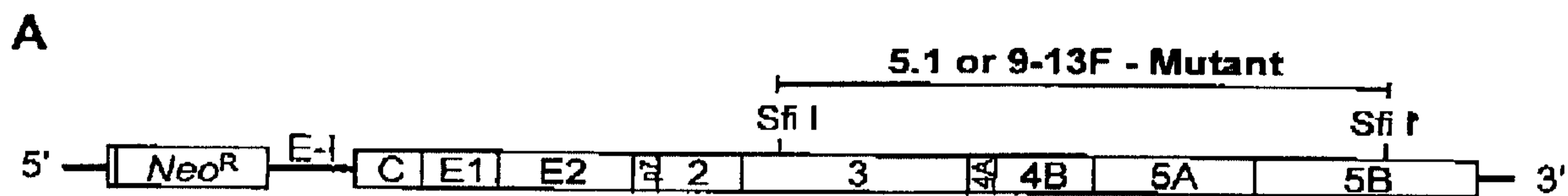


Fig. 10



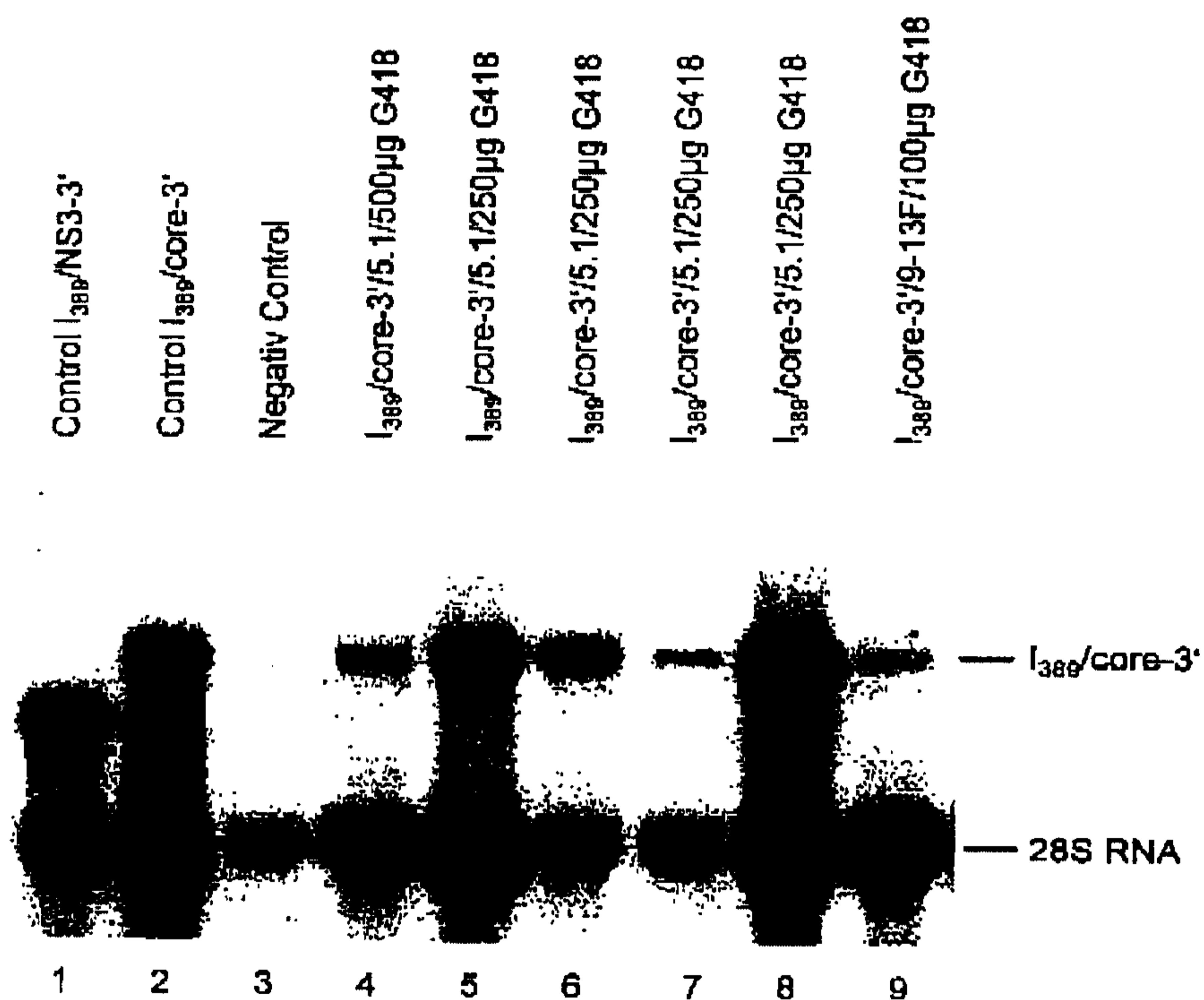
**Fig. 11**



**B**

G418-Concentration:	500 µg/ml	250 µg/ml	100 µg/ml
Number of Cell Colonies:	1	23	75

**C**



**Fig. 12**

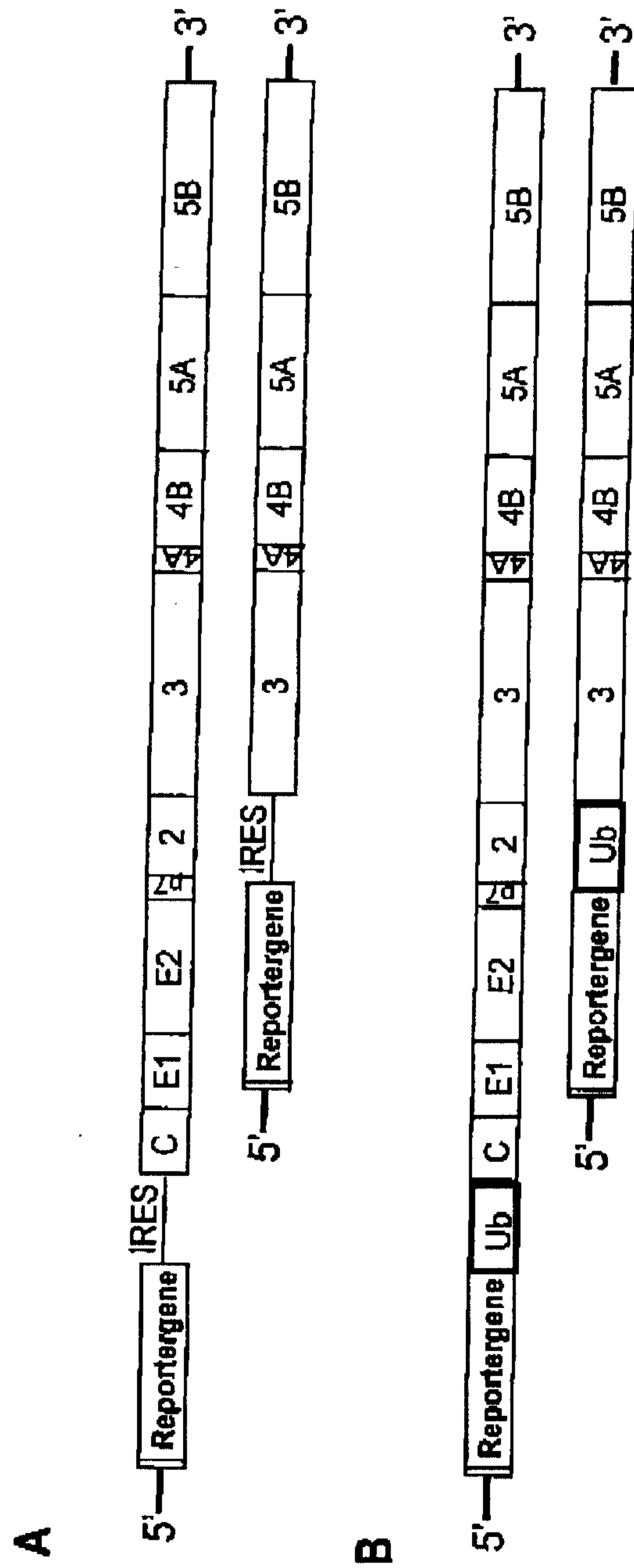


Fig. 13



Fig. 14

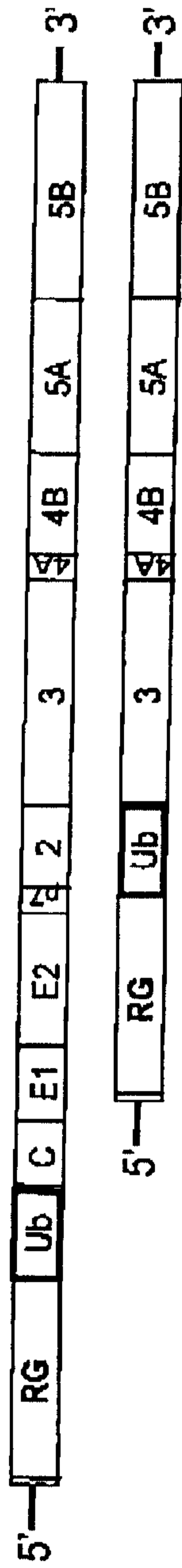


Fig. 15

