UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus de São José do Rio Preto

ANA CAROLINA GOMES JARDIM

Análise das quasiespécies do vírus da hepatite C genótipo 1 por meio da região genômica NS5A

São José do Rio Preto - SP
2011

ANA CAROLINA GOMES JARDIM

Análise das quasiespécies do vírus da hepatite C genótipo 1 por meio da região genômica NS5A

Tese apresentada para obtenção do título de Doutor em Genética

Orientadora: Profa. Dra. Paula Rahal
Co-orientadora: Profa. Dra. Isabel M. V. G. Carvalho-Mello

São José do Rio Preto - SP
2011

Jardim, Ana Carolina Gomes.
Análise das quasiespécies do vírus da hepatite C genótipo 1 por meio da região genômica NS5A/ Ana Carolina Gomes Jardim. – São José do Rio Preto: [s.n.], 2011.
87 f. : il ; 30 cm.

Orientador: Paula Rahal
Co-orientador: Isabel Maria Vicente Guedes de Carvalho-Mello
Tese (doutorado) – Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Genética molecular. 2. Genômica. 3. Hepatite C. 4. Vírus – Aspectos genéticos. I. Rahal, Paula. II. Carvalho-Mello, Isabel Maria Vicente Guedes de. III. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. IV. Título.

CDU – 578.53

ANA CAROLINA GOMES JARDIM

Análise das quasiespécies do vírus da hepatite C genótipo 1 por meio da região genômica NS5A

Tese apresentada para obtenção do título de Doutor em Genética, junto ao Programa de Pós-graduação em Genética do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual "Júlio de Mesquita Filho", Campus de São José do Rio Preto.

BANCA EXAMINADORA

Profa. Dra. Paula Rahal
UNESP – São José do Rio Preto

Profa. Dra. Camila Malta Romano
USP – São Paulo

Prof. Dr. Jonny Yokosawa
UFU – Ubelândia

Prof. Dr. Maurício Lacerda Nogueira
FAMERP – São José do Rio Preto

Profa. Dra. Fátima Pereira de Souza
UNESP – São José do Rio Preto

São José do Rio Preto, 25 de fevereiro de 2011

Dedico este trabalho...

A minha mãe Rosângela e minha avó Antônia.

Obrigada pelo amor, apoio, confiança, exemplo, presença e companheirismo.
Sem vocês nada seria possível e nada teria sentido!
Amo muito vocês!

AGRADECIMENTOS

- À minha orientadora Profa. Dra. Paula Rahal por todos os ensinamentos profissionais e pessoais que me fizeram crescer e amadurecer ao longo destes anos de convivência. Obrigada pela confiança, dedicação e carinho durante esta jornada, pelo apoio em todos os momentos, e pela oportunidade de fazer parte da família genoma;
- À minha co-orientadora Profa. Dra. Isabel M. V. G. Carvalho-Mello por toda dedicação, carinho e entusiasmo. Obrigada pelos ensinamentos e por estar sempre disposta a melhor a qualidade deste trabalho;
- À minha ex-aluna de co-orientação Renata Prandini de Adum Matos, pela dedicação e pelo aprendizado recíproco durante a nossa convivência. Obrigada pela paciência e companheirismo;
- À Profa. Dra. Cláudia Carareto pelo carinho e pela prontidão em auxiliar com as análises deste trabalho. Obrigada pelos ensinamentos e pela disponibilidade;
- A amiga Cíntia pelo companheirismo de todos esses anos de amizade e cumplicidade, o que nos tornou uma dupla dinâmica em muitos aspectos. Obrigada pelas discussões científicas, pelas aulas de evolução, pelas correções de inglês, pelo otimismo, e pela torcida;
- À amiga Lilian Hiromi Tomonari Yamasaki, pela ajuda neste e outros trabalhos, pela amizade e pelas pérolas ou Miukajens que nos divertia diariamente;
- À Carol Bonfim, pela dedicação como auxiliou para esse trabalho;
- À amiga Marília, com quem sempre pude contar para me ajudar com questões profissionais e pessoais desde o início de minhas atividades científicas. Obrigada pela amizade, carinho e atenção de todos esses anos;
- Aos amigos que compartilharam a bancada, a sala de pós-graduandos, as discussões científicas, os problemas, as longas e cansativas horas de preparo de células competentes, os momentos de distração e alegria no laboratório de Estudos Genômicos. Cíntia, Marília, Renata, Lílian Yamasaki, Carol Bonfim, Paola, Bruno, Érica, Jucimara, Marina, Tasso, Natalia, André, Ana Cláudia, Mariana, Plínio, Tamires, Henrique,

Fernanda, Marcelo, Paulo, Paulinho, Sérgio, Lílian Campos e Fátima, obrigada pelo apoio e pelo crescimento pessoal que o convívio com vocês me proporcionou;

- À amiga Lenira, pela amizade, carinho e atenção;
- Aos amigos do Laboratório de Virologia da Universidade de Leeds Zsofia, Ana, Jamel, Barney, Mair, Toshana, Doug, Lisy, Carsten, Petra, Sian, Bjorn and Cherly pelo carinho e receptividade nos meses que trabalhamos juntos;
- Ao Prof. Dr. Mark Harris pela oportunidade de estagiar em seu laboratório, pelos ensinamentos, paciência e dedicação;
- Ao Dr. Maurício Lacerda Nogueira por sempre recorrer com carinho quando necessária a utilização de seu laboratório;
- À minha mãe e avó pelo amor e amizade de toda uma vida;
- À minha família que sempre acreditou em mim e me apoiou em todos os momentos, mesmo sem entender completamente as singularidades da pesquisa científica;
- Ao meu namorado Victor D'Addario pelo apoio incondicional, pela compreensão e companheirismo. Obrigada por estar sempre presente, me ajudar nos momentos difíceis e por todo carinho e conforto;
- Às queridas amigas de longa data Marília, Marcella, Carol, Cíntia, Lisandra, Danielle, Symara e Karina, por todos os anos de amizade, companheirismo e cumplicidade;
- Aos amigos Paula, Carmélia, Maysa, Mayra, Marcela, Talita, Fernando, Aninha, Gislaine, Michele, Adolfo pela amizade;
- À Lourdes, Gilberto, Márcia, Rafael e André pelo carinho e receptividade com que me receberam em sua família;
- Aos alunos da pós-graduação em genética, pela amizade, pelos momentos de tensão compartilhados, e pelo companheirismo nos ideais que batalhando juntos;
- Aos professores da pós-graduação em genética, pelos ensinamentos, pela disponibilidade e dedicação;
- Ao Prof. Dr. Carlos Roberto Ceron por todo carinho e toda confiança. Obrigada por me mostrar o caminho para a pesquisa científica;
- À Profa. Dra. Fátima Pereira de Souza, pela atenção e ensinamentos no início de minhas atividades científicas. Obrigada pela confiança e carinho;

- À coordenação da pós-graduação em genética, pela prontidão em esclarecer todas as minhas dúvidas, escutar minhas reivindicações e ajudar no que foi possível;
- À sessão de pós-graduação, especialmente à Rosemar e Silvinha, por me auxiliarem com todas as dúvidas e toda a burocracia em todos esses anos de pós-graduação. Obrigada por me ajudarem em várias situações difíceis;
- Ao Laboratório de Estudos Genômicos e ao Instituto de Biociências, Letras e Ciências Exatas, pela infra-estrutura de ensino e pesquisa, essenciais para a realização deste trabalho;
- À FAPESP pelo apoio financeiro disponibilizado, essencial para a realização deste trabalho;
- Aos meus gatos Bebê, Dinky e Théo por me fazerem companhia em todos os longos dias e noites de trabalho, mesmo quando dormiam em cima dos papers ou digitavam suas contribuições científicas nos arquivos em que eu trabalhava. A minha cachorra Babi por todos os anos de companheirismo;
- A todas as pessoas que não foram citadas, mas que participaram da minha vida e contribuíram direta ou indiretamente para minha formação pessoal e profissional, meu muito obrigada;
- A Deus, por tudo.

"Não sei o que possa parecer aos olhos do mundo, mas aos meus pareço apenas ter sido como um menino brincando à beira-mar, divertindo-me com o fato de encontrar de vez em quando um seixo mais liso ou uma concha mais bonita que o normal, enquanto o grande oceano da verdade permanece completamente por descobrir à minha frente."

Issac Newton

RESUMO

A composição de quasiespécies do vírus da Hepatite C (HCV) pode ter implicações importantes com relação à persistência viral e à resposta a terapia baseada em Interferon. A região NS5A completa foi analisada para avaliar se a composição de quasiespécies do HCV 1a/1b está relacionada à resposta ao tratamento combinado de interferon peguilado (PEG-IFN) e ribavirina. Seiscentos e noventa seqüências correspondentes a região não estrutural 5A (NS5A) completa foram geradas a partir de amostras coletadas antes, durante a após a administração da terapia de pacientes respondedores, não respondedores e respondedores ao final do tratamento. Este estudo apresenta evidências de que a homogeneidade da composição de quasiespécies, e a baixa complexidade e diversidade da região NS5A em amostras pré-terapia estão associados à resposta virológica sustentada. Portanto, a alta diversidade e complexidade de quasiespécies podem fornecer ao vírus melhores oportunidades de evadir a terapia antiviral. Análises filogenéticas não demonstraram o agrupamento das seqüências de acordo os padrões específicos de resposta ao tratamento. Contudo, o agrupamento distinto de seqüências pré e pós-terapia foi observado, sugerindo que um processo adaptativo ocorreu durante o período analisado. Adicionalmente, a dinâmica evolutiva da composição de quasiespécies demonstrou estar sob pressão seletiva purificadora ou purificadora relaxada, o que é condizente com a população de quasiespécies diversificada no pré-terapia, seguida de um aumento em freqüência de quasiespécies predominantes nas amostras pós-tratamento, provavelmente devido a conferirem alguma vantagem ao vírus. Estes resultados sugerem que a diversidade de quasiespécies da região NS5A pode ser importante para o entendimento dos mecanismos de baixa resposta virológica sustentada em pacientes com Hepatite C crônica.

Palavras-chave: Vírus da hepatite C; quasiespécies; proteína não estrutural 5A; terapia de interferon e ribavirina; análise genética e evolutiva.

ABSTRACT

The quasispecies composition of Hepatitis C virus (HCV) could have important implications with regard to viral persistence and response to interferon-based therapy. The complete NS5A was analyzed to evaluate whether the composition of NS5A quasispecies of HCV 1a/1b is related to responsiveness to combined interferon pegylated (PEG-IFN) and ribavirin therapy. Six hundred and ninety full-length NS5A sequences were generated from samples collected before, during and after treatment from virological sustained responder, non-responder and the end-of-treatment responder patients. This study provides evidence that homogeneity of quasispecies composition, low diversity and less complexity of the NS5A region pre-therapy are associated with viral clearance. Therefore, higher diversity and complexity of quasispecies could offer the virus a better opportunity of evading anti-viral therapy. Phylogenetic relationships concerning complete NS5A sequences obtained from patients did not demonstrate clustering associated with specific response patterns. However, distinctive clustering of pre/post-therapy sequences was observed, suggesting that an evolutionary process occurred during the time course examined. In addition, the evolution of quasispecies over time was subjected to purifying or relaxed purifying selection. This could explain the initial diversified composition of quasispecies at baseline, followed by an increase in the frequency of a predominant quasispecies in 'after treatment' samples of non-responders and end-of-treatment responders, probably because it offers some advantage for the virus. These results suggest that quasispecies diversity of the NS5A region could be important for elucidating the mechanism underlying treatment failure in patients infected with chronic hepatitis C.

Key-words: Hepatitis C virus; quasispecies; non-structural 5A protein; pegylated interferon and ribavirin therapy; genetic and evolutionary analysis.

SUMÁRIO

INTRODUÇÃO

1. INTRODUÇÃO

História e Patologia

Em 1989, Choo e colaboradores isolaram a partir do plasma de um chimpanzé cronicamente infectado, parte do genoma do vírus causador de hepatites não-A e não-B (NANB). Este vírus foi então denominado de vírus da Hepatite C (HCV) e foi associado à cerca de 90% dos casos de hepatite NANB (CHOO et al., 1989; CHOO et al., 1990). Desde sua identificação, tornou-se evidente que um considerável número de pessoas está infectada cronicamente com este vírus. O HCV é o principal causador de doenças hepáticas e um dos mais importantes problemas de saúde pública mundial (MUNIR et al., 2010).

O fígado é o órgão alvo primário da infecção pelo HCV. A infecção aguda é frequentemente assintomática ou associada a sintomas leves e não específicos, dificultando o diagnóstico precoce. Somente 15 – 30% dos indivíduos infectados eliminam o vírus durante a fase aguda, sendo que os demais pacientes desenvolvem Hepatite C crônica, uma tendência característica da infecção pelo HCV (CHISARI, 2005).

A hepatite C crônica pode causar fibrose hepática severa e resultar no desenvolvimento de cirrose hepática em 1/5 dos indivíduos infectados após aproximadamente 25 – 30 anos . Conseqüentemente, está associada com um aumento do risco de falência hepática e carcinoma hepatocelular (MORENO-OTERO, 2005; VOLK, 2010). A infecção crônica pelo HCV é atualmente a maior causa de indicação de transplante de fígado, compreendendo 40-50% de indivíduos que esperam a realização de transplante e pacientes que já realizaram um transplante prévio (BROWN, 2005; DE MARTIN et al., 2010).

Epidemiologia e Transmissão

A estimativa da prevalência global da infecção por HCV é de 170 milhões de indivíduos infectados, com aproximadamente 3 milhões de novas infecções a cada ano (LAVANCHY, 2009). No Brasil, a estimativa é de 1.5 a 1.7% para a população em geral e doadores de sangue (LYRA; FAN; DI BISCEGLIE, 2004), e de 17.7 a 25.6% para pacientes co-infectados com HIV (REICHE et al., 2008). Nas diferentes regiões brasileiras as prevalências são: 0,9 a 2,4% no Norte, 1,7 a 3,4% no Nordeste, 1,0 a 1,4% no Centro-Oeste, 0,8 a 2,8% no Sudeste, e 1,1 a 2,1% no Sul (CAMPIOTTO et al., 2005).

Durante décadas, a fonte de infecção por HCV mais importante foi a transmissão por exposição parenteral de sangue contaminado, ou pelo uso ilícito de drogas injetáveis. A transmissão via transfusional foi significantemente reduzida como conseqüência da introdução dos testes de triagem de doadores de sangue para anti-HCV no início da década de 1990 e pela utilização de procedimentos de inativação viral na produção de fatores coagulantes sanguíneos. Atualmente, o uso de drogas injetáveis é a principal via de transmissão do HCV, sendo responsável por mais de 40% das infecções (Kew et al., 2004). No Brasil, 5.8 a 36.2 % dos usuários de drogas estão infectados pelo HCV (LOPES et al., 2009).

Os fatores de risco mais frequentemente relacionados à transmissão de HCV são as transfusões sanguíneas, usuários de drogas injetáveis, injeções terapêuticas não seguras, hemodiálise ou tatuagem (ALTER, 2007). No entanto, 15-40% dos pacientes infectados com HCV não apresentam um fator de risco evidente. Nesses casos, a transmissão do vírus por via ocupacional, perinatal, intrafamiliar e/ou sexual é questionada (ACKERMAN, 1998; DE CARVALHO-MELLO et al., 2010).

A transmissão ocupacional da infecção por HCV é altamente relacionada com trabalhadores da saúde que tem sustentado uma contaminação por injúrias com agulhas, sendo a taxa de infecção nessas circunstâncias entre 0 a 3% (CHUNG et al., 2003). É estimado que a transmissão de HCV através de via perinatal ocorra em 2,7 a 6% dos nascimentos de mães infectadas (DAL MOLIN et al., 2002; FERRERO et al., 2003). Estudos têm demonstrado a transmissão por contato intrafamiliar em grupos com fatores de transmissão não convencionais (CAVALHEIRO NDE et al., 2009; HYDER et al., 2009). A taxa de transmissão foi de 8,9% e sugeriu que o uso compartilhado de objetos domésticos pode ter um importante na transmissão do HCV (LA TORRE et al., 2006).

Em contraste com a consistente evidência da transmissão parenteral, a taxa de transmissão sexual permanece controversa. Vários estudos sugerem que o risco de transmissão sexual do HCV é mínimo ou inexistente, com sua incidência variando de 0 a 3% (BROOK, 2002; FLAMM, 2003).

Devido à grande variedade de atividades humanas que envolvem a exposição parenteral de sangue ou derivados de sangue, existem muitos modos biologicamente plausíveis de transmissão do HCV, além dos associados claramente com a infecção. Esses modos de transmissão incluem procedimentos cosméticos e religiosos, praticas culturais, tais como tatuagem e piercing, barbearia comercial, circuncisão e acupuntura (SHEPARD; FINELLI; ALTER, 2005).

Um estudo recente avaliou fatores epidemiológicos que poderiam influenciar a transmissão do HCV, e sugeriu a associação de métodos de transmissão ao comportamento social, idade e subtipo viral. Neste estudo, foi reportado que a transmissão do vírus foi observada mais frequentemente em pacientes mais jovens e com alta conectividade social (ROMANO et al., 2010).

O vírus da Hepatite C

O vírus da hepatite C (HCV) é um vírus de RNA fita simples, com polaridade positiva, pertencente à família Flaviviridae (KUIKEN; SIMMONDS, 2009). A partícula viral mede aproximadamente 50 nm. É formada por um envelope viral derivado das membranas do hospedeiro, onde estão inseridas as glicoproteínas virais E1 e E2, um capsídeo protéico formado por proteínas do core, e pelo genoma viral constituído de uma molécula de RNA fita simples linear, constituída de 9600 nucleotídeos, e de polaridade positiva (KAITO et al., 1994) (Figura 1).

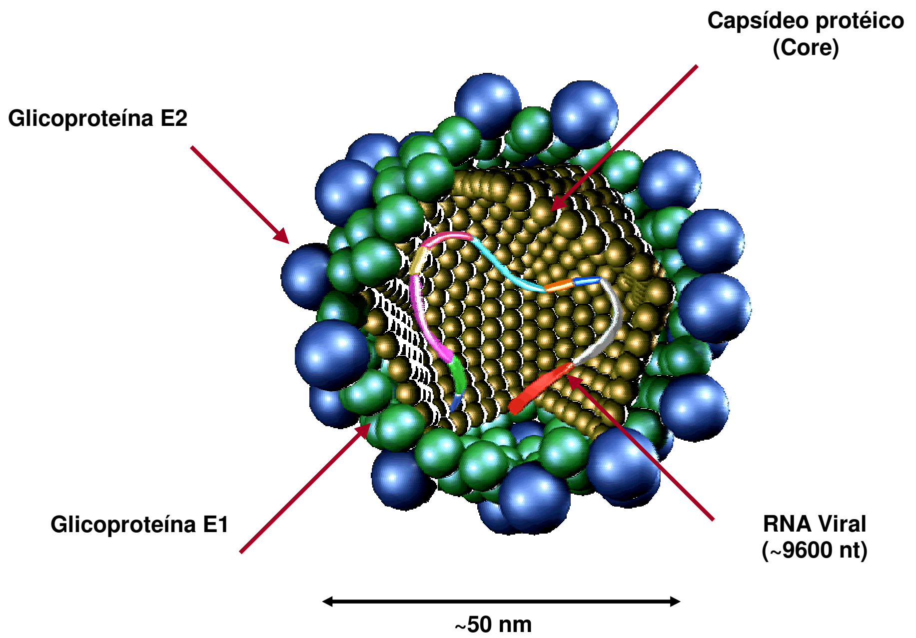


Figura 1. Morfologia do vírus da hepatite C: partícula viral. Fonte: adaptada de James, 2001.

O genoma do HCV possui uma única fase aberta de leitura (“open reading frame” ORF), flanqueada por regiões não traduzidas (“untranslated region” UTR) nas extremidades 5’ e 3’, com aproximadamente 340 e 230 nucleotídeos, respectivamente (PENIN et al., 2004). A região 5’UTR possui seis domínios estruturais em forma de alça (*stem loop*), dos quais os domínios II, III e IV formam o sítio de entrada do ribossomo (IRES "internal ribossomal entry site"), que facilita a tradução do RNA do HCV. Esta região também contém sinais de replicação essenciais para a formação de fitas de RNA de polaridade negativa, que tem função de intermediário replicativo. A região 3'UTR apresenta uma estrutura tripartida e é essencial para replicação. É composta de uma seqüência variável, seguida de uma região poli U (polipirimidina), e uma seqüência única de 98 nucleotídeos que é altamente conservada (TANG; GRISE, 2009) (Figura 2A).

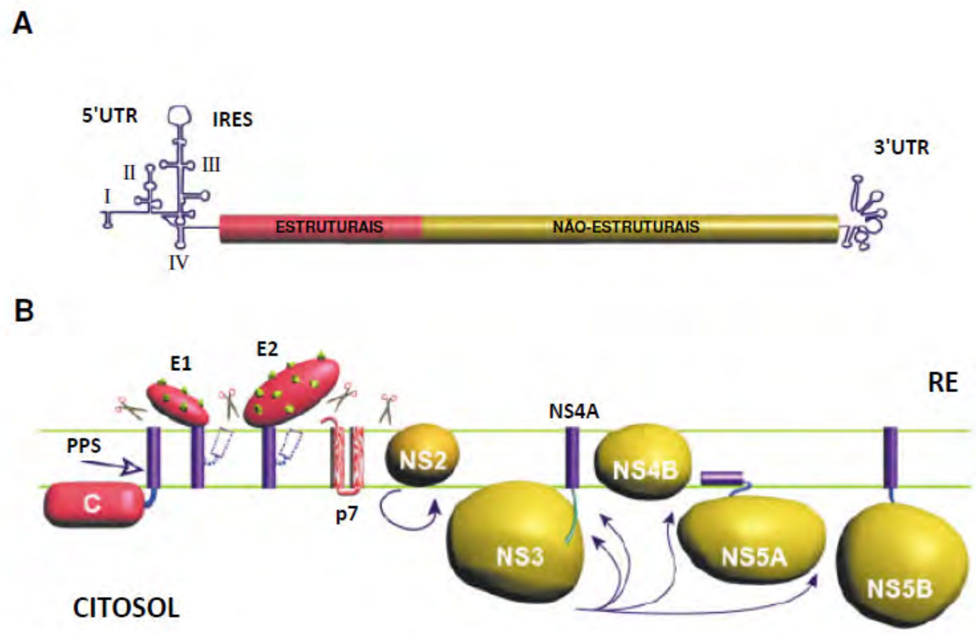


Figura 2. Estrutura genômica e proteômica do HCV. (A) Organização do genoma do HCV. Região aberta de leitura única flanqueada por regiões não traduzidas 5'UTR e 3'UTR. (B) Processamento da poliproteína e localização das 10 proteínas virais na membrana do retículo endoplasmático. Tesouras indicam os sítios de processamento de peptídeos de sinal no RE; seta cíclica, processamento autocatalítico da junção NS2/NS3; setas longas, sítios de processamento pelo complexo protease NS3/NS4A; seta intramembrânica, processamento por peptidases de sinal. Fonte: adaptada de Penin et al., 2004.

A tradução da ORF do genoma do HCV produz uma poliproteína precursora de aproximadamente 3000 aminoácidos, que é processada proteoliticamente em 10 proteínas virais por proteases celulares e viral. A região amino-terminal da poliproteína codifica as proteínas estruturais do vírion: a proteína do core (C), e as glicoproteínas E1 e E2. Após a região estrutural, uma pequena proteína integral de membrana é traduzida, a p7. Em seguida, encontram-se as proteínas não estruturais (NS) NS2, NS3, NS4A, NS4B, NS5A e NS5B, que estão envolvidas nos processos intracelulares do ciclo de replicação do vírus. Peptidases de sinal celulares processam proteoliticamente peptídeos de sinal entre as proteínas C/E1, E1/E2 e E2/p7, e também a junção p7/NS2. O processamento na região NS ocorre pela ação de duas enzimas virais: a NS2 e a NS3-4A serino protease (LINDENBACH; RICE, 2005; SHARMA, 2010) (Figura 2B).

A estratégia replicação do HCV é parecida com outros vírus de RNA de polaridade positiva. O primeiro estágio da infecção pelo HCV consiste da ligação do vírus à superfície da célula hospedeira, seguido da interação específica entre as glicoproteínas do envelope viral e receptores celular, como o CD81, a mais bem caracterizada molécula receptora para a entrada do HCV da célula (TANG; GRISE, 2009). Posteriormente, o nucleocapsídeo do HCV é liberado no citoplasma celular como resutado da fusão entre as menbranas virais e celulares. Após a entrada do vírus e liberação do material genético, o IRES promove a iniciação da tradução da poliproteína que segue o processamento das proteínas virais. As proteínas estruturais se associam (core) ou se integram (E1, E2 e p7) com a membrana do retículo endoplasmático (RE) (Figura 3) e formam oligômeros funcionais que promoverão a montagem da nova partícula viral. As proteínas não estruturais se associam do lado citoplasmático da membrana do RE onde interagem entre si e com as proteínas hospedeiras para formar a maquinaria de replicação viral. Essa maquinaria usa seu próprio genoma como molde para transcrição de fita complementar negativa de RNA. A fita negativa ou dupla fita, por sua vez, serve como uma molécula replicativa intermediária na síntese de uma nova molécula de RNA de polaridade positiva que pode ser usada para tradução, replicação ou então ser empacotada para constituir novos vírus (DE FRANCESCO et al., 2003; CHEVALIEZ; PAWLOTSKY, 2006).

Diversidade genética do genoma viral

A variabilidade genética do HCV existe em diferentes níveis. A comparação de seqüências nucleotídicas de variantes de HCV obtidas de diferentes indivíduos, em regiões geográficas diversas, demonstrou a existência de pelo menos 7 grupos geneticamente distintos (genótipos 1 a 7). Considerando o genoma completo, os 7 grupos diferem em aproximadamente 30 % de sítios de nucleotídeos, com maior variabilidade em algumas regiões, como E1, E2 e V3, e a região de menor variabilidade no genoma do HCV é encontrada na 5'UTR. Cada um dos sete genótipos possui vários subtipos, identificados por letras minúsculas (a, b, c etc) (SIMMONDS et al., 2005; MURPHY et al., 2007; KUIKEN; SIMMONDS, 2009).

Os genótipos 1, 2 e 3 apresentam distribuição universal, mas a prevalência de cada um destes genótipos varia de acordo com a área geográfica. O genótipo 1 é o mais prevalente mundialmente, sendo que a sua freqüência varia de 40 a 80%, dependendo da região. O genótipo 4 é endêmico no Egito e encontrado em outros países da África e Oriente Médio. Genótipos 5 e 6 são encontrados na África do Sul e Hong Kong, respectivamente (SIMMONDS et al., 1994; SIMMONDS, 2004; MARTINS et al., 2006). No Brasil, o padrão de distribuição dos genótipos segue o padrão mundial, sendo o genótipo 1 o mais prevalente (65%), seguido do genótipo 3 (30%), genótipo 2 (4%), e outros genótipos (1%) (Campiotto et al., 2005; Martins et al., 2006).

Como outros vírus de RNA, o HCV existe em um hospedeiro como um conjunto de variantes distintas geneticamente, mas altamente relacionadas, que são referidas como quasiespecies (MARTELL et al., 1992; DOMINGO et al., 2006). As quasiespecies diferem em sítios nucleotídicos no genoma viral em menos de 10% (ZHOU et al., 2007). Múltiplas variantes genômicas presentes simultaneamente e a alta taxa em que novas variantes são geradas, estão relacionadas à falta de atividade corretiva da NS5B RdRp durante a replicação do genoma viral e a abundância de vírions que são produzidos diariamente. Estes dados indicam que a taxa de mutação no HCV é de 1.5 – 2.0 X 10^{-3} substituições de bases, por sítio do genoma, por ano (LE GUILLOU-GUILLEMETTE et al., 2007).

Como conseqüência das substituições incorporadas no genoma do HCV, uma proporção substancial de quasiespecies será defectiva devido ao potencial deletério de determinadas mutações. Em contraste, mutações não deletérias que são acumuladas durante os ciclos de replicação, são transmitidas para a progênie viral e poderão conferir vantagens ou desvantagens para cada quasiespecies, de acordo com o ambiente replicativo (PAWLOTSKY, 2003). A variação nas quasiespecies representa um grande problema para os indivíduos infectados, devido às implicações do potencial adaptativo do HCV na evasão e controle da resposta do hospedeiro à infecção, e na sensibilidade diferencial à terapia baseada em IFN (GALE; FOY, 2005).

A persistência viral pode ser determinada por muitos fatores que incluem alta heterogeneidade do vírus. A alta diversidade do HCV pode favorecer e proporcionar melhores chances na emergência de variantes resistentes ao tratamento, e vários estudos relacionaram a composição de quasiespecies com resistência viral ao tratamento. Entretanto, esses dados ainda são controversos, pois diferentes resultados foram obtidos por meio de diferentes metodologias e regiões virais distintas (TOYODA et al., 1997; PAWLOTSKY, PELLERIN; et al., 1998; UEDA et al., 2004; PUIG-BASAGOITI et al., 2005; SALMERON et al., 2006; JAIN et al., 2009; SALUDES et al., 2010).

Proteína Não-estrutural 5A

A NS5A é uma fosfoproteína multifuncional de aproximadamente 447 resíduos, que pode ser encontrada em uma forma basal fosforilada (56 kDa) e hiperfosforilada (58 kDa). A hiperfosforilação é um processo dependente de resíduos de serina na região central (224, 228 e 231) e da associação entre a NS5A e a NS4A nos resíduos 162-166 (TANJI, HIJIKATA; et al., 1995; TANJI, KANEKO; et al., 1995; ASABE et al., 1997). A proteína completa é dividida em três domínios intercalados por sequências de baixa complexidade: o domínio I (resíduos 01 a 213), o domínio II (resíduos 250 a 342) e o domínio III (resíduos 356 a 447) (TELLINGHUISEN et al., 2004; TELLINGHUISEN; MARCOTRIGIANO; RICE, 2005) (Figura 3).

Os primeiros 30 aminoácidos no aminoterminal da NS5A são correspondentes a um domínio de ligação à membrana e pode ser responsável pela localização da proteína no retículo endoplasmático. Esta região foi predita formar uma alfa-hélice anfipática altamente conservada e ser necessária e suficiente para mediar a associação da NS5A com a membrana do Retículo endoplasmático (MACDONALD; HARRIS, 2004; HE; A.; TAN, 2006). A função de ancoragem à membrana é essencial para a replicação viral (SAUTER et al., 2009). Os 180 aminoácidos seguintes formam o restante do domínio I, e constitui uma região altamente conservada e apresenta a estrutura do sítio de ligação ao zinco, processo essencial para a replicação do HCV (GALE et al., 1998; TELLINGHUISEN et al., 2004) (Figura 3).

Os domínios II e III apresentam enovelamento não-definido e não são tão bem caracterizados quanto o domínio I (LIANG et al., 2007; HANOULLE et al., 2009). Proteínas sem enovelamento demonstram alta flexibilidade, característica que pode permitir a adaptação da estrutura local ou global ao ambiente celular. Essa flexibilidade intrínseca da estrutura permite o reconhecimento de um grande número de ligantes sem comprometer a especificidade protéica (WRIGHT; DYSON, 1999). O domínio III indica não ser essencial para a o processo de replicação do RNA viral, mas foi demonstrado que participa da montagem do vírion. Além disso, foram identificados diversos sítios de fosforilação neste domínio (APPEL et al., 2008; HUGHES et al., 2009).

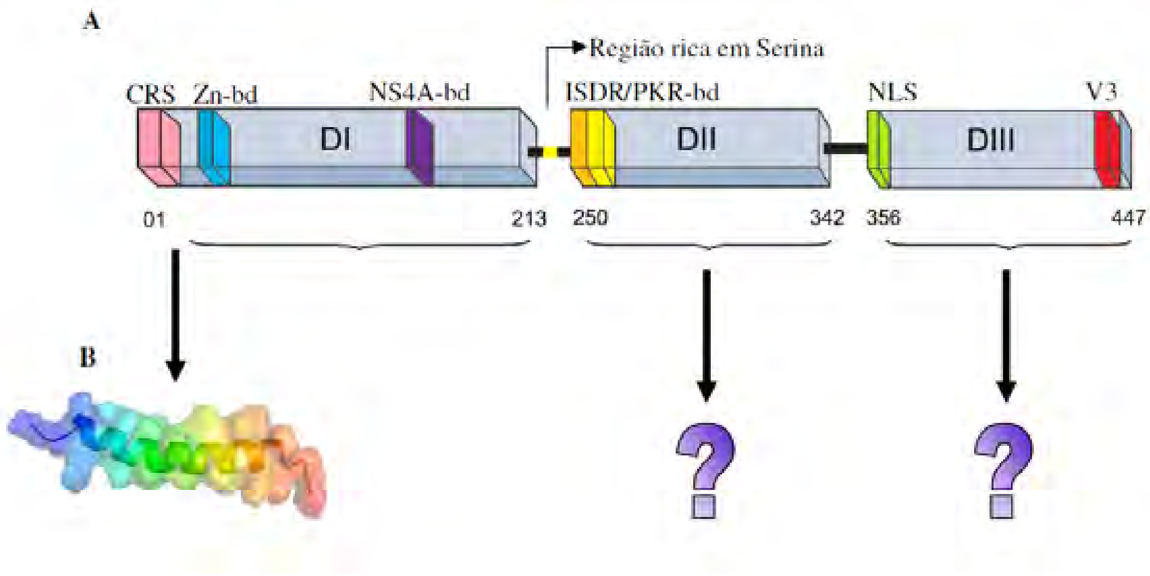


Figura 3. A) Esquema estrutural da NS5A, mostrando as principais regiões conhecidas. CRS: âncora de membrana, Zn-bd: região de ligação ao Zinco, NS4A-bd: região de ligação à NS4A, ISDR: região de sensibilidade ao Interferon, PKR-bd: região de ligação à PKR, NLS: sinal de localização nuclear. Os números estão de acordo com o protótipo HCV-J (HOFMANN; ZEUZEM; SARRAZIN, 2005). B) Esquema em *cartoon* e superfície da CRS, de acordo com o arquivo do PDB 1R7C.

Um evento pós-traducional observado na NS5A é o processamento proteolítico para ativar um sinal de localização nuclear NLS (aa 2326-2334 do genoma completo ou aa 354-362 da NS5A) (Figura 3). Embora a proteína NS5A completa esteja localiza no citoplasma, formas da NS5A de 31 kDa (aa 2127-2361 do genoma completo ou aa 155-389 da NS5A), resultante de clivagens por caspases celulares, foram localizadas no núcleo, sugerindo que estas formas da NS5A podem migrar para dentro do núcleo e agir como ativadores transcricionais potentes. Outros estudos indicam que a presença da região N-terminal de 27 aa, denominada sinal de retenção citoplasmática CRS (aa 1973-1999 do genoma completo ou 1-27 da NS5A) (Figura 3) é capaz de manter a proteína NS5A no citoplasma, mostrando uma atividade dominante sobre a NLS (Reyes, 2002).

A proteína NS5A é a proteína mais relacionada à resistência ao IFN. Seqüências genômicas completas do HCV de pacientes respondedores e não-respondedores ao IFN-α (Interferon-α) foram analisadas, resultando na identificação de uma região heterogênea na NS5A descrita como região determinante de sensibilidade ao Interferon (ISDR) (Figura 3). O aumento do número de mutações nesta região foi correlacionado ao aumento da taxa de resposta virológica sustentada, sugerindo uma possível relação da NS5A em conferir sensibilidade ou resistência à terapia com IFN-α (ENOMOTO et al., 1995; ENOMOTO et al., 1996). Entretanto, o fator preditivo das mutações na ISDR na resposta ao tratamento com IFN foi questionado por outros estudos, que observaram resultados contrastantes (CHAYAMA et al., 1997; KOMATSU et al., 1997; KUROSAKI et al., 1997). Posteriormente, dois estudos baseados na análise de 675 e 1230 seqüências ISDR publicadas individualmente em banco de dados demonstraram uma forte correlação da NS5A ISDR com a resposta ao IFN (WITHERELL; BEINEKE, 2001; PASCU et al., 2004). A resistência ao IFN seria explicada pela habilidade da NS5A ligar-se a quinase indutora de interferon ativada por RNA dupla fita (PKR), levando a sua inibição (GALE et al., 1997).

Apesar da intensiva investigação no papel da NS5A na resposta ao tratamento, uma resposta conclusiva ainda não foi obtida. Alguns estudos reportaram que a baixa complexidade e variabilidade da região NS5A completa, ou sub-regiões, de variantes virais observadas em amostras coletadas antes do início do tratamento de pacientes infectados cronicamente com C, estão associadas à resposta virológica sustentada (PUIG-BASAGOITI et al., 2005; JAIN et al., 2009). Jain e colaboradores, concluíram em seu estudo que o efeito antiviral inicial do interferon é determinado, pelo menos em parte, pela composição viral no início da terapia, quando analisada a região NS5A (JAIN et al., 2009). Por fim, evidencias sugerem que certas características do HCV, como a alta variabilidade genética viral da região

NS5A, podem ser indicativas da resistência ao tratamento baseado em IFN (PAWLOTSKY, GERMANIDIS; et al., 1998).

Tratamento

O Interferon α (IFN-α) foi o primeiro tratamento que mostrou ter um efeito benéfico nos pacientes com hepatite C crônica. Porém, a administração de IFN-α como monoterapia apresentou uma taxa de resposta virológica sustentada em 16 – 20% dos pacientes tratados, e em 35 - 40% dos pacientes submetidos a tratamento combinados de IFN e ribavirina (FELD; HOOFNAGLE, 2005; PAWLOTSKY, 2006).

Posteriormente, terapias baseadas na administração de interferons modificados pela adição de uma molécula de polietileno glicol demonstraram ser mais eficientes que as terapias que envolvem a utilização do interferon convencional. O maior efeito do PEG-IFN é o retardamento na eliminação da droga, possibilitando a manutenção de uma concentração estável no sangue, com a administração da droga uma vez por semana. (MORENO-OTERO, 2005; HAYASHI; TAKEHARA, 2006).

Atualmente, a combinação de PEG-IFN e ribavirina é a melhor terapia para o tratamento da hepatite C crônica. Em estudos com pacientes cronicamente infectados, que não apresentavam cirrose hepática, a taxa de resposta virológica sustentada foi de 76 – 84% em pacientes com infecção pelo genótipo 2 ou 3 do HCV, e de 42 – 52% em pacientes infectados com HCV do genótipo 1 (WOHNSLAND; HOFMANN; SARRAZIN, 2007; KLENERMAN; FLEMING; BARNES, 2009). Entretanto, o tratamento com PEG-IFN é caro e apresenta efeitos colaterais diversos levando a complicações durante o tratamento. Além disso, devido ao alto custo da terapia, o IFN convencional ainda se mantém como terapia principal em países subdesenvolvidos (MUNIR et al., 2010).

A resposta virológica sustentada (SVR, *sustained virological response*) é definida pela ausência de detecção de RNA do HCV por no mínimo 6 meses após o término da terapia, com limite de detecção de 50 unidades internacionais/mL (LINDSAY, 1997). Outro padrão de resposta a terapia é a resposta ao final do tratamento (ETR, *end of treatment response*) que ocorre em 10 - 25 % dos pacientes, e é determinada pela ausência de RNA de HCV ao final do tratamento com posterior detecção no período pós-tratamento. Não respondedores (NR, *non-responders*) compreendem um terço dos pacientes infectados tratados, e são pacientes no qual é observada a detecção continua do RNA de HCV durante e após o tratamento (FELD; HOOFNAGLE, 2005; HAYASHI; TAKEHARA, 2006) (Figura 4).

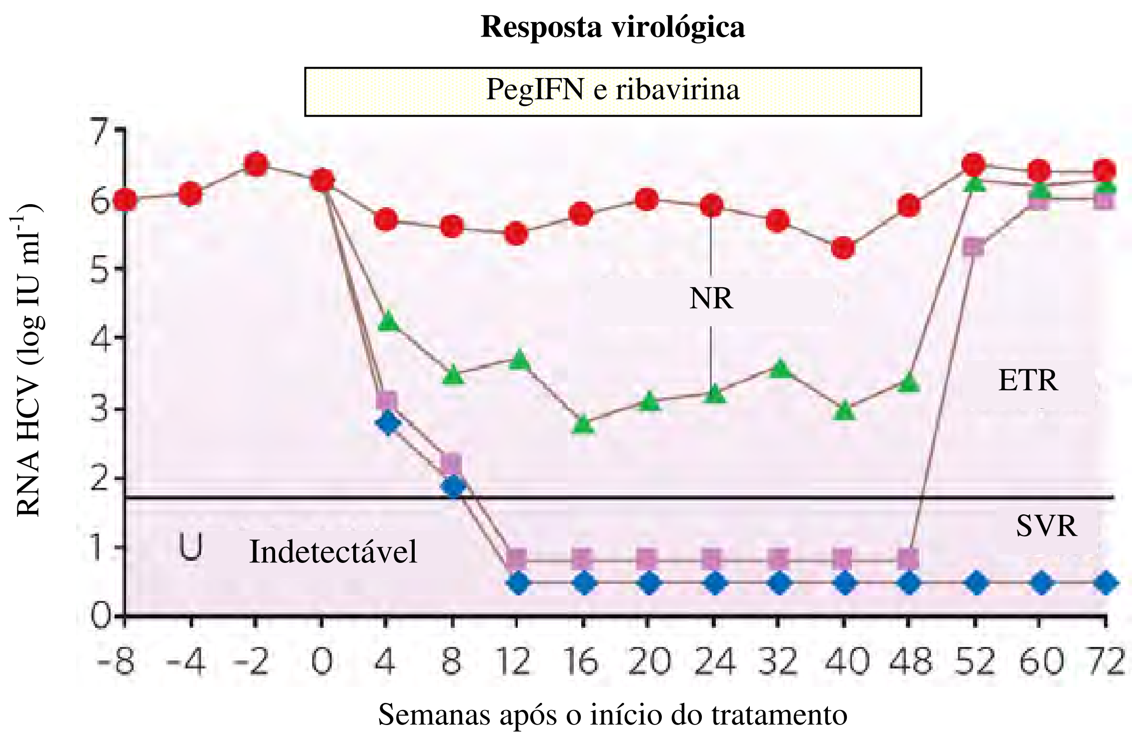


Figura 4. Resposta virológica à terapia de peginterferon e ribavirina. SVR: resposta virológica sustentada, NR: não-respondedor, ETR: Resposta ao final do tratamento. Fonte: adaptada de Feld & Hoofnagle, 2005.

Alguns fatores estão relacionados com a resposta à terapia baseada em IFN. Estes fatores compreendem características do hospedeiro, tais como sexo, idade, peso corporal, raça, co-infecção com vírus da hepatite B, duração da infecção e doença avançada do fígado, bem como características virais, tais como genótipo, carga viral e quasiespecies (FELD; HOOFNAGLE, 2005; MORENO-OTERO, 2005; SALMERON et al., 2006; WOHNSLAND; HOFMANN; SARRAZIN, 2007).

Embora um melhor conhecimento sobre os aspectos biológicos do HCV tenha ocorrido nos últimos anos, uma terapia eficaz para o tratamento da hepatite C permanece uma problemática para a maioria dos pacientes infectados (JAWAID; KHUWAJA, 2008; CHATTOPADHYAY et al., 2009). Estratégias como a formulação de novas moléculas de IFN como o Albuferon, formada pela fusão de IFNα e albumina humana, inibidores da protease NS3/4A e a polimerase NS5B viral, interferência de RNA etc estão em desenvolvimento (SHIMAKAMI; LANFORD; LEMON, 2009; MALLET; VALLET-PICHARD; POL, 2010). Entretanto, a natureza altamente replicativa e a capacidade viral do HCV em sustentar mudanças nucleotídicas no genoma viral de forma eficiente, produzindo variantes resistentes à drogas apresentam-se como grandes barreiras para o desenvolvimento de terapias bem sucedidas para o tratamento da hepatite C (KITAZATO; WANG; KOBAYASHI, 2007).

OBJETIVO

2. OBJETIVO

Analisar a dinâmica evolutiva das quasiespécies do HCV genótipo 1 por meio da região NS5A em amostras de pacientes tratados com PEG-IFN e ribavirina, e avaliar a importância da diversidade genética na resposta ao tratamento.

ARTIGO CIENTÍFICO

ARTIGO CIENTÍFICO

Quasispecies variability in the NS5A region of Hepatitis C virus genotype 1 is indicative of non-response to peginterferon and ribavirin therapy

---

Artigo científico submetido na revista científica *PLoS ONE*

(eISSN-1932-6203)

TITLE

Quasispecies variability in the NS5A region of Hepatitis C virus genotype 1 is indicative of non-response to peginterferon and ribavirin therapy

AUTHORS

Ana Carolina Gomes Jardim[a], Cíntia Bittar[a], Renata Prandini Adum de Matos [a], Lílian Hiromi Tomonari Yamasaki[a], Rafael Alves da Silva[b], João Renato Rebello Pinho[c], Roberta Maria Fachini[d], Claudia Marcia Aparecida Carareto[a], Isabel Maria Vicente Guedes de Carvalho-Mello[b], Paula Rahal[a*]

[a]Departament of Biology, Institute of Bioscience, Language and Exact Science, São Paulo State University, São José do Rio Preto, SP, Brazil

[b]Viral Imonology laboratory, Butantan Institute, São Paulo, SP, Brazil

[c]Departament of Gastroenterology, São Paulo Institute of Tropical Medicine, School of Medicine, University of São Paulo, SP, Brazil.

[d]Department Hepatology, São José do Rio Preto School of Medicine, SP, Brazil

*correspondence to: Paula Rahal, PhD, Departamento de Biologia, Instituto de Biociências, Letras e Ciências Exatas -Universidade Estadual Paulista, Rua Cristóvão Colombo, 2265, 15054-000, São José do Rio Preto-SP, Brasil. Fax: +55-17-32212390. E-mail: prahal@ibilce.unesp.br

FUNDING: This study was funded by "Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP" project with process number 07/52073-0 and grant with process number 06/60012-9. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

COMPETING INTERESTS: The authors have declared that no competing interests exist.

ABSTRACT

The quasispecies composition of Hepatitis C virus (HCV) could have important implications with regard to viral persistence and response to interferon-based therapy. The complete NS5A was analyzed to evaluate whether the composition of NS5A quasispecies of HCV 1a/1b is related to responsiveness to combined interferon pegylated (PEG-IFN) and ribavirin therapy. Six hundred and ninety full-length NS5A sequences were generated from samples collected before, during and after treatment from virological sustained responder, non-responder and the end-of-treatment responder patients. This study provides evidence that homogeneity of quasispecies composition, low diversity and less complexity of the NS5A region pre-therapy are associated with viral clearance. Therefore, higher diversity and complexity of quasispecies could offer the virus a better opportunity of evading anti-viral therapy. Phylogenetic relationships concerning complete NS5A sequences obtained from patients did not demonstrate clustering associated with specific response patterns. However, distinctive clustering of pre/post-therapy sequences was observed, suggesting that an evolutionary process occurred during the time course examined. In addition, the evolution of quasispecies over time was subjected to purifying or relaxed purifying selection. This could explain the initial diversified composition of quasispecies at baseline, followed by an increase in the frequency of a predominant quasispecies in ‘after treatment’ samples of non-responders and end-of-treatment responders, probably because it offers some advantage for the virus. These results suggest that quasispecies diversity of the NS5A region could be important for elucidating the mechanism underlying treatment failure in patients infected with chronic hepatitis C.

INTRODUCTION

Hepatitis C virus (HCV) is the major etiological agent of chronic hepatitis worldwide and approximately 130 million people are infected with this virus [1]. Chronic infection frequently progresses to liver cirrhosis with an elevated risk for the development of hepatocellular carcinoma [2,3].

The only effective treatment for chronic hepatitis C is based on interferon (IFN) or pegylated IFN in combination with ribavirin (RBV), leading to a sustained virological response in approximately 50% of patients infected with genotypes 1a/1b [4,5]. Several host parameters, disease characteristics and virus-related factors are relevant to the possibility of permanent viral clearance after therapy [6,7].

A member of the family *Flaviviridae*, HCV is an enveloped virus with a positive, single-stranded RNA genome approximately 9.5 kb in length, encoding a single polyprotein of approximately 3000 amino acids that is co- and post-translationally cleaved by viral and cellular proteases into structural and non structural proteins [8].

HCV is classified into seven genotypes, HCV-1 to -7, with each genotype being further subdivided into subtypes such as HCV-1a and 1b [9,10]. Furthermore, in infected individuals, HCV circulates as a population of several closely related viral variants referred to as "quasispecies" [11,12]. New variants are continuously generated during viral replication as a result of errors made by the viral RNA-dependent RNA polymerase, which lacks proofreading activity, during rapid replication [13]. The quasispecies nature of HCV could have important implications for viral persistence, pathogenicity and resistance to anti-viral agents [14,15].

The non-structural 5A (NS5A) protein is implicated in interferon resistance [16]. Enomoto et al. [17,18] suggested that the genetic heterogeneity of a specific domain of the NS5A region of HCV, termed the IFN sensitivity determining region (ISDR), was related to treatment responses in Japanese patients with HCV genotype 1b infection. This is a controversial issue, but analysis of published information supports the hypothesis that a relationship exists between NS5A heterogeneity and response to therapy [19,20,21,22].

NS5A is an RNA binding phosphorylated protein comprising three domains separated by trypsin-sensitive low complexity sequences (LCS I and LCS II) and an N-terminal amphipathic alpha-helix that anchors the protein to intracellular membranes [23,24,25]. Full-length NS5A protein appears to be located exclusively in the cytoplasm. However, N-terminal deleted forms have been found in the nucleus, suggesting that they can move to the nucleus and act as potent transcriptional activators [26,27,28].

The NS5A region has generated much interest owing to its association with responses to therapy. In the present study, the evolution of the complete NS5A region of HCV genotypes 1a and 1b was examined in patients undergoing pegylated IFN plus ribavirin therapy. The results demonstrate that quasispecies diversity and complexity of the NS5A region could help elucidate the mechanism underlying treatment failure in patients infected with chronic hepatitis C.

RESULTS

Patient characteristics

The characteristics of the patients are presented in Table 1. The population was predominantly male (NR 75% vs. SVR 67% vs. ETR 100%); 55% of patients were infected with HCV genotype 1a (NR 25% vs. SVR 67% vs. ETR 75%) and 45% of patients were infected with HCV genotype 1b (NR 75% vs. SVR 33% vs. ETR 25%). The mean age of the patients was 43 ± 13.9 years (NR 46.3 ±10.6 vs. SVR 39.7 ± 2.6 vs. 42.3 ± 6.2). No statistically significant differences were observed between groups.

The levels of viremia in baseline samples did not differ significantly among the patient groups (NR 6.52 vs. SVR 6.39 vs. ETR 6.17). Patients who did not respond to therapy had a significant decrease in viremia after 12 weeks of treatment (4.99 ± 0.94 log IU/ml; *p=0.002*), with recovering levels in samples collected after therapy was complete (6.30 – 6.86 log IU/ml). The viral load values varied among the samples from end of treatment responders. The mean viremia levels for these patients ranged from 4.99 to 5.90 log IU/ml (Table 1).

Nonsense mutations in the NS5A region are observed in vivo

The dataset of 690 entire NS5A sequences was submitted for genetic analysis. Eight nonsense mutations were detected in the NS5A sequence, two of them being common to more than one sequence. Codon 9 nonsense mutation was observed in two clones (before treatment samples from patients PCZ and EAC) and codon 84 in three clones (two clones in the EAO6m and one clone in the AAF4m sample). Nonsense mutations were also observed in codons 47 (before treatment sample from RMO), 233 (CSM5m) and 399 (ACC12wt) (Figure 1). The majority of nonsense mutations were detected in Domain I of the NS5A region (codons 47 and 84), with only one being located in domain III (codon 399), one in the Alpha Helix (codon 9) region and one in the Low Complexity Sequence I Region (codon 233). The numeration of codons was based on the NS5A protein, the first codon being the first amino acid coded for in the NS5A genomic region.

Part of this nonsense mutation could provide a functional NS5A protein. Therefore, the sequences at the N-terminal region of NS5A in which the nonsense mutation occurred were analyzed to locate the next methionine residue. The methionines are highlighted in figure 1a. The nonsense mutation 399 was not investigated owing to the presence of a methionine after it. This sequence could provide functional proteins despite missing a C-terminal portion (Figure 1b).

Low genetic complexity and diversity of NS5A are indicative of sustained response

Shannon entropy analysis showed that the genetic complexity of the NS5A region at baseline was significantly lower for the SVR group (mean value 0.00706) than for the NR (0.017529) and ETR groups (0.017427) (*p=0.0253* and *p=0.0265*, respectively) (Figure 2a). The genetic diversity, calculated using the genetic distance, was also significantly lower for the SVR group (0.0050) than for the NR (0.0133) and ETR (0.0148) groups (*p=0.021* and *p=0.022*, respectively; Figure 2b). These results demonstrated that quasispecies of SVR samples at baseline are less complex and diverse than those of NR and ETR samples. The mean entropy and genetic distance of during-treatment samples for NR and after the completion of therapy for NR and ETR samples were comparable when the results of diversity and complexity of each patient were analyzed individually, but no profile was detected for patients in the same response group. Generally, values had a propensity to decrease in samples collected during and after the end of treatment. However, these differences were not significant (Table 1 and Figure 2).

NS5A quasispecies composition is dynamic over time

In order to evaluate the number of NS5A quasispecies in each sample, all sequences in this study were analyzed using LOCQSPEC 1.0 software. Figure 3 presents LOCQSPEC 1.0 analyses for ETR and NR patients. The percentages of different quasispecies in ETR, NR and SVR samples are presented in table 2.

Concerning ETR patients, there was no specific profile of quasispecies diversity. However, samples showed a more varied quasispecies composition before treatment (PCZ 87%, CSM 80%; AAF 93% and RMO 100%) and a more homogeneous composition in several of the after-treatment samples (PCZ5m 33%; CSM5m 54% and RMO5m 73%, for example). Despite this, the composition of quasispecies in the samples collected after treatment demonstrated different dynamics. PCZ and CSM samples presented with an oscillation in the quasispecies composition. There was a decrease in diversity in the 2m samples (PCZ2m and CSM2m 47%), an increase in the 4m samples (PCZ4m 67% and CSM4m 73%), a decrease in the 5m samples (PCZ5m 33% and CSM4m 54%) and an increase in 6m samples (PCZ6m 47% and CSM6m 73%). Patient RMO presented with a significant decrease in the 4m (60%) and increase in the 6m sample (100%), and patient AAF showed a degree of homogenization in the last sample analyzed (AAF6m 87%) Table 2.

In some cases, the homogenization of quasispecies composition appeared to be related to the appearance of predominant quasispecies. For patient PCZ, the predominant quasispecies was identified two months after the end of treatment (40% of sequences) and was sustained (PCZ3m 40%; PCZ4m 26%; PCZ5m 73%; PCZ6m 60%). The same dynamic was observed for patient CSM (CSM 2m 60%, CSM 3m 60%; CSM 4m 33%; CSM 5m 53%; CSM 6m 33%). Patient AAF revealed the same quasispecies throughout the after treatment samples but the frequency was low (AAF4m/5m 6%; AAF6m 13%). A quasispecies was identified across RMO samples (RMO3m/4m 6%; RMO5m 27%), but the same quasispecies was not identified at RMO6m (Figure 3a).

The NR samples presented with a more heterogeneous quasispecies composition than the ETR patient samples (Table 2). Patients EAO and ACS showed a decrease in quasispecies diversity in samples collected during treatment (EAO12wt/ACS12wt). A specific quasispecies was observed across the samples but did not appear to be related to homogeneity. The ICC was the only patient in whom the same quasispecies were identified in before- and after-treatment samples (Figure 3b).

The SVR samples contained the most homogeneous quasispecies composition when compared with ETR and NR samples collected before treatment (Figure 3c).

NS5A quasispecies experience genetic evolution over time

Phylogenetic trees were reconstructed from 690 full-length NS5A sequences and reference sequences as described in the Methods section. Sequences clustered according to genotype, presenting monophyletic clusters with genotype 1a or 1b references. Viral isolates corresponding to the same patient samples also clustered, showing they were more closely related than other patient isolates (Figure 4). The topology of sequences from each patient was individually analyzed to evaluate the phylogenetic relationship among isolates. In general, the phylogenetic analysis suggested clustering of isolates of samples collected before treatment from ETR patients (Supporting Information Fig S1-S4). All quasispecies from samples collected before treatment were grouped in a monophyletic group for patients AAF and PCZ. The isolates of samples collected at different times after treatment were mixed and grouped into another cluster. The topology of clusters was 80% sustained by the bootstrap value for PCZ sequences. Fourteen of the fifteen quasispecies identified in samples collected before treatment were grouped together in a monophyletic group (67% sustained by the bootstrap value) in the CSM sample analysis. Patient RMO topology demonstrated a monophyletic group of 11 'before-treatment' quasispecies (89% sustained by the bootstrap value). Before treatment, quasispecies clustering was observed for EAO from NR patients, where 13 of the 15 isolates were grouped as a monophyletic cluster. Sequences from samples collected before treatment from patients ACC, ACS and ICC tended to group together with quasispecies identified after 12 weeks of treatment. However, bootstrap values only sustained clusters for patient ACS (bootstrap value of 69%). (Supporting Information Fig S5-S8)

In some cases, the phylogenetic analysis suggested that isolates identified in samples collected after treatment demonstrated the longest distances from the main nodes (Supporting Information).

NS5A quasispecies are under relaxed purifying selective pressure

Phylogenetic trees were reconstructed using DNA alignment of pre and post-therapy NS5A sequences corresponding to each patient in the study. The set of sequences relative to a specific patient was subjected to phylogenetic reconstruction by maximum likelihood. As a result, eight individual phylogenetic trees were generated (Supporting Information Fig S1-S8). In order to test whether selection pressure varied among different clades, one or two clades of interest were selected for testing. Clades with predominant quasispecies and those in which isolates identified in samples collected after treatment that demonstrated the longest distances from the main nodes were selected, making the assumption that the clade of interest would demonstrate different values of ω (dN/dS) from other branches of the tree. For patient CSM, it was not possible to select any specific clade; the detailed results are presented in Table 3. The values of ω ranged among patients from $\omega = 0.1101$ to $\omega = 0.7177$, and the majority of values indicated a purifying selection or relaxed purifying selective pressure. In Figure 5, the ω values illustrate this variation. For patients RMO and ACS, the foreground ω values were statistically different from the background ones. For these patients, two clades were selected for analysis using phylogenetic topology. For patient RMO these were clade one, where post-therapy quasispecies demonstrated the longest distances, and clade two, where predominant quasispecies were grouped. Accordingly, their ω values were different (clade 1: $\omega F = 0.2487$, $2\Delta l = 6.4343$, and clade 2: $\omega F = 0.7177$, $2\Delta l = 21.4983$). For patient ACS, clade one, where post-therapy quasispecies were located, and clade two, where quasispecies from the first sample collected after the end of therapy were grouped, were selected for analysis. The ω values were different for each of these clades (clade 1 $\omega F = 0.2685$, $2\Delta l = 9.264$, and clade 2 $\omega F = 0.5783$, $2\Delta l = 11.8750$). These results indicate a relaxed purifying selection, with the exception of $\omega F = 0.7177$, which approximates to neutral evolution.

DISCUSSION

Hepatitis C virus (HCV) has been the subject of research and clinical investigations because of its worldwide prevalence and major role in chronic liver disease. The only approved therapy is interferon-based in combination with ribavirin, but it results in low rates of sustained virological response (SVR). Many viral and host factors are predictive of the virological response [6,7]. Viral factors other than genotype and viral load may be involved in the response to combined Pegylated IFN and ribavirin therapy.

Quasispecies composition appears to be important in the IFN resistance mechanisms [29], and genetic variability has been studied in various regions of the HCV genome, mainly in the E2 and NS5A regions [30,31,32,33]. However, few reports have investigated the complete NS5A region [30,31,34,35,36,37] and/or the quasispecies evolution during Pegylated Interferon and Ribavirin anti-viral therapy [21,38]. Additional investigation of NS5A region variability may be relevant as heterogeneity in this region before treatment has been related to IFN-monotherapy responsiveness. However, studies have presented controversial data and the mechanism underlying how NS5A interferes with responsiveness to treatment has yet to be elucidated.

In the present study, the genetic variability of the complete NS5A region in patients infected with HCV genotype 1 was analyzed. Six hundred and ninety sequences of the entire NS5A genomic region from 11 patients' samples were analyzed. No specific profile of quasispecies diversity for all patients from each response type could be identified. However, samples showed less complexity and lower diversity of quasispecies composition pre-therapy (baseline) for sustained virological responders (SVR) than for non-responders (NR) and patients who relapsed after the end of therapy (ETR).

These findings are consistent with previous observations made on patients undergoing anti-HCV therapy. It was suggested that less complexity and lower diversity of HCV quasispecies at the baseline is associated with viral clearance [19,21,39,40,41]. In addition, evidence indicates that HCV characteristics including the high genetic variability could be responsible for the low efficiency of anti-HCV treatment [22]. Jain et al. observed that the initial anti-viral effect of interferon is influenced by the genetic composition of the viral quasispecies at the time of treatment initiation, and patients who demonstrated high viral diversity are less likely to respond to combined treatment [21]. Therefore, it is possible that higher diversity of quasispecies provides the virus with a better opportunity of evading anti-viral therapy.

In agreement with the genetic distance and entropy results, the number of viral strains at baseline was lower for SVR patients than NR and ETR patients, showing that quasispecies composition was more homogeneous than other groups of responses. A recent study evaluating the clinical and virological parameters that could be associated with or predictive of therapy outcome at baseline demonstrated that a higher number of quasispecies variants in the E1 E2 region was significantly associated with treatment failure. The authors inferred that independent factors cannot provide a consistent prediction of therapy response, and that host and virus characteristics are involved in treatment failure [42].

Analyzing ETR samples, it was evident that quasispecies composition, analyzed by tracking the number of strains in post-therapy samples, was more homogeneous than in other samples. This homogeneity was most dramatic in patients CSM and PCZ, who presented with homogeneous quasispecies composition from the rebound and sustained it (last processed sample 24 weeks after the end of therapy). As with ETR patients, the HCV RNA is undetectable by the end of therapy, suggesting that the selective pressure of therapy acts on quasispecies diversity during therapy until HCV becomes undetectable in samples. It is plausible that during or after the end of therapy a predominant quasispecies rises in frequency owing to its ‘better fitness’, while others decrease progressively in frequency or are eliminated. The persistence of this predominant variant could be favorable for sustaining the infection, causing it to become chronic again.

Farci at al. evaluated the quasispecies composition of the E1/E2 region in samples from patients who were untreated or treated with conventional IFN combined with ribavirin. Analyzing samples of relapse patients, they observed greater viral diversity at baseline (no strain comprising more than 25% of the population), and an emergence at relapse of a new dominant strain, or the emergence of a minor dominant population [43]. They inferred that the emergence of a new viral strain after rebound implies that the majority of baseline variants were sensitive to IFN and that the origin of these new strains is uncertain. The authors suggested several hypotheses for the origin of these new strains including that they could be new mutants, variants that were present at undetectable levels before its onset, or variants that pre-existed in viral reservoirs that were less accessible to IFN. The authors concluded that it is probable that very low levels of virus replication continued to occur despite the disappearance of viremia [43].

A study focusing on breakthrough response pattern, defined by patients who have an initial response followed by reactivation whilst receiving IFN therapy, did not identify quasispecies variants from breakthrough in baseline samples [38]. This study assumed that HCV variants sampled at the time of breakthrough represent drug-resistant quasispecies, and if present at baseline, they must be present very low frequency. According to the authors it is possible that viral breakthrough could be attributed to the selection of pre-existing drug-resistant variants or the emergence of a different quasispecies with reduced sensitivity to IFN, explaining the new viral variants evident at breakthrough. The study indicated that selection is mostly responsible for appearance of drug resistant quasispecies at the breakthrough point [38]. As breakthrough and relapse indicate rebound of infection after undetectable HCV RNA levels, is reasonable that the results presented herein are in agreement when considering the quasispecies composition and diversity at breakthrough and relapse time.

The data collected from non-responder samples revealed a very diverse composition of viral strains over the time course of the investigation. As described previously, the high variability of the quasispecies population in these patient samples might represent a continuous process of adaptation [44], as these variants were continuously eliminated during and after the end of treatment in most cases. Alternatively, it could offer the virus some advantage in sustaining the infection as a large number of quasispecies may indicate a better opportunity for virus persistence [45,46].

Pawlotsky et al. compared pre- and post-therapy NS5A amino acid sequences from non-responder patients and demonstrated that most variants from after-treatment samples were not detected before treatment, whereas most pre-treatment variants were no longer evident after treatment [22]. The data presented here agree with their findings, with the exception of samples from patient ICC. In the ICC samples, when investigating quasispecies composition, a variant was identified at baseline (13% of NS5A amino acid sequences) that was detected after 12 weeks of treatment and two months after therapy was completed. Several studies have showed that variants resistant to therapy may be present before treatment, and it is proposed that the persistence of these variants in patients who fail to respond suggests the existence of virus strains with inherent resistance to IFN [43,47].

Consistent with previous observations, phylogenetic analysis of virus sequences obtained from patients failed to show any clustering associated with specific response [35,36,43,48]. However, the data suggested that pre- and post-therapy isolates from samples of ETR patients tended to group in different branches. For NR samples, before-treatment quasispecies variants tended to group with quasispecies found after 12 weeks of treatment, and post-therapy sequences grouped together in another branch. The distinctive clustering of pre/post-therapy sequences had been described previously [22,38,49], and suggests that an evolutionary process occurred over time. In some cases, the phylogenetic analysis demonstrated that isolates identified in after-treatment samples had the longest distances from the main nodes. This may suggest that changes can improve the fitness of quasispecies in order to establish a persistent infection after treatment selection pressure.

In order to investigate whether quasispecies in samples collected after treatment or predominant quasispecies identified over the time course were under differential selective pressure, the set of sequences of each NR and ETR patient underwent selective pressure analysis. Anti-viral drugs could put selective pressure on regions of the virus genome, resulting in predominance of more adaptive variants. In general, the results showed that a purifying selection is driving the evolution of quasispecies over time. This may explain the initial diversified composition of quasispecies at baseline for some ETR and NR patients, and an increase of a predominant quasispecies probably offers some advantage for the virus. The purifying selection was relaxed in some cases (clades selected for patients ACS and RMO), and allows non-predominant and less advantageous variants to survive at a lower frequency. Relaxed purifying selection may be indicative of some codons of the entire protein being under positive selection. A study analyzing PKR binding domain and the V3 domain in the NS5A region identified positively selected sites but failed to detect a pattern mechanism for the inefficient response to anti-viral treatment. The strategies that HCV applies to escape selective pressure of treatments and the host immune system depend on a set of factors that confer resistance so that the virus persists [50].

Genetic analyses identified nonsense mutations in the NS5A sequence of some HCV variants. Nonsense mutations in samples from HCV-infected patients have been described [35,36,51], but the structural and functional implications of such mutations detected *in vivo* are unclear. Most of the nonsense mutations detected were located in the N-terminal region of NS5A

(Alpha Helix, Domain I or Low Complexity Sequence I regions of the NS5A); only one was located in the C-terminal region (domain III). It is known that N-terminal mutants of NS5A are preferentially located in the nucleus and are reported to function as transcriptional regulators. A recent study using a HCV replicating cell system revealed that during the life cycle a variety of N-terminally truncated NS5A fragments are generated [52]. Tests on several of these truncated NS5A fragments demonstrated that the truncated forms NS5A 31-449, NS5A 221-449 and NS5A 301-449 were preferentially located in the nucleus, and NS5A 105-449 was equally distributed between the cytoplasm and nucleus. The full length NS5A (1-449) was located in the extra-nuclear compartment as previously described. However, truncated forms impaired HCV replication. In contrast, domain III of the C-terminal NS5A region can be deleted with no or minimal effect on RNA replication [53,54], but the C-terminal region residing between amino acid residues 2404 and 2435 of domain III is crucial for virus production [55]. In order to elucidate the effects of nonsense mutations identified in the present study, further analysis with HCV replicating cells is necessary. Truncated NS5A forms observed in previous studies show drawbacks in terms of replication or assembly, but these effects could be overcome by the diversity of quasispecies composition. Furthermore, the generation of N-terminally truncated forms of NS5A that interfere with the vital process of virus replication may represent an alternative mechanism to regulate HCV replication [52].

In conclusion, this study demonstrates that the analysis of quasispecies diversity of the NS5A region could provide important information for elucidating the mechanism underlying treatment failure in patients with chronic hepatitis C infection. The results indicate that heterogeneous diversity of quasispecies pre-therapy could be indicative of a low response to IFN-based therapy, and that homogeneity of quasispecies composition at baseline is associated with viral clearance. In addition, quasispecies distribution of NS5A was variable and distinctly clustered over time, indicating an evolutionary process driven by the pressure of therapy. Therefore, the quasispecies composition and evolution over time are relevant factors to be considered in terms of patient outcome after combined therapy for chronic hepatitis C.

MATERIALS AND METHODS

Patients

Eleven naïve patients infected chronically with HCV RNA genotypes 1a or 1b were enrolled from the Hepatology Department of the São José do Rio Preto School of Medicine. Patients with other concomitant liver diseases [hepatitis B virus (HBV) or other hepatotropic virus infections, alcohol abuse, autoimmune hepatitis and hereditary liver diseases] were excluded. The Ethics Committee of the São José do Rio Preto School of Medicine approved the study (protocol number 2025/2004) under the resolution CNS 196/96 of this committee. Written informed consent was obtained from all patients. Plasma samples were collected before, during and after the end of treatment. During the 48-week treatment, patients received PEG-IFN-α-2b (according to body weight) subcutaneously once a week and RBV daily, taken orally at a dose of 600-1,200 mg (according to body weight).

The patients were classified into three groups according to their response to therapy: three patients demonstrated sustained virological response (SVR); four patients were non-responders (NR) and four patients were end-of-treatment virological responders (ETR) (Table 1). SVR was defined as absence of HCV RNA in plasma using qualitative PCR six months after the end of therapy. NR was defined as continued presence of HCV RNA in plasma during treatment and six months after the end of treatment. Patients who were HCV RNA-negative at the end of therapy but experienced a relapse were classified as ETR.

For all patients, one sample before treatment was analyzed. Additional samples were collected during therapy (12 weeks of therapy) and after the end of therapy (14 days, two and six months) for non-responders, and at the relapse time and then monthly for end-of-treatment responders. Samples analyzed for each patient are detailed in Table 1.

RNA extraction, RT-PCR and NS5A amplification

Total RNA was extracted from 140 μl of plasma using the commercially available QIAamp Viral RNA Kit (Qiagen, Uniscience). RNA was reverse-transcribed into cDNA using a High-Capacity cDNA Archive kit (Applied Biosystems, Foster City, CA, USA) and random primers. The mixture was incubated at 37°C for two hours. For amplification of the entire NS5A region of the HCV genome, a nested polymerase chain reaction (PCR) was performed using the primers described previously [35].

The viral load was quantified using the Cobas TaqMan HCV Test according to the manufacturer's instructions.

Cloning and sequencing

PCR products of approximately 1.7 kb were purified and ligated into the pCR-XL-TOPO-vector using the TOPO XL PCR cloning kit (InvitrogenTM Life Technologies, Carlsbad, CA, USA). The ligation products were transformed into *Escherichia coli* competent cells (InvitrogenTM Life Technologies, Carlsbad, CA, USA). Fifteen transformants were randomly chosen for further studies, and plasmid DNA was isolated from a 3.0-ml broth culture using the GeneJET plasmid Miniprep Kit (Fermentas). Recombinant pCR-XL-TOPO-NS5A clones were sequenced using dideoxy terminator automated sequencing (ABI Prism Ready Reaction Mix; Applied Biosystems, Foster City, CA, USA) using an ABI Prism 377 and ABI 3130XL sequencers, according to the manufacturer's instructions (Applied Biosystems Inc, Foster City, CA, USA.). Eight to ten sequencing reactions were performed for each clone, using flanking primers, M13 Forward and M13 Reverse (Invitrogen TM Life Technologies, Carlsbad, CA, USA), internal forward primers (1aF1 5'CACCAGTGGATAASCTCGGA3', 1aF2 5'CCCATYAATGCCTACACCAC3', 1aF3 5'CTGTCYGCTCCATCTCTCA3' and 1aF4 5'GARTCAGARAACAAAGTGGTG3' for genotype 1a; 1bF1 5'ATCCTCTCHAR-CCTTACCAT3', 1bF2 5' GRACATTCCCCRTCAACGC3', 1bF3 5' 5'GTCCTRACAG-AATCCACMGTG3' and 1bF4 5'CCARTTGTCTGCGCCTTC3' for genotype 1b) and internal reverse primers (1aR1 5'ACATWGAGCAACACACGAC3', 1aR2 5'GTTCCC-TTGAGAGATGGAGC3' and 1aR3 5'TAGGCATTRATGGGGAAGGT3' for genotype 1a; 1bR1 5'ARCARCAGACGACGTCCTC3', 1bR2 5'AGCGGGTCGAAAGAGTCCA 3' and 1bR3 5' GAACCGTTTTGACATGTCC3' for genotype 1b).

Genetic Analysis of NS5A and Phylogenetic reconstruction

All sequences were subjected to Phred-Phrap programs [56,57,58]. These programs analyze the quality of the sequences and align them in complete NS5A contigs. Six hundred and ninety full-length NS5A sequences were generated in this study, corresponding to 11 patients. The nucleotide sequence contigs (1344 nucleotides for genotype 1a and 1341 for genotype 1b) were aligned using the Clustal X program (version 1.81) [59] and amino acid sequences were obtained. Primer sequences were removed from all sequences using the BioEdit program (version 7.0.5.3) [60]. The genetic distance between pairs of sequences were calculated with MEGA version 4 using the p-distance method [61]. The variability of amino acid residues in each position (i) was measured by calculating the Shannon entropy [22]. The entropy value is a measure of the lack of information at each position of the amino acid sequence. The Shannon Entropy was calculated at the amino acid level as follows: $H(i) = -(\Sigma f(b,i)\log(\text{base } 2)f(b,i))$, where $f(b,i)$ is the frequency with which each residue b appears in position i of the protein. When different groups of virological response were compared, the value was normalized by Sn=H/logN, where N is the total number of sequences analyzed in each sample [60].

To construct the phylogenetic tree of NS5A variants obtained from the patient samples, the PAUP* version 4 program was used [62]. Phylogenetic trees were constructed using the maximum likelihood method [63] with the model of substitution, as determined by hierarchical likelihood ratio test score criteria in Modeltest 3.06 [64]. Base frequency, gamma distribution and transition/transversion ratios were determined (from the data) by Modeltest 3.06. A thousand replicates were used to test the reliability of the tree topology, and bootstrap values >70 were considered significant [65]. Genetic and phylogenetic analyses were performed using the standard genotype 1a sequence H77 (NC_004102.1) as a reference and genotype 1b sequence HCV-J (D90208.1), obtained from GenBank.

Analysis of variable selective pressure

The ratio between the relative rate of non-synonymous substitution to the relative rate of synonymous substitution (ω=dN/dS) measures the strength of selection acting on a protein-coding gene. Assuming synonymous mutations are subjected to almost strictly neutral selection, ω<1, ω=1, and ω>1 represent negative selection, neutral evolution, and positive Darwinian selection, respectively [66]. Maximum likelihood analysis of the sequence evolution was performed using the CODEML program in the PAML 3.15 software package [67]. Initially, phylogenetic trees were reconstructed using the maximum likelihood method and the HKY model of substitution, as determined by hierarchical likelihood ratio test score criteria in Modeltest 3.06 for each dataset of NS5A sequences corresponding to each non-responder and end-of-treatment responder in this study. The ratios of global synonymous changes per site (dS) versus replacement changes per site (dN) for each tree were calculated using two models described by Yang [68]. The one-ratio model assumed an equal ω ratio for all branches in the phylogeny. The two-ratio model assumed two ω ratios: one branch for the background (ω B), one for the foreground branch (ω F = branch of interest) leading to a specific clade of the phylogenetic trees, specified in the Supporting Information. This analysis makes different assumptions about the dN/dS ratios for branches of interest relative to the background dN/dS ratio for all other branches. For instance, the ‘‘two-ratio’’ model assumes that the branches of interest have a dN/dS ratio that is different from the background ratio.

Analysis of quasispecies composition

For analyzing quasispecies variability, software LOCQSPEC 1.0 was designed [69]. This is a tool built in C++ programming language that allows a set of sequences to be compared, finding similar or different sequences as results. All sequences of complete NS5A generated in this study were subjected to the program, and the contigs that presented the same nucleotide or amino acid sequences were grouped together. Each group corresponds to a set of identical sequences that were identified more than once. The construction of each group started with an identity verification based on parity among sequences that were grouped. The first sequence of the set was directly associated with the first group. For the other sequences, their identity was checked with regard to the entire set of existing groups. There was an optimization step in this process to stop the verification when a mismatch was identified. Otherwise there was identity, the sequence was associated with the group, and the next sequence started to be verified. If all existing groups were verified and no identity was found, a new group was formed and the new sequence was associated with it. Therefore, it was possible to determine the number of varied and identical quasispecies in the same sample.

Statistical Analysis

The results are presented as mean ± SD, or as percentages. Comparisons among the sustained virological responders, end of treatment responders and non-responders were determined using Tukey's or Fisher's statistical test after performing an one-way ANOVA or Chi square. In all tests, a P value less than 0.05 was considered significant.

Accession numbers

All sequences obtained in this study were submitted to the GenBank nucleotide sequence database (http://www.ncbi.nlm.nih.gov/genbank/) under the follow accession numbers: HQ823765 - HQ824349, EU309511 - EU309525, EU309586 - EU309599 and EU309600 - EU309614.

ACKNOWLEDGEMENTS

We are grateful to FAPESP for grants that supported this work.

AUTHOR CONTRIBUTIONS

Conceived and designed the study: ACGJ CB IMVGCM PR. Performed the experiments: ACGJ RPAM. P Analyzed the data: ACGJ CB LHTY RAS CMAC IMVGCM PR. Contributed to acquisition of data/ reagents/ material/ analysis tool: RMF JRRP PR. Wrote the paper: ACGJ. Critically revised the paper: CB CMAC IMVGCM RP.

REFERENCES

1. GBD (2004) Global burden of disease (GBD) for hepatitis C. J Clin Pharmacol 44: 20-29.
2. Chevaliez S, Pawlotsky JM (2007) Hepatitis C virus: virology, diagnosis and management of antiviral therapy. World J Gastroenterol 13: 2461-2466.
3. Saito I, Miyamura T, Ohbayashi A, Harada H, Katayama T, et al. (1990) Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. Proc Natl Acad Sci U S A 87: 6547-6549.

4. Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, et al. (2002) Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 347: 975-982.
5. Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, et al. (2001) Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet 358: 958-965.
6. Pawlotsky JM (2003) Mechanisms of antiviral treatment efficacy and failure in chronic hepatitis C. Antiviral Res 59: 1-11.
7. Pawlotsky JM (2005) Current and future concepts in hepatitis C therapy. Semin Liver Dis 25: 72-83.
8. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, et al. (1989) Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 244: 359-362.
9. Murphy DG, Willems B, Deschenes M, Hilzenrat N, Mousseau R, et al. (2007) Use of sequence analysis of the NS5B region for routine genotyping of hepatitis C virus with reference to C/E1 and 5' untranslated region sequences. J Clin Microbiol 45: 1102-1112.
10. Simmonds P, Bukh J, Combet C, Deleage G, Enomoto N, et al. (2005) Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatology 42: 962-973.
11. Martell M, Esteban JI, Quer J, Genesca J, Weiner A, et al. (1992) Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. J Virol 66: 3225-3229.
12. Domingo E, Martin V, Perales C, Grande-Perez A, Garcia-Arriaza J, et al. (2006) Viruses as quasispecies: biological implications. Curr Top Microbiol Immunol 299: 51-82.
13. Pawlotsky JM (2003) Hepatitis C virus genetic variability: pathogenic and clinical implications. Clin Liver Dis 7: 45-66.
14. Forns X, Bukh J (1999) The molecular biology of hepatitis C virus. Genotypes and quasispecies. Clin Liver Dis 3: 693-716, vii.
15. Forns X, Purcell RH, Bukh J (1999) Quasispecies in viral persistence and pathogenesis of hepatitis C virus. Trends Microbiol 7: 402-410.
16. Le Guillou-Guillemette H, Vallet S, Gaudy-Graffin C, Payan C, Pivert A, et al. (2007) Genetic diversity of the hepatitis C virus: impact and issues in the antiviral therapy. World J Gastroenterol 13: 2416-2426.
17. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, et al. (1995) Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. J Clin Invest 96: 224-230.
18. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, et al. (1996) Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. N Engl J Med 334: 77-81.
19. Puig-Basagoiti F, Forns X, Furcic I, Ampurdanes S, Gimenez-Barcons M, et al. (2005) Dynamics of hepatitis C virus NS5A quasispecies during interferon and ribavirin therapy in responder and non-responder patients with genotype 1b chronic hepatitis C. J Gen Virol 86: 1067-1075.
20. Witherell GW, Beineke P (2001) Statistical analysis of combined substitutions in nonstructural 5A region of hepatitis C virus and interferon response. J Med Virol 63: 8-16.
21. Jain MK, Yuan HJ, Adams-Huet B, Reeck A, Shelton J, et al. (2009) Pegylated interferon and ribavirin promote early evolution of nonstructural 5A protein in individuals with hepatitis C who demonstrate a response to treatment. J Infect Dis 200: 866-876.

22. Pawlotsky JM, Germanidis G, Neumann AU, Pellerin M, Frainais PO, et al. (1998) Interferon resistance of hepatitis C virus genotype 1b: relationship to nonstructural 5A gene quasispecies mutations. J Virol 72: 2795-2805.
23. Brass V, Bieck E, Montserret R, Wolk B, Hellings JA, et al. (2002) An amino-terminal amphipathic alpha-helix mediates membrane association of the hepatitis C virus nonstructural protein 5A. J Biol Chem 277: 8130-8139.
24. Penin F, Brass V, Appel N, Ramboarina S, Montserret R, et al. (2004) Structure and function of the membrane anchor domain of hepatitis C virus nonstructural protein 5A. J Biol Chem 279: 40835-40843.
25. Tellinghuisen TL, Marcotrigiano J, Gorbalenya AE, Rice CM (2004) The NS5A protein of hepatitis C virus is a zinc metalloprotein. J Biol Chem 279: 48576-48587.
26. Ide Y, Zhang L, Chen M, Inchauspe G, Bahl C, et al. (1996) Characterization of the nuclear localization signal and subcellular distribution of hepatitis C virus nonstructural protein NS5A. Gene 182: 203-211.
27. Pawlotsky JM, Germanidis G (1999) The non-structural 5A protein of hepatitis C virus. J Viral Hepat 6: 343-356.
28. Satoh S, Hirota M, Noguchi T, Hijikata M, Handa H, et al. (2000) Cleavage of hepatitis C virus nonstructural protein 5A by a caspase-like protease(s) in mammalian cells. Virology 270: 476-487.
29. Salmeron J, De Rueda PM, Ruiz-Extremera A, Casado J, Huertas C, et al. (2006) Quasispecies as predictive response factors for antiviral treatment in patients with chronic hepatitis C. Dig Dis Sci 51: 960-967.
30. Duverlie G, Khorsi H, Castelain S, Jaillon O, Izopet J, et al. (1998) Sequence analysis of the NS5A protein of European hepatitis C virus 1b isolates and relation to interferon sensitivity. J Gen Virol 79 (Pt 6): 1373-1381.
31. Pascu M, Martus P, Hohne M, Wiedenmann B, Hopf U, et al. (2004) Sustained virological response in hepatitis C virus type 1b infected patients is predicted by the number of mutations within the NS5A-ISDR: a meta-analysis focused on geographical differences. Gut 53: 1345-1351.
32. Gaudy C, Lambele M, Moreau A, Veillon P, Lunel F, et al. (2005) Mutations within the hepatitis C virus genotype 1b E2-PePHD domain do not correlate with treatment outcome. J Clin Microbiol 43: 750-754.
33. Gerotto M, Dal Pero F, Pontisso P, Noventa F, Gatta A, et al. (2000) Two PKR inhibitor HCV proteins correlate with early but not sustained response to interferon. Gastroenterology 119: 1649-1655.
34. Nousbaum J, Polyak SJ, Ray SC, Sullivan DG, Larson AM, et al. (2000) Prospective characterization of full-length hepatitis C virus NS5A quasispecies during induction and combination antiviral therapy. J Virol 74: 9028-9038.
35. Jardim AC, Yamasaki LH, de Queiroz AT, Bittar C, Pinho JR, et al. (2009) Quasispecies of hepatitis C virus genotype 1 and treatment outcome with peginterferon and ribavirin. Infect Genet Evol 9: 689-698.
36. Bittar C, Jardim AC, Yamasaki LH, de Queiroz AT, Carareto CM, et al. (2010) Genetic diversity of NS5A protein from hepatitis C virus genotype 3a and its relationship to therapy response. BMC Infect Dis 10: 36.
37. Veillon P, Payan C, Le Guillou-Guillemette H, Gaudy C, Lunel F (2007) Quasispecies evolution in NS5A region of hepatitis C virus genotype 1b during interferon or combined interferon-ribavirin therapy. World J Gastroenterol 13: 1195-1203.
38. Xu Z, Fan X, Xu Y, Di Bisceglie AM (2008) Comparative analysis of nearly full-length hepatitis C virus quasispecies from patients experiencing viral breakthrough during

antiviral therapy: clustered mutations in three functional genes, E2, NS2, and NS5a. J Virol 82: 9417-9424.
39. Pawlotsky JM, Pellerin M, Bouvier M, Roudot-Thoraval F, Germanidis G, et al. (1998) Genetic complexity of the hypervariable region 1 (HVR1) of hepatitis C virus (HCV): influence on the characteristics of the infection and responses to interferon alfa therapy in patients with chronic hepatitis C. J Med Virol 54: 256-264.
40. Toyoda H, Kumada T, Nakano S, Takeda I, Sugiyama K, et al. (1997) Quasispecies nature of hepatitis C virus and response to alpha interferon: significance as a predictor of direct response to interferon. J Hepatol 26: 6-13.
41. Ueda E, Enomoto N, Sakamoto N, Hamano K, Sato C, et al. (2004) Changes of HCV quasispecies during combination therapy with interferon and ribavirin. Hepatol Res 29: 89-96.
42. Saludes V, Bracho MA, Valero O, Ardevol M, Planas R, et al. (2010) Baseline prediction of combination therapy outcome in hepatitis C virus 1b infected patients by discriminant analysis using viral and host factors. PLoS One 5: e14132.
43. Farci P, Strazzera R, Alter HJ, Farci S, Degioannis D, et al. (2002) Early changes in hepatitis C viral quasispecies during interferon therapy predict the therapeutic outcome. Proc Natl Acad Sci U S A 99: 3081-3086.
44. Figlerowicz M, Jackowiak P, Formanowicz P, Kedziora P, Alejska M, et al. Hepatitis C virus quasispecies in chronically infected children subjected to interferon-ribavirin therapy. Arch Virol.
45. Ramirez S, Perez-del-Pulgar S, Carrion JA, Coto-Llerena M, Mensa L, et al. Hepatitis C virus superinfection of liver grafts: a detailed analysis of early exclusion of non-dominant virus strains. J Gen Virol 91: 1183-1188.
46. Ray SC, Wang YM, Laeyendecker O, Ticehurst JR, Villano SA, et al. (1999) Acute hepatitis C virus structural gene sequences as predictors of persistent viremia: hypervariable region 1 as a decoy. J Virol 73: 2938-2946.
47. Figlerowicz M, Jackowiak P, Formanowicz P, Kedziora P, Alejska M, et al. (2010) Hepatitis C virus quasispecies in chronically infected children subjected to interferon-ribavirin therapy. Arch Virol 155: 1977-1987.
48. Zekri AR, El-Din HM, Bahnassy AA, Khaled MM, Omar A, et al. (2007) Genetic distance and heterogenecity between quasispecies is a critical predictor to IFN response in Egyptian patients with HCV genotype-4. Virol J 4: 16.
49. Pellerin M, Lopez-Aguirre Y, Penin F, Dhumeaux D, Pawlotsky JM (2004) Hepatitis C virus quasispecies variability modulates nonstructural protein 5A transcriptional activation, pointing to cellular compartmentalization of virus-host interactions. J Virol 78: 4617-4627.
50. Cuevas JM, Torres-Puente M, Jimenez-Hernandez N, Bracho MA, Garcia-Robles I, et al. (2008) Genetic variability of hepatitis C virus before and after combined therapy of interferon plus ribavirin. PLoS One 3: e3058.
51. Fan W, Zhu W, Wei L, Wang Q, Yin L, et al. (2005) Nonstructural 5A gene variability of hepatitis C virus (HCV) during a 10-year follow up. J Gastroenterol 40: 43-51.
52. Sauter D, Himmelsbach K, Kriegs M, Carvajal Yepes M, Hildt E (2009) Localization determines function: N-terminally truncated NS5A fragments accumulate in the nucleus and impair HCV replication. J Hepatol 50: 861-871.
53. Appel N, Pietschmann T, Bartenschlager R (2005) Mutational analysis of hepatitis C virus nonstructural protein 5A: potential role of differential phosphorylation in RNA replication and identification of a genetically flexible domain. J Virol 79: 3187-3194.

54. Moradpour D, Evans MJ, Gosert R, Yuan Z, Blum HE, et al. (2004) Insertion of green fluorescent protein into nonstructural protein 5A allows direct visualization of functional hepatitis C virus replication complexes. J Virol 78: 7400-7409.
55. Appel N, Zayas M, Miller S, Krijnse-Locker J, Schaller T, et al. (2008) Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. PLoS Pathog 4: e1000035.
56. Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. Genome Res 8: 195-202.
57. Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8: 186-194.
58. Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8: 175-185.
59. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876-4882.
60. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98.
61. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599.
62. Swofford DL (2002) PAUP* Phylogenetic Analysis Using Parsimony (*and Others Methods). Sunderland, Massachusetts.
63. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-425.
64. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818.
65. McCormack GP, Clewley JP (2002) The application of molecular phylogenetics to the analysis of viral genome diversity and evolution. Rev Med Virol 12: 221-238.
66. Yang Z (2002) Inference of selection from multiple species alignments. Curr Opin Genet Dev 12: 688-694.
67. Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 13: 555-556.
68. Yang Z (1998) Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol 15: 568-573.
69. Marucci EA, Zafalon GF, Jardim AC, Yamasaki LH, Bittar C, et al. (2008) Routine libraries for pattern recognition in quasispecies. Genet Mol Res 7: 970-981.

LEGEND TO FIGURES

Figure 1. Alignment of sequence representation demonstrating the nonsense mutations observed in this study. A) Nonsense mutation at amino-terminal of the NS5A genomic region, positions 9, 47 and 84. B) One nonsense mutation located at the middle of the NS5A genomic region position 233, and one nonsense mutation located at the carboxy-terminal of the NS5A genomic region, position 399. Pink boxes indicate nonsense mutations. The next methionines are highlighted in gray. References sequences for genotypes 1a (H77_NS5A1a) and 1b (HCVJ_NS5A1b) are indicated.

Figure 2. Genetic variability of HCV quasispecies in samples collected over time. A) Genetic complexity calculated using normalized entropy (Shannon entropy). B) Diversity of HCV quasispecies calculated using genetic distance (p-distance). Patterns of response to therapy are represented by colours: sustained virological responders in red, non-responders in blue and end-of-treatment responders in green. P=value calculated using Fisher's test.

Figure 3. Quasispecies diversity within the NS5A protein based on the number of strains. A) The end-of-treatment responders. B) The non-responders. C) The sustained virological responders. Vertical bars represent viral variants. Inside the bars are the numbers of identical clones and, consequently, the number of identical quasispecies. The predominant quasispecies identified from the samples of each patient are identified by specific colors. The time of sample collection are indicated below the bars.

Figure 4. Unrooted phylogenetic trees reconstructed from the complete NS5A region. Phylogenetic tree reconstructed using the maximum-likelihood (ML) method, a heuristic search with Nearest Neighbor Interchange (NNI) branch-swapping algorithm. A) HCV

genotype 1a tree was based on the GTR+G substitution model. B) HCV genotype 1b tree was based on the TVMef+G substitution model. References are utilized in both trees. Patterns of response are represented by colors: sustained virological responders in red, non-responders in blue and end-of-treatment responders in green.

Figure 5. Selection pressure in particular clades of NS5A quasispecies. Omega values (ω) were analysed for sequences referent to each patient by individual reconstruction of phylogenetic trees, using the maximum-likelihood (ML) method based on the HKY substitution model, and analysis of specific clades (Supporting Information) using CODEML. Patients ICC, ACS, EAO and ACC are non-responders and PCZ, AAF and RMO are end-of-treatment responders.

Table 1. Clinical and virological characteristics of patients

Patient and type of treatment response	Sex	Age	HCV Genotype	Sample	Viral Load log UI/ml	Genetic Distance	Shannon Entropy
SVR							
MAG	F	44	1b	●	6.04	0.006 ± 0.001	0.0096
ELRB	M	40	1a	●	6.25	0.005 ± 0.001	0.0074
EAC	M	35	1a	●	6.57	0.004 ± 0.001	0.0042
NR							
ICC	F	41	1a	●	5.80	0.008 ± 0.001	0.0126
				12wt	4.59	0.009 ± 0.001	0.0113
				14d	6.18	0.011 ± 0.001	0.0157
				2m	6.46	0.008 ± 0.001	0.0103
				6m	6.40	0.007 ± 0.001	0.0156
ACS	M	72	1b	●	6.85	0.017 ± 0.001	0.0235
				12wt	3.71	0.007 ± 0.001	0.0140
				14d	6.27	0.005 ± 0.001	0.0134
				2m	6.73	0.007 ± 0.001	0.0132
				6m	>7.2	0.012 ± 0.002	0.0192
EAO	M	21	1a	●	6.92	0.012 ± 0.002	0.0202
				12wt	5.54	0.006 ± 0.001	0.0097
				14d	6.43	0.012 ± 0.001	0.0172
				2m	6.76	0.012 ± 0.001	0.0177
				6m	6.96	0.008 ± 0.001	0.0119
ACC	M	51	1b	●	6.50	0.016 ± 0.002	0.0139
				12wt	6.13	0.019 ± 0.002	0.0311
				14d	6.24	0.014 ± 0.001	0.0206
				2m	7.18	0.014 ± 0.001	0.0216
				6m	6.83	0.014 ± 0.001	0.0216
ETR							
PCZ	M	39	1a	●	5.95	0.011 ± 0.001	0.0131
				28d	4.99	0.004 ± 0.001	0.0094
				2m	5.39	0.003 ± 0.001	0.0058
				3m	4.74	0.003 ± 0.001	0.0055
				4m	5.03	0.002 ± 0.001	0.0073
				5m	5.26	0.001 ± 0.001	0.0033
				6m	5.08	0.012 ± 0.000	0.0038
CSM	M	27	1a	●	6.00	0.005 ± 0.001	0.0137
				2m	6.63	0.001 ± 0.001	0.0055
				3m	6.83	0.002 ± 0.001	0.0038
				4m	6.30	0.003 ± 0.001	0.0066
				5m	6.10	0.009 ± 0.001	0.0074
				6m	6.58	0.014 ± 0.001	0.0109
AAF	M	56	1b	●	6.66	0.016 ± 0.002	0.0161
				4m	6.24	0.014 ± 0.002	0.0162
				5m	6.28	0.017 ± 0.002	0.0141
				6m	6.34	0.022 ± 0.002	0.0145
RMO	M	47	1a	●	6.05	0.033 ± 0.002	0.0268
				2m	5.67	0.027 ± 0.003	0.0466
				3m	5.92	0.003 ± 0.002	0.0391
				4m	5.57	0.011 ± 0.001	0.0072
				5m	5.87	0.009 ± 0.002	0.0197
				6m	2.93	0.011 ± 0.002	0.0171

[a]SVR, sustained virological response; NR, non-response; ETR, end-of-treatment response
●=before treatment; wt=weeks of treatment; d=days after the end of treatment; m=months after the end of treatment

Table 2. Composition of quasispecies in patients' samples.

Patients	Samples collection points								
	●	**12wt**	**14d**	**28d**	**2m**	**3m**	**4m**	**5m**	**6m**
SVR									
MAG	80% (12)	------	------	------	------	------	------	------	------
ELRB	67% (10)	------	------	------	------	------	------	------	------
EAC	47% (7)	------	------	------	------	------	------	------	------
NR									
ICC	87% (13)	87% (13)	100% (15)	------	73% (11)	------	------	------	93% (14)
ACS	93% (14)	73% (11)	100% (15)	------	87% (13)	------	------	------	100% (15)
EAO	87% (13)	73% (11)	73% (11)	------	80% (12)	------	------	------	87% (13)
ACC	87% (13)	100% (15)	100% (15)	------	93% (14)	------	------	------	100% (15)
ETR									
PCZ	87% (13)	------	------	67% (10)	47% (7)	47% (7)	67% (10)	33% (5)	47% (7)
CSM	80% (12)	------	------	------	47% (7)	47% (7)	73% (11)	54% (8)	73% (11)
AAF	93% (14)	------	------	------	------	------	100% (15)	100% (15)	87% (13)
RMO	100% (15)	------	------	------	100% (15)	100% (15)	60% (9)	73% (11)	100% (15)

SVR=
sustained virological responder; NR=non-responder; ETR=end-of-treatment responder.
●=before treatment; wt=weeks of treatment; d=days after the end of treatment; m=months after the end of treatment.

Table 3. Selection pressure analysis within particular clades of NS5A quasispecies.

Patient	Model	l	2Δl	Estimates
ICC	one-ratio	-4404.76677		**ωB = 0.1692**
	two-ratio	-4404.28734	0.9587	ωB = 0.1521
				ωF = 0.1872
ACS[a]	one-ratio	-4558.41000		ωB = 0.1835
	two-ratio	-4552.47248	**11.8750	ωB = 0.1650
				ωF = 0.5783
	two-ratio	-4553.77795	**9.264	ωB = 0.1419
				ωF = 0.2685
EAO	one-ratio	-4755.95340		**ωB = 0.1705**
	two-ratio	-4755.92063	0.0822	ωB = 0.1727
				ωF = 0.1618
	two-ratio	-4754.96926	1.9683	ωB = 0.1818
				ωF = 0.1246
ACC	one-ratio	-6356.59235		**ωB = 0.1433**
	two-ratio	-6356.28458	0.6155	ωB = 0.1338
				ωF = 0.1521
PCZ	one-ratio	-3854.07081		**ωB = 0.2003**
	two-ratio	-3853.33452	1.4726	ωB = 0.1977
				ωF = 0.2350
AAF	one-ratio	-5111.91233		**ωB = 0.1101**
	two-ratio	-5111.91231	0.0004	ωB = 0.1102
				ωF = 0.1101
RMO[a]	one-ratio	-5455.19420		ωB = 0.2013
	two-ratio	-5451.97704	*6.4343	ωB = 0.1604
				ωF = 0.2487
	two-ratio	-5444.44505	**21.4983	ωB = 0.1776
				ωF = 0.7177

Non-responder: ICC, ACS, EAO, ACC; End-of-treatment responder: PCZ, AAF, RMO
X^2 critical values, 1 df: *: 3.84; **: 6.63; $2\Delta \ell = 2 (l_1 - l_0)$
The ω values considered are presented in bold.
[a] Two clades selected for analyses.

Figure 1

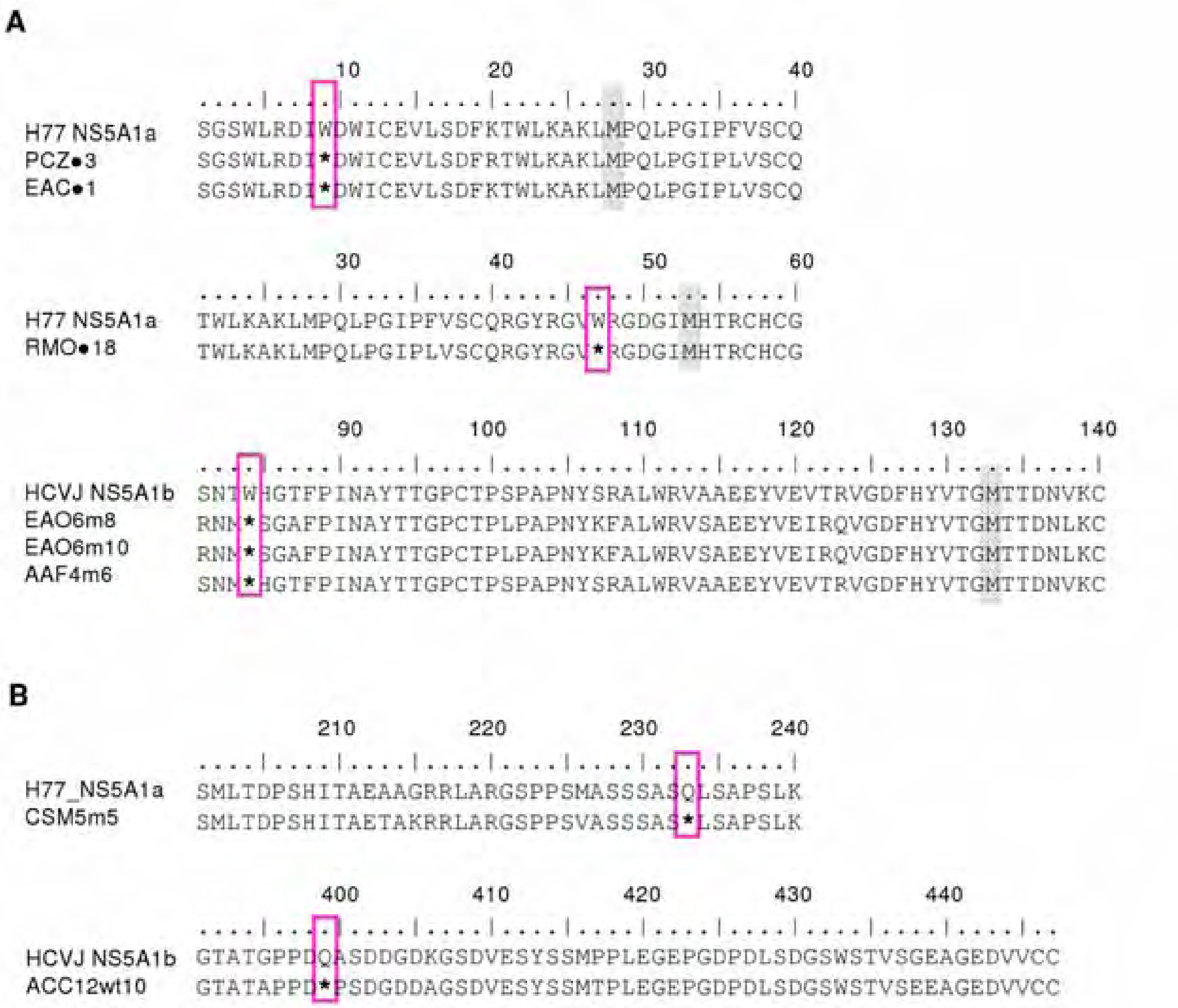

A
H77 NS5A1a SGSWLRDIWDWICEVLSDFKTWLKAKLMPQLPGIPFVSCQ
PCZ•3 SGSWLRDI*DWICEVLSDFRTWLKAKLMPQLPGIPLVSCQ
EAC•1 SGSWLRDI*DWICEVLSDFKTWLKAKLMPQLPGIPLVSCQ
H77 NS5A1a TWLKAKLMPQLPGIPFVSCQRGYRGVWRGDGIMHTRCHCG
RMO•18 TWLKAKLMPQLPGIPLVSCQRGYRGV*RGDGIMHTRCHCG
HCVJ NS5A1b SNTWHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKC
EAO6m8 RNN*SGAFPINAYTTGPCTPLPAPNYKFALWRVSAEEYVEIRQVGDFHYVTGMTTDNLKC
EAO6m10 RNN*SGAFPINAYTTGPCTPLPAPNYKFALWRVSAEEYVEIRQVGDFHYVTGMTTDNLKC
AAF4m6 SNN*HGTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKC
B
H77_NS5A1a SMLTDPSHITAEAAGRRLARGSPPSMASSSASQLSAPSLK
CSM5m5 SMLTDPSHITAETAKRRLARGSPPSVASSSAS*LSAPSLK
HCVJ NS5A1b GTATGPPDQASDDGDKGSDVESYSSMPPLEGEPGDPDLSDGSWSTVSGEAGEDVVCC
ACC12wt10 GTATAPPD*PSDGDDAGSDVESYSSMTPLEGEPGDPDLSDGSWSTVSEEAGEDVVCC

Figure 2

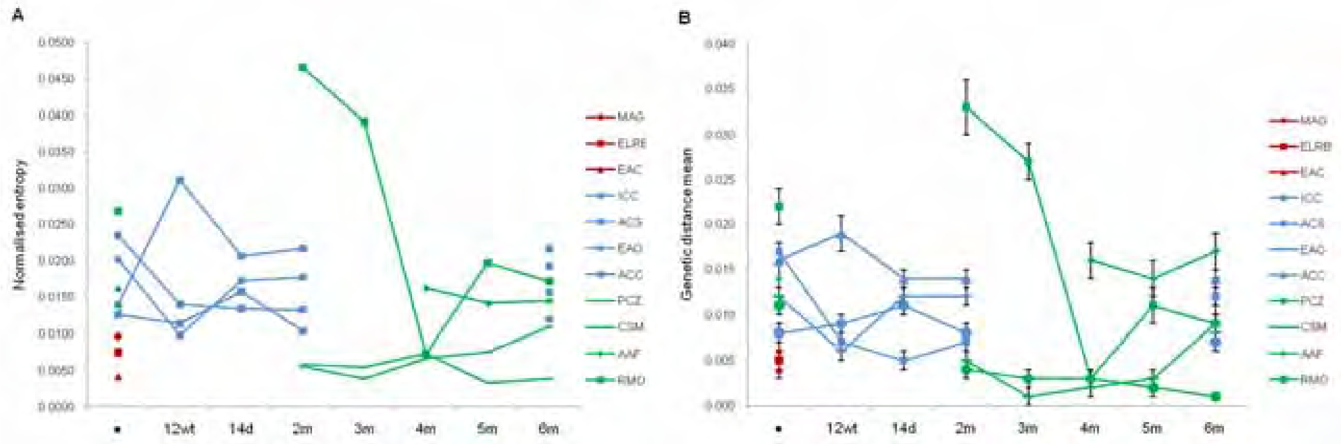

A
Normalised entropy
0.0000
0.0050
0.0100
0.0150
0.0200
0.0250
0.0300
0.0350
0.0400
0.0450
0.0500
12wt
14d
2m
3m
4m
5m
6m
MAG
ELRB
EAC
ICC
ACS
EAO
ACC
PCZ
CSM
AAF
RMO
B
Genetic distance mean
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
0.040

Figure 3

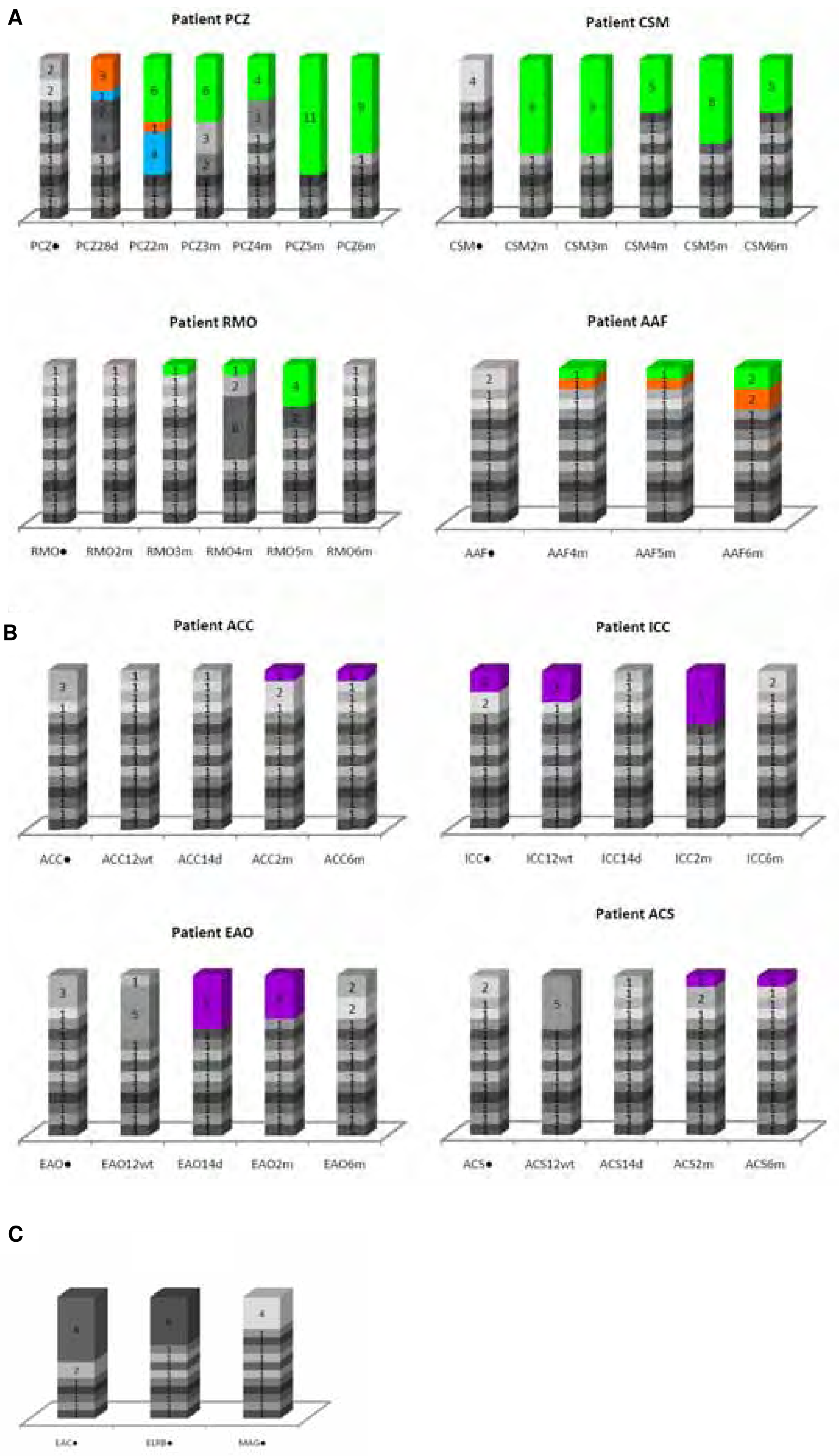

A
Patient PCZ
PCZ● PCZ28d PCZ2m PCZ3m PCZ4m PCZ5m PCZ6m
Patient CSM
CSM● CSM2m CSM3m CSM4m CSM5m CSM6m
Patient RMO
RMO● RMO2m RMO3m RMO4m RMO5m RMO6m
Patient AAF
AAF● AAF4m AAF5m AAF6m
B
Patient ACC
ACC● ACC12wt ACC14d ACC2m ACC6m
Patient ICC
ICC● ICC12wt ICC14d ICC2m ICC6m
Patient EAO
EAO● EAO12wt EAO14d EAO2m EAO6m
Patient ACS
ACS● ACS12wt ACS14d ACS2m ACS6m
C

Figure 4

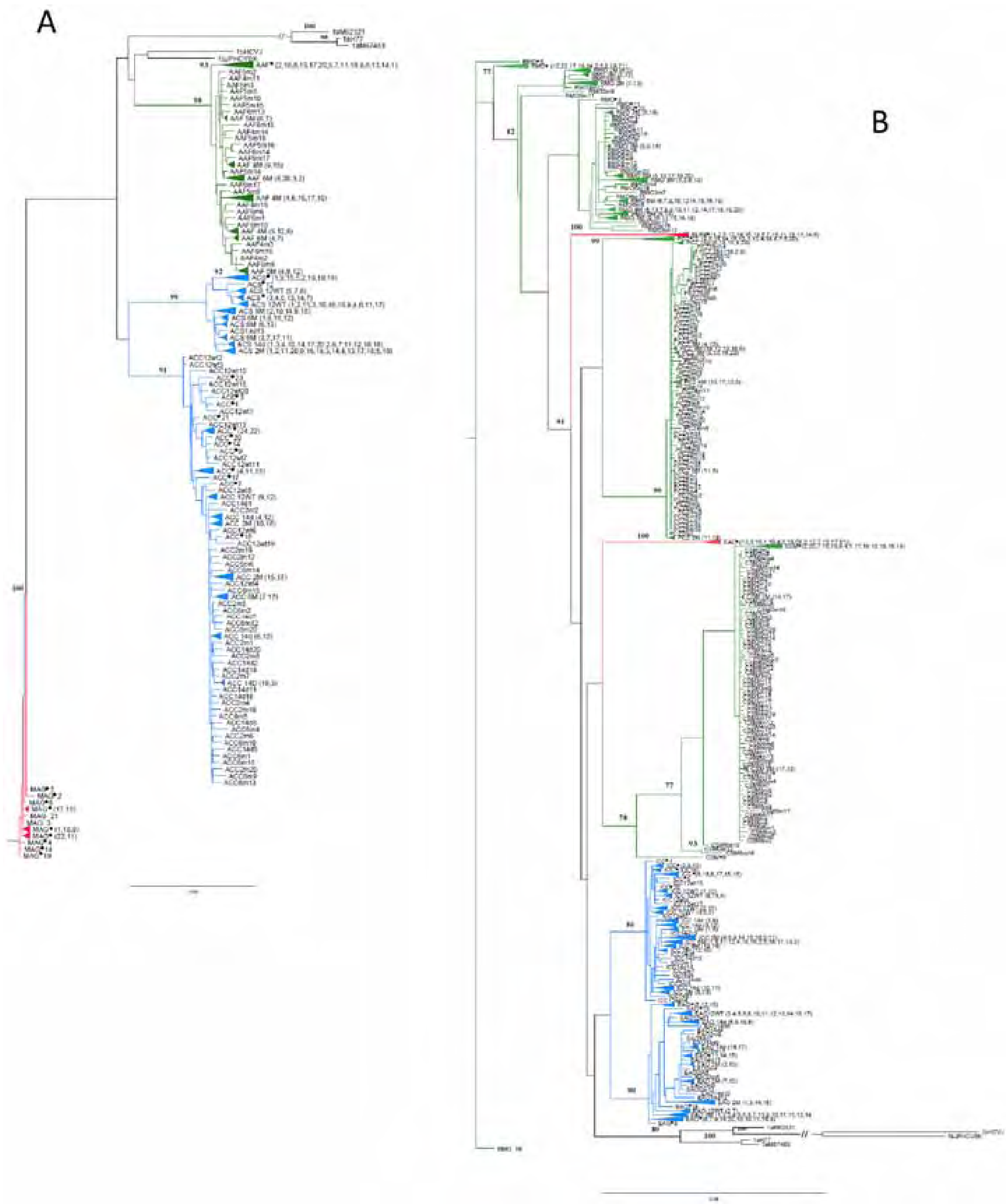
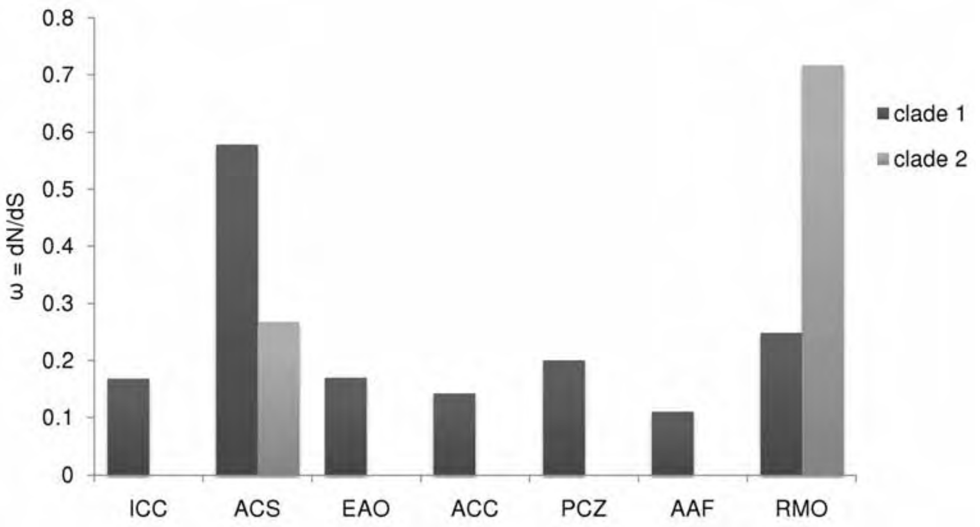


Figure 5


0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
ω = dN/dS
ICC
ACS
EAO
ACC
PCZ
AAF
RMO
clade 1
clade 2

SUPPORTING INFORMATION

Figure S1. Maximum likelihood tree constructed from full length NS5A region sequences obtained from samples from patient AAF (ETR) plus reference sequence of genotype 1b HCV-J. The number of 1000 permuted trees supporting a clade indicated when that proportion was greater than 70%. The same quasispecies are colored in red or pink. A sequence with nonsense mutation is colored in blue. The clade selected for selective pressure analysis is indicated by a gray line.

Figure S2. Maximum likelihood tree constructed from full length NS5A region sequences obtained from samples of PCZ (ETR) plus reference sequence of genotype 1b HCV-J. The number of 1000 permuted trees supporting a clade indicated when that proportion was greater than 70%. The same quasispecies are colored in red. Sequence with nonsense mutation is colored in blue. The clade selected for selective pressure analysis is indicated by a gray line.

Figure S3. Maximum likelihood tree constructed from full length NS5A region sequences obtained from samples of RMO (ETR) plus reference sequence of genotype 1b HCV-J. The number of 1000 permuted trees supporting a clade indicated when that proportion was greater than 70%. The same quasispecies are colored in red. Sequence with nonsense mutation is colored in blue. The clades selected for selective pressure analysis are indicated by a gray line (clade 1, $\omega = 0.2487$ and clade 2, $\omega = 0.7177$).

Figure S4. Maximum likelihood tree constructed from full length NS5A region sequences obtained from samples of CSM (ETR) plus reference sequence of genotype 1b HCV-J. The number of 1000 permuted trees supporting a clade indicated when that proportion was greater than 70%. The same quasispecies are colored in red. Sequence with nonsense mutation is colored in blue.

Figure S5. Maximum likelihood tree constructed from full length NS5A region sequences obtained from samples of ACS (NR) plus reference sequence of genotype 1b HCV-J. The number of 1000 permuted trees supporting a clade indicated when that proportion was greater than 70%. The same quasispecies are colored in red. The clades selected for selective pressure analysis are indicated by a gray line (clade 1, $\omega = 0.5783$ and clade 2, $\omega = 0.2685$).

Figure S6. Maximum likelihood tree constructed from full length NS5A region sequences obtained from samples of ACC (NR) plus reference sequence of genotype 1b HCV-J. The number of 1000 permuted trees supporting a clade indicated when that proportion was greater than 70%. The same quasispecies are colored in red. Sequence with nonsense mutation is colored in blue. The clade selected for selective pressure analysis is indicated by a gray line.

Figure S7. Maximum likelihood tree constructed from full length NS5A region sequences obtained from samples of EAO (NR) plus reference sequence of genotype 1b HCV-J. The number of 1000 permuted trees supporting a clade indicated when that proportion was greater than 70%. The same quasispecies are colored in red. Sequence with nonsense mutation is colored in blue. The clade selected for selective pressure analysis is indicated by a gray line.

Figure S8. Maximum likelihood tree constructed from full length NS5A region sequences obtained from samples of ICC (NR) plus reference sequence of genotype 1b HCV-J. The number of 1000 permuted trees supporting a clade indicated when that proportion was greater than 70%. The same quasispecies are colored in red. The clade selected for selective pressure analysis is indicated by a gray line.

Figure S1

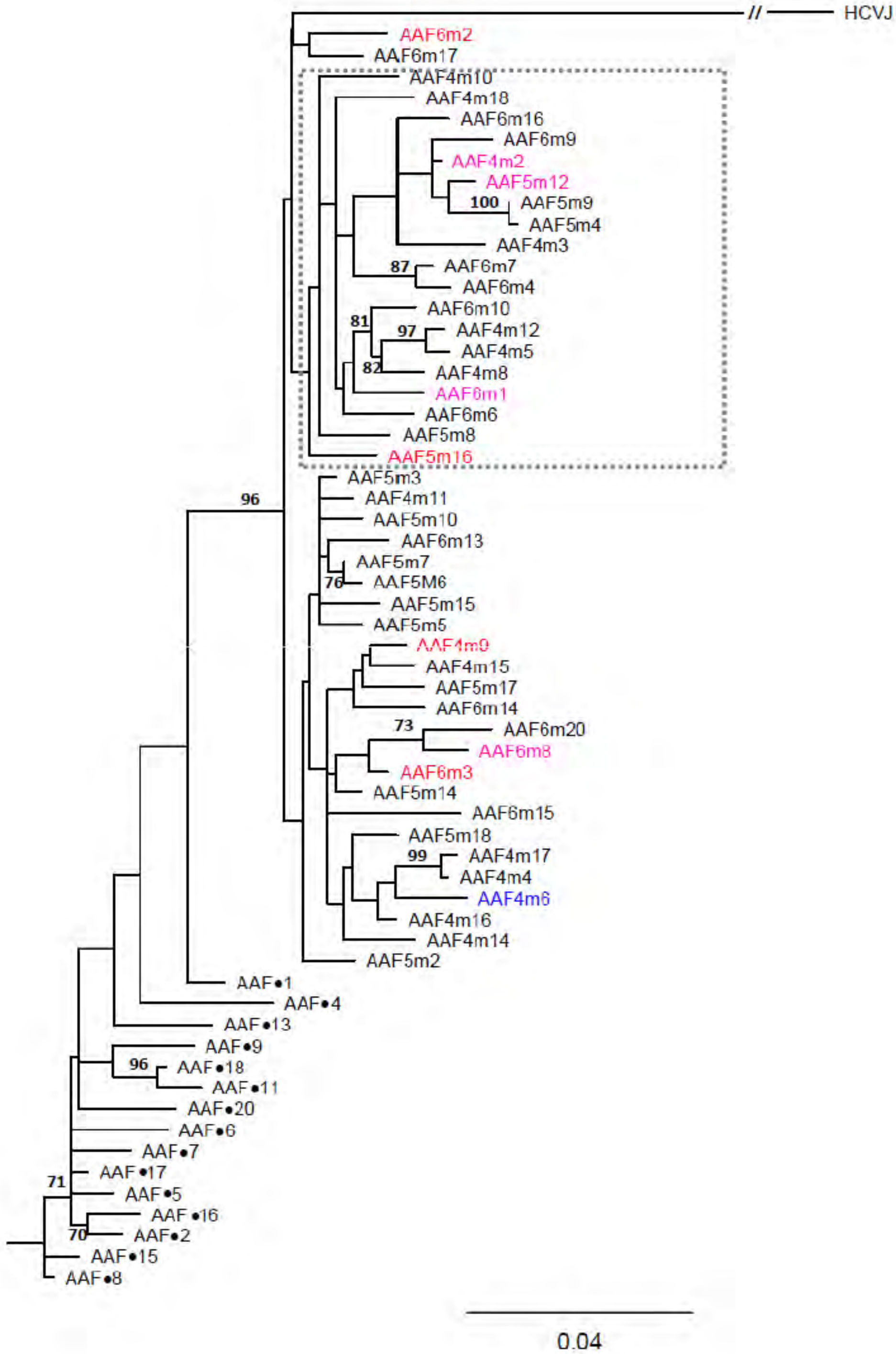

HCVJ
AAF6m2
AAF6m17
AAF4m10
AAF4m18
AAF6m16
AAF6m9
AAF4m2
AAF5m12
100
AAF5m9
AAF5m4
AAF4m3
87
AAF6m7
AAF6m4
AAF6m10
81
97
AAF4m12
AAF4m5
82
AAF4m8
AAF6m1
AAF6m6
AAF5m8
AAF5m16
AAF5m3
96
AAF4m11
AAF5m10
AAF6m13
AAF5m7
76
AAF5M6
AAF5m15
AAF5m5
AAF4m9
AAF4m15
AAF5m17
AAF6m14
73
AAF6m20
AAF6m8
AAF6m3
AAF5m14
AAF6m15
AAF5m18
99
AAF4m17
AAF4m4
AAF4m6
AAF4m16
AAF4m14
AAF5m2
AAF•1
AAF•4
AAF•13
AAF•9
96
AAF•18
AAF•11
AAF•20
AAF•6
AAF•7
AAF•17
71
AAF•5
AAF•16
70
AAF•2
AAF•15
AAF•8
0.04

Figure S2


H77
PCZ•20
PCZ•7
PCZ•13
PCZ•14
PCZ•10
PCZ•18
PCZ•8
80
100
PCZ5m19
71
74
82
PCZ28d20
PCZ6m5
PCZ5m13
PCZ4m19
PCZ4m2
PCZ4m8
PCZ4m9
70
PCZ3m18
PCZ5m18
85
PCZ3m2
PCZ4m11
PCZ4m13
0.02

Figure S3


RMO3m11
RMO2m7
RMO3m12
RMO2m13
RMO2m2
RMO2m3
RMO5m17
RMO2m15
RMO3m10
RMO3m6
RMO3m3
RMO3m1
RMO3m13
RMO5m18
RMO5m15
RMO5m9
RMO5m13
RMO3m2
RMO2m14
RMO2m11
RMO2m17
RMO6m5
RMO5m20
RMO2m4
RMO6m12
RMO6m9
RMO4m15
RMO3m4
RMO3m7
RMO4m20
RMO4m7
65
100
79
69
91
90
72
81
89
99
97
2
1
0.02

Figure S4

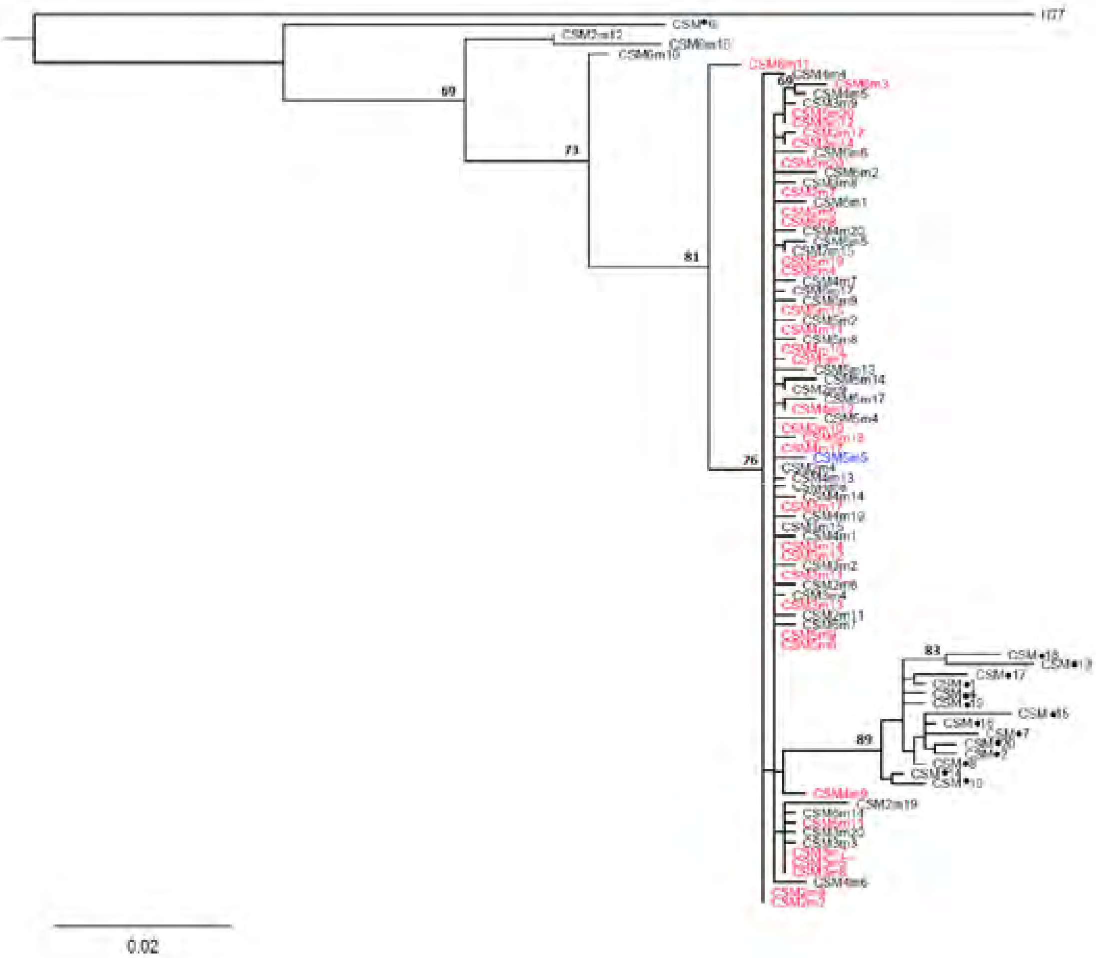


Figure S5

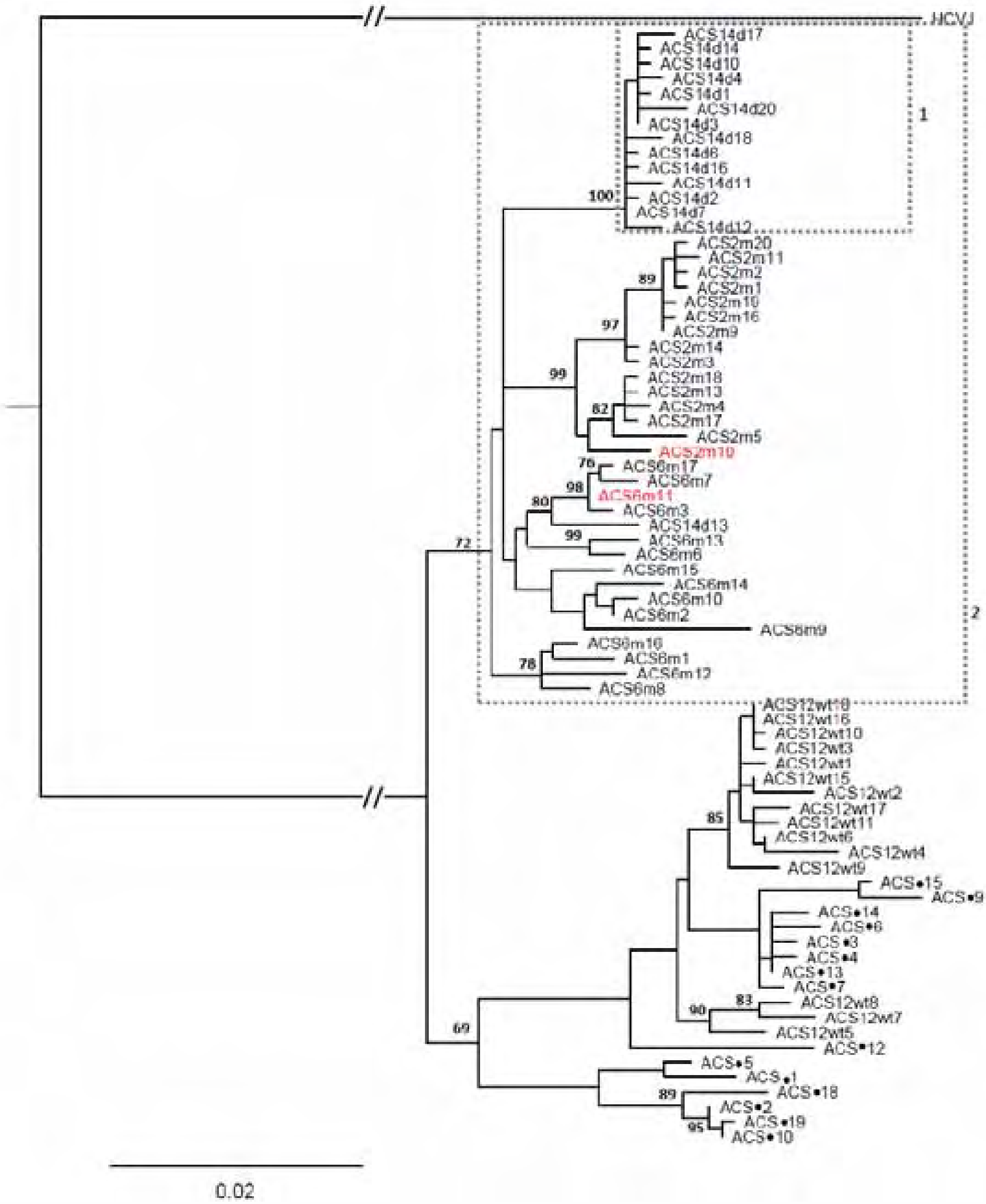

HCVJ
ACS14d17
ACS14d14
ACS14d10
ACS14d4
ACS14d1
ACS14d20
ACS14d3
ACS14d18
ACS14d6
ACS14d16
ACS14d11
ACS14d2
ACS14d7
ACS14d12
1
100
ACS2m20
ACS2m11
ACS2m2
ACS2m1
ACS2m10
ACS2m16
ACS2m9
ACS2m14
ACS2m3
ACS2m18
ACS2m13
ACS2m4
ACS2m17
ACS2m5
ACS2m19
89
97
99
82
ACS6m17
ACS6m7
ACS6m11
ACS6m3
ACS14d13
ACS6m13
ACS6m6
ACS6m15
ACS6m14
ACS6m10
ACS6m2
ACS6m9
76
98
80
99
72
2
ACS6m16
ACS6m1
ACS6m12
ACS6m8
78
ACS12wt18
ACS12wt16
ACS12wt10
ACS12wt3
ACS12wt1
ACS12wt15
ACS12wt2
ACS12wt17
ACS12wt11
ACS12wt6
ACS12wt4
ACS12wt9
85
ACS•15
ACS•9
ACS•14
ACS•6
ACS•3
ACS•4
ACS•13
ACS•7
ACS12wt8
ACS12wt7
ACS12wt5
ACS•12
83
90
69
ACS•5
ACS•1
ACS•18
ACS•2
ACS•19
ACS•10
89
95
0.02

Figure S6

HCVJ
92
ACC12wt3
ACC12wt2
ACC12wt13
ACC12wt11
ACC12wt7
ACC12wt10*
ACC12wt6
ACC12wt19
ACC14d11
ACC6m15
ACC6m5
69
ACC14d18
ACC6m16
ACC6m1
ACC14d5
ACC6m13
ACC2m20
ACC14d8
ACC6m9
ACC6m4
ACC6m10
ACC2m6
ACC2m4
ACC6m20
ACC6m12
ACC6m2
ACC2m5
ACC6m14
ACC2m17
ACC2m15
ACC2m10
ACC14d1
ACC2m19
ACC2m18
ACC2m2
ACC6m17
ACC6m7
ACC12wt4
ACC6m6
ACC2m8
ACC14d20
ACC2m1
ACC2m7
ACC14d14
96
ACC14d19
ACC14d3
ACC14d2
ACC14d10
ACC14d6
ACC14d12
ACC2m16
ACC14d4
ACC12wt12
ACC12wt9
ACC12wt5
94
ACC12wt20
ACC12wt16
ACC12wt1
0.03

Figure S7

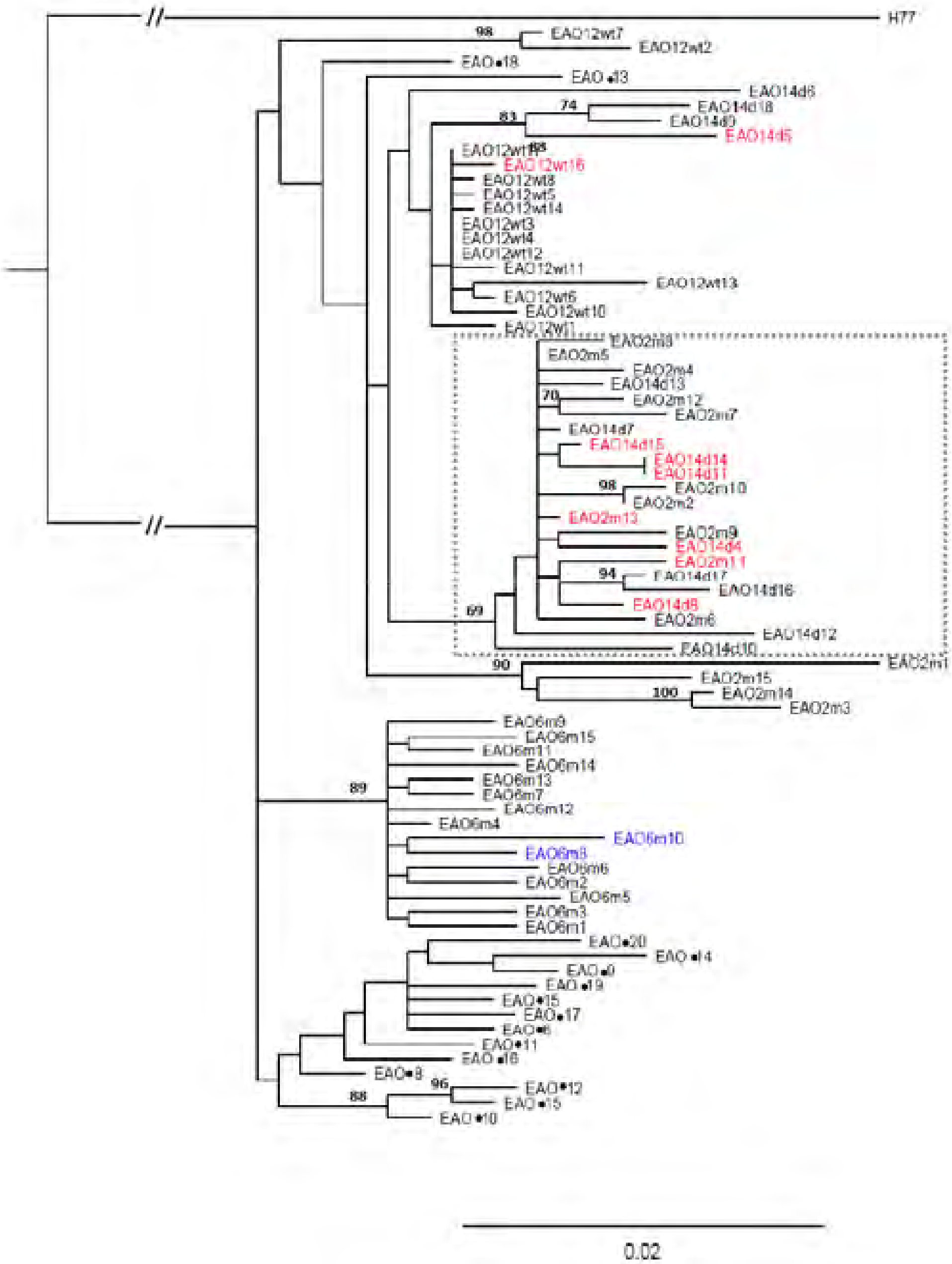
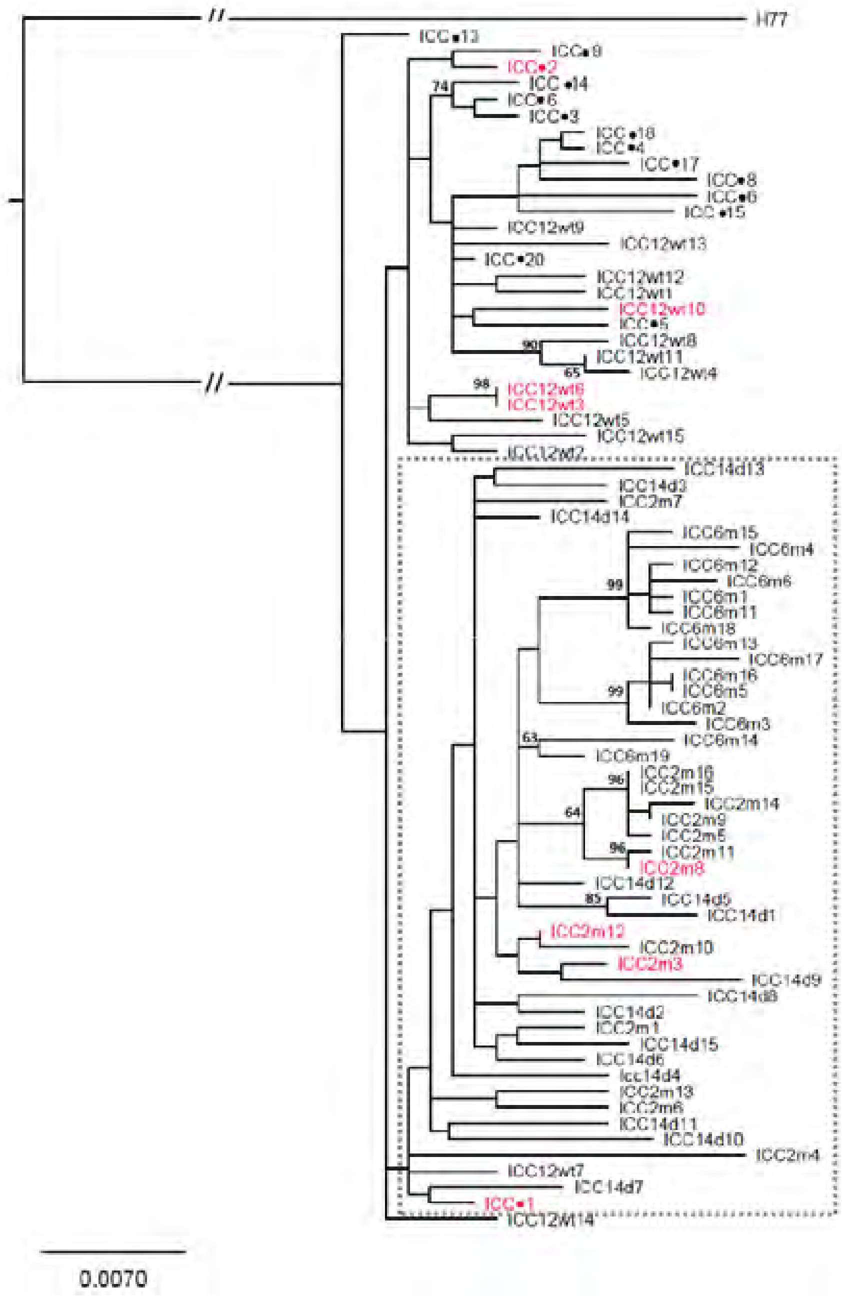

H77
98
EAO12wt7
EAO12wt2
74
EAO14d6
EAO14d18
83
EAO14d9
EAO14d5
EAO12wt16
EAO12wt8
EAO12wt5
EAO12wt14
EAO12wt3
EAO12wt4
EAO12wt12
EAO12wt11
EAO12wt13
EAO12wt6
EAO12wt10
EAO2m8
EAO2m5
EAO2m4
EAO14d13
70
EAO2m12
EAO2m7
EAO14d7
EAO14d15
EAO14d14
EAO14d11
98
EAO2m10
EAO2m2
EAO2m13
EAO2m9
EAO14d4
EAO2m11
94
EAO14d17
EAO14d16
EAO14d8
69
EAO2m6
EAO14d12
EAO14d10
EAO2m1
90
EAO2m15
100
EAO2m14
EAO2m3
EAO6m9
EAO6m15
EAO6m11
EAO6m14
89
EAO6m13
EAO6m7
EAO6m12
EAO6m4
EAO6m10
EAO6m8
EAO6m6
EAO6m2
EAO6m5
EAO6m3
EAO6m1
96
88
0.02

Figure S8


H77
ICC•13
ICC•9
ICC•2
74
ICC •14
ICC•6
ICC•3
ICC •18
ICC•4
ICC•17
ICC•8
ICC•6
ICC •15
ICC12wt9
ICC12wt13
ICC•20
ICC12wt12
ICC12wt1
ICC12wt10
ICC•5
90
ICC12wt8
ICC12wt11
65
ICC12wt4
98
ICC12wt6
ICC12wt3
ICC12wt5
ICC12wt15
ICC12wt2
ICC14d13
ICC14d3
ICC2m7
ICC14d14
ICC6m15
ICC6m4
ICC6m12
ICC6m6
99
ICC6m1
ICC6m11
ICC6m18
ICC6m13
ICC6m17
ICC6m16
ICC6m5
99
ICC6m2
ICC6m3
63
ICC6m14
ICC6m19
96
ICC2m16
ICC2m15
64
ICC2m14
ICC2m9
ICC2m5
96
ICC2m11
ICC2m8
ICC14d12
85
ICC14d5
ICC14d1
ICC2m12
ICC2m10
ICC2m3
ICC14d9
ICC14d8
ICC14d2
ICC2m1
ICC14d15
ICC14d6
Icc14d4
ICC2m13
ICC2m6
ICC14d11
ICC14d10
ICC2m4
ICC12wt7
ICC14d7
ICC•1
ICC12wt14
0.0070

CONCLUSÕES

4. CONCLUSÕES

- Pacientes que adquiriram resposta virológica sustentada apresentaram composições de quasiespecies do HCV menos diversa e complexa em amostras pré-tratamento quando avaliada a região NS5A. Contrariamente, alta variabilidade genética nesta região foi relacionada à baixa eficiência a resposta a terapia anti-HCV;

- O número de variantes virais nas amostras pré-tratamento também foi menor para pacientes respondedores demonstrando populações mais homogêneas quando comparadas com as populações heterogêneas de pacientes não respondedores e respondedores ao final do tratamento;

- Quasiespecies correspondentes a amostras pré e pós-terapia agruparam-se separadamente em análises filogenéticas enfatizando o processo evolucionário que influenciou a dinâmica de quasiespecies em diferentes momentos da infecção;

- Pacientes respondedores ao final do tratamento apresentaram homogeneização da composição de quasiespecies após o final da terapia, e emergência de quasiespecies predominantes que se mantiveram no restabelecimento da infecção crônica;

- Pacientes não respondedores apresentaram populações heterogêneas durante todo o curso da infecção;

- A dinâmica evolutiva das quasiespecies ao longo do tempo foi, no geral, dirigida por seleção purificadora ou purificadora relaxada, concordando com a produção e eliminação contínua de quasispecies em pacientes não-respondedores e o estabelecimento de predominantes em pacientes respondedores ao final do tratamento.

REFERÊNCIAS BIBLIOGRÁFICAS

5. REFERÊNCIAS BIBLIOGRÁFICAS

ACKERMAN, Z. Hepatitis C virus in various human body fluids: a systematic review. **Hepatology Research**, v. 11, p. 26-40, 1998.

ALTER, M. J. Epidemiology of hepatitis C virus infection. **World J Gastroenterol**, v. 13, n. 17, p. 2436-41, May 7 2007.

APPEL, N. et al. Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. **PLoS Pathog**, v. 4, n. 3, p. e1000035, Mar 2008.

ASABE, S. I. et al. The N-terminal region of hepatitis C virus-encoded NS5A is important for NS4A-dependent phosphorylation. **J Virol**, v. 71, n. 1, p. 790-6, Jan 1997.

BROOK, M. G. Sexually acquired hepatitis. **Sex Transm Infect**, v. 78, n. 4, p. 235-40, Aug 2002.

BROWN, R. S. Hepatitis C and liver transplantation. **Nature**, v. 436, n. 7053, p. 973-8, Aug 18 2005.

CAMPIOTTO, S. et al. Geographic distribution of hepatitis C virus genotypes in Brazil. **Braz J Med Biol Res**, v. 38, n. 1, p. 41-9, Jan 2005.

CAVALHEIRO NDE, P. et al. Hepatitis C: sexual or intrafamilial transmission? Epidemiological and phylogenetic analysis of hepatitis C virus in 24 infected couples. **Rev Soc Bras Med Trop**, v. 42, n. 3, p. 239-44, May-Jun 2009.

CHATTOPADHYAY, D. et al. Recent advancements for the evaluation of anti-viral activities of natural products. **N Biotechnol**, v. 25, n. 5, p. 347-68, Jun 2009.

CHAYAMA, K. et al. Pretreatment virus load and multiple amino acid substitutions in the interferon sensitivity-determining region predict the outcome of interferon treatment in patients with chronic genotype 1b hepatitis C virus infection. **Hepatology**, v. 25, n. 3, p. 745-9, Mar 1997.

CHEVALIEZ, S.; PAWLOTSKY, J. M. HCV Genome and Life Cycle. 2006.

CHISARI, F. V. Unscrambling hepatitis C virus-host interactions. **Nature**, v. 436, n. 7053, p. 930-2, Aug 18 2005.

CHOO, Q. L. et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. **Science**, v. 244, n. 4902, p. 359-62, Apr 21 1989.

CHOO, Q. L. et al. Hepatitis C virus: the major causative agent of viral non-A, non-B hepatitis. **Br Med Bull**, v. 46, n. 2, p. 423-41, Apr 1990.

CHUNG, H. et al. Risk of HCV transmission after needlestick injury, and the efficacy of short-duration interferon administration to prevent HCV transmission to medical personnel. **J Gastroenterol**, v. 38, n. 9, p. 877-9, 2003.

DAL MOLIN, G. et al. Mother-to-infant transmission of hepatitis C virus: rate of infection and assessment of viral load and IgM anti-HCV as risk factors. **J Med Virol**, v. 67, n. 2, p. 137-42, Jun 2002.

DE CARVALHO-MELLO, I. M. et al. Molecular evidence of horizontal transmission of hepatitis C virus within couples. **J Gen Virol**, v. 91, n. Pt 3, p. 691-6, Mar 2010.

DE FRANCESCO, R. et al. Approaching a new era for hepatitis C virus therapy: inhibitors of the NS3-4A serine protease and the NS5B RNA-dependent RNA polymerase. **Antiviral Res**, v. 58, n. 1, p. 1-16, Mar 2003.

DE MARTIN, E. et al. Fibrosis progression and the pros and cons of antiviral therapy for hepatitis C virus recurrence after liver transplantation: a review. **Transplant Proc**, v. 42, n. 6, p. 2223-5, Jul-Aug 2010.

DOMINGO, E. et al. Viruses as quasispecies: biological implications. **Curr Top Microbiol Immunol**, v. 299, p. 51-82, 2006.

ENOMOTO, N. et al. Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. **J Clin Invest**, v. 96, n. 1, p. 224-30, Jul 1995.

ENOMOTO, N. et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. **N Engl J Med**, v. 334, n. 2, p. 77-81, Jan 11 1996.

FELD, J. J.; HOOFNAGLE, J. H. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. **Nature**, v. 436, n. 7053, p. 967-72, Aug 18 2005.

FERRERO, S. et al. Prospective study of mother-to-infant transmission of hepatitis C virus: a 10-year survey (1990-2000). **Acta Obstet Gynecol Scand**, v. 82, n. 3, p. 229-34, Mar 2003.

FLAMM, S. L. Chronic hepatitis C virus infection. **JAMA**, v. 289, n. 18, p. 2413-7, May 14 2003.

GALE, M., JR. et al. Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. **Mol Cell Biol**, v. 18, n. 9, p. 5208-18, Sep 1998.

GALE, M., JR.; FOY, E. M. Evasion of intracellular host defence by hepatitis C virus. **Nature**, v. 436, n. 7053, p. 939-45, Aug 18 2005.

GALE, M. J., JR. et al. Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein. **Virology**, v. 230, n. 2, p. 217-27, Apr 14 1997.

HANOULLE, X. et al. Domain 3 of non-structural protein 5A from hepatitis C virus is natively unfolded. **Biochem Biophys Res Commun**, v. 381, n. 4, p. 634-8, Apr 17 2009.

HAYASHI, N.; TAKEHARA, T. Antiviral therapy for chronic hepatitis C: past, present, and future. **J Gastroenterol**, v. 41, n. 1, p. 17-27, Jan 2006.

HE, Y.; A., S. K.; TAN, S.-L. HCV NS5A: A Multifunctional Regulator of Cellular Pathways and Virus Replication. In: TAN, S.-L. (Ed.). **Hepatitis C Viruses Genomes and Molecular Biology**. Indianapolis, USA, 2006.

HOFMANN, W. P.; ZEUZEM, S.; SARRAZIN, C. Hepatitis C virus-related resistance mechanisms to interferon alpha-based antiviral therapy. **J Clin Virol**, v. 32, n. 2, p. 86-91, Feb 2005.

HUGHES, M. et al. A conserved proline between domains II and III of hepatitis C virus NS5A influences both RNA replication and virus assembly. **J Virol**, v. 83, n. 20, p. 10788-96, Oct 2009.

HYDER, Q. et al. Non-conventional transmission of hepatitis C: a true possibility ignored. **J Pak Med Assoc**, v. 59, n. 7, p. 430-3, Jul 2009.

JAIN, M. K. et al. Pegylated interferon and ribavirin promote early evolution of nonstructural 5A protein in individuals with hepatitis C who demonstrate a response to treatment. **J Infect Dis**, v. 200, n. 6, p. 866-76, Sep 15 2009.

JAWAID, A.; KHUWAJA, A. K. Treatment and vaccination for hepatitis C: present and future. **J Ayub Med Coll Abbottabad**, v. 20, n. 1, p. 129-33, Jan-Mar 2008.

KAITO, M. et al. Hepatitis C virus particle detected by immunoelectron microscopic study. **J Gen Virol**, v. 75 (Pt 7), p. 1755-60, Jul 1994.

KITAZATO, K.; WANG, Y.; KOBAYASHI, N. Viral infectious disease and natural products with antiviral activity. **Drug discoveries & therapeutics**, v. 1, n. 1, 2007.

KLENERMAN, P.; FLEMING, V.; BARNES, E. What are the prospects for controlling hepatitis C? **PLoS Med**, v. 6, n. 6, p. e1000096, Jun 16 2009.

KOMATSU, H. et al. Mutations in the nonstructural protein 5A gene and response to interferon therapy in young patients with chronic hepatitis C virus 1b infection. **J Med Virol**, v. 53, n. 4, p. 361-5, Dec 1997.

KUIKEN, C.; SIMMONDS, P. Nomenclature and numbering of the hepatitis C virus. **Methods Mol Biol**, v. 510, p. 33-53, 2009.

KUROSAKI, M. et al. Analysis of genotypes and amino acid residues 2209 to 2248 of the NS5A region of hepatitis C virus in relation to the response to interferon-beta therapy. **Hepatology**, v. 25, n. 3, p. 750-3, Mar 1997.

LA TORRE, G. et al. Correlates of HCV seropositivity among familial contacts of HCV positive patients. **BMC Public Health**, v. 6, p. 237, 2006.

LAVANCHY, D. The global burden of hepatitis C. **Liver Int**, v. 29 Suppl 1, p. 74-81, Jan 2009.

LE GUILLOU-GUILLEMETTE, H. et al. Genetic diversity of the hepatitis C virus: impact and issues in the antiviral therapy. **World J Gastroenterol**, v. 13, n. 17, p. 2416-26, May 7 2007.

LIANG, Y. et al. Domain 2 of nonstructural protein 5A (NS5A) of hepatitis C virus is natively unfolded. **Biochemistry**, v. 46, n. 41, p. 11550-8, Oct 16 2007.

LINDENBACH, B. D.; RICE, C. M. Unravelling hepatitis C virus replication from genome to function. **Nature**, v. 436, n. 7053, p. 933-8, Aug 18 2005.

LINDSAY, K. L. Therapy of hepatitis C: overview. **Hepatology**, v. 26, n. 3 Suppl 1, p. 71S-77S, Sep 1997.

LOPES, C. L. et al. Prevalence, risk factors and genotypes of hepatitis C virus infection among drug users, Central-Western Brazil. **Rev Saude Publica**, v. 43 Suppl 1, p. 43-50, Aug 2009.

LYRA, A. C.; FAN, X.; DI BISCEGLIE, A. M. Molecular biology and clinical implication of hepatitis C virus. **Braz J Med Biol Res**, v. 37, n. 5, p. 691-5, May 2004.

MACDONALD, A.; HARRIS, M. Hepatitis C virus NS5A: tales of a promiscuous protein. **J Gen Virol**, v. 85, n. Pt 9, p. 2485-502, Sep 2004.

MALLET, V.; VALLET-PICHARD, A.; POL, S. New trends in hepatitis C management. **Presse Med**, v. 39, n. 4, p. 446-51, Apr 2010.

MARTELL, M. et al. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. **J Virol**, v. 66, n. 5, p. 3225-9, May 1992.

MARTINS, R. M. et al. Distribution of hepatitis C virus genotypes among blood donors from mid-west region of Brazil. **Rev Inst Med Trop Sao Paulo**, v. 48, n. 1, p. 53-5, Jan-Feb 2006.

MORENO-OTERO, R. Therapeutic modalities in hepatitis C: challenges and development. **J Viral Hepat**, v. 12, n. 1, p. 10-9, Jan 2005.

MUNIR, S. et al. Hepatitis C treatment: current and future perspectives. **Virol J**, v. 7, p. 296, 2010.

MURPHY, D. G. et al. Use of sequence analysis of the NS5B region for routine genotyping of hepatitis C virus with reference to C/E1 and 5' untranslated region sequences. **J Clin Microbiol**, v. 45, n. 4, p. 1102-12, Apr 2007.

PASCU, M. et al. Sustained virological response in hepatitis C virus type 1b infected patients is predicted by the number of mutations within the NS5A-ISDR: a meta-analysis focused on geographical differences. **Gut**, v. 53, n. 9, p. 1345-51, Sep 2004.

PAWLOTSKY, J. M. Hepatitis C virus genetic variability: pathogenic and clinical implications. **Clin Liver Dis**, v. 7, n. 1, p. 45-66, Feb 2003.

PAWLOTSKY, J. M. Therapy of hepatitis C: from empiricism to eradication. **Hepatology**, v. 43, n. 2 Suppl 1, p. S207-20, Feb 2006.

PAWLOTSKY, J. M. et al. Interferon resistance of hepatitis C virus genotype 1b: relationship to nonstructural 5A gene quasispecies mutations. **J Virol**, v. 72, n. 4, p. 2795-805, Apr 1998.

PAWLOTSKY, J. M. et al. Genetic complexity of the hypervariable region 1 (HVR1) of hepatitis C virus (HCV): influence on the characteristics of the infection and responses to interferon alfa therapy in patients with chronic hepatitis C. **J Med Virol**, v. 54, n. 4, p. 256-64, Apr 1998.

PENIN, F. et al. Structural biology of hepatitis C virus. **Hepatology**, v. 39, n. 1, p. 5-19, Jan 2004.

PUIG-BASAGOITI, F. et al. Dynamics of hepatitis C virus NS5A quasispecies during interferon and ribavirin therapy in responder and non-responder patients with genotype 1b chronic hepatitis C. **J Gen Virol**, v. 86, n. Pt 4, p. 1067-75, Apr 2005.

REICHE, E. M. et al. Epidemiological, immunological and virological characteristics, and disease progression of HIV-1/HCV-co-infected patients from a southern Brazilian population. **Int J Mol Med**, v. 21, n. 3, p. 387-95, Mar 2008.

ROMANO, C. M. et al. Social networks shape the transmission dynamics of hepatitis C virus. **PLoS One**, v. 5, n. 6, p. e11170, 2010.

SALMERON, J. et al. Quasispecies as predictive response factors for antiviral treatment in patients with chronic hepatitis C. **Dig Dis Sci**, v. 51, n. 5, p. 960-7, May 2006.

SALUDES, V. et al. Baseline prediction of combination therapy outcome in hepatitis C virus 1b infected patients by discriminant analysis using viral and host factors. **PLoS One**, v. 5, n. 11, p. e14132, 2010.

SAUTER, D. et al. Localization determines function: N-terminally truncated NS5A fragments accumulate in the nucleus and impair HCV replication. **J Hepatol**, v. 50, n. 5, p. 861-71, May 2009.

SHARMA, S. D. Hepatitis C virus: molecular biology & current therapeutic options. **Indian J Med Res**, v. 131, p. 17-34, Jan 2010.

SHEPARD, C. W.; FINELLI, L.; ALTER, M. J. Global epidemiology of hepatitis C virus infection. **Lancet Infect Dis**, v. 5, n. 9, p. 558-67, Sep 2005.

SHIMAKAMI, T.; LANFORD, R. E.; LEMON, S. M. Hepatitis C: recent successes and continuing challenges in the development of improved treatment modalities. **Curr Opin Pharmacol**, v. 9, n. 5, p. 537-44, Oct 2009.

SIMMONDS, P. Genetic diversity and evolution of hepatitis C virus--15 years on. **J Gen Virol**, v. 85, n. Pt 11, p. 3173-88, Nov 2004.

SIMMONDS, P. et al. A proposed system for the nomenclature of hepatitis C viral genotypes. **Hepatology**, v. 19, n. 5, p. 1321-4, May 1994.

SIMMONDS, P. et al. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. **Hepatology**, v. 42, n. 4, p. 962-73, Oct 2005.

TANG, H.; GRISE, H. Cellular and molecular biology of HCV infection and hepatitis. **Clin Sci (Lond)**, v. 117, n. 2, p. 49-65, Jul 2009.

TANJI, Y. et al. Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. **J Virol**, v. 69, n. 3, p. 1575-81, Mar 1995.

TANJI, Y. et al. Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. **J Virol**, v. 69, n. 7, p. 3980-6, Jul 1995.

TELLINGHUISEN, T. L. et al. The NS5A protein of hepatitis C virus is a zinc metalloprotein. **J Biol Chem**, v. 279, n. 47, p. 48576-87, Nov 19 2004.

TELLINGHUISEN, T. L.; MARCOTRIGIANO, J.; RICE, C. M. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. **Nature**, v. 435, n. 7040, p. 374-9, May 19 2005.

TOYODA, H. et al. Quasispecies nature of hepatitis C virus and response to alpha interferon: significance as a predictor of direct response to interferon. **J Hepatol**, v. 26, n. 1, p. 6-13, Jan 1997.

UEDA, E. et al. Changes of HCV quasispecies during combination therapy with interferon and ribavirin. **Hepatol Res**, v. 29, n. 2, p. 89-96, Jun 2004.

VOLK, M. L. Antiviral therapy for hepatitis C: why are so few patients being treated? **J Antimicrob Chemother**, v. 65, n. 7, p. 1327-9, Jul 2010.

WITHERELL, G. W.; BEINEKE, P. Statistical analysis of combined substitutions in nonstructural 5A region of hepatitis C virus and interferon response. **J Med Virol**, v. 63, n. 1, p. 8-16, Jan 2001.

WOHNSLAND, A.; HOFMANN, W. P.; SARRAZIN, C. Viral determinants of resistance to treatment in patients with hepatitis C. **Clin Microbiol Rev**, v. 20, n. 1, p. 23-38, Jan 2007.

WRIGHT, P. E.; DYSON, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. **J Mol Biol**, v. 293, n. 2, p. 321-31, Oct 22 1999.

ZHOU, D. et al. Separation of near full-length hepatitis C virus quasispecies variants from a complex population. **J Virol Methods**, v. 141, n. 2, p. 220-4, May 2007.

ANEXO

ANEXO 1

Ana Carolina Gomes <caroljardim@gmail.com>

Fwd: Submission Confirmation for Quasispecies variability in the NS5A region of Hepatitis C virus genotype 1 is indicative of non-response to peginterferon and ribavirin therapy

1 mensagem

prahal@ibilce.unesp.br <prahal@ibilce.unesp.br> **16 de fevereiro de 2011 07:47**
Para: caroljardim@gmail.com

This message was sent using IMP, the Internet Messaging Program.

-------- Mensagem encaminhada --------
From: "PLoS ONE" <plosone@plos.org>
To: "Paula Rahal" <prahal@ibilce.unesp.br>
Date: 16 Feb 2011 01:26:26 -0500
Subject: Submission Confirmation for Quasispecies variability in the NS5A region of Hepatitis C virus genotype 1 is indicative of non-response to peginterferon and ribavirin therapy
Dear Ms Rahal,

Your submission entitled "Quasispecies variability in the NS5A region of Hepatitis C virus genotype 1 is indicative of non-response to peginterferon and ribavirin therapy" has been received by PLoS ONE. You will be able to check on the progress of your paper by logging on to Editorial Manager as an author. The URL is http://pone.edmgr.com/.

Your manuscript will be given a reference number once an Editor has been assigned.

We would also like to forewarn you that we are experiencing a residual slowdown incurred by the last end-of-year holiday period, so there could be a delay in the review of your manuscript. Our editorial board and reviewers are comprised of faculty and staff from universities around the world, and many of these individuals took vacation during that period. We will do our utmost to process your manuscript in a prompt manner, but wanted you to be aware of these delays that occurred outside of our control. For more information, see our blog post at: http://blogs.plos.org/everyone/2010/11/22/holiday-service-update/.

Thank you for submitting your work to this journal.

Kind regards,

PLoS ONE

Autorizo a reprodução xenográfica para fins de pesquisa.

São José do Rio Preto, 04 de março de 2011

____

Assinatura