SUPPLEMENTARY MATERIAL

Supplementary Materials and Methods

Plasmids encoding cell culture infectious HCV recombinants

Original cell culture infectious genotype 1a strain TN (TNcc), genotype 2a strain J6 (J6cc) and genotype 3a strain DBN (DBN3acc) recombinants were developed previously.[26–28] High-yield HI-recombinants were engineered based on generated subclones harboring the desired mutations using In-Fusion technology. In brief, 50ng of gel purified PCR amplicons generated using the subclones with the sequences of interest as templates were incubated with 2µl of 5x In-Fusion HD enzyme (Takara Bio) and H_2O added to a total volume of 10µl. The reaction was incubated for 15min at 50°C and 4µl of the mix was transformed in 50µl Stellar Competent Cells (Takara) according to the manufacturers´ protocol.

Previously developed cell culture infectious HCV recombinants with genotype(isolate) 1a(TN), 1b(J4), 2a(J6), 2b(J8), 3a(S52), 3a(DBN), 4a(ED43), 5a(SA13), 6a(HK6a)[24,25,30] specific core-NS2 and remaining sequences of genotype 2a isolate JFH1 were used in *in vitro* neutralization assays. Except for 6a(HK6a) containing two vital cell culture adaptive substitutions in the envelope proteins, these viruses reflected the consensus envelope protein sequences determined *in vivo*.

For all plasmid preparations used in experiments the complete HCV sequence was verified by Sanger sequencing (Macrogen).

Maintenance of Huh7.5 and HEK283 cells

The human hepatoma cell line Huh7.5[25] and the human embryonic kidney cell line HEK293 were cultured with Dulbecco's modified Eagle medium (DMEM) (Invitrogen) containing 10% fetal bovine serum (Sigma), penicillin 100U/ml and streptomycin 100µg/ml (P/S) (Gibco/Invitrogen) at 37°C and 5% CO_2; cultures were split every 2–3 days using trypsin (ThermoFisher Scientific) as described.[25]

Transfection of HCV *in vitro* RNA transcripts in Huh7.5 cells

In vitro HCV RNA transcripts were generated and transfections were carried out as previously described.[25] Briefly, plasmids were linearized with Xbal (New England Biolabs) following the sequence encoding the 3´end of the HCV genome, and RNA *in vitro* transcription reactions were carried out using T7 RNA polymerase (Promega). RNA concentrations were measured using the Qubit RNA BR Assay Kit (Thermo Fisher Scientific); for recombinants directly compared in the same graph, equal amounts of RNA transcripts were used. Prior to the day of transfection, 400,000

Alzua GP, *et al. Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326323

cells were seeded per well of a 6 well plate (Nunc); 7.6µg of TNcc or TNcc-HI, 7.5µg of J6cc or J6cc-HI, and 5µg of DBN3acc or DBNcc-HI RNA transcripts were combined with 5µl Lipofectamine2000 (Invitrogen) in a total of 500µl Opti-MEM (Gibco/Invitrogen) and cells were incubated with the resulting transfection complexes in a total of 2ml Opti-MEM for 4-6h prior to washing with PBS and addition of serum containing medium (DMEM+10%FBS+P/S). Transfected cultures were split every 2-3 days. When cells were split replicate cell cultures were plated in chamber slides for immunostaining to monitor the % of infected cells. In addition, supernatant was collected and stored at -80°C for determination of HCV infectivity titers.

Infection of Huh7.5 cells with HCV

400,000cells/well in 6 well plates plated the previous day, were inoculated at the specified multiplicity of infection (MOI) using supernatants derived from the transfection experiment at the peak of infection, as determined by immunostaining and infectivity titration. Infected cultures were split every 2-3 days. When cells were split, replicate cell cultures were plated in chamber slides for immunostaining to monitor the % of infected cells. In addition, supernatant was collected and stored at -80°C for determination of HCV infectivity titers.

Serial passage of HCV in Huh7.5 cells

For serial passage of cell culture infectious HCV recombinants, 10^6 cells were seeded into T25 flasks (ThermoFisher Scientific) and the following day inoculated with up to 1ml cell culture supernatant derived from the previous passage at the peak of infection as determined by immunostaining. The first passage culture was inoculated with up to 1ml of supernatant derived from a prior transfection culture at the peak of infection. When cells were split, replicate cell cultures were plated in chamber slides for immunostaining of HCV antigen to monitor the % of infected cells and supernatant was collected and stored at -80°C for determination of HCV infectivity titers.

Generation of virus stocks

For generation of HCV stocks for neutralization assays and of seed stocks for vaccine virus production, $6x10^6$ cells were plated in T175 flasks and the next day inoculated with cell culture supernatant derived from the peak of infection. When cells were split, replicate cell cultures were plated in chamber slides for immunostaining to monitor the % of infected cells. In addition, supernatant was collected and stored at -80°C for determination of HCV infectivity titers.

Supernatants derived on days where a high % of HCV infected cells was recorded by immunostaining, and/or a high infectivity titer was recorded, were pooled and used as virus stocks. For virus stocks used in neuralization assays, the envelope protein sequence was confirmed by Sanger sequencing. For virus seed stocks for vaccine virus production, the complete HCV open reading frame (ORF) was analyzed by next generation sequencing (NGS).

Immunostaining of HCV antigens in cell culture

Immunostaining of HCV antigens was done on replicate cultures derived when HCV infected cells were split using primary monoclonal anti-HCV core antibody C7-50 ((EnzoLifeSciences diluted 1:5000 in BSK (PBS supplemented with 0.2% skim milk (Easis) and 1% Bovine Serum Albumin, (Sigma)) and monoclonal anti-HCV NS5A antibody 9E10[24] diluted 1:5000 in BSK.[25] After at least 1h of incubation and subsequent washing, secondary antibody Alexa Flour 488 goat anti mouse IgG (H+L) (Invitrogen) diluted 1:500 in BSK was added and cell nuclei were counterstained with Hoechst reagent (Invitrogen).[25,34] The % of infected cells was scored from 0% to 100%, using intervals of 10%, by fluorescence microscopy with a Zeiss Axio Vert.A1 microscope.

Determination of HCV infectivity titers

HCV infectivity titers were determined in cell culture supernatant as focus-forming units (FFU)/ml as previously described.[25] In brief, 6000 cells were seeded per well of 96 well plates (Nunc) and infected the following day with serially diluted cell culture supernatants, testing each dilution in triplicate. After 48h of incubation, cells were fixed using methanol and stained for HCV antigens using primary anti-HCV core antibody C7-50 (EnzoLifeSciences) diluted 1:1000 in BSK and anti-HCV NS5A antibody 9E10 diluted 1:3000 in BSK, followed by secondary antibody ECL sheep anti-mouse IgG diluted 1:500 in BSK.[34] The HCV antigen positive cells were visualized and FFU were automatically counted using an Immunospot series 5 UV analyzer (CTL Europe GmbH) as described.[25]

Sequence analysis of cell culture derived HCV

HCV RNA was extracted from cell culture supernatants using Trizol LS (Life technology) and the RNeasy MinElute kit (QIAGEN) as described.[33] RNA was eluted and used for reverse transcription with Maxima H Minus Reverse Transcriptase (ThermoScientific) and genotype specific reverse primers binding to the HCV 3´UTR variable region as described[33,32] (Supplementary Table 1).

Polymerase chain reaction (PCR)-based amplification of the complete open reading frame (ORF) was carried out using Q5 Hot start High-Fidelity DNA polymerase and genotype specific primers as described[33,32] (Supplementary Table 2). Alternatively, for virus stocks used for neutralization assays, a PCR amplicon spanning the envelope proteins was generated (Supplementary Table 2). The sequence of the resulting PCR amplicons was either determined by Sanger sequencing or by NGS. NGS was carried out as described.[33] In brief, PCR amplicons were loaded on a gel with SYBR safe DNA gel stain (Invitrogen), visualized with blue light and gel extracted with the Large Fragment DNA recovery kit (Zymo Research). Size selection was done with magnetic beads and adaptors with unique barcodes with the use of TruSeq Nano DNA library kit (Illumina). Finally, paired-end sequencing was performed on the Miseq platform. Viruses from serial passage experiments, first passage kinetic experiments and virus seed stocks used for vaccine production, were subjected to NGS of the complete ORF. Virus stocks for neutralization assays were subjected to Sanger sequencing of genome regions encoding E1 and E2.

Subclonal analysis

For serially passaged viruses, PCR amplicons spanning the complete ORF were subcloned using the TOPO-XL Cloning kit (Invitrogen) following the manufacturers´ instructions. The HCV sequence of the resulting plasmids was determined by Sanger sequencing. TNcc-PP-10, TNcc-PP-18, TNcc-PP-38.1, J6cc-PP-35 and DBNcc-PP-16 were analyzed like this. For TNcc-PP-10, TNcc-PP-18 and TNcc-PP-38.1, phylogenetic analysis was carried out using MAFFT for aligning sequences and subsequently PhyML for building the phylogeny visualized by FigTree v1.4.3.

HCV production for vaccine generation

HCV production for immunization studies was done using previously developed protocols.[35] In brief, $18x10^6$ cells were plated in T500 triple layer cell culture flasks (ThermoFisher Scientific). The next day, cells were infected at MOI 0.003 with genotype 1a, 2a and 3a seed stocks. On day 1 post infection, around $1.8x10^8$ cells were plated in 10-layer cell-factories (ThermoFisher Scientific). When 80% of cells were estimated to be infected by monitoring of a replicate T25 cell culture, cells in the cell-factory were washed with pre-warmed PBS and DMEM was replaced by Adenovirus Expression Medium (AEM) supplemented with P/S.[34] From the 10-layer cell-factories, supernatant was harvested every 2–3 days. Each factory yielded 5 harvests of 800ml each. For each HCV recombinant,

4 cell-factories were done, collecting a total volume of ~16l of virus containing supernatant. Harvested supernatant was stored at -80°C until further processing.

Processing of HCV for vaccine generation

HCV was processed for vaccine development using previously established protocols.[35] In brief, initial clarification of ~16l virus containing supernatant harvested from cell-factories was carried out through a 5µm capsule filter, followed by a 0.65µm Sartopure® PP3 filter (Sartorius). Afterwards, two tangential flow filtration (TFF) steps were carried out using a fiber with a surface area (SA) of 2600cm^2 and a molecular weight cut off (MWCO) of 500kDa and subsequently a fiber with a SA of 790cm^2 and a MWCO of 500kDa (MINIKROS 65cm 500KD MPES, MINIKROS 20cm 500KD MPES, Repligen), concentrating ~16l to ~350ml and subsequently to ~35ml. For a subsequent 3-cushion ultracentrifugation (UC) step, 3 cushions containing 1ml of 60%, 28% and 10% Optiprep Density Gradient Medium (Sigma), respectively, were layered on top of each other in 6 UC tubes. To each tube, ~6ml of virus sample were added and the total volume was adjusted to 11ml with PBS. After centrifugation with a Beckman SW-41 rotor at 40,000 revolutions per minute (rpm) at 4°C for 2h, fraction 1 (8ml), fraction 2 (1.2ml) and fraction 3 (2ml) were collected from the top of each tube. Fraction 2 was pooled from 6 tubes and adjusted to a total volume of 20ml with PBS prior to TFF with a fiber with a SA of 20cm^2 and a MWCO of 500kDa (MICROKROS 20CM 500K MPES 0.5MM, Repligen) resulting in concentration to ~2ml. For a subsequent equilibrium density gradient UC step one day prior to the experiment a semi-continuous 40% to 10% iodixanol gradient was prepared by layering 2.5ml of 40%, 30%, 20% and 10% Optiprep Density Gradient Medium on top of each other followed by equilibration at 4°C. The ~2ml sample resulting from the previous step was added on top of the gradient and ultracentrifugated at 40,000rpm with a Beckman SW-41 rotor for 6h at 4°C. Eighteen fractions of ~ 550µl were collected from the bottom of the tube and each fraction was weighed for determination of buoyant density. The three fractions with a density closest to 1.1g/ml were pooled, obtaining a final volume of ~1.5ml. For a subsequent size exclusion chromatography step 1.54g of Sephadex G-100 (Sigma Aldrich) was mixed with 35ml sterile water 3 days prior to the experiment and added to a chromatography PD-10 column (PD-10 reservoirs, GE Healthcare Life Science). The ~1.5ml of sample resulting from the previous step were then added to the column, and 12 fractions of ~1ml were eluted with NaCl (9 mg/ml). Based on absorbance determined with a NanoDrop (Thermo Scientific) at an OD of 230nm, 5 fractions were pooled, obtaining a final volume of close to 5ml. Finally, the resulting preparation was subjected to UV-

irradiation with a UVG-54 Handheld UV lamp (240nm UV, 6 watt) (Analytik Jena) for 25min with frequent agitation using a 6 well plate with 1.25ml of sample per well. To confirm inactivation, cells seeded one day prior to the experiment were inoculated with 100µl of UV inactivated sample, and cell cultures were followed for 3 weeks by splitting and immunostaining every 2-3 days.

Immunization of mice

To evaluate immunogenicity of the developed vaccine candidates, 6-8 week-old female BALB/c mice were acquired from Taconic Farms, Denmark, and were housed in certified animal facilities at the University of Copenhagen. Animals were resting for at least one week following arrival to the animal facility. Animals were subcutaneously immunized four times every 3 weeks with processed and inactivated HCV or as a control with 100µg EndoFit OVA (Invitrogen) formulated with adjuvant AddaVax 50%/50% (v/v). Genotype 1a, 2a and 3a HCV vaccines contained an equivalent of 6.8, 8.6, and 8.6 log10 FFU, respectively, determined prior to inactivation. Each experimental group including the control group consisted of 3 animals. Thus, a total of 12 animals were used. Animals were randomly assigned to the different groups and treatments were administered in random order. The size of the groups was determined based on availability of vaccine antigen. Two weeks after the last immunization mice were sacrificed and serum was obtained. IgG was purified from serum of each animal and was evaluated in *in vitro* neutralization assays as the primary outcome measure; in addition, IgG pools derived from each group were evaluated in *in vitro* neutralization assays as the primary outcome measure and in HCV E1/E2 complex ELISA assays as a secondary outcome measure. No animals were excluded from the study. All data obtained from all animals were reported. During the experiments, persons handling animals and samples were blinded to the identity of the vaccine antigen. Experiments were conducted in accordance with national Danish guidelines regarding animal experiments (Amendment # 1306 of November 23, 2007), approved by the Danish Animal Experiments Inspectorate, Ministry of Justice, permission numbers 2015-15-0201-00623.

Patient samples

Sera or plasma from patients with chronic hepatitis C (CHC) were collected between May 2011 and August 2021 in biobanks attached to the Danish Database for Hepatitis B and C and the HCV Tandem cohort at the Department of Infectious Diseases, Copenhagen University Hospital, Hvidovre. Patients were ≥18 years, had no previous history of treatment for CHC, no co-infection with human immunodeficiency virus or hepatitis B virus, and no recent intravenous drug use. The HCV Tandem

cohort was approved by the Regional Ethical Committee (H-21004361), and the Danish Data Protection Agency (2012-58-0004); written informed consent was provided by all patients.

Purification, concentration and quantification of IgG from mouse serum and patient serum or plasma

IgG from individual mouse serum or individual patient serum / plasma samples was purified with the Amicon® Pro Affinity Concentration Kit Protein G with 50kDa Amicon® Ultra-0.5 Device (Merck Millipore), in accordance to the manufacturers´ instructions. Briefly, for mouse samples 200µl and for patient samples 600µl of Protein G resin was added to the column, followed by a wash step. Next, serum or plasma was added and incubated for 1h at room temperature on a shaker; 300-500µl of mouse serum or 600µl of patient serum or plasma was processed. Following the second wash, elution and neutralization buffer were added to obtain IgG. Concentration of IgG was done with the Vivaspin 500, 30,000 MWCO (GE Lifescience) kit according to the manufacturers´ instructions. Briefly, IgG resulting from the previous step was diluted in PBS to a total volume of 500µl and centrifugated at 14,000g until a volume of ~80µl was obtained. Concentration of the resulting mouse IgG preparations was determined with the IgG (TOTAL) mouse uncoated ELISA Kit (ThermoFisher) according to the manufacturers´ instructions. Briefly, 96 well plates were coated with capture antibody and incubated overnight, followed by incubation with blocking buffer. After several washes, concentrated and serially diluted IgG was added in duplicates together with the provided standards. Then, the detection antibody was added and incubated for two hours on a shaker. Next, substrate was added followed by stop solution and absorbance at an optical density (OD) of 490nm was measured (FLUOstar OPTIMA, BMG Labtech). The generated standard curve was used to calculate the IgG concentration. Concentration of the resulting patient IgG preparations was determined with the Cedex Bio Analyzer (Roche) according to the manufacturers´ instructions.

In vitro neutralization assay

Neutralization with human derived monoclonal antibodies AR3A and AR4A,[17,18]and polyclonal IgG C211[19] was done as described in.[25] In brief, these antibodies were serially diluted in cell growth medium (DMEM+10%FBS+P/S), added to the virus and incubated in a total of 100µl for 1h at 37°C. Each antibody concentration was tested in triplicate. Virus only controls were prepared by mixing virus with cell growth medium. Virus-antibody mixes, virus-medium mixes or medium only were then added to Huh7.5 cells, which had been seeded at 6000cells/well in 96 well poly-D-Lysine coated plates (Nunc) the day prior to the experiment. Following 3.5h incubation, cells were washed with

PBS and 100µl cell growth medium was added per well. After 48h of incubation at 37°C, plates were fixed with methanol and stained for HCV core and NS5A antigen as described for infectivity titrations.

Neutralization with purified mouse serum IgG or patient serum / plasma IgG was done using previously established assays.[37] In brief, these IgG were serially diluted in cell growth medium in a total of 3µl and added to virus diluted in 7µl to a total of 10µl and incubated for 1.5h at 37°C. Each antibody concentration was tested in triplicate. Virus only controls were prepared by mixing virus with cell growth medium. Subsequently, 30µl of cell growth medium were added to virus-IgG or virus-medium mixes and the resulting samples were added to Huh7.5 cells, which had been seeded at 6000cells/well in 96 well poly-D-Lysine coated plates (Nunc) the day prior to the experiment. Following 4.5h incubation at 37°C, cells were washed with PBS and 100µl of cell growth medium were added per well. As a positive neutralization control, C211 antibody and a well-defined genotype 5a virus[31] were mixed and added to each 96-well plate. After 48h of incubation at 37°C, plates were fixed with methanol and BSK was added at 300µl/well. Following 1h incubation, BSK was removed and 100µg/ml Fab Fragment Goat anti-mouse IgG (Jackson ImmunoResearch) diluted in PBS was added and incubated for 1h at room temperature. Next, plates were washed and 50µl of monoclonal primary antibody 9E10 diluted 1:5000 in BSK was added per well. The remaining staining steps were done as described above for infectivity titrations.

For all neutralization assays, the percentage of neutralization was calculated relating the number of FFU in experimental wells to the mean number of FFU in virus only wells. For all neutralization assays, half maximal effective concentration (EC50) was calculated based on concentration-response curves, using top and bottom constrains of 0 and 100% and the formula $y=100/(1+10^{(\log 10EC50-X)\times hillslope})$ using GraphPad Prism, as described.[34] Fold-changes in neutralization sensitivity given in Figure 4 were calculated as [(EC50 of original virus)/(EC50 of virus with envelope substitutions)]; these calculated values are not meant to reflect data accuracy.

HCV E1/E2 complex ELISA

Expression plasmids harboring E1 and E2 envelope protein sequences of TNcc-HI, J6cc-HI and DBNcc-HI were constructed using the In-Fusion HD cloning kit (Takara Bio) and 50ng of PCR amplicons harboring E1 and E2 sequences of the TNcc-HI, J6cc-HI or DBNcc-HI plasmid and an HCV pseudo-particle expression plasmid. The E1 and E2 sequence in the resulting plasmids was sequence confirmed (Macrogen). Plasmids were transfected in HEK293 cells, plated at

$8x10^5$cells/well in 6-well plates 1 day prior to the experiment, using 5µl Lipofectamine2000 (Invitrogen) and 5µg plasmid in a total volume of 2ml of Opti-MEM (Invitrogen). Following 6h of incubation, Opti-MEM was replaced with DMEM+10%FBS+P/S. In a replicate culture, after 48h of incubation at 37°C, transfection efficiency was determined by immunostaining for E1/E2 protein using the monoclonal antibody AR4A[17] diluted 1:5000 in BSK for 1h at room temperature, followed by application of the anti-human Alexa Fluor 488 coupled secondary antibody, as described above. Following confirmation of expression, cell growth medium was discarded, and total protein was collected in lysate buffer (1% Triton X-100, 50nM Tris-HCl, 150nM NaCl, pH8.0). Next, the cell lysate was treated with Benzonase endonuclease (Sigma) and 2nM of $MgCl_2$ and finally, the generated protein was measured with BCA protein assay kit (Pierce™) according to the manufacturers´ instructions.

For ELISA assays, 96-well plates (Nunc) were coated with *Galanthus nivalis* lectin (Medicago) and the next day the plate was washed with PBS and blocked with 200µl of PBS-5% non-fat milk (Easis). The next day, the plates were washed with PBS containing 0.1% Tween (PBST) and 50µl of E1/E2 complexes (200µg/ml) were added to the plates and incubated at 4°C overnight. After several washes with PBST, purified mouse serum IgG or immune-sera were serially diluted in PBST-1% non-fat milk, dilutions were added to plates in duplicates, and plates were incubated for 2h at room temperature, followed by a washing step with PBST. Binding of antibodies to E1/E2 complexes was detected by secondary antibody ECL sheep anti-mouse IgG horseradish-peroxidase linked whole antibody (GE Healthcare) diluted 1:1000 in PBST-1% non-fat milk and incubated for 1h at room temperature, followed by a washing step with PBST. TMB substrate (3,3′, 5,5′-tetramethylbenzidine, Thermo Scientific) was then added for 10min followed by ELISA Stop Solution (Invitrogen). Absorbance was determined at 450nm with the use of an ELISA plate reader (BIO-TEK Instruments, Inc.). As positive controls, mouse anti-E2 antibody AP33[39] was used to bind TNcc-HI E1/E2 and DBNcc-HI E1/E2 complexes and H77.39[40] was used to bind J6cc-HI E1/E2 complexes. As negative control, only secondary antibody, ECL sheep anti-mouse IgG horseradish-peroxidase linked whole antibody (GE Healthcare) was added to the E1/E2 complexes omitting prior addition of mouse serum IgG, immune-sera or control antibodies.

Supplementary Results

Phylogenetic analysis of subclones of polyclonal passaged genotype 1a HCV from an initial passage line

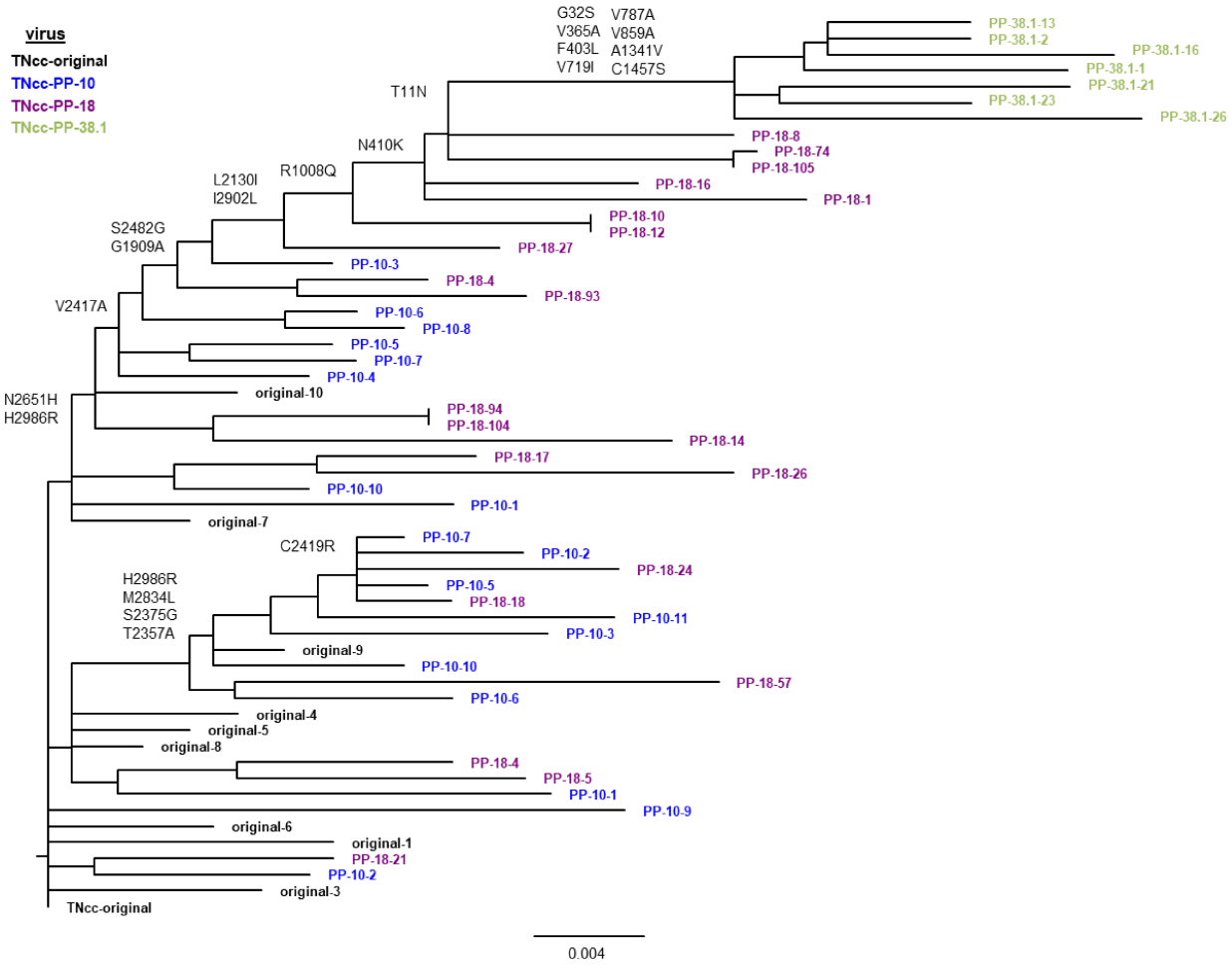
For the original TNcc, TNcc-PP-10, TNcc-PP-18 and TNcc-PP-38.1, generated subclones were subjected to phylogenetic analysis using PhyML; subclones for the original TNcc were generated and reported previously.[33] Results of this analysis are shown in Supplementary Figure 1. In comparison to the TNcc plasmid sequence, in passage 10 and 18 viruses, this analysis revealed two main quasispecies populations, designated A and B. The A population was mainly characterized by signature substitutions N2651H and H2986R. The B population was mainly characterized by signature substitutions T2357A, S2375G, M2834L and H2986R. In the passage 10 subclones, the A and B population had similar frequency, with 8/18 subclones belonging to A and 7/18 subclones to B, respectively. In the passage 18 subclones, the A population appeared to prevail over the B population with 15/21 and 3/21 subclones belonging to the A and B population, respectively. TNcc-PP-38.1 was characterized by population A signature substitutions and acquisition of additional substitutions corresponding to the substitutions found with >80% frequency in TNcc-PP-38.1 (Supplementary Table 4). Thus, subclonal analysis described in this section reflected results obtained by NGS analysis of polyclonal cell culture virus (Supplementary Table 3, 4 and 7).

Downstream processing of HCV for vaccine experiments

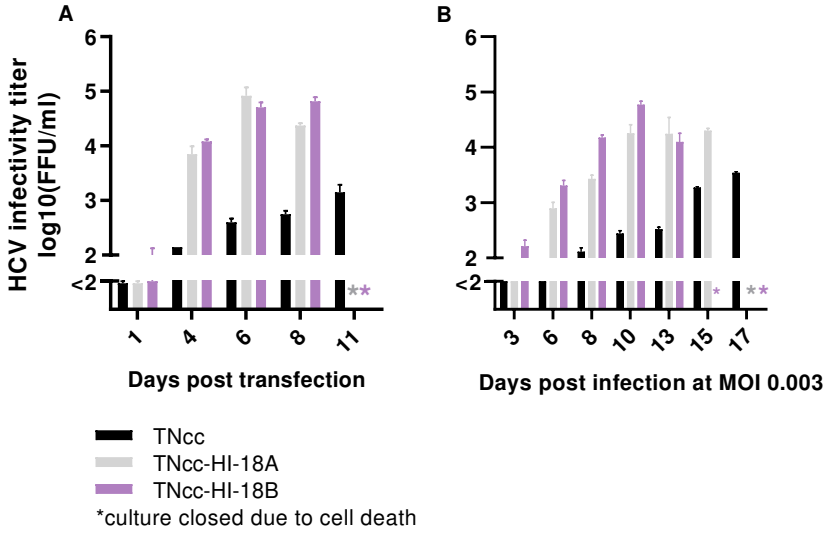
While the general protocol for HCV downstream processing is described in Materials and Methods / Supplementary Materials and Methods, this section describes the results obtained during downstream processing of genotype 1a, 2a and 3a vaccine viruses. Downstream processing was initiated with a first filter clarification step using filters with a pore size of 5µm and 0.65µm, followed by two TFF steps with two hollow fibers with different surface areas, which resulted in a volume reduction from ~16l to ~35ml. The resulting material had HCV infectivity titers of 7.5, 9.2 and 8.8 log10 FFU/ml for genotype 1a, 2a and 3a HCV, respectively (Supplementary Figure 4). The resulting material was distributed into six different ultracentrifugation vials which were subjected to 3-cushion ultracentrifugation resulting in three fractions. Fraction 2 from each of the six vials (~1.2ml each) were pooled amounting to a total volume of ~7.5ml with HCV infectivity titers of 7.8, 9.5 and 9.1log10 FFU/ml for genotype 1a, 2a and 3a HCV, respectively (Supplementary Figure 4). This was followed by TFF using a small hollow fiber reducing the volume to ~2ml with infectivity titers of 8.5, 11 and 9.6log10 FFU/ml for genotype 1a, 2a and 3a HCV, respectively (Supplementary Figure

4). These samples were subjected to gradient ultracentrifugation; 18 fractions (~550µl each) were collected and weighed to determine their buoyant densities. The three fractions with densities closest to 1.1g/ml were pooled amounting to a total volume of ~1.5ml. These pools yielded infectivity titers of 7.8, 9.9 and 9.6 log10 FFU/ml for genotype 1a, 2a and 3a HCV, respectively (Supplementary Figure 4). These pools were then subjected to Sephadex chromatography, where 12 fractions (~1ml each) were collected. Five of these fractions were pooled based on absorbance at OD of 230nm amounting to a total volume of ~4.8ml with infectivity titers of 7.4, 9.2 and 9.1 log10 FFU/ml for genotype 1a, 2a and 3a HCV, respectively (Supplementary Figure 4). Thus, recoveries in downstream processing were 9%, 17% and 30% for genotype 1a, 2a and 3a HCV, respectively. The pools resulting from chromatography were UV inactivated to yield the final vaccine antigens. To confirm virus inactivation, replicate cell cultures were inoculated and maintained for three weeks; cells were split every 2-3 days, when replicate cultures for immunostaining for HCV antigen were plated, to confirm inactivation based on absence of HCV antigen positive cells (Supplementary Figure 5).

Supplementary Figures

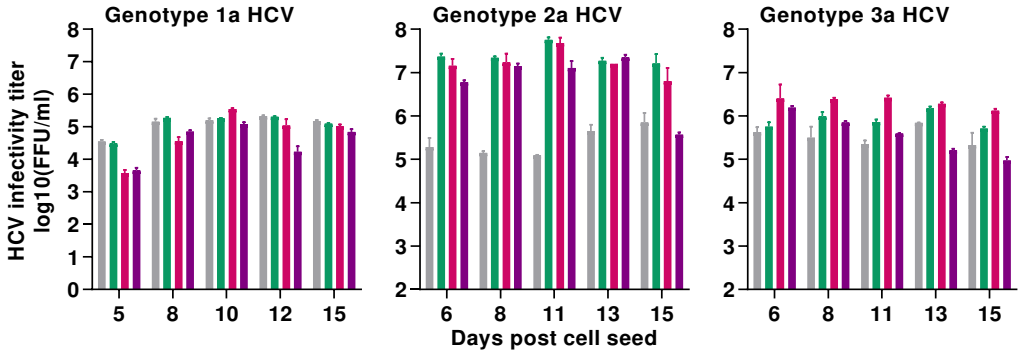


Supplementary Figure 1. Phylogenetic analysis of subclones of full ORF amplicons of polyclonal passaged genotype 1a HCV**.** Phylogenetic analysis and ancestral reconstruction of subclones of RT-PCR amplicons spanning the entire ORF of the original TNcc used as starting material for serial passage experiments, TNcc-PP-10, TNcc-PP-18 and TNcc-PP-38.1 (Figure 1 and 2 and Supplementary Table 3, 4 and 7). Respective subclones are designated PP-10-X, PP-18-X and PP-38.1-X, with "X" being the subclone number.



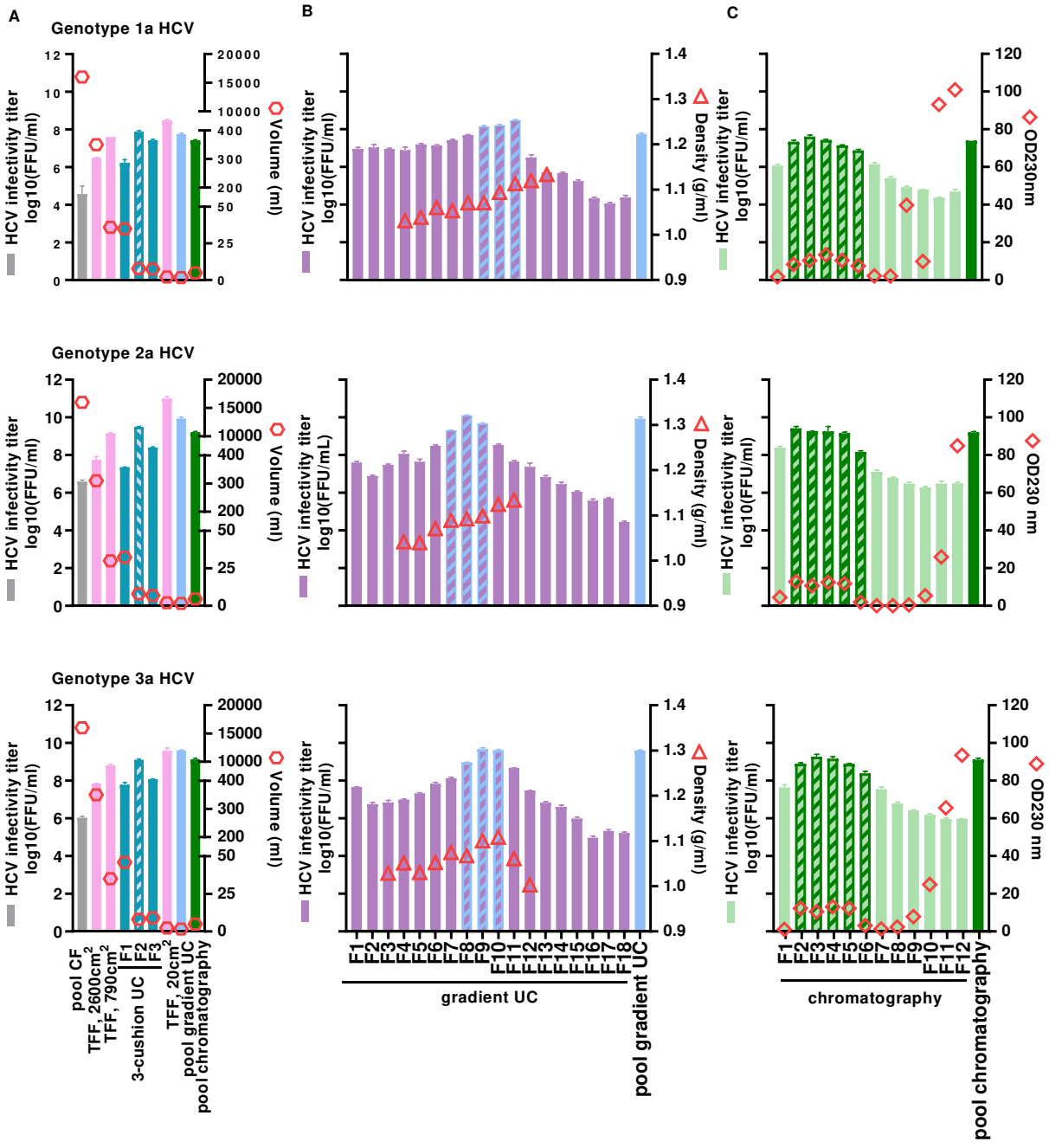
Supplementary Figure 2. HCV genotype 1a recombinants engineered based on an initial passage line had increased viral fitness in cell culture. (A) Original TNcc and newly engineered TNcc-HI-18A and TNcc-HI-18B HCV recombinants were transfected into Huh7.5 cells using the same amount of HCV RNA *in vitro* transcripts. (B) First passage kinetic experiments were inoculated at MOI 0.003 with TNcc and HI-viruses using supernatants derived from the transfection experiment when peak infectivity titers were observed. Cultures were split every 2-3 days. On these days HCV infectivity titers given in log10 FFU/ml were determined with an infectivity titration assay as means of three replicates with standard deviation (SD). *, cell culture was closed due to HCV induced cell death.

346



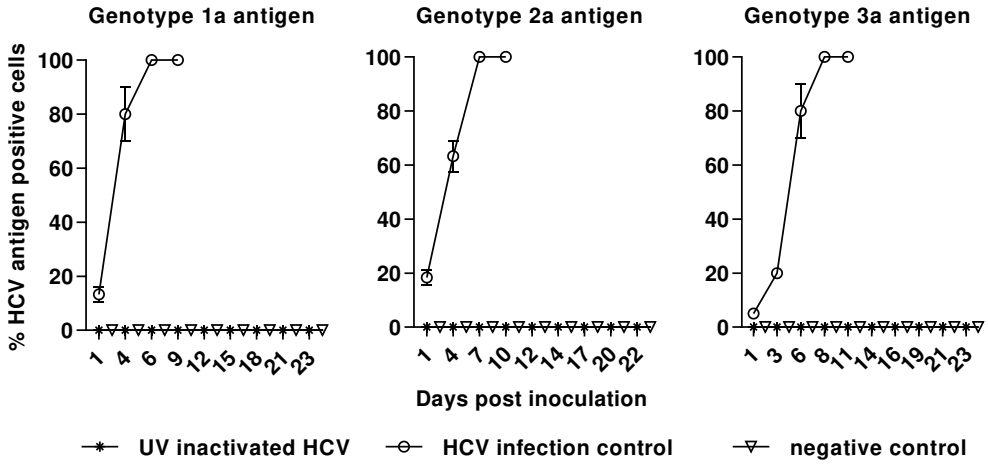
347

348 Supplementary Figure 3. Production of genotype 1a, 2a, and 3a vaccine viruses. Huh7.5 cells were
349 infected with genotype 1a, 2a and 3a HCV seed stocks at MOI 0.003 and transferred to 10-layer cell-
350 factories. For each genotype four cell-factories were processed as indicated by different colors. Per
351 cell-factory, five harvests of 800ml supernatant were collected at the indicated days post cell seed
352 into the cell-factory. HCV infectivity titers in harvests were determined as log10 FFU/ml and are
353 means of three replicates with SD. Thus, per virus 16l HCV containing supernatant were produced
354 and pooled for further processing.

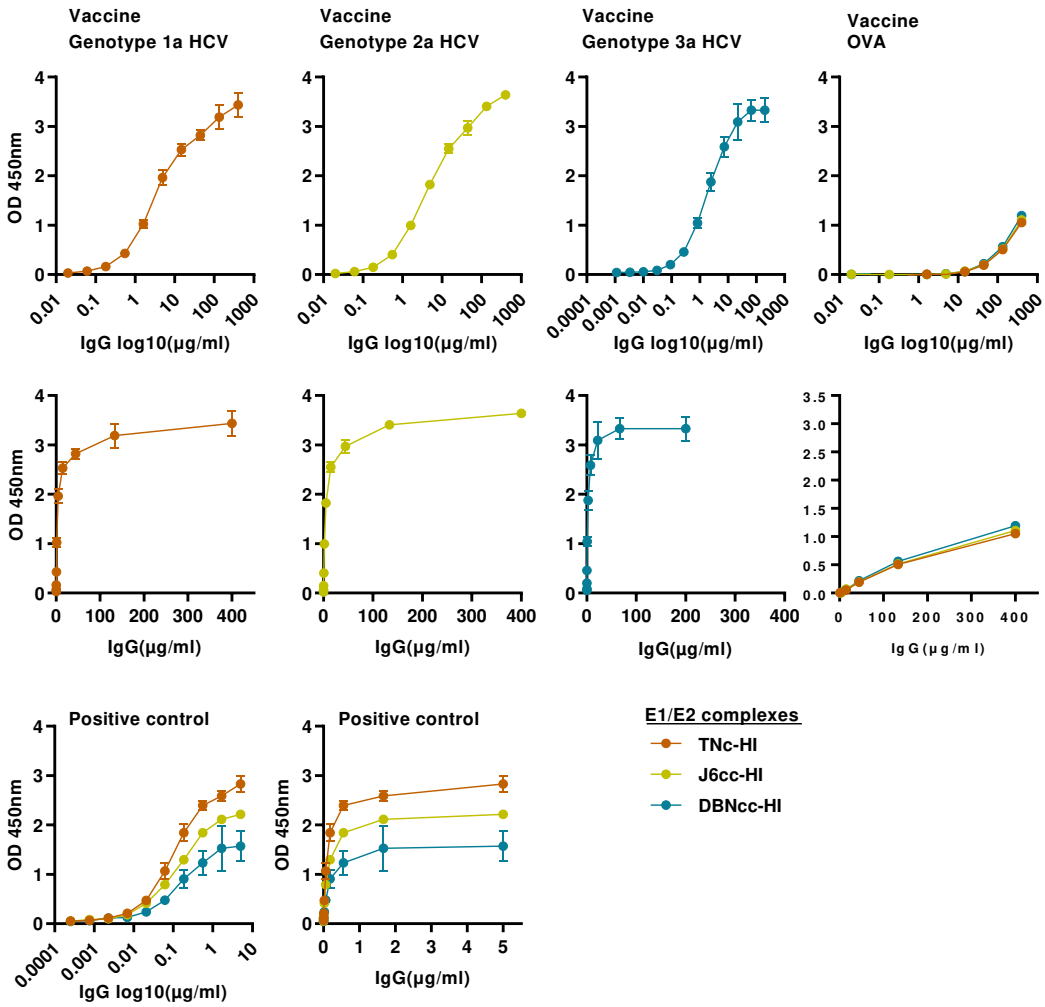


Supplementary Figure 4. Downstream processing of cell culture produced genotype 1a, 2a and 3a HCV. For each of genotype 1a, 2a and 3a HCV, ~16l of pooled virus containing cell culture supernatant produced in cell-factories were concentrated for vaccine production. (A) HCV infectivity titers (left y-axis) and volumes (right y-axis) of material from the different processing steps: pool CF,

pooled supernatant from cell-factories; TFF, 2600cm^2 and TFF, 790cm^2, TFF with hollow fibers with specified surface area; 3-cushion UC F1, F2 and F3, fraction 1, 2 and 3 from 3-cushion ultracentrifugation with the striped bar indicating F2 subjected to further processing; TFF, 20cm^2, TFF with hollow fiber with specified surface area; pool gradient UC, pool of 3 fractions selected from fraction 7-11 shown in B; pool chromatography, pool of fractions selected from fraction 2-6 shown in C. (B) HCV infectivity titers (left y-axis) and buoyant densities (right y-axis) of 18 fractions collected following equilibrium gradient ultracentrifugation (gradient UC). Striped bars indicate the fractions that were pooled for further processing (pool gradient UC, same samples also shown in A). (C) HCV infectivity titers (left y-axis) and OD230nm determinations (right y-axis) for 12 fractions collected during Sephadex chromatography. Striped bars indicate the fractions that were pooled prior to UV inactivation (pool chromatography, samples also shown in A). (A-C) HCV infectivity titers are given as log10 FFU/ml and as means of three replicates with SD. For buoyant density and OD determination, a single measurement was done.

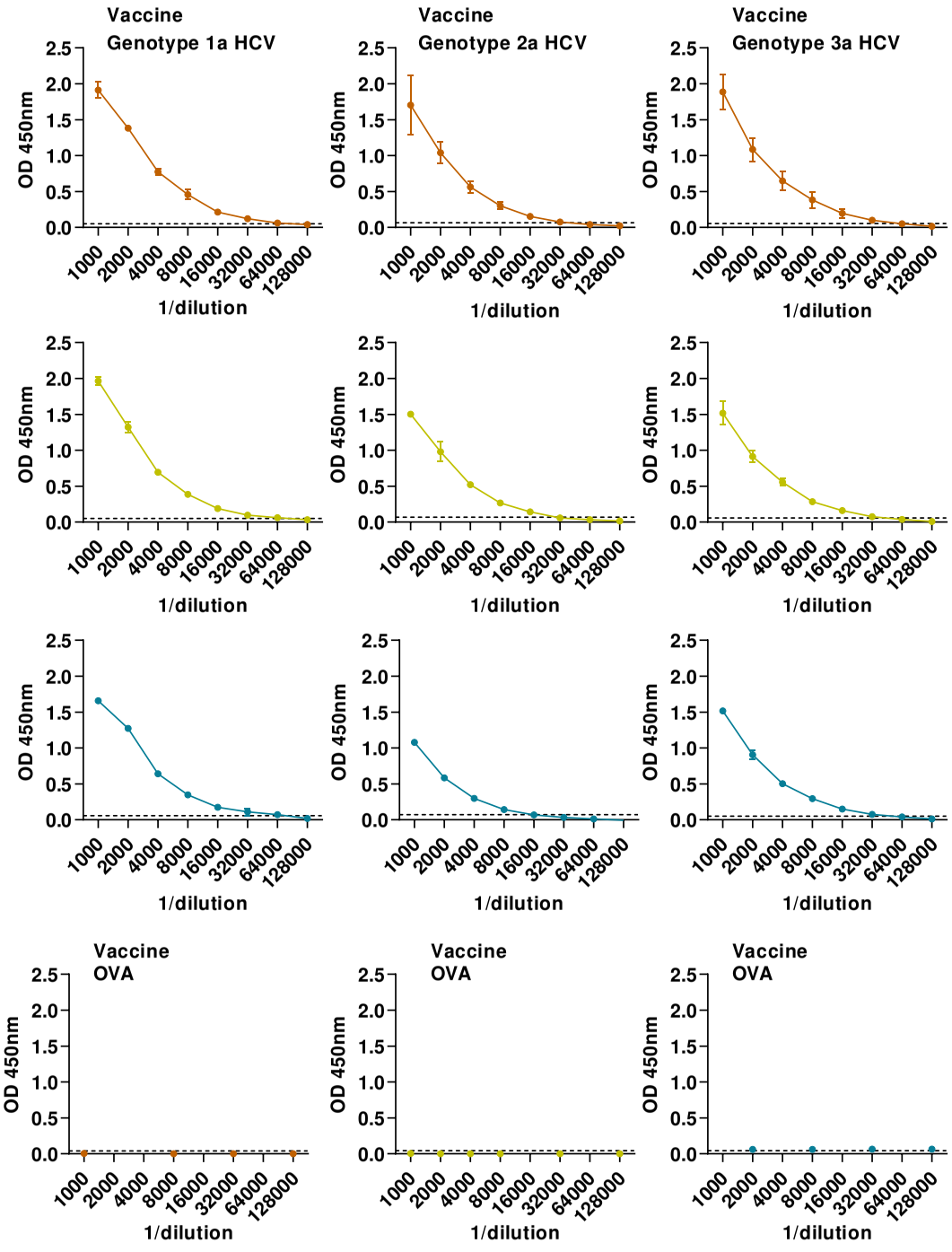


Supplementary Figure 5. Absence of HCV antigen positive cells in cell cultures inoculated with UV inactivated HCV. Three Huh7.5 cell cultures plated the previous day at 80000 cells per well in a 24-well dish, were inoculated with 100µl of UV irradiated HCV of the specified genotype. Positive HCV infection control cultures were inoculated with a genotype 5a HCV[31] at an MOI that resulted in robust infection. Negative control cultures were not inoculated. Cultures inoculated with irradiated HCV and negative control cultures were followed for 3 weeks by splitting every 2-3 days and determination of the percentage of HCV antigen positive cells by immunostaining for HCV core and NS5A. HCV infection control cultures were followed until peak of infection. Percentages of HCV antigen positive cells are means of immunostainings from three replicate cell cultures with SD. For the negative control datapoints are nudged by 0.5 units in the x-axis direction.



Supplementary Figure 6. Immunization with inactivated genotype 1a, 2a or 3a HCV elicited IgG binding to HCV envelope proteins approaching a binding plateau. Binding capacity of pooled purified serum IgG from mice immunized with inactivated genotype 1a, 2a or 3a HCV or OVA to E1/E2 complexes of the homologous HI-recombinants was evaluated by ELISA. Values are OD reads at 450nm following subtraction of mean OD of 8 negative controls. Datapoints are means of two replicates with SD. Positive controls: Instead of serum IgG, well-characterized primary antibodies were used: AP33[38] for binding to TNcc-HI and DBNcc-HI E1/E2 and H77.39[39] for binding to J6cc-HI E1/E2. Negative controls: No IgG was used and TNcc-HI, J6cc-HI and DBNcc-HI E1/E2 were incubated with secondary antibody only; for negative controls, OD reads were ~0.05. Data are shown

either with logarithmic or linear x axis. Compared to Figure 8, higher IgG concentrations were used to evaluate if a binding plateau could be approached.



396

Alzua GP, *et al*. *Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326323

Supplementary Figure 7. Immune-sera binding to HCV envelope proteins and determination of endpoint titers. Binding capacity of immune-sera from mice immunized with inactivated genotype 1a, 2a or 3a HCV or OVA to E1/E2 complexes of the specified HI-recombinants was evaluated by ELISA. Values are OD reads at 450nm following subtraction of mean OD of 8 negative controls. Datapoints are means of two replicates with SD. Negative controls: No immune-sera were used and TNcc-HI, J6cc-HI and DBNcc-HI E1/E2 were incubated with secondary antibody only; for negative controls, OD reads were ~0.05. Broken-lines: 2-fold mean OD of negative controls, used as cut-off for determination of endpoint titers.

405 **Supplementary Tables**

406 Supplementary Table 1. Primers for reverse transcription for generation of full-length HCV ORF
407 amplicons.

3' cDNA primers	Target virus	Primer sequence (5'-3')
1a-9405-RT	1a(TNcc, TNcc-HI)	TAAGAGGCCGGAGTGTTTAC
2a-9481-RT	2a(J6cc, J6cc-HI)	CTATGGAGTGTACCTAGTGTGTGC
3a-9435-RT	3a(DBNcc, DBNcc-HI)	AAAAGAATGGAGTGTTAT

408

409 Supplementary Table 2. Primers for PCR for generation of full-length HCV ORF amplicons and
410 amplicons spanning E1/E2.

Primers	Target virus	Primer sequence (5'-3')
1a-209-F	1a(TNcc, TNcc-HI)	TGCCTGATAGGGTGCTTGCG
1a-9402-R	1a(TNcc, TNcc-HI)	AGGCCGGAGTGTTTACCCCA
1a-3285R	1a(TNcc, TNcc-HI)	TGGTCTCCATCTGGGAAAAG
2a-303-F	2a(J6cc, J6cc-HI)	CTTGCGAGTGCCCCGGGAGG
2a-9467-R	2a(J6cc, J6cc-HI)	TGGAGTGTACCTAGTGTGTGCCGCTC
2a-3774-R	2a(J6cc, J6cc-HI)	GGGATGACATCAGCGTTCCGCGTGAC
3a-293-F	3a(DBNcc, DBNcc-HI)	GATAGGGTGCTTGCGAGTGCC
3a-9432-R	3a(DBNcc, DBNcc-HI)	AGAATGGAGTGTTATCCTACCAGCTCA
3a-3694R	3a(DBNcc, DBNcc-HI)	CTGGCCACCCAACRAGRTCCT

411

412 Supplementary Table 3. Coding nucleotide changes identified in polyclonal passaged genotype 1a
413 HCV from an initial passage line.

HCV protein[a]	Nucleotide change[b]			Allele frequency (%)[c]			AA change TNcc reference[d]	Engineered construct[e]	
	TNcc position	TNcc reference	change	TNcc-PP-10	TNcc-PP-18	Gt 1a HCV seed stock		TNcc-HI-18A	TNcc-HI-18B
core	373	C	A	16	26	41	T11N		
	430	T	C	7	12	11	I30T		
E2	1571	C	G	18	27	42	N410K		
	2464	G	A	11	29	16	S708N		
NS2	2822	G	A	5	16	24	M827I		
	2895	A	G	17	30	15	R852G		
	2935	A	C	1	14	21	N865T		
	3364	G	A	18	28	44	R1008Q		
NS3	3804	C	G	20	33	15	P1155A		
	3978	A	T	2	10	23	N1213Y		
	4848	G	A	5	10	10	A1503T		
NS4B	5812	G	A	15	5	12	G1824D		
	6067	G	C	25	36	59	G1909A		
NS5A	6729	C	A	4	24	40	L2130I		
	7296	C	T	5	15	23	P2319S		
	7410	A	G	24	34	17	T2357A		x
	7464	A	G	29	45	27	S2375G		x
	7588	A	G	9	20	23	D2416G		
	7591	T	C	32	34	46	V2417A		
	7596	T	C	25	43	29	C2419R		x
NS5B	7785	A	G	18	29	43	S2482G	x	
	8292	A	C	61	60	80	N2651H	x	
	8841	A	T	29	38	20	M2834L		x
	8901	A	C	5	11	11	I2854L		
	8985	A	C	9	18	24	I2882L		
	9045	A	C	7	27	43	I2902L	x	
	9298	A	G	72	96	98	H2986R	x	x

414

415 [a], HCV protein, in which specified change was located.

416 [b], Nucleotide change specified by nucleotide position and identity relating to the TNcc genome
417 (Genbank accession number JX993348) as a reference.[26] Coding changes with allele frequency of at
418 least 10% in one of the analyzed samples are listed.

419 [c], Allele frequency of the identified nucleotide change. Polyclonal passage 10 (TNcc-PP-10) and 18
420 (TNcc-PP-18) viruses from the initial passage line (Figure 1), and the seed stock used for vaccine
421 production (Gt 1a HCV seed stock) were analyzed.

422 [d], Amino acid change specified by amino acid position and identity relating to the TNcc polyprotein;
423 amino acid position numbers are identical to position numbers in relation to the reference H77
424 polyprotein (Genbank accession number AF009606).
425 [e], x indicates that the respective nucleotide change was engineered for generation of TNcc-HI-18A
426 and TNcc-HI-18B, respectively.

Alzua GP, *et al. Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326323

Supplementary Table 4. Coding nucleotide changes identified in polyclonal passaged genotype 1a HCV from a later passage line**.**

HCV protein[a]	Nucleotide change[b]			Allele frequency (%)[c]	AA change TNcc reference[d]	Engineered construct[e]
	TNcc position	TNcc reference	change			
				TNcc-PP-38.1		TNcc-HI
core	373	C	A	97	T11N	x
	435	G	A	48	G32S	x
	756	C	A	17	L139I	
E1	1435	T	C	99	V365A	x
E2	1548	T	C	99	F403L	x
	1571	C	G	100	N410K	x
	2496	G	A	100	V719I	x
p7	2701	T	C	92	V787A	x
NS2	2917	T	C	100	V859A	x
	3364	G	A	100	R1008Q	x
NS3	4267	C	T	15	A1309V	
	4363	C	T	100	A1341V	x
	4711	G	C	99	C1457S	x
NS4B	6067	G	C	100	G1909A	x
NS5A	6729	C	A	100	L2130I	x
	7591	T	C	100	V2417A	x
NS5B	7785	A	G	99	S2482G	x
	8292	A	C	99	N2651H	x
	9045	A	C	100	I2902L	x
	9298	A	G	99	H2986R	x

[a], HCV protein in which specified change was located.

[b], Nucleotide change specified by nucleotide position and identity relating to the TNcc genome (Genbank accession number JX993348) as a reference.[26] Coding changes with allele frequency of at least 10% in one of the analyzed samples are listed.

[c], Allele frequency of the identified nucleotide change. Polyclonal passage 38.1 (TNcc-PP-38.1) virus from the later passage line was analyzed (Figure 1).

[d], Amino acid change specified by amino acid position and identity relating to the TNcc polyprotein; amino acid position numbers are identical to position numbers in relation to the reference H77 polyprotein (Genbank accession number AF009606).

[e], x indicates that the respective nucleotide change was engineered for generation of TNcc-HI.

Supplementary Table 5. Coding nucleotide changes identified in polyclonal passaged genotype 2a HCV.

HCV Protein[a]	Nucleotide change[b] J6cc position	J6cc reference	change	Allele frequency (%)[c] J6cc-PP-35	Gt 2a HCV seed stock	AA change J6cc reference[d]	AA position H77 reference[e]	Engineered construct[f] J6cc-HI
core	572	A	G	99	99	K78E	78	x
	758	G	C	81	81	V140L	140	x
E1	1325	A	T	100	100	T329S	329	x
E2	1583	A	G	14	14	N415D	415	
	1640	C	A	100	100	H434N	434	x
	2063	G	A	83	83	A575T	573	x
p7	2618	G	A	94	94	A760T	756	x
	2658	T	C	94	94	V773A	769	x
NS2	2823	T	C	93	93	V828A	824	x
	2913	A	C	17	17	E858A	854	
NS3	3777	C	T	82	82	A1146V	1142	x
	4328	A	G	100	100	I1330V	1326	x
	4839	T	C	93	93	L1500P	1496	x
NS4B	5843	G	A	82	82	V1835I	1831	x
NS5A	6338	A	G	93	93	T2000A	1996	x
	7428	T	C	100	92	L2363P	2363	x
	7439	G	C	100	100	A2367P	2367	x
	7661	T	C	45	45	C2441R	2419	
NS5B	8687	G	C	81	81	E2783Q	2761	x
	9378	T	C	95	45	L3013S	2991	x

[a], HCV protein in which specified change was located.

[b], Nucleotide change specified by nucleotide position and identity relating to the J6cc genome (Genbank accession number JQ745650) as a reference.[27] Coding changes with allele frequency of at least 10% in one of the analyzed samples are listed.

[c], Allele frequency of the identified nucleotide change. Polyclonal passage 35 (J6cc-PP-35) virus (Figure 1) and the seed stock used for vaccine production (Gt 2a HCV seed stock) were analyzed.

[d], Amino acid change specified by amino acid position and identity relating to the J6cc polyprotein (Genbank accession number JQ745650).

[e], Amino acid position relating to the H77 reference polyprotein (Genbank accession number AF009606).

[f], x indicates that the respective nucleotide change was engineered for generation of J6cc-HI.

455 Supplementary Table 6. Coding nucleotide changes identified in polyclonal passaged genotype 3a
456 HCV.

HCV protein[a]	Nucleotide change[b]			Allele frequency (%)[c]		AA change DBN3acc reference[d]	AA position H77 reference[e]	Engineered construct[f]
	DBN3acc position	DBN3acc reference	change					
				DBNcc-PP-16	Gt 3a HCV seed stock			DBNcc-HI
E2	1522	G	A	97	100	G395R	395	x
NS2	2905	G	A	96	100	A856T	849	x
	2962	G	A	97	100	G875R	868	x
NS3	5105	A	G	97	100	N1589S	1582	x
NS4B	5597	T	G	96	100	I1753S	1746	x
NS5A	7622	T	C	97	100	V2428A	2417	x
NS5B	8759	A	G	97	100	D2807G	2796	x

457

458 [a], HCV protein in which specified change was located.

459 [b], Nucleotide change specified by nucleotide position and identity relating to the DBN3acc genome
460 (Genbank accession number KX280714) as a reference.[28] Coding changes with allele frequency of at
461 least 10% in one of the analyzed samples are listed.

462 [c], Allele frequency of the identified nucleotide change. Polyclonal passage 16 (DBNcc-PP-16) virus
463 (Figure 1) and the seed stock used for vaccine production (Gt 3a HCV seed stock) were analyzed.

464 [d], Amino acid change specified by amino acid position and identity relating to the DBN3acc
465 polyprotein (Genbank accession number KX280714).

466 [e], Amino acid position relating to the H77 reference polyprotein (Genbank accession number
467 AF009606).

468 [f], x indicates that the respective nucleotide change was engineered for generation of DBNcc-HI.

469 Supplementary Table 7. Subclonal analysis of polyclonal passaged genotype 1a HCV from a later
470 passage line.

HCV protein[a]	**core**				**E1**		**E2**						
Nucleotide position[b]	373	435	463	756	1188	1435	1497	1512	1548	1571	1842	2232	2496
TNcc reference	C	G	G	C	T	T	T	A	T	C	A	A	G
Nucleotide change	A	A	C	A	C	C	C	G	C	G	G	G	A
Allele frequency (%)[c]	97	48	nd	17	nd	99	6	9	99	100	nd	nd	100
# Subclones[d]													
1	x	x				x			x	x			x
2	x	x				x			x	x			x
3	x	x			x	x			x	x			x
4	x	x				x			x	x	x		x
5	x			x		x		x	x	x		x	x
6	x	x				x		x	x	x			x
7	x	x	x	x		x	x		x	x			x
AA change TNcc reference[e]	T11N	G32S	G41A	L139I	S283P	V365A	Y386H	S391G	F403L	N410K	S501G	M631V	V719I
Engineered construct[f]	x	x				x			x	x			x

HCV protein[a]	**p7**		**NS2**			**NS3**							
Nucleotide position[b]	2650	2701	2917	2959	3364	3876	3907	4071	4138	4267	4363	4711	5130
TNcc reference	T	T	T	T	G	T	T	T	C	C	C	G	G
Nucleotide change	C	C	C	C	A	A	G	C	G	T	T	C	A
Allele frequency (%)[c]	1	92	100	8	100	nd	nd	nd	nd	15	100	99	nd
# Subclones[d]													
1		x	x		x	x					x	x	
2		x	x	x	x						x	x	
3	x	x	x		x		x	x			x	x	
4		x	x		x						x	x	
5		x	x		x						x	x	x
6		x	x		x						x	x	
7		x	x		x				x		x	x	
AA change TNcc reference[e]	V770A	V787A	V859A	V873A	R1008Q	L1179I	V1189G	Y1244H	A1266G	A1309V	A1341V	C1457S	A1597T
Engineered construct[f]		x	x		x						x	x	

HCV protein[a]	**NS4B**		**NS5A**								
Nucleotide position[b]	5575	6067	6315	6729	7288	7386	7404	7407	7422	7500	7591
TNcc reference	T	G	A	C	T	A	T	T	A	T	T
Nucleotide change	C	C	G	A	A	T	C	C	T	C	C
Allele frequency (%)[c]	1	100	nd	100	nd	nd	1	6	0.4	nd	100
# Subclones[d]											
1		x		x			x				x
2		x		x		x					x
3		x	x	x							x
4		x		x						x	x
5		x		x	x						x
6		x		x				x			x
7	x	x		x					x		x
AA change TNcc reference[e]	V1745A	G1909A	K1992E	L2130I	L2316H	T2349S	S2355P	S2356P	T2361S	S2387P	V2417A
Engineered construct[f]		x		x							x

HCV protein[a]	**NS5B**					
Nucleotide position[b]	7785	7993	8292	8779	9045	9298
TNcc reference	A	T	A	C	A	A
Nucleotide change	G	C	C	T	C	G
Allele frequency (%)[c]	99	nd	99	nd	100	99
# Subclones[d]						
1	x	x	x		x	x
2	x		x		x	x
3	x		x		x	x
4	x		x		x	x
5	x		x		x	x
6	x		x		x	x
7	x		x	x	x	x
AA change TNcc reference[e]	S2482G	V2551A	N2651H	A2813V	I2902L	H2986R
Engineered construct[f]	x		x		x	x

[a], HCV protein, in which specified change was located.

[b], Nucleotide position, identity and change relating to the TNcc genome (Genbank accession number JX993348).[26] Coding changes occurring in at least one subclone are listed.

[c], Allele frequency of the identified nucleotide change in TNcc-PP-38.1 as determined by NGS (Supplementary Table 4). nd, not detected in NGS.

[d], Number of subclone. x, indicates presence of the respective change. Grey shadings indicate which subclone fragments were used for construction of the engineered construct.

479 [e], Amino acid change specified by amino acid position, identity and change relating to the TNcc
480 polyprotein (Genbank accession number JX993348).
481 [f], x, indicates that the respective nucleotide change was engineered for generation of TNcc-HI.

482 Supplementary Table 8. Subclonal analysis of polyclonal passaged genotype 2a HCV.

HCV protein[a]	**core**										**E1**		
Nucleotide position[b]	382	383	405	527	572	758	830	845	872	908	1030	1118	1199
J6cc reference	C	A	T	C	A	G	T	T	T	T	G	A	G
Nucleotide change	A	C	C	del	C	C	C	A	G	A	C	G	T
Allele frequency (%)[c]	nd	nd	nd	nd	99	80	nd	nd	nd	nd	nd	nd	7
# Subclones[d]													
2					x	x							
10					x	x							x
11			x		x	x						x	
13					x			x					
14					x	x							
18					x	x							
19					x	x				x			
20					x	x							
22	x	x		x	x	x			x				x
26					x	x					x		
27					x	x	x						
AA change J6cc reference[e]	N14K	T15P	V22A		K78E	V140L	F164L	L169I	L178V	S190T	E230D	T260A	A287S
Engineered construct[f]					x	x							

HCV protein[a]	**E1**				**E2**									
Nucleotide position[b]	1318	1325	1418	1439	1536	1583	1590	1640	1940	2063	2126	2174	2228	2454
J6cc reference	G	A	G	G	T	A	A	C	A	G	A	C	A	T
Nucleotide change	A	T	del	C	C	G	G	A	G	A	G	A	G	A
Allele frequency (%)[c]	nd	100	nd	nd	4	12	4	99,8	8	82	2	nd	nd	3
# Subclones[d]														
2		x						x		x				
10		x						x		x				
11		x		x				x		x				
13		x						x		x				
14	x	x						x		x				
18		x			x			x		x				
19		x				x		x		x				
20		x						x		x		x		
22		x					x	x	x					x
26		x						x		x			x	
27		x	x			x		x		x	x			
AA change J6cc reference[e]	W326*	T329S	na	A367P	L399P	N415D	N417I	N434H	N532D	A575T	T596A	L612M	I630V	F705Y
Engineered construct[f]		x						x		x				

483

HCV protein[a]	**p7**					**NS2**								
Nucleotide position[b]	2616	2618	2658	2685	2735	2823	2879	2904	2913	2984	3011	3160	3212	3295
J6cc reference	A	G	T	T	T	T	T	C	A	G	G	G	A	G
Nucleotide change	G	A	C	ins	C	C	G	T	C	del	A	C	G	C
Allele frequency (%)[c]	nd	93	93	nd	nd	92	nd	nd	15	nd	nd	nd	nd	nd
# Subclones[d]														
2		x	x			x						x		
10		x	x			x			x					
11		x	x			x								
13		x	x		x	x	x			x				
14	x	x	x			x								
18		x	x	x		x								x
19		x	x			x			x				x	
20		x	x			x								
22								x			x			
26		x	x			x								
27		x	x			x			x					
AA change J6cc reference[e]	H759R	A760T	V773A	na	S799P	V828A	L846V	T855I	E858A	na	D891N	M940I	T958A	K985N
Engineered construct[f]		x	x			x								

HCV protein[a]	**NS3**												**NS4A**
Nucleotide position[b]	3702	3705	3777	3870	4173	4328	4694	4839	5154	5157	5157	5208	5432
J6cc reference	C	A	C	G	T	A	T	T	C	C	C	T	G
Nucleotide change	T	G	T	A	C	G	C	C	T	del	ins	C	A
Allele frequency (%)[c]	nd	nd	81	nd	nd	100	nd	93	nd	nd	nd	nd	0.2
# Subclones[d]													
2			x			x		x	x				x
10			x	x		x		x					
11	x		x			x		x				x	
13			x		x	x		x					
14			x			x		x					
18			x			x		x		x			
19		x	x			x		x			x		
20			x			x		x					
22			x			x		x					
26			x			x		x					
27			x			x	x	x					
AA change J6cc reference[e]	T1121I	K1122R	A1146V	R1177K	I1278T	I1330V	F1452L	L1500P	P1605L	na	na	V1623A	V1698I
Engineered construct[f]			x			x		x					

484

HCV protein[a]	NS4B								NS5A				
Nucleotide position[b]	5562	5563	5571	5787	5796	5843	6078	6210	6338	6509-6694	6563	6581	6596
J6cc reference	A	G	C	T	G	G	G	T	A		G	A	A
Nucleotide change	G	C	G	ins	A	A	A	C	G	del	T	T	G
Allele frequency (%)[c]	nd	0.1	nd	nd	nd	81	nd	nd	92	nd	nd	9	nd
# Subclones[d]													
2						x			x				
10						x	x		x				
11					x	x			x				
13						x			x		x		
14						x			x				
18				x		x			x				
19			x			x			x			x	
20	x	x				x			x			x	x
22						x			x	x			
26						x			x				
27						x		x	x				
AA change J6cc reference[e]	Q1741R	Q1741H	S1744C	na	W1819*	V1835I	G1913E	I1957T	T2000A	na	V2075L	N2081Y	I2086V
Engineered construct[f]						x			x				

HCV protein[a]	NS5A												
Nucleotide position[b]	6639	6861	6951	7026	7059	7107	7122	7187	7204	7335	7358	7428	7436
J6cc reference	A	T	G	A	A	T	A	C	G	C	A	T	G
Nucleotide change	G	C	T	C	G	C	G	T	T	G	G	C	T
Allele frequency (%)[c]	2	nd	nd	nd	nd	nd	nd	1	nd	nd	nd	92	nd
# Subclones[d]													
2					x						x	x	
10		x										x	
11								x				x	
13		x		x								x	
14			x									x	
18							x					x	
19												x	
20	x											x	
22									x			x	x
26												x	
27						x				x		x	
AA change J6cc reference[e]	H2100R	V2174A	S2204I	D2229A	D2240G	L2256P	E2261G	P2283S	W2289C	T2332R	I2340V	L2363P	G2366W
Engineered construct[f]												x	

485

HCV protein[a]	NS5A							NS5B							
Nucleotide position[b]	7439	7523	7586	7625	7640	7649	7661	7908	8209	8559	8639	8687	9017	9308	9378
J6cc reference	G	C	C	T	G	G	T	G	A	A	G	G	G	T	T
Nucleotide change	C	ins	del	C	A	C	C	A	del	G	A	C	ins	C	C
Allele frequency (%)[c]	100	nd	nd	nd	6	nd	39	nd	nd	nd	nd	81	nd	nd	95
# Subclones[d]															
2	x		x				x	x				x			x
10	x						x					x	x		x
11	x		x				x				x	x			x
13	x				x							x			x
14	x						x					x			x
18	x		x		x							x			x
19	x						x					x			x
20	x	x					x			x		del			x
22	x					x						x		x	x
26	x								x			x			x
27	x			x								x			x
AA change J6cc reference[e]	A2367P	na	na	W2429R	E2434K	D2437H	C2441R	R2523K	na	K2740R	E2767K	E2783Q	na	S2990P	L3013S
Engineered construct[f]	x											x			x

[a], HCV protein, in which specified change was located.

[b], Nucleotide position, identity and change relating to the J6cc genome (Genbank accession number JQ745650).[27] Coding changes as well as insertions (ins) or deletions (del) occurring in at least one subclone are listed.

[c], Allele frequency of the identified nucleotide change in J6cc-PP-35 as determined by NGS (Supplementary Table 5). nd, not detected in NGS.

[d], Number of subclone. x, indicates presence of the respective change. Grey shadings indicate which subclone fragments were used for construction of the engineered construct.

[e], Amino acid change specified by amino acid position, identity and change relating to the J6cc polyprotein (Genbank accession number JQ745650). na, not applicable. *, stop codon.

[f], x, indicates that the respective nucleotide change was engineered for generation of J6cc-HI.

498 Supplementary Table 9. Subclonal analysis of polyclonal passaged genotype 3a HCV.

HCV protein[a]	core				E1	E2				p7
Nucleotide position[b]	386	529	618	755	1453	1501	1522	1544	2068	2731
DBN3acc reference	T	C	G	T	G	A	G	T	G	A
Nucleotide change	C	A	A	C	T	G	A	C	A	G
Allele frequency (%)[c]	nd	nd	nd	nd	nd	7	97	nd	0.6	nd
# Subclones[d]										
1							x		x	
2				x			x			
3			x				x			
4	x	x			x		x	x		
5						x	x			
6							x			
7				x			x			
8							x			
9							x			
10							x			x
AA change DBN3acc reference[e]	I16T	P64T	W93*	L139P	A372S	I388V	G395R	L402S	E577K	T798A
Engineered construct[f]							x			

HCV protein[a]	NS2				NS3					NS4A		
Nucleotide position[b]	2905	2941	2962	3152	4093	4588	4771	5105	5158	5549	5589	5597
DBN3acc reference	G	G	G	G	G	A	T	A	C	A	A	T
Nucleotide change	A	del	A	A	C	G	ins	G	T	G	T	G
Allele frequency (%)[c]	96	nd	97	0.7	nd	0.4	nd	97	nd	0.6	0.4	96
# Subclones[d]												
1	x		x					x	x			x
2						x	x	x				x
3	x		x					x		x		x
4	x	x	x					x				x
5	x		x					x			x	x
6	x		x	x	x			x				x
7						x	x	x				x
8	x		x					x				x
9	x		x					x				x
10	x		x					x				x
AA change DBN3acc reference[e]	A856T	na	G875R	G938E	V1252L	M1417V	na	N1589S	P1607S	E1737G	Q1750H	I1753S
Engineered construct[f]	x		x					x				x

499

HCV protein[a]	**NS4B**			**NS5A**					**NS5B**				
Nucleotide position[b]	5726	5872	6427	6800	7048	7409	7465	7622	8552	8571	8612	8759	8824
DBN3acc reference	C	G	G	T	A	A	A	T	G	T	A	A	G
Nucleotide change	T	C	A	T	G	T	T	C	C	del	C	G	A
Allele frequency (%)[c]	nd	nd	nd	ins	nd	nd	nd	97	nd	nd	nd	97	3
# Subclones[d]													
1								x		x	x	x	
2				x			x	x				x	
3								x				x	
4		x						x				x	
5			x					x	x			x	
6					x			x				x	x
7				x			x	x				x	
8	x							x				x	
9						x		x				x	
10								x				x	
AA change DBN3acc reference[e]	S1796F	A1845P	G2030S	na	N2237D	K2357I	T2376S	V2428A	G2738A	na	D2758A	D2807G	E2829K
Engineered construct[f]								x				x	

[a], HCV protein, in which specified change was located.

[b], Nucleotide position, identity and change relating to the DBN3acc genomes (Genbank accession number KX280714).[28] Coding changes as well as insertions (ins) or deletions (del) occurring in at least one subclone are listed.

[c], Allele frequency of the identified nucleotide change in DBNcc-PP-16 as determined by NGS (Supplementary Table 6). nd, not detected in NGS.

[d], Number of subclone. x, indicates presence of the respective change. Grey shadings indicate which subclone fragments were used for construction of the engineered construct.

[e], Amino acid change specified by amino acid position, identity and change relating to the DBN3acc polyprotein (Genbank accession number KX280714). na, not applicable. *, stop codon

[f], x indicates that the respective nucleotide change was engineered for generation of DBNcc-HI.

512 Supplementary Table 10. Substitutions acquired during serial passage of HCV in this study and
513 previously reported cell culture infectious HCV recombinants.

AA position H77 reference[a]	Substitutions in this study[b]			Substitutions in previous studies[c]						References[d]
	TNcc-P38	J6cc-P35	DBNcc-P16	H77cc	HCV1cc	H77-JFH1	J4-JFH1	J6-JFH1	JFH1	
32	G-S								G-S	Aligeti et al., *J Virol.* 2015;22:11523-11533
78		K-E						K-E		Bungyoku et al., *J Gen Virol.* 2009;7:1681-1691
787	V-A					V-A				Scheel et al., *Proc Natl Acad Sci USA.* 2008;3:997-1002
1326		I-V			I-V					Li et al., *J Virol.* 2015;1:811-823
1496		L-P					Q-L			Gottwein et al, *Hepatology.* 2009;49:364-377
1909	G-A			G-S						Li et al., *J Virol.* 2015;1:811-823
2417	V-A		V-A	V-A						Li et al., J Virol. 2015;1:811-823

514

515 [a], Amino acid position, relating to the H77 reference polyprotein (Genbank accession number
516 AF009606).

517 [b], Amino acid substitutions identified in the TNcc-P38.1, J6cc-P35 and DBNcc-P16 polyproteins that
518 were subsequently engineered.

519 [c], Amino acid substitutions reported in previously developed cell culture infectious HCV
520 recombinants.

521 [d], literature references