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

Evolução das quasiespécies da proteína NS5A doVírus da Hepatite C genótipo 3a

CÍNTIA BITTAR OLIVA

São José do Rio Preto – SP

2012

CÍNTIA BITTAR OLIVA

Evolução das quasiespécies da proteína NS5A doVírus da Hepatite C genótipo 3a

Tese apresentada para obtenção do Título de Doutora em Genética

Orientadora: Profa. Dra. Paula Rahal

Co-orientadora: Profa. Dra. Isabel Maria Vicente Guedes de Carvalho Mello

São José do Rio Preto - SP

2012

Oliva, Cíntia Bittar.
Evolução das quasiespécies da proteína NS5A do Vírus da Hepatite C genótipo 3a/ Cíntia Bittar Oliva. – São José do Rio Preto: [s.n.], 2012
65f. :il ; 30 cm.

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

1. Vírus. 2. Vírus da hepatite C. 3. Quasiespécies 4. Evolução. 5. NS5A. I. Rahal, Paula. II. Mello, Isabel Maria Vicente Guedes de Carvalho. III. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. IV. Título.

CDU - 578.53

CÍNTIA BITTAR OLIVA

Evolução das quasiespécies da proteína NS5A doVírus da Hepatite C genótipo 3a

Tese apresentada para obtenção do Título de Doutora 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
Instituto de Biociências, Letras e Ciências Exatas – UNESP
São José do Rio Preto

Profª. Drª. Flora Maria de Campos Fernandes
Universidade Federal da Bahia
Salvador

Drª Camila Malta Romano
Universidade de São Paulo - Instituto de Medicina Tropical de São Paulo

Prof.Dr. Adriano Mondini
Faculdade de Ciências Farmacêuticas – UNESP
Araraquara

Prof[a.] Dr[a.] Maria Tercília Vilela de Azeredo Oliveira
Instituto de Biociências, Letras e Ciências Exatas – UNESP
São José do Rio Preto

São José do Rio Preto, 24 de fevereiro de 2012.

Dedico este trabalho ...

A minha mãeCatrin Bittar,

Que sempre esteve presente em todos os momentos da minha vida, dos mais simples aos mais importantes, me ajudando e incentivando.

Mais uma vez, obrigada pelo seu amor, carinho e dedicação.

Ao meu pai José Luiz Silva Oliva,

Que me ama, respeita e sempre acreditou em mim.

Obrigada pelo seu amor e carinho.

Amo Vocês!

Agradecimentos

Acredito que realizar um doutorado vai muito além dos quatro anos de desenvolvimento da Tese, é o final de um longo caminho que percorremos durante nossas vidas descobrindo onde e porque queremos chegar. Fizeram parte desse caminho diversas pessoas como familiares, professores e amigos e todos eles contribuem de uma forma ou de outra para aminha formação e me auxiliaram a tomar as decisões que me trouxeram a este momento em que finalizo minha Tese de Doutorado. Agradeço a todos que passaram em algum momento pela minha vida, todos contribuíram para minha formação pessoal e profissional e me ajudaram a chegar até aqui.

Faço aqui alguns agradecimentos especiais...

À minha orientadora Profª. Drª. Paula Rahal pelas oportunidades, por confiar em meu trabalho e pelo carinho e amizade

⭐

À minha co-orientadora Profª. Drª.Isabel Maria Vicente Guedes de Carvalho-Mello pela amizade, carinho e paciência

⭐

Aos amigos do Laboratório de Estudos Genômicos, obrigada por fazerem do ambiente de trabalho um lugar leve e divertido

⭐

Ã Lenira pela amizade, obrigada por toda ajuda e carinho

⭐

À Carol Jardim eMarília Calmon por serem grandes amigas em todos os momentos

⋆

A Carol Bonfim pela amizade e pelo auxílio no desenvolvimento desse trabalho

⋆

À Lílian Yamasaki pela amizade e por continuar me socorrendo quando preciso

⋆

À Renata Mattos pelo auxílio no desenvolvimento deste trabalho

⋆

Aos colaboradores deste trabalho pelo auxílio para desenvolvê-lo

⋆

Aos meus bebês Maluk, Spike, Adam, Preta, Julieta, Bambi, Debby, Llyr, Biba, Brisa, Lugh, Baby, Rusty, Joey, Lolla ...vocês fazem a vida mais leve com seu carinho e amor incondicional.... amo vocês!

⋆

À FAPESP pelo auxílio financeiro

⋆

Aos membros da banca por se disponibilizarem em avaliar e enriquecer este trabalho

⋆

Aos mestres do IBILCE …

Durante meu doutorado tive a oportunidade de realizar um estagio no exterior. Durante esse período tive contato com outros pesquisadores, diversos assuntos e técnicas, precisei buscar soluções rápidas, devido ao curto período de tempo, para os problemas encontrados. Nem sempre os assuntos eram parte dos conhecimentos específicos que adquiri nos últimos anos. Mas percebi que consegui lidar com todos com relativa facilidade. Acredito que devido ao distanciamento da rotina, percebi, não que já não soubesse, que a facilidade em enfrentar os obstáculos e encontrar soluções mesmo em assuntos que não eram de meu inteiro domínio até então, se devia em grande parte a base que recebi de meus professores durante a graduação. Cada um de vocês forneceu ferramentas... que mesmo depois de algum tempo .. ainda estão presentes e ficam guardadas para quando forem necessárias. Eu mesma não fazia idéia das coisas que ainda sabia! Esse é o maior tesouro que não só um pesquisador, mas que uma pessoa pode ter! Durante este período também escrevi a versão final deste trabalho que apesar de ser um trabalho científico tem um ligeiro toque pessoal...e sou grata por ter aprendido com vocês que é permitido a um trabalho científico ser elegante e carregar um pouco de nossa personalidade. Enfim, agradeço a todos vocês por terem me ensinado não só a teoria, mas principalmente por terem deixado a personalidade de vocês fazerem parte desse aprendizado.

SUMÁRIO

PARTE I

PARTE II

PARTE I

Resumo, Introdução e Objetivos

LISTA DE FIGURAS

PARTE 1

PARTE 2

Supplementary material

LISTA DE TABELAS

LISTA DE ABREVIATURAS, SÍMBOLOS E SIGLAS

HCV	*Hepatitis C Vírus* (Vírus da Hepatite C)
HCC	Hepatocellular carcinoma (Carcinoma Hepatocelular)
DNA	Ácido desoxiribonucleico
RNA	Ácido ribonucléico
kb	Quilobases
C	Proteína Core
E1	Proteína do envelope 1
E2	Proteína do envelope 2
p7	Proteína p7
NS2	Proteína não estrutural 2
NS3	Proteína não estrutural 3
NS4A	Proteína não estrutural 4A
NS4B	Proteína não estrutural 4B
NS5A	Proteína não estrutural 5A
NS5B	Proteína não estrutural 5B
UTR	*Untranslatedregion* – Região não traduzida
IRES	*Internalribossomalentry site* (Sítio interno de entrada ribossomal)
IFN	Interferon
IFN-α	Interferon alfa

PEG-IFN	Interferon modificado por polietilenoglicol
RVS	Resposta virológica sustentada
FDA	US Foods and drugs administration
TVR	Telaprevir
BOC	Boceprovir
DAA	*Direct-acting antiviral drug*s(Drogas antivirais de ação direta)
CRS	*Cytoplasmicretentionsignal* (Sinal de retenção citoplasmático)
NLS	*Nuclear localizationsignal*(Sinal de localização nuclear)
PKR	*Proteinkinase RNA activated*(Quinase indutora de interferon ativada por RNA dupla fita)
ISDR	*Interferonsensitivitydeterminingregion* (Região Determinante de Sensibilidade ao Interferon)
IL8	Interleucina 8
V3	Regiãovariável 3
w	weeks
d	days
m	months
ω	Selective pressure estimation
dN	Number of non-synonymous substitutions per non-synonymous site
dS	Number of synonymous substitutions per synonymous site
BT	Before treatment

NR	Non responder
SVR	Sustained virological response
ETR	End of treatment response
Gln	Glutamin
Trp	Tryptophan
Lys	Lysin
HIV	Human Immunodeficiency Virus
cDNA	Complementary DNA
PCR	Polymerase chain reaction
µl	Microliter
$MgCl^2$	Magnesium Chloride
DMSO	Dimethyl sulfoxide
mM	Millimolar
pmol	Picomole
°C	Degrees Celsius
HKY	Hasegawa, Kishino and Yano 1985 model
G	Gamma distribution

RESUMO

RESUMO

A Hepatite C é uma doença presente em todo o mundo. O vírus da Hepatite C (HCV), o agente etiológico dessa doença, é um vírus de RNA de fita simples positiva. Seu genoma codifica uma única poliproteína precursora que após processamento origina dez proteínas virais. A NS5A, uma das proteínas virais não estruturais, esta associada com a resposta ao tratamento baseado em Interferon, tratamento aprovado para Hepatite C no Brazil.O HCV tem uma alta taxa de mutação levando a uma alta variabilidade, fator importante para a evasão da resposta imune e a resposta ao tratamento. O objetivo deste trabalho foi analisar a evolução das quasiespécies antes, durante e após o tratamento em pacientes infectados com HCV genótipo 3a que apresentaram diferentes respostas ao tratamento. O RNA viral foi extraído, o cDNA sintetizado, a região NS5A amplificada e clonada e 15 clones de cada ponto de coleta foram seqüenciados. As sequências foram analisadas com relação a história evolutiva, diversidade genética e seleção. Nossas análises mostram que a população viral que persiste após o tratamento na maioria dos pacientes não respondedores está presente em amostras pré-tratamento sugerindo uma aptidão para evadir o tratamento. Ainda a maioria das amostras pré-tratamento de pacientes respondedores ao final do tratamento ou não apresentou a população encontrada nas amostras pós-tratamento ou apresentou em menor freqüência. As exceções ilustram a característica única do processo evolutivo e conseqüentemente o processo de resistência ao tratamento em cada paciente. A evolução do vírus da Hepatite C ao longo do tratamento aparenta ser o resultado de uma relação evolutiva única entre as cepas virais e cada hospedeiro humano, levando a persistência do vírus ou a resposta ao tratamento.

ABSTRACT

ABSTRACT

Hepatitis C is a disease spread throughout the world. Hepatitis C virus (HCV), the etiological agent of this disease, is a single-stranded positive RNA virus. Its genome encodes a single precursor protein that yields ten proteins after processing. NS5A, one of the non-structural viral proteins, is most associated with interferon-based therapy response, the approved treatment for hepatitis C in Brazil. HCV has a high mutation rate and therefore high variability, which may be important for evading the immune system and response to therapy. The aim of this study was to analyze the evolution of NS5A quasispecies before, during, and after treatment in patients infected with HCV genotype 3a who presented different therapy responses.Viral RNA was extracted, cDNA was synthesized, the NS5A region was amplified and cloned, and 15 clones from each time-point were sequenced. The sequences were analyzed for evolutionary history, genetic diversity and selection. Our analysis shows that the viral population that persists after treatment for most non-response patients are is present in before-treatment samples, suggesting it is fitted to evasion of treatment. Accordingly, most before-treatment samples from end-of-treatment response patients either did not show the population found after the relapse or showed it in low abundance. The exceptions illustrate the uniqueness of the evolutionary process, and therefore the treatment resistance process, in each patient.Hepatitis C virus evolution throughout treatment appears to be the result of a unique evolutionary relationship between viral strains and each human host, leading to either persistence or clearance.

INTRODUÇÃO

1.INTRODUÇÃO

1.1 Histórico e Epidemiologia da Hepatite C

Originalmente acreditava-se que as Hepatites virais consistiam em dois tipos, Hepatite A e Hepatite B. Na ausência de testes diagnósticos que pudessem distinguir os agentes causadores, o diagnóstico ocorria pela identificação da via de transmissão e pelo tempo de incubação da doença(Krugman, Giles *et al.*, 1967). Com o isolamento dosvírus das Hepatites A e B amostras de sangue estocadas foram analisadas evidenciando que algumas amostras eram negativas para ambos os vírus identificados. Este fato levou a utilização do termo hepatite não-A não-B para os casos não identificados (Feinstone, Kapikian *et al.*, 1975). Em 1989 Chooet al isolou parte do genoma deste vírus desconhecido,a partir do plasma de um chimpanzé cronicamente infectado e este ficou conhecido como Vírus da Hepatite C (HCV), (Choo, Kuo *et al.*, 1989).

O HCV é um vírus de distribuição mundial sendo aprevalência global da infecção aproximadamente 2 a 3%, correspondendo a 130 a 170 milhões de pessoas soropositivas no mundo (Lavanchy, ; Shepard, Finelli *et al.*, 2005; Alter, 2007).

1.2 Fatores de Risco e Transmissão do vírus da Hepatite C

O HCV é transmitido pelo contato percutâneo com sangue contaminado. Os fatores de risco associados à infecção incluem transfusão de sangue e seus derivados, transplante de órgãos, usuários de drogas injetáveis, injeções terapêuticas não seguras, procedimentos percutâneos como piercing, tatuagem e circuncisão(com material inadequadamente esterilizado),exposição ocupacional a sangue, perinatal, e sexual (Who, 2000; Alter, 2007).A transmissão sexual é assunto ainda controversodevido à dificuldade de determinar se ocorreu por relações sexuais ou algum outro tipo de

contato próximo, sendo alvo de diversos estudos (Squadrito, Orlando *et al.*, 1999; Nakayama, Sugai *et al.*, 2005; Tahan, Karaca *et al.*, 2005; Mello, 2006; Urbanus, Van De Laar *et al.*, 2009; Nakamura, Tanaka *et al.*, 2011).

Antes da identificação do vírus causador da Hepatite C, a transfusão de sangue era a principal forma de transmissão da doença. Após a implementação dos testes-anti HCV em bancos de sangue a transmissão via transfusão foi praticamente eliminada nos países que adotamesse procedimento(Busch, Glynn *et al.*, 2005). Contudo, em países que não possuem recursos para testar os doadores essa ainda é uma importante via de transmissão juntamente com injeções terapêuticas não seguras(Hauri, Armstrong *et al.*, 2004; Hladik, Kataaha *et al.*, 2006; Alter, 2011).

Outra importante forma de transmissão do HCV é o uso de drogas injetáveis, sendo juntamente com a transfusão de sangue contaminado os dois principais fatores de risco. Usuários de drogas injetáveis constituem o grupo de risco com maior prevalência de infecção por HCV (Alter, 2011). Usuários com mais de trinta anos de idade que utilizam drogas há mais de cinco anos apresentam prevalência de 75% a 90%. Em usuários mais novos (abaixo de 30 anos) a prevalência nos primeiros 2 a 3 anos de uso reduziu de 80% no fim da década de 1980 para 30%desde o fim da década de 1990(Alter, 2011).

1.3 Sintomatologia e Prognóstico

O período de infecção, antes do aparecimento de sintomas clínicos, é de 15 a 150 dias. Na infecção aguda os sintomas mais comuns são fadiga e icterícia, porém a maioria dos casos, de 60% a 70%, é assintomática, mesmo quando evoluem para infecção crônica (Who, 2000).

Estima-se que a resolução espontânea da infecção ocorra em 14% a 29% dos casos (Alter, Purcell *et al.*, 1989; Aach, Stevens *et al.*, 1991; Alter, Conry-Cantilena *et al.*, 1997; Locasciulli, Testa *et al.*, 1997; Thomas, Astemborski *et al.*, 2000). Contudo uma taxa de resolução espontânea mais alta foi relatada em pessoas jovens variando de 42% a 45%(Kenny-Walsh, 1999; Vogt, Lang *et al.*, 1999; Casiraghi, De Paschale *et al.*, 2004; Seeff, 2009).

As pessoas infectadas que evoluem para a fase crônica possuem um risco de apresentar doença do fígado progressiva caracterizada pela inflamação continua levando a fibrose, cirrose e o carcinoma hepatocelular (HCC)(Seeff, 2000; , 2009). Dos indivíduos com infecção persistente estima-se que 20% irão desenvolver cirrose hepática e o carcinoma hepatocelular ocorre em até 2,5% dos casos. (Bowen e Walker, 2005). A princípio acreditava-se que a grande incidência de HCC em pacientes infectados por HCV era uma consequência indireta da infecção viral. A infecção pelo vírus da Hepatite C leva em muitos casos a cirrose hepática, fator que predispõe o desenvolvimento de câncer (Liang, Ye *et al.*, 2007). Outro importante fator indireto é a alta taxa de inflamação seguida de regeneração decorrente da infecção, combinado ao potencial de quebra de DNA dupla fita causada por estresse oxidativo(Nakamoto, Guidotti *et al.*, 1998; Okuda, Li *et al.*, 2002; Bartsch e Nair, 2004). Contudo hoje se acredita as proteínas virais possuam um papel direto na hepatocarcinogênese. As proteínas virais que vem sendo relacionadas com a hepatocarcinogênese são Core, E2, NS3 e NS5A (Levrero, 2006; Tsai e Chung, 2010).

1.4 O Vírus

O vírus da Hepatite C (HCV) é um vírus de RNA de fita simples, com polaridade positiva, pertencente à família Flaviviridae. O genoma do HCVpossui 9,6kb

e codifica uma grande poliproteína precursora de aproximadamente 3000 resíduos de aminoácidos, sendo essa clivada pelas proteases virais e do hospedeiro para gerar, pelo menos, 10 proteínas na seguinte ordem, iniciando pela extremidade amino terminal: core (C), envelope (E1; E2) e p7, não estruturais (NS2, NS3, NS4A, NS4B, NS5A, NS5B) (Kato, 2001). (Figura 1)

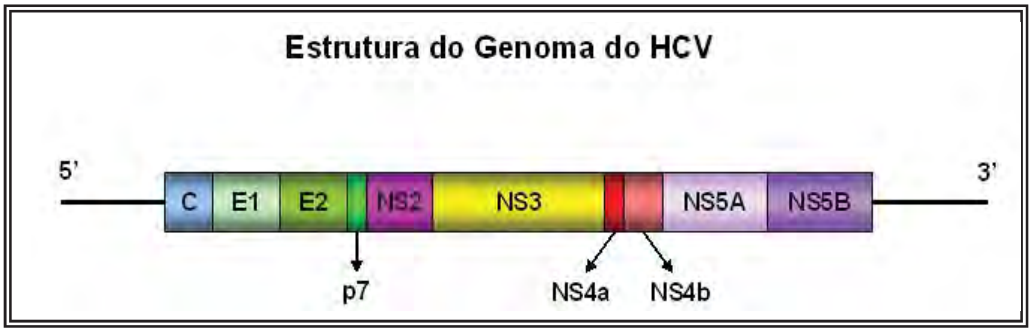


Figura 1: Representação esquemática da estrutura do genoma do vírus da Hepatite C (HCV). Proteínas estruturais: C – Core; E1 – Envelope 1; E2 – Envelope 2; p7; Não Estruturais:; NS2; NS3; NS4A; NS4B; NS5A; NS5B.

A região 5' não-traduzida(5´UTR) é altamente conservada e contém um sítio de entrada ribossomal (IRES) essencial para a tradução do RNA viral(Lyra, Fan *et al.*, 2004). A região 3' UTR é composta por uma pequena região variável, uma cauda poli (U/UC) e uma região altamente conservada de 98 nucleotídeos essenciais para a replicação viral (Moradpour e Blum, 2004).

A proteína Core,a primeira da poliproteína, é uma proteína de ligação a RNA e está envolvida na formação do nucleocapsídeo. Ela é removida da poliproteína por peptidases de sinal do hospedeiro que clivam na região C-terminal e liberam a forma imatura desta proteína (Santolini, Migliaccio *et al.*, 1994). Posteriormente o peptídeo de sinal presente na região C-terminal da proteína Core é processado por peptidases do hospedeiro resultando na proteína madura(Mclauchlan, 2000).

As proteínas de envelope E1 e E2 são glicoproteínas liberadas da poliproteína viral por peptidases de sinal do hospedeiro(Dubuisson, Penin *et al.*, 2002). Essas proteínas são essências para a entrada do vírus na célula e para a montagem das partículas infecciosas (Wakita, Pietschmann *et al.*, 2005; Bartosch e Cosset, 2006; Cocquerel, Voisset *et al.*, 2006).Por estarem envolvidas na entrada do vírus na célula constituem importante alvo no desenvolvimento de moléculas antivirais que bloqueiem a entrada do HCV (Helle, Wychowski *et al.*, 2006).

O polipeptídeop7 é localizado na poliproteína entre as proteínas estruturais e não estruturais, sendo liberada da poliproteína por peptidases do hospedeiro(Lin, Lindenbach *et al.*, 1994; Mizushima, Hijikata *et al.*, 1994; Dubuisson, Penin *et al.*, 2002). É uma pequena proteína de membrana (7KDa) composta por dois domínios transmembrânicos tendo as extremidades amino- e carboxi- terminal orientada em direção ao lúmen do retículo endoplasmático, contudo sua localização celular ainda não esta clara(Carrere-Kremer, Montpellier*et al.*, 2004). Foi demonstrado que este polipeptídeo pode atuar como canal de íons e que sua presença é essencial para a infectividade em chimpanzés(Griffin, Beales *et al.*, 2003; Pavlovic, Neville *et al.*, 2003; Sakai, Claire *et al.*, 2003; Griffin, Harvey *et al.*, 2004).

A proteína NS2 é uma proteína de membrana não essencial a formação do complexo replicativo viral (Lohmann, Korner*et al.*, 1999; Blight, Kolykhalov *et al.*, 2000). A função da NS2 em sua forma madura é desconhecida, contudo no contexto da poliproteína, ela participa da clivagem entre NS2-NS3 (Dubuisson, 2007). Estudos sugerem que esteja relacionada a expressão e liberação da partícula viral (Lin, Lindenbach *et al.*, 1994; Kalinina, Norder *et al.*, 2002; Pietschmann, Kaul *et al.*, 2006).Embora a proteína NS2 não seja comumente relacionada à hepatocarcinogênese,

ela foi reportada como anti-apoptótica (via CIDE-B) e tendo um possível envolvimento na inibição da transcrição de genes celulares

A NS3 é uma proteína multifuncional, que possui um domínio aminoterminalserina-protease e um domínio carboxiterminalhelicase. A NS3 forma um complexo não covalente com a NS4A que é um polipeptídeo ancorado na membrana. A NS4A age como um co-fator da NS3 e também a estabiliza(Kim, Morgenstern *et al.*, 1996; Love, Parge *et al.*, 1996). A proteína NS3 interage com diversas proteínas virais e pode estar envolvida na hepatocarcinogênese(Love, Parge *et al.*, 1996; Deng, Nagano-Fujii *et al.*, 2006).

A proteína NS4B, não está ainda muito bem caracterizada, porém estudos sugerem que esta proteína provoca alterações específicas na membrana produzindo um ambiente favorável à replicação viral(Egger, Wolk *et al.*, 2002).

A proteína NS5A é uma fosfoproteína que tem atraído atenção considerável por seu potencial papel de modulação da resposta ao interferon (IFN) e em outros processos da célula do hospedeiro(Le Guillou-Guillemette, Vallet *et al.*, 2007).

A NS5B é uma proteína enzimática com função de RNA polimerase dependente de RNA, sendo, portanto responsável pela replicação viral(Bartenschlager, Frese *et al.*, 2004).

A figura 2 mostra uma representação esquemática da organização do genoma viral (topo) e a localização das proteínas na membrana do reticulo endoplasmático(Dubuisson, 2007).

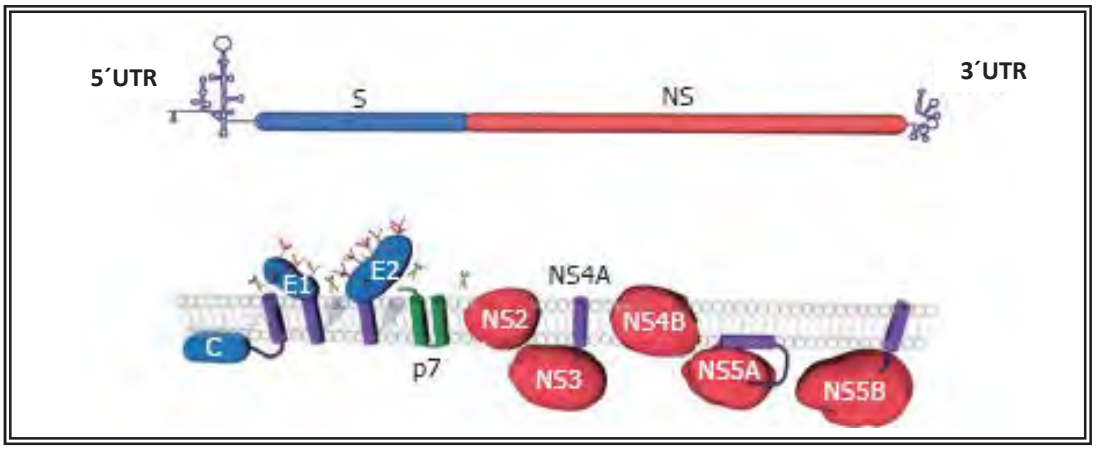


Figura 2: Representação esquemática da organização do genoma viral (topo) e a localização das proteínas na membrana do reticulo endoplasmático. (Extraída de Dubuisson, 2007)

1.5 Variabilidade Genética

Os vírus de RNA são organismos que apresentam altas taxas de mutação. Isso ocorre devido aos erros acumulados durante sua replicação, em decorrência da falta de atividade corretiva da RNA polimerase dependente de RNA viral(Domingo, Escarmis *et al.*, 1996).

A comparação de sequências de nucleotídeos de HCV de variantes coletadas de diferentes indivíduos e regiões geográficas, revelou a existência de pelo menos 7 grandes genótipos. Numerados de 1 a 7,muitos contém variantes mais relacionadas, os subtipos, classificados alfabeticamente (1a, 1b, 3a, 3e, etc)(Simmonds, Bukh*et al.*, 2005; Murphy, Chamberland *et al.*, 2007). Em média, os genótipos diferem, no genoma completo, de 30% a 35% com relação aos nucleotídeos enquanto os subtipos diferem de 20% a 25% (Simmonds, Holmes *et al.*, 1993; Murphy, Chamberland *et al.*, 2007).

Os diferentes genótipos influem na evolução da doença e seu tratamento. Estudos mostram relação entre paciente infectados com genótipo 3 e o desenvolvimento de esteatose, tendo sido identificada uma sequência de aminoácidos específica na proteína Core deste genótipo relacionada a esteatose(Mihm, Fayyazi*et al.*, 1997; Rubbia-Brandt, Quadri *et al.*, 2000; Rubbia-Brandt, Leandro *et al.*, 2001; Sharma, Balan *et al.*, 2004; Hissar, Goyal *et al.*, 2006; Hourioux, Patient *et al.*, 2007; Qiang, Yang *et al.*, 2009). A resposta ao tratamento também esta relacionada ao genótipo sendo o genótipo 1 mais resistente ao tratamento com IFN que os demais genótipos (Kanai, Kako *et al.*, 1992; Yoshioka, Kakumu *et al.*, 1992; Chayama, Tsubota *et al.*, 1993).

Os genótipos 1, 2 e 3 e seus subtipos têm distribuição mundial(Kershenobich, Razavi *et al.*, ; Bukh, Purcell *et al.*, 1993; Simmonds, Holmes *et al.*, 1993; Smith, Pathirana *et al.*, 1997; Cornberg, Razavi *et al.*, 2011; Sievert, Altraif *et al.*, 2011). Já os genótipos 4 e 5 estão presentes na África e o genótipo 6, principalmente na Ásia (Sievert, Altraif*et al.*, 2011).

No Brasil, um estudo, que analisou 1688 amostras de todo o país, mostrou a presença dos genótipos 1, 2, 3, 4 e 5. As freqüências gerais encontradas foram de 64,9% para o genótipo 1, 4,6% para o genótipo 2, 30,2% para o genótipo 3, 0,2% para o genótipo 4 e 0,1% para o genótipo 5 (Campiotto, Pinho *et al.*, 2005). Em outro estudo realizado no interior do estado de São Paulo, para um total de 1018 amostras foram encontrados em 62,9% HCV genótipo 1, 2,1% genótipo 2, 34,5% genótipo 3 e 0,2% genótipo 5(Corvino, Henriques *et al.*, 2006).

1.5.1 Quasiespécies

O HCV, assim como outros vírus de RNA, devido às mutações aleatórias decorrentes da falta de atividade corretiva da RNA polimerase viral, a proteína NS5B,

circula no hospedeiro como um conjunto de variantes virais geneticamente semelhantes chamadas de quasispecies (Martell, Esteban *et al.*, 1992). Essa característica possibilita a existência de um conjunto de variantes possivelmente benéficas ao vírus quando consideradas as populações. Essas variantes podem contribuir para a rápida evolução viral, sendo as populações de vírus capazes de mudar em um curto período de tempo. As mutações aleatórias decorrentes de erros durante a replicação, ocorrem ao acaso maspodem conferir um melhor *fitness* ao organismo, permitindo que o gene mutado persista e eventualmente seja predominante na população onde ele confere alguma vantagem adaptativa (Simmonds, 2004).A característica do HCV de circular como quasiespécies confere ao vírus um maior potencial de se adaptara cada novo hospedeiro e aos desafios da infecção(Eigen e Biebricher, 1988; Eigen, 1993; Coffin, 1995; Domingo e Holland, 1997).

Os impactos biológicos da variabilidade incluem o aparecimento de mutações de escape, alteração do tropismo celular e mudanças na virulência e na abrangência de hospedeiros. O acúmulo de mutações também é importante no desenvolvimento de resistência a agentes anti-virais(Forns e Bukh, 1999).

Contudo a seleção “Darwiniana” ou seleção natural não é o único processo que guia a evolução do vírus da Hepatite C. Na verdade, as trocas de nucleotídeos neutras são responsáveis pela variabilidade genéticanão só do HCV, mas dos genomas em geral (Kimura, 1968). Trocas de nucleotídeos que possuem pouco ou nenhum efeito no *fitness* do organismo podem ou não ser fixadas por acaso e a freqüência de fixação de trocas neutras em um gene ao longo do tempo é relativamente constante(Kimura, 1983). Evidências de ambas, evolução neutra e Darwiniana podem ser encontradas no genoma do HCV. Regiões como E2 apresentam rápida evolução em sua região hipervariavel para evitar o reconhecimento por anticorpos (seleção natural), enquanto a variação

apresentada entre os genótipos, da ordem de 30%, não confere alteração de fenótipo viral (evolução neutra) (Simmonds, 2004). Sendo assimo genoma evolui em diferentes taxas variando de acordo com a região e a pressão a que ela esta submetida. Regiões como a 5'UTR e core são altamente conservadas, as proteínas NS2, NS3, NS5B e 3'UTR são relativamente variáveis, enquanto as proteínas de envelope E1 e E2, NS4 e NS5A exibem a maior diversidade de sequência(Simmonds, 2004; Le Guillou-Guillemette, Vallet*et al.*, 2007).

1.6Tratamento

O primeiro tratamento para a Hepatite C surgiu antes mesmo de sua descoberta. Em 1986, Hoofnagaleet al, reportou que IFN-α era efetivo no tratamento da Hepatite não-A, não-B(Hoofnagle, Mullen *et al.*, 1986). A efetividade doIFN-α, no entanto era limitada correspondendo de 20% a 30% dos casos. Outros estudos levaram à utilização do IFN-α juntamente com a ribavirina que produzia resposta virológica sustentada (RVS) em aproximadamente 40% dos casos, variando de 30% para genótipo 1 a 60% para os genótipos 2 e 3 (Kato, 2001). Mais recentemente a utilização de uma forma de interferon-alfa modificado por polietilenoglicol (PEG-IFN) em combinação com ribavirinaaumentou em 40% a resposta virológica sustentada (Manns, Mchutchison *et al.*, 2001; Fried, Shiffman *et al.*, 2002).

A decisão de tratar os pacientes com Hepatite C crônica depende de diversos parâmetros, incluindo uma avaliação da severidade da doença e de seu prognóstico, a presença de contra-indicações à terapia e à disponibilidade do paciente em ser tratado. A genotipagem viral deve ser realizada antes do inicio por ser determinante da duração do tratamento e do procedimento de monitoramento virológico(Chevaliez e Pawlotsky, 2007).

No Brasil, atualmente o tratamento adotado para a Hepatite C é a combinação de interferon-alfa (IFN-α) com ribavirinapara genótipos 2 e 3 e uma forma de interferon-alfa modificado por polietilenoglicol (PEG-IFN) juntamente com ribavirina para genótipo 1(Brasil, 2011).

O ano de 2011 inaugurou um novo capítulo na história do tratamento da Hepatite C.A FDA (U.S. FoodandDrugAdministration), agência que regulamenta o uso de novas drogas, aprovou o uso de duas drogas antivirais de ação direta (DAA – *direct-acting antiviral dru*gs). As primeiras DAAs aprovadas são os inibidores de protease Telaprevir (TVR) e Boceprovir (BOC)(Poordad, 2010; Vachon e Dieterich, 2011).As estratégias para o tratamento com esses novos medicamentos incluem a combinação com PEG-IFN e Ribavirina(Poordad, 2010).

1.7 NS5A

A proteína NS5A é a proteína mais relacionada à resistência ao IFN. É uma proteína pleoitrópica envolvida na replicação viral e em muitas interações com vias de sinalização celular (Le Guillou-Guillemette, Vallet *et al.*, 2007). Esta proteína parece ter um importante papel, apesar de incerto, como modulador direto da atividade da RNA polimerase dependente de RNA (Shirota, Luo *et al.*, 2002) e interfere nas vias celulares induzidas por IFN (Polyak, 2003).

Em sua região N-terminal, os primeiros 27 aminoácidos constituem o sinal de retenção citoplasmático (CRS – *Cytoplasmicretentionsignal*) que é capaz de manter a proteína NS5A no citoplasma (Reyes, 2002). A NS5A completa parece estar localizada exclusivamente no citoplasma, contudo esta proteína possui um sinal de localização nuclear (NLS – *Nuclear localizationsignal*) e formas com a região N-terminal deletadas

foram encontradas no núcleo sugerindo que tais formas atuariam como potentes ativadores transcricionais(Pawlotsky e Germanidis, 1999).

A resistência ao IFN seria explicada pela habilidade da NS5A ligar-se a quinase indutora de interferon ativada por RNA dupla fita (PKR – *Proteinkinase RNA activated*), levando a sua inibição (Gale, Korth *et al.*, 1997). A região responsável pela resistência ao IFN, uma seqüência conservada de 40 aminoácidos dentro da PKR, localiza-se no centro da NS5A, e foi denominada região determinante de resistência ao IFN (ISDR – *Interferonsensitivitydeterminingregion*). (Figura 3) Em 1995 Enomotoet al descreveram uma correlação entre um maior número de mutações presentes nesta região e a resposta virológica sustentada, em pacientes infectados com genótipo 1 (Enomoto, Sakuma *et al.*, 1995). Outros estudos de grupos japoneses encontraram resultados semelhantes também para genótipo 1b (Enomoto, Sakuma *et al.*, 1996; Chayama, Tsubota *et al.*, 1997; Kurosaki, Enomoto *et al.*, 1997; Arase, Ikeda *et al.*, 1999; Nakano, Fukuda *et al.*, 1999). Contudo estudos europeus e americanos não corroboraram estes resultados (Zeuzem, Lee *et al.*, 1997; Duverlie, Khorsi *et al.*, 1998; Chung, Monto *et al.*, 1999; Squadrito, Orlando *et al.*, 1999; Berg, Mas Marques *et al.*, 2000; Nousbaum, Polyak *et al.*, 2000; Sarrazin, Berg *et al.*, 2000; Murphy, Rosen *et al.*, 2002). Em 2004 um trabalho de Pascu e colaboradores (2004) estudaram seqüências de ISDR de genótipo 1b de pacientes japoneses e europeus, onde foi confirmada a relevância da ISDR para predizer a resposta ao tratamento. Este estudo também encontrou maior probabilidade das mutações na ISDR levarem a RVS em pacientes japoneses quando comparados aos europeus, sugerindo uma diferença geográfica (Pascu, Martus *et al.*, 2004).

Ainda, alguns estudos também com genótipo 1 mostram evidências de efeitos da NS5A na atividade do IFN, independentes da PKR. Um exemplo é a regulação positiva,

pela NS5A, da interleucina8 (IL8), que tem a propriedade de atenuar as propriedades antivirais do IFN, sugerindo que o produção endógena de IL8 possa facilitar a infecção viral (Polyak, Khabar*et al.*, 2001; Girard, Shalhoub *et al.*, 2002).

Na região C-terminal, o domínio V3 com 27 aminoácidos, tem alta taxa de variabilidade. Duverlieet al observaram relação entre o nível de mutações em V3 e a resistência à terapia de interferon. Cepas virais resistentes são mais conservadas em comparação às sensíveis(Duverlie, Khorsi *et al.*, 1998).

As implicações da variabilidade na resistência aos agentes antivirais têm sido discutidas, contudo poucos estudos focam a proteína NS5A inteira com relação às quasiespécies (Le Guillou-Guillemette, Vallet *et al.*, 2007).

OBJETIVOS

2. OBJETIVOS

2.1 Geral

Estudar a evolução de quasiespécies da região NS5A de HCV genótipo 3 em amostras de pacientes Não-respondedores ao Tratamento e Respondedores ao Final do Tratamento com IFN e Ribavirina, durante e após o tratamento.

2.2 Específicos

1. Clonar e Sequenciar a região codificante da proteína NS5A do Vírus da Hepatite C genótipo 3 de amostras coletadas durante e após o tratamento de pacientes Não-respondedores ao Tratamento e Respondedores ao Final do Tratamento.

2. Analisar a evolução do perfil de quasiespécies durante e após o tratamento.

3. Comparar o perfil de quasiespécies encontrado nos pontos analisados com o perfil observado nas amostras pré-tratamento.

PARTE II

Artigo Científico

Title: On Hepatitis C virus evolution: the interaction between virus and host towards treatment outcome.

Authors: Cíntia Bittar[1]; Ana Carolina Gomes Jardim[1]; Lilian Hiromi Tomonari Yamasaki[1]; Claudia Márcia Aparecida Carareto[1]; João Renato Rebello Pinho[2,3]; Philippe Lemey[4]; Isabel Maria Vicente Guedes de Carvalho-Mello*[5]; Paula Rahal[1]

Address of the institution at which the work was performed:
UNESP – São Paulo State University – IBILCE – Institute of Bioscience, Language & Literature and Exact Science – Department of Biology
Rua Cristóvão Colombo, 2265
Bairro Jardim Nazareth - CEP 15054-010
São José do Rio Preto – São Paulo - Brasil

Author's Affiliation:

1- UNESP – São Paulo State University – IBILCE – Institute of Bioscience, Language & Literature and Exact Science – Department of Biology
Rua Cristóvão Colombo, 2265 - Bairro Jardim Nazareth - CEP 15054-010 - São José do Rio Preto – São Paulo – Brasil

Cíntia Bittar – cibittar@gmail.com

Ana Carolina Gomes Jardim - caroljardim@gmail.com

Lilian Hiromi Tomonari Yamasaki - lilian.hty@gmail.com

Claudia Márcia Aparecida Carareto - carareto@ibilce.unesp.br

Paula Rahal - rahalp@yahoo.com.br

2- USP – São Paulo University – Faculty of Medicine – Department of Gatroenterology – Laboratory of Hepatology and Gastroenterology from Institute of Tropical Medicine
Av. Dr. Arnaldo, 455 - Cerqueira César - CEP: 01246903 - São Paulo - SP – Brasil

João Renato Rebello Pinho - jrrpinho@usp.br

3- Albert Einstein Israeli Hospital - Department of Clinical Pathology
Av. Albert Einstein, 627/701 - São Paulo - SP - Brazil - CEP 05652-000

João Renato Rebello Pinho - jrrpinho@usp.br

4- Katholieke Universiteit Leuven - Laboratory of Clinical and Epidemiological Virology (Rega Institute)
Kapucijnenvoer 33 blok i - box 7001 - 3000 Leuven - Belgium

Philippe Lemey - philippe.lemey@rega.kuleuven.be

5- Butantan Institute – Viral Immunology Laboratory
Av Vital Brasil nº 1500, CEP 05503-900**,** Butantã, São Paulo –SP – Brazil

Isabel Maria Vicente Guedes de Carvalho-Mello - imvgcmello@butantan.gov.br

***Correspondent author information:**

Instituto Butantan

Av Vital Brasil nº 1500, CEP 05503-900**,** Butantã, São Paulo –SP – Brazil

Laboratório de Imunologia Viral

Fone +55 11 37267222 ramal 2105

FAX: +55 11 37261505

e-mail: imvgcmello@butantan.gov.br

On Hepatitis C virus evolution: the interaction between virus and host towards treatment outcome

Abstract

Background: Hepatitis C is a disease spread throughout the world. Hepatitis C virus (HCV), the etiological agent of this disease, is a single-stranded positive RNA virus. Its genome encodes a single precursor protein that yields ten proteins after processing. NS5A, one of the non-structural viral proteins, is most associated with interferon-based therapy response, the approved treatment for hepatitis C in Brazil. HCV has a high mutation rate and therefore high variability, which may be important for evading the immune system and response to therapy. The aim of this study was to analyze the evolution of NS5A quasispecies before, during, and after treatment in patients infected with HCV genotype 3a who presented different therapy responses.

Results: Viral RNA was extracted, cDNA was synthesized, the NS5A region was amplified and cloned, and 15 clones from each time-point were sequenced. The sequences were analyzed for evolutionary history, genetic diversity and selection. Our analysis shows that the viral population that persists after treatment for most non-response patients are is present in before-treatment samples, suggesting it is fitted to evasion of treatment. Accordingly, most before-treatment samples from end-of-treatment response patients either did not show the population found after the relapse or showed it in low abundance. The exceptions illustrate the uniqueness of the evolutionary process, and therefore the treatment resistance process, in each patient.

Conclusion: Hepatitis C virus evolution throughout treatment appears to be the result of a unique evolutionary relationship between viral strains and each human host, leading to either persistence or clearance.

Key words: Hepatitis C virus, NS5A, evolution

Introduction

Hepatitis C is a world-wide disease. The World Health Organization estimate is that 3% of the world's population has been infected with the hepatitis C virus (HCV), which is the etiological agent of this disease [1].

HCV is a single-stranded positive RNA virus member of the Flaviviridae family. Its genome is 9.6Kb long and encodes for a single precursor protein with approximately 3,000 amino acid residues. This polypeptide is processed by viral and host proteins, resulting in 10 individual proteins: core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B[2].

NS5A, one of the viral non-structural proteins, is the protein most associated with interferon-based therapy, the accepted treatment for hepatitis C. NS5A is multifunctional, and although not all of its functions are clearly established yet, it is known to be involved in viral replication and interactions with cell signaling pathways [3-5].

Some studies have identified regions in NS5A that have specific functions. At the amino terminal end there is a 27 amino acid cytoplasmic retention signal (CRS) responsible for keeping NS5A in the cytoplasm [6]. Another region, called the PKR-binding region, is responsible for binding to cellular protein kinase R (PKR), inhibiting it, and ultimately resulting in the suppression of interferon (IFN) antiviral activity [7]. Within the PKR-binding region there is an IFN sensitivity-determining region (ISDR), which some studies correlated the accumulation of mutations to therapy response [8, 9]. Despite the CRS, NS5A also contains a nuclear localization signal (NLS) that, in the absence of CRS, results in the translocation of NS5A to the nucleus [6, 10]. In the C-terminal portion of NS5A the genetic variability of a region called V3 is also associated with IFN therapy response [11, 12].

Owing to the lack of proof-reading activity of the viral RNA-dependent RNA polymerase (NS5B protein), HCV has a high mutation rate and therefore high variability. This variation happens at different levels, including genotypes (30% to 35% difference), subtypes (20% to 25% difference) and different but closely related genomes called quasispecies [13-15]. Although mutations can be prejudicial for the virus when they result in non-viable strains, having a pool of slightly different strains with

mutations that are initially neutral or quasi-neutral can increase fitness after changes in the initial condition, and be useful for evading the immune response and treatment [16-19].

The aim of this study was to analyze the evolution of NS5A quasispecies before, during, and after treatment of patients infected with HCV genotype 3a, which presents different therapy responses. This is the first work to analyze the evolutionary history, genetic diversity and selection of full-length NS5A quasispecies from HCV genotype 3a along treatment.

Results

Sequences

This study generated 435 sequences of full length NS5A (1,356 bp) from samples collected during and after interferon plus ribavirin treatment from 8 patients. The sequences were submitted to GenBank; their accession numbers are **JN689511-JN689930**.

Quasispecies Analysis

We analyzed the sequences to address quasispecies variability at each time point in each patient. Figure 1 shows a graphic representation of the results; each tile represents a different amino acid sequence and colored tiles represent a persistent sequence through time points. The viral quasispecies analysis from non-response patient samples shows that amino acid sequences that were circulating before treatment were also found during and after treatment in several patients, including patient P75 at 24 weeks (w) of treatment and in patient P145 at 28 days (d) after treatment (Figure 1). Patients P07, P60 and P75 had sequences that were sampled at more than one time-point after treatment. Patient P07 presented two strains that were sampled at the 28 day and 5 month (m) time-points, patient P60 presented one sequence that was found at both 2 months and 5 months, and patient P75 had the same sequence at the 21 day and 2 month time-points. Patient P145 presented one sequence that was first sampled at 12 weeks during the treatment, then later sampled at 5 months.

In contrast, most end-of-treatment response patients had a different profile. No quasispecies from before treatment were found during or after treatment. However, it was shown that a quasispecies that arises at the relapse time-point, or close to the relapse point, such as in patient P109 (2m), is sampled until the end of the follow-up, with the exception of patient P82. In two patients, P109 and P119, this sequence increased in frequency with increasing time (Figure 1).

Genetic distance

Genetic distance data show that the non-responder patients presented a more homogeneous population of quasispecies during treatment than the end-of-treatment response patients, except for patient P60 which the genetic distance was 0.0081 before treatment, and then rose through follow-up to 0.0205 at 5 months (Figure 1). Quasispecies in the end-of-treatment response patients showed variable genetic distances and different behaviors depending upon the patient. Patient P20, for example, although presenting quasispecies with similar distances in before-treatment (0.0289) and 5 month time-points (0.0296), showed much lower distances at 3 and 4 months (0.0046 and 0.0069, respectively) (for statistics see Table ST-1A to ST-1H, supplementary material). In patient P82, the genetic distance rose from 0.0097 to 0.0207, stabilizing at 3 and 5 months; in contrast, in patient P119 it decreased from 0.0144 to 0.0034. Patient P109 also showed a decrease in genetic distance from the before-treatment (0.0211) to 5 month (0.0047) time-points. However, at 21 days after treatment (relapse point) the distance was higher (0.0261) than before treatment; at 2 months it started falling (0.0170) and remained constant until 4 months, then it showed a drastic decrease at the 5 month time-point (0.0047).

Global and site-by-site selective pressure (ω) estimation

The estimates of the overall selective pressure (ω=dN/dS) of each of the response groups showed that non-responder patients have a similar ω (0.2828) to the end-of-treatment responders (0.2234). This analysis was also carried out for the sequences from the sustained virological response group (**EU826189 to EU826218; EU826249 to EU826263; EU826294 to EU826307**), which presented lower ω values (0.1494) (Figure 2), although the differences were not statistically significant.

When ω was analyzed by patient, the values varied from 0.20 to 0.46, suggesting that NS5A is under a relaxed purifying selection in the quasispecies circulating in the patients. However, by analyzing each time-point separately, variations could be detected. Patient P145 was the only individual who showed evidence of positive selection at 2 months, with ω=1.63, but by 5 months this had decreased to 0.7. Evidence of neutral evolution was found in time-points from two patients, P82 – BT (Before treatment) and P07 – 2m, with ω values of 1.03 and 1.01 respectively (Figure 3).

The estimation of the dN/dS rates for each site showed negatively selected sites in the NS5A sequences of all patients, in agreement with the above results. Figure 4 indicates the locations of these sites in the sequence and their percentages in each patient. However, evidence for positive selection was found in two amino acids, 122 and 408 of the quasispecies from patient P109. Negatively selected sites were identified in all patients. The results are shown in Figure 4. End-of-treatment response patients had more negatively selected sites (n-163; 9.02%) than non-response patients (n-101; 5.59%), and the difference was statistically significant ($p<0.001$). Some sites were negatively selected in more than one sequence.

Stop codons

During the analysis of all the sequences generated in this study and considered together with those obtained by Bittar *et al* (2010), 23 stop codons were detected [20]. The majority of them, 73.9% (17/23), were found in NR sequences, and no stop codon was identified in SVR. Some of these non-sense mutations were found at the same site in different patients and, in one case (P145), through different time-points. The sites where the stop codons occurred are shown in Figure 5. Most of the stop codons (15) were TAG, followed by TGA (5) and TAA (3). The amino acids that were most associated with mutating to a stop codon were glutamine (Gln) at four sites, and tryptophan (Trp) and lysine (Lys) at three sites each; these mutations correspond to 20 of the 23 mutations found. All of the sites where these codons were found were highly conserved. Some stop codons at the same site in different patient sequences were encoded by different codons (Figure 5).

Relative genetic diversity

The relative genetic diversity analyses can be seen in the skyride plots shown in Figure 6. In these plots we can see that the quasispecies diversity in most patients shows an exponential growth before the beginning of the treatment, except for patient P119, who seems to be constant, and patient P60, who shows variation with time. Around the beginning of the treatment (represented by the red line for the approximate BT time-point), all patients show a major oscillation in quasispecies diversity, with the exception of patient P82 samples, which remained almost constant with only a slight variation. Most patients presented a decrease in genetic diversity after the beginning of the treatment (P07, P75, P20, P109 and P119). Patient P145 showed an initial increase in diversity followed by a fluctuation, which is in accordance with the distance data for this patient.

Phylogenetic relationships and correlation with other data

A phylogenetic tree was constructed with the sequences from this study together with before treatment sequences from the previous study [20] (Accession numbers: EU826174 to EU826233 and from EU826249 to EU826352) and the reference sequence for genotype 3a NZL1 (Accession number: D17763) (Figure 7). All sequences from the same patient grouped on a monophyletic branch with significant supporting values (supplementary material - figures SR-1 to SR-9), showing there was no cross-contamination among patients. Sequences from two patients, P109 and P119, were grouped in a monophyletic branch with a high bootstrap (100) (Figure SR-9, supplementary material). These indicate a closer relationship between the viruses from these patients, suggesting an epidemiological link. From the epidemiological questionnaire, which patients filled in when they enrolled in the study, we could not establish any direct link between the patients and their routes of transmission. The only inference that can be made is that their viruses had the same origin.

Sequences from P60 and P82 from before the treatment clustered together, away from the other time-points, in a monophyletic branch with bootstrap of 98 and 86 respectively (Figures SR-3 and SR-5, supplementary material).

From the phylogenetic tree for patient P20 (Figure SR-2, supplementary material), the viral population at 5 months seems to be derived from two separate

before-treatment populations. Interestingly, one of these populations, which was the most prevalent before treatment, is not present in any of the 3 and 4 months time-point samples. This data supports the genetic distance data that show a major decrease in 3 and 4 month time-point sequences when only one population was sampled. At 5 months, the second viral population reappears, which increases the genetic distance. The ω values at 5 months can be explained not by an increase in selective pressure, but is probably a result of an increase in the synonymous changes due to the presence of two different populations. Ultimately patient P20 viral population changed converging towards a previously existing, but not predominant, population and, after evading treatment, both populations were able to reestablish themselves.

The phylogenetic tree also demonstrated some interesting evolutionary behaviors of the viral population of patient P60 (Figure SR-3, supplementary material). After-treatment sequences from patient P60 were grouped on two different branches with significant support. The first group comprises one sequence from the 28 day time-point, one at 2 months, and nine sequences from the 5 month time-point. The second group, which contains all other P60 after-treatment sequences time-points together with the six remaining sequences from the 5 month time point, grouped on a terminal monophyletic branch along with three sequences from the 28 day time-point and eight sequences from 2 months. The first group had a low prevalence in the first two time-points after treatment, having only one quasispecies in each, but became the most prevalent at the last time point. This scenario corroborates the genetic distance results where P60 showed an increase at 5 months, probably due to the existence of quasispecies that evolved from two different lineages.

The phylogenetic tree for patient P07 shows all of the sequences from the 2 month time-point clustering as a monophyletic group, with a bootstrap of 93, indicating a recent common ancestor with a before-treatment strain (Figure SR-1, supplementary material). All samples from the 28 day time-points are clustered in two different branches together with some 5 month strains. The remaining strains from the 5 month time-point are grouped on a third outer monophyletic branch separated from all other after-treatment time-points that share a common ancestor with a before-treatment strain. These data explain why the genetic distance rises in the fifth month (Figure 1).

Phylogenetic data on patient P109 show the sequences clustered on two major branches (Figure SR-6, supplementary material). This indicates that the majority of the after-treatment population is derived from a population that is not sampled in a before-treatment time-point. Conversely, all after-treatment strains from patient P119 are derived from the before-treatment samples.

Discussion

Our study characterizes hepatitis C evolution, based on NS5A protein, throughout treatment in patients infected with the HCV genotype 3a who show different treatment outcomes.

The high mutation rate of RNA viruses, such as HCV, provides a pool of closely related variants, the quasispecies, which provide the virus with many possibilities for evading the immune system and therapy. It is interesting that in two of the non-response patients, the same quasispecies that were sampled before the treatment were also sampled during (P75) or after treatment (P145). This is unlikely to be since these samples were collected more than six months apart, and HCV is estimated to produce 10^{12} copies of its genome per day [21] with a high mutation rate, suggesting that these strains provided some advantage to the virus enabling it to survive the treatment, and were therefore selected.

In contrast, the virus circulating in end-of-treatment response patients was not initially successful in surviving the treatment, so the virus was not detected using the PCR technique. Presumably, in order to reestablish the infection, the virus replicates at low rates and at some point strains present with a better fitness to the new condition, and because they are therefore capable of evading the treatment, are selected, resulting in the relapse. Relapsing indicates that the virus managed to survive the battle against treatment and to evade the immune response, and the relapse point is the closest representation of this scenario. After the best-fitted pool of strains for the new condition has been selected, the different sequences would be expected to converge to the most favorable genetic composition for survival. This is the case for patients P109 and P119. In the distance data we observe that at the time-point where a predominant quasispecies arises, the genetic divergence between the sequences falls, even though it still has a low

frequency. It is also expected that through time the virus population would recover its variability.

The analysis of selective pressure by site showed some sites that are negatively selected in more than one patient, suggesting a functional constraint of the protein. However, as this does not apply for all patients, and one site, site 408, was both positively and negatively selected in two different patients (P109 and P20, respectively), it seems that the evolutionary processes are unique in each patient and are guided by the interaction of virus and host. Viral samples from patients who had an end-of-treatment response have more sites undergoing negative selection than samples found in non-response patients, and this difference is statistically significant. Since this group passed through a recent bottleneck due to a strong selective event (Interferon/Ribavirin administration), it probably conserved the amino acids that proved most effective in evading both treatment and the immune system, thus allowing the infection to be reestablished. It is interesting, however, that the ETR group, despite having more negatively selected sites, presented an overall ω close to that of the NR group. A possible explanation for this is that the remaining sites have more relaxed constraints.

The occurrence of stop codons at the same site in different patients, and in one case persisting through all time-points of the same patient, is quite interesting. One explanation could be that during translation the ribosome jumps or reads through the stop codon. However, this hypothesis does not explain why it occurs at the same site in a considerable number in different patients (12 out of 23). Since NS5A is a multifunctional protein and its functions are not fully understood, the proteins resulting from this RNA that encode stop codons could have a specific function during the infection, different from the one achieved by full-length NS5A. One of the known functions of NS5A is an important role in viral replication, and it has been shown that defective NS5A impairs HCV replication. Though it is also known that NS5A is the only HCV protein that can act in *trans* complementation, meaning it is acting on a viral RNA other than the one from which it has been translated [22, 23]. These characteristics make it possible that the defective genomes could persist by being encapsulated with the help of normal NS5A acting in *trans.* Other studies have shown that defective genomes circulate during HCV infection and also during other *Flavivirus* infections [24-26]. Another point to consider is that having NS5A sequences with stop codons do not seem to be disadvantageous for viral survival. No nonsense mutations were identified in

sequences from SVR patients, while 73.9% of them were found in sequences from non-responders [20]. These defective variants could give the virus some advantage and play an important role in evading the therapy response. Another issue that suggests some kind of viral fitness enhancement by these mutations is that most of the stop codons that were found at the same site in different patients were not encoded by the same codon. For example, site 111 (originally TGG) mutated into three different stop codons in three different patients, two of them changed to TGA and the other to TAG. In the case of site 166 (originally AAG), two clones from the same patient and same time-point presented two different stop codons (TAA and TGA), and in two other patients the nonsense mutation was due to TAG. This site (166) presented all three codons that stop translation. Site 386 had mutated from CAG to TAG in patient P60 at 5 months and from CAA to TAA in patient P119. These sites converged to nonsense mutations in different ways both in the same patient and in different patients, suggesting that stopping translation at these points provides some kind of evolutionary advantage to the virus.

Some insights on evolution

Viruses are microorganisms incapable of replicating, translating or assembling viral particles outside their hosts. In short, they cannot survive outside their host. This is one of the reasons why some do not consider the virus as a form of life. However, since this is a work on the evolution of a genomic region of the hepatitis C virus, and we understand that only living organisms are capable of evolving, we will use terms such as surviving, living and dying in relation to viruses in this section.

Well-adapted viruses are the ones capable of surviving in terms of population for a long time; not only surviving, but also having an efficient way of infecting, replicating and being transmitted from host to host. Viruses co-evolve with their hosts, with the virus trying to evade the host defenses at the same time as the host is trying to clear the infection. Well-adapted viruses do not kill their hosts, at least not in the short term, as by doing so they would ultimately be killing themselves and have less time to spread viral particles to other hosts. The Human Immunodeficiency Virus (HIV), for example, has adapted to infect, survive and replicate in humans, but it also can kill its host very quickly. As HIV is considered a virus of recent origin, thought to have been introduced into humans between 1884 and 1924 [27], it has probably had insufficient time to

establish the most fitted scenario through co-evolution. The hepatitis C virus, on the other hand, can be considered a successful virus in evolutionary terms. It causes a persistent infection and remains silent, presenting no clinical symptoms, for decades. Hepatitis C virus genotypes have been co-evolving with the human host for a long time; they are thought to have diverged 500 to 2000 years ago, leading to a better fitness [28].

The samples used in this study were obtained from chronically infected patients who had been through interferon and Ribavirin treatment and had not responded. These viruses not only evaded the host immune system, but also evaded therapy. What we can see in this study is that, although HCV is adapted to infect *Homo sapiens* in general, each individual specimen represents a different environment, leading to different evolutionary processes. In some cases the viral population from before treatment belongs to a separate monophyletic group from the other time-points (P60 and P82), or most samples from other time-points (P109). This means that persistence is achieved by strains that were not initially the best fitted for survival, and therefore were not sampled before the treatment started, but after a new selective environment – the treatment - was introduced the strain gained fitness. We can also find the opposite, where all after-treatment strains are derived from a unique before-treatment population (P119). In another case (P20) viral population changed by converging to a previously existing, but not predominant, population and, after evading treatment, both populations were reestablished. Finally, the after-treatment strains were all derived from before-treatment samples in some patients (P07, P75 and P145).

From this analysis we can notice that before-treatment samples from most of the non-response patients (except P60) presented a viral population that persisted after treatment, suggesting they were fitted to evade the treatment. Accordingly, most before -treatment samples from end-of-treatment response patients (except P119) either did not show or showed in a low frequency the population that was identified after the relapse. The exceptions illustrate the uniqueness of the evolutionary process, and therefore the treatment resistance process in each patient.

Conclusion

Hepatitis C virus is a highly variable virus, which confers a range of possibilities for it to evolve and adapt to new conditions. The evolution of this virus during and after treatment is linked to the environment to which it is subjected, that is the patient. Each patient has their own immune response and therefore the virus infecting each patient must find its own specific way to evade at the same time as the host is finding a way to clear the infection. The patients enrolled in this study were chronically infected with HCV, which means that the virus had already succeeded in the first step, evading the immune system, and was now facing the second one, the treatment. It seems that the viruses that resisted the treatment (non-responder samples) already had strains that could evade therapy before treatment started. No specific pattern could be found in viral strains that could determine therapy response. Finally, hepatitis C virus evolution throughout treatment appears to be the result of a unique evolutionary relationship between viral strains and each human host, leading either to persistence or to clearance.

Acknowledgements

This work was financially supported by FAPESP.

Materials and methods

Population and samples: Blood samples were collected from 12 HCV genotype 3-positive patients at the Blood Center of São José do Rio Preto, State of São Paulo, Brazil. All samples were collected before treatment, during (12 weeks or 24 weeks) and after treatment (21 days, 28 days, 2 months, 3 months, 4 months and 5 months) to provide treatment response data. Patients were classified into three groups according to treatment response: four patients were sustained virological responders (SVR), i.e. the virus was not detected after treatment or during the 6-months follow-up; four were non-responders (NR), since they showed no virological response; and four were end-of-treatment responders (ETR), i.e. a virological response was detected, but at the 6-months follow-up there was a relapse. Samples from the SVR group and before treatment samples from most patients enrolled in this study were analyzed in a previous

study and were used in this study for comparative analyses of the evolutionary behavior (**EU826189 to EU826218; EU826249 to EU826263; EU826294 to EU826307**) [20]. The exception was patient P82; before-treatment samples from patient P82 were analyzed in this work as they had not been analyzed in the previous study. All patients were infected with HCV genotype 3a. Treatment consisted of INF-α and Ribavirin administration for 24 weeks. For the non-response patients, one time-point sample was taken during treatment and three were taken after treatment. Owing to patient availability, some samples could not be collected. End-of-treatment response time-points were considered the relapse point, and all patients provided samples up to 5 months. Time-points for each patient were obtained as follows: NR (P07 – 28 days, 2 months and 5 months; P60 -28 days, 2 months and 5 months; P75 – 24 weeks, 21 days, 2 months and 5 months; P145 – 12 weeks, 28 days, 2 months and 5 months) ETR (P20 – 3 months, 4 months and 5 months; P82 – before treatment, 2 months, 3 months and 5 months; P109 - 21 days, 2 months, 3 months, 4 months and 5 months; P119 -3 months, 4 months and 5 months) (Supplementary Figure SM-1). Patients with a history of alcoholism or infection with another agent that could cause liver damage were excluded. This study was approved by the Ethics Committee of the Hospital de Base from São José do Rio Preto, and all participants signed an informed consent.

Extraction of HCV RNA and amplification of the NS5A region: Viral RNA was extracted from blood serum samples using a QIAamp Viral RNA Mini Kit (QIAgen), and cDNA was synthesized using a High-Capacity cDNA Archive Kit (Applied Biosystems) according to the manufacturer's instructions. For the PCR and NESTED-PCR reactions, three sets of forward and reverse primers for the NS5A region were designed (Table ST-2, supplementary material). The amplification mix for both PCR and NESTED-PCR reactions contained 1µl (5U) of a proofreading polymerase (Long PCR Enzyme Mix; Fermentas) together with 5µl of 10X Long PCR Buffer with $MgCl_2$, 4µl of DMSO, 1µl of dNTP (10mM), 1µl of each primer (20pmol), 10µl of synthesized cDNA for the PCR reaction and 5µl of PCR product for the NESTED-PCR reaction, plus nuclease-free water provided with the enzyme kit to a final volume of 50µl. The amplified products were analyzed on a 1% agarose gel.

Cloning and sequencing: Cloning was performed using a TOPO XL Cloning TM Kit (Invitrogen). Fragments of 15 clones at each time-point from each patient were purified

using a GeneJET Plasmid Purification Kit (Fermentas). The entire NS5A region was sequenced using eight primers: the vector primers *M13F* and *M13R*, and six inner primers: three sense and three anti-sense (Sense: *H.NS5AI-F1, H.NS5AI-F2* and *H.NS5AI-F3*, antisense: *H.NS5AI-R1*, *H.NS5AI-R2* and *H.NS5AI-R3*) [20]. The sequencing reaction was performed with BigDye Terminator (Applied Biosystems) and the products were sequenced in an ABI 3130XL sequencer (Applied Biosystems). The reaction mixture consisted of 2.2 µl of Milli-Q autoclaved water, 2µl of 5x Sequencing Buffer (Applied Biosystems), 0.8µl of primer (5pmol/µl), 2µl of Big ET Dye Terminator, plus 3µl of sample. Sequencing reactions consisted of a "hot start "step of 10 min at 95°C, followed by 96°C for 1 minute, then 25 cycles of 96°C for 15s, 50°C for 15s and 60°C for 4 min.

Sequence analysis: In order to analyze viral evolution over time, sequences from samples collected before treatment, and previously published (Accession numbers: EU826174 to EU826233 and from EU826249 to EU826352), were also used in the analyses [20]. The sequences obtained in this study were subjected to BioMol - Electropherogram quality analysis (http://adenina.biomol.unb.br/phph/) [29], a phred phrap [30, 31] analysis site, for quality checking and contig construction. The contigs obtained for each clone were aligned, along with the reference sequence NZL1 for genotype 3 (GenBank accession number D17763), using *Clustal W* software nested in the BioEdit 7.0.9.0 package [32, 33]. All sequences were edited on *Bio Edit* [32] to remove the vector fragments, leaving only the complete sequence of the NS5A region.

Evolutionary analysis

Quasispecies were analyzed using *LOCSPEQ 1.0* software [34]. The genetic distances were calculated using MEGA 5.0 software [35]. The overall dN/dS ratio (ω) and ω of each site were calculated with HyPhy [36]. Bayesian skyride plots were performed using the BEAST package [37]. Phylogenetic analysis was performed using *PhyML* [38]. A Maximum-Likelihood tree was constructed using the HKY85 substitution model including a Gamma distribution parameter (HKY+G). Bootstrap was performed with 500 replicates for the tree containing all the sequences and 1000 replicates for individual trees from each patient. Values above 70% were considered significant.

Statistics: Statistical analyses for the distance and ω data were performed by Mann-Whitney tests, and a Chi-square test was used for negatively selected sites. p-values lower than 0.005 were considered significant.

Author's Contributions

CB, ACGJ, IMVGCM , PR: contributed to the study design, carried out the molecular biology experiments, sequence alignment, evolutionary analysis and manuscript design.

LHTY: carried out the molecular biology experiments and statistical analysis.

CMAC: contributed with significant evolutionary knowledge on data analysis.

PL: carried out evolutionary analysis.

JRRP: contributed to the study design.

References

1. WHO: **World Health Organization** 2000.
2. Kato N: **Molecular virology of hepatitis C virus**. *Acta Med Okayama* 2001, **55**(3):133-159.
3. Le Guillou-Guillemette H, Vallet S, Gaudy-Graffin C, Payan C, Pivert A, Goudeau A, Lunel-Fabiani F: **Genetic diversity of the hepatitis C virus: impact and issues in the antiviral therapy**. *World J Gastroenterol* 2007, **13**(17):2416-2426.
4. Shirota Y, Luo H, Qin W, Kaneko S, Yamashita T, Kobayashi K, Murakami S: **Hepatitis C virus (HCV) NS5A binds RNA-dependent RNA polymerase (RdRP) NS5B and modulates RNA-dependent RNA polymerase activity**. *J Biol Chem* 2002, **277**(13):11149-11155.
5. Polyak SJ: **Hepatitis C virus--cell interactions and their role in pathogenesis**. *Clin Liver Dis* 2003, **7**(1):67-88.
6. Satoh S, Hirota M, Noguchi T, Hijikata M, Handa H, Shimotohno K: **Cleavage of hepatitis C virus nonstructural protein 5A by a caspase-like protease(s) in mammalian cells**. *Virology* 2000, **270**(2):476-487.
7. Gale MJ, Jr., Korth MJ, Tang NM, Tan SL, Hopkins DA, Dever TE, Polyak SJ, Gretch DR, Katze MG: **Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein**. *Virology* 1997, **230**(2):217-227.
8. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, Izumi N, Marumo F, Sato C: **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* 1995, **96**(1):224-230.
9. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, Ogura Y, Izumi N, Marumo F, Sato C: **Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection**. *N Engl J Med* 1996, **334**(2):77-81.
10. Ide Y, Zhang L, Chen M, Inchauspe G, Bahl C, Sasaguri Y, Padmanabhan R: **Characterization of the nuclear localization signal and subcellular distribution of hepatitis C virus nonstructural protein NS5A**. *Gene* 1996, **182**(1-2):203-211.
11. Inchauspe G, Zebedee S, Lee DH, Sugitani M, Nasoff M, Prince AM: **Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates**. *Proc Natl Acad Sci U S A* 1991, **88**(22):10292-10296.
12. Duverlie G, Khorsi H, Castelain S, Jaillon O, Izopet J, Lunel F, Eb F, Penin F, Wychowski C: **Sequence analysis of the NS5A protein of European hepatitis C virus 1b isolates and relation to interferon sensitivity**. *J Gen Virol* 1998, **79 (Pt 6)**:1373-1381.
13. Martell M, Esteban JI, Quer J, Genesca J, Weiner A, Esteban R, Guardia J, Gomez J: **Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution**. *J Virol* 1992, **66**(5):3225-3229.
14. Simmonds P, Holmes EC, Cha TA, Chan SW, McOmish F, Irvine B, Beall E, Yap PL, Kolberg J, Urdea MS: **Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region**. *J Gen Virol* 1993, **74 (Pt 11)**:2391-2399.
15. Bowen DG, Walker CM: **The origin of quasispecies: cause or consequence of chronic hepatitis C viral infection?** *J Hepatol* 2005, **42**(3):408-417.
16. Coffin JM: **HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy**. *Science* 1995, **267**(5197):483-489.
17. Domingo E, Holland JJ: **RNA virus mutations and fitness for survival**. *Annu Rev Microbiol* 1997, **51**:151-178.

18. Eigen M: **Viral quasispecies**. *Sci Am* 1993, **269**(1):42-49.
19. Eigen M, Biebricher C: **Sequence space and quasispecies distribution**, vol. 3. Boca Raton, FL: CRC Press; 1988.
20. Bittar C, Jardim AC, Yamasaki LH, de Queiroz AT, Carareto CM, Pinho JR, de Carvalho-Mello IM, Rahal P: **Genetic diversity of NS5A protein from hepatitis C virus genotype 3a and its relationship to therapy response**. *BMC Infect Dis* 2010, **10**:36.
21. Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, Layden TJ, Perelson AS: **Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy**. *Science* 1998, **282**(5386):103-107.
22. Appel N, Herian U, Bartenschlager R: **Efficient rescue of hepatitis C virus RNA replication by trans-complementation with nonstructural protein 5A**. *J Virol* 2005, **79**(2):896-909.
23. Tong X, Malcolm BA: **Trans-complementation of HCV replication by non-structural protein 5A**. *Virus Res* 2006, **115**(2):122-130.
24. Aaskov J, Buzacott K, Thu HM, Lowry K, Holmes EC: **Long-term transmission of defective RNA viruses in humans and Aedes mosquitoes**. *Science* 2006, **311**(5758):236-238.
25. Craig S, Thu HM, Lowry K, Wang XF, Holmes EC, Aaskov J: **Diverse dengue type 2 virus populations contain recombinant and both parental viruses in a single mosquito host**. *J Virol* 2003, **77**(7):4463-4467.
26. Jardim AC, Yamasaki LH, de Queiroz AT, Bittar C, Pinho JR, Carareto CM, Rahal P, Mello IM: **Quasispecies of hepatitis C virus genotype 1 and treatment outcome with Peginterferon and Ribavirin**. *Infect Genet Evol* 2009, **9**(4):689-698.
27. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, Muyembe JJ, Kabongo JM, Kalengayi RM, Van Marck E *et al*: **Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960**. *Nature* 2008, **455**(7213):661-664.
28. Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncales FL, Jr., Haussinger D, Diago M, Carosi G, Dhumeaux D *et al*: **Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection**. *N Engl J Med* 2002, **347**(13):975-982.
29. **BioMol - Electropherogram quality analysis** [http://www.biomol.unb.br/phph/]
30. Ewing B, Green P: **Base-calling of automated sequencer traces using phred. II. Error probabilities**. *Genome Res* 1998, **8**(3):186-194.
31. Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using phred. I. Accuracy assessment**. *Genome Res* 1998, **8**(3):175-185.
32. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symposium Series* 1999, **41**:95-98.
33. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools**. *Nucleic Acids Res* 1997, **25**(24):4876-4882.
34. Marucci EA, Zafalon GF, Jardim AC, Yamasaki LH, Bittar C, Rahal P, Machado JM: **Routine libraries for pattern recognition in quasispecies**. *Genet Mol Res* 2008, **7**(3):970-981.
35. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: **MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods**. *Mol Biol Evol*.
36. Pond SL, Frost SD, Muse SV: **HyPhy: hypothesis testing using phylogenies**. *Bioinformatics* 2005, **21**(5):676-679.
37. Drummond AJ, Rambaut A: **BEAST: Bayesian evolutionary analysis by sampling trees**. *BMC Evol Biol* 2007, **7**:214.
38. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood**. *Syst Biol* 2003, **52**(5):696-704.

Figures

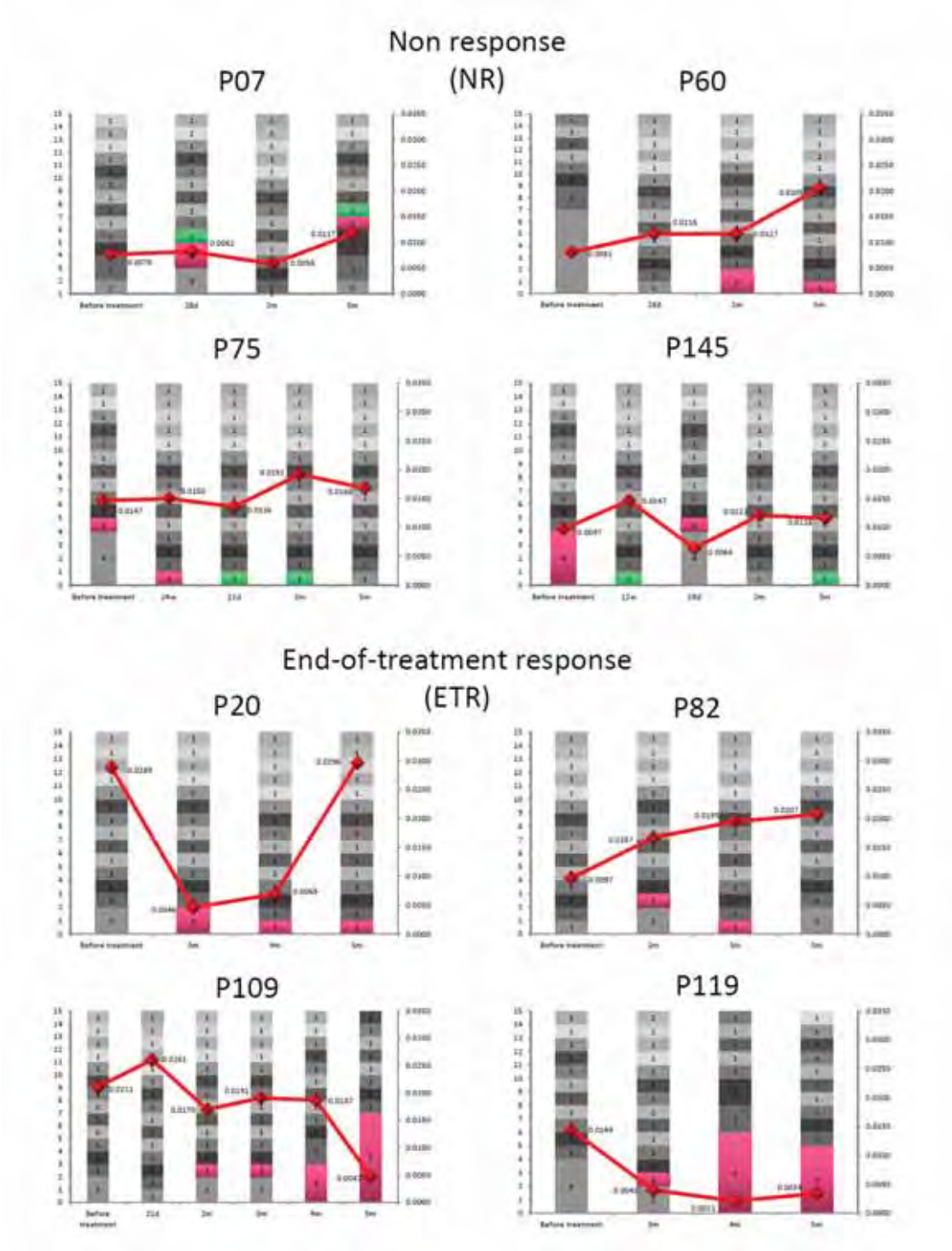


Figure 1: Schematic representation of distance versus quasispecies**.** Left Y axis - graphic representation of different amino acid quasispecies. Tiles in shades of gray represent sequences that were sampled in only one time-point. Colored tiles represent the same quasispecies in the same or different time-points. Right Y axis - genetic distance data. W-weeks; d- days; m-months

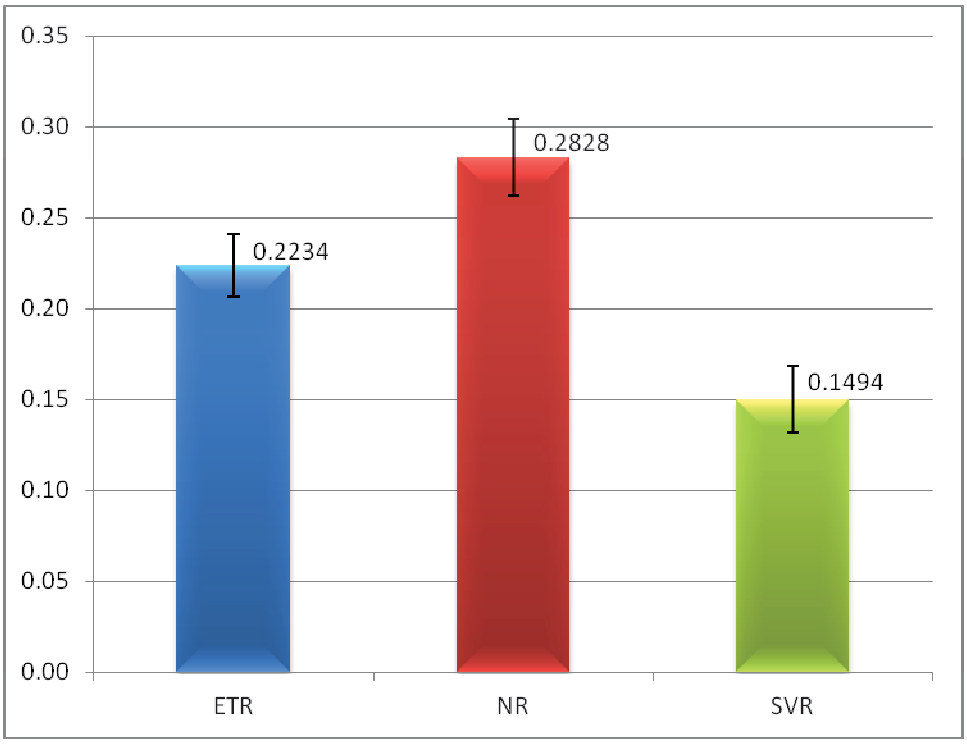


Figure 2: Overall ω of each response group. ETR: end-of-treatment response; NR: non-response; SVR: sustained virological response

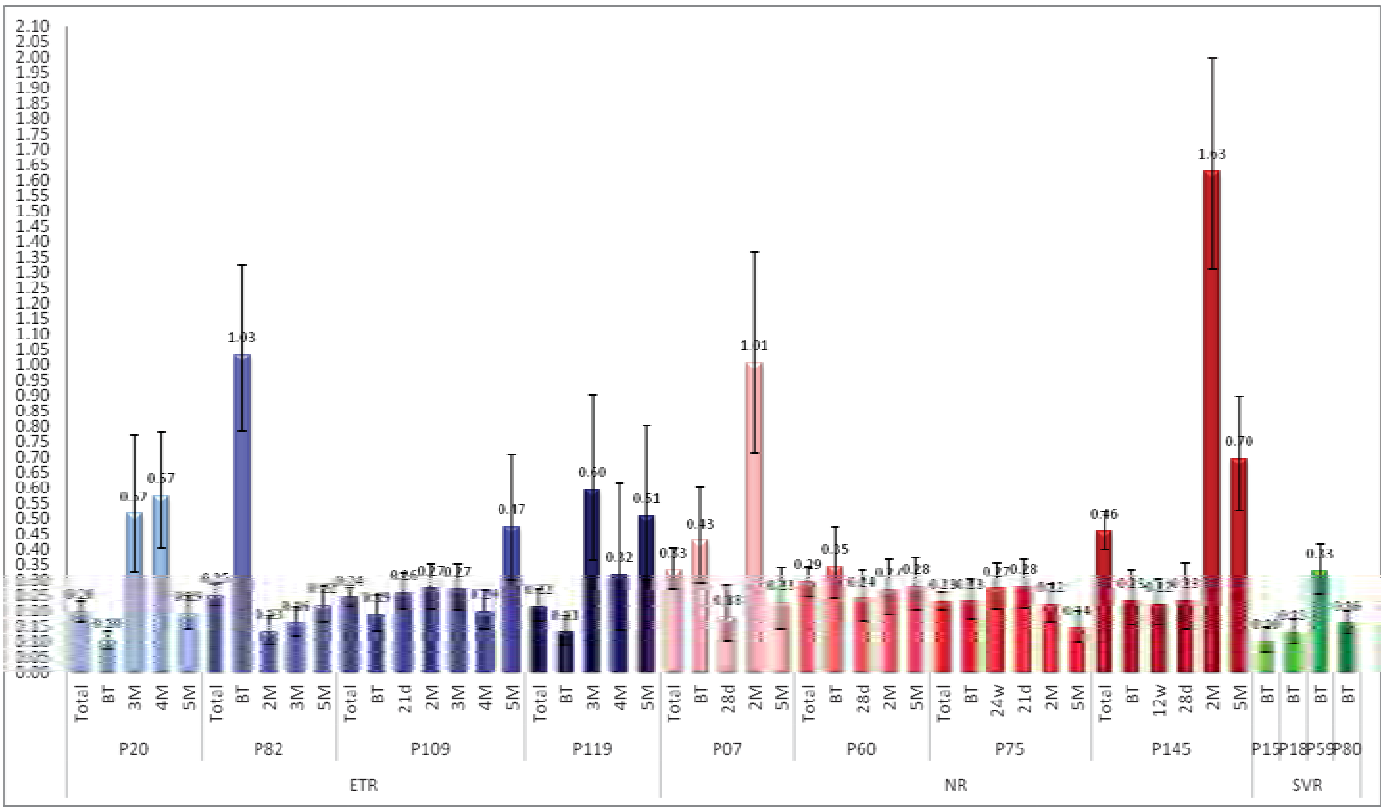
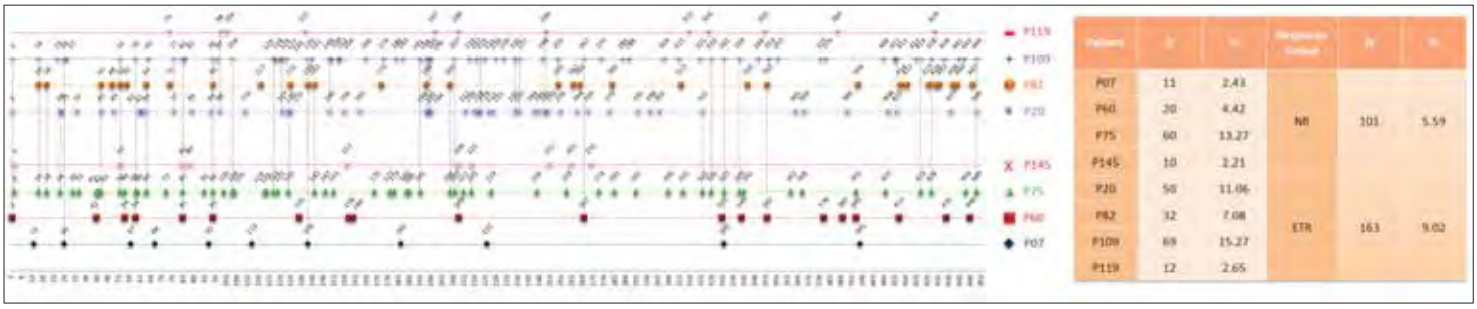
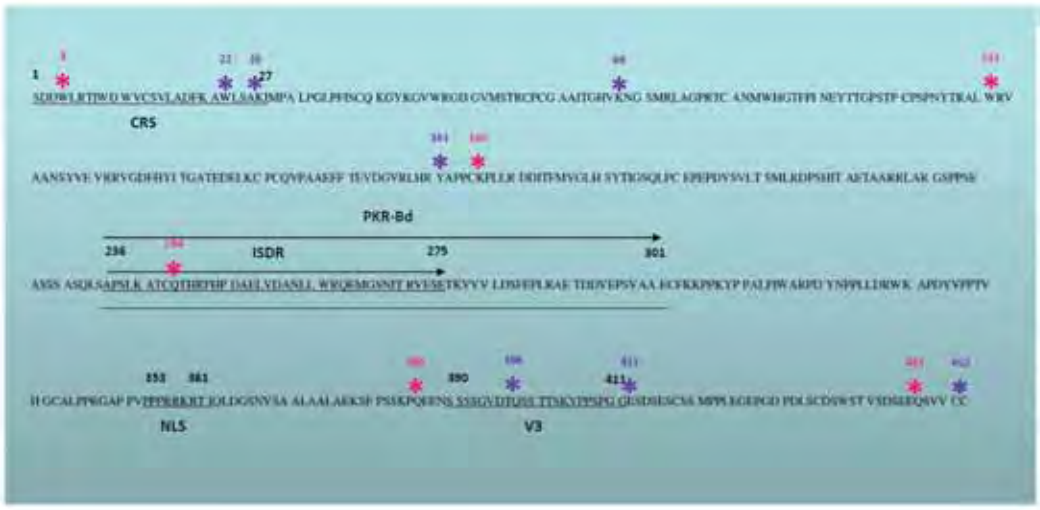


Figure 3: ω rates for each patient (Total) and for each time point. BT-Before treatment; w-weeks; d-days; m-months.



Patient	N	%	Response Cohort	N	%
P07	11	2.43			
P60	20	4.42	NR	101	5.59
P75	60	13.27			
P145	10	2.21			
P20	50	11.06			
P82	32	7.08	ETR	163	9.02
P109	69	15.27			
P119	12	2.65			

Figure 4: Sites under negative selection for each patient. Measured by ω rates considering p<0.1. Vertical lines link sites negatively selected in more than one patient.



4	P145	Before treatment	TGG/TGA
		12 weeks	TGG/TGA
22	P145	5 months	TGG/TAG
26	P145	2 months	AAG/TAG
98	P145	2 months	AAG/TAG
	P07	2 months	TGG/TGA
111	P60	5 months	TGG/TGA
	P109	21 days	TGG/TAG
181	P82	5 months	TAC/TAA
	P60	Before treatment (2 clones)	AAG/TAA AAG/TGA
186	P75	5 months	AAG/TAG
	P145	Before treatment	AAG/TAG
244	P82	2 months	CAG/TAG
309	P60	5 months	CAG/TAG
	P110	3 months	CAA/TAA
366	P75	24 weeks	CAG/TAG
412	P145	2 months	GAG/TAG
	P07	Before treatment	CAG/TAG
447	P75	2 months	CAG/TAG
	P109	Before treatment	CAG/TAG

Figure 5: Graphic representation of sites where stop codons were identified. Pink – more than one occurrence; purple-one occurrence. NZL1 (GenBank accession number D17763) – reference sequence for genotype 3 was used for the graphic representation. CRS – cytoplasmic retention site; PKR-Bd – PKR binding region; ISDR – IFN sensitivity determining region; NLS – nuclear localization signal; V3 – variable region 3.

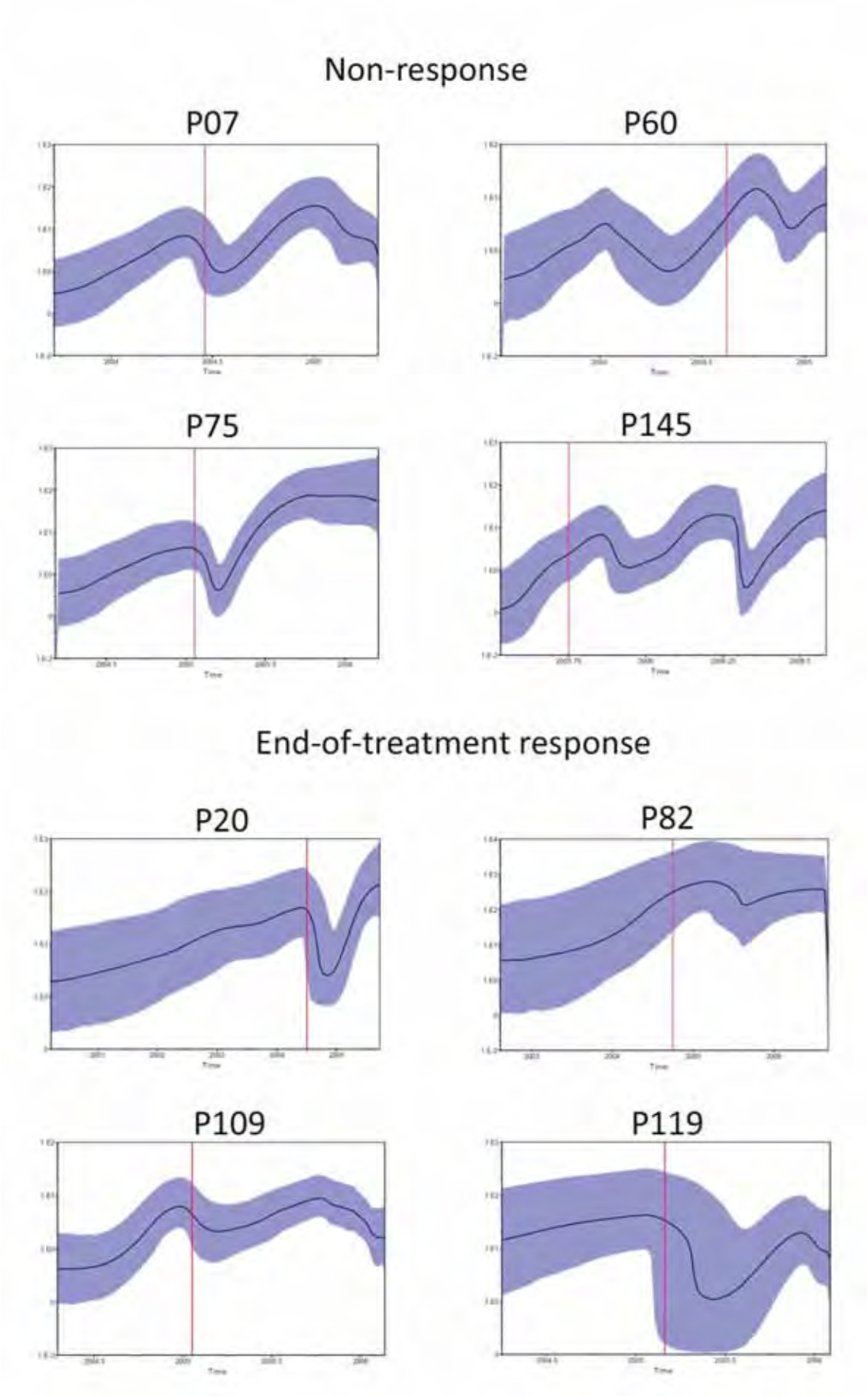


Figure 6: Bayesian Skyride plots showing the relative genetic diversity. Black solid lines represent the median posterior distribution, blue shaded areas are the 95% Bayesian credible intervals, vertical red lines represent the approximate time at which the treatment started (BT time-point).

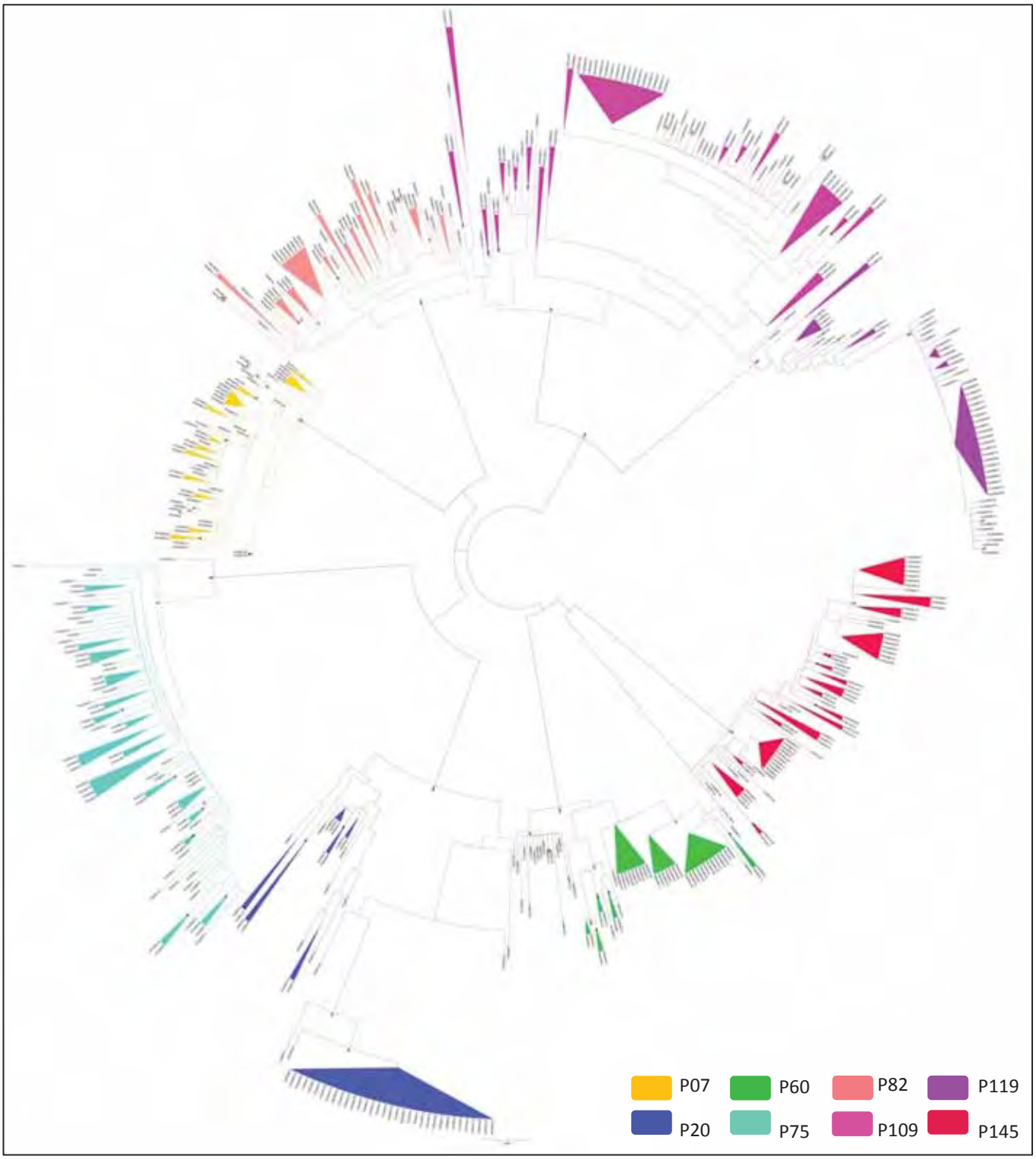


Figure 7: Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 500 replicates. Values above 70% were considered significant.

Supplementary material

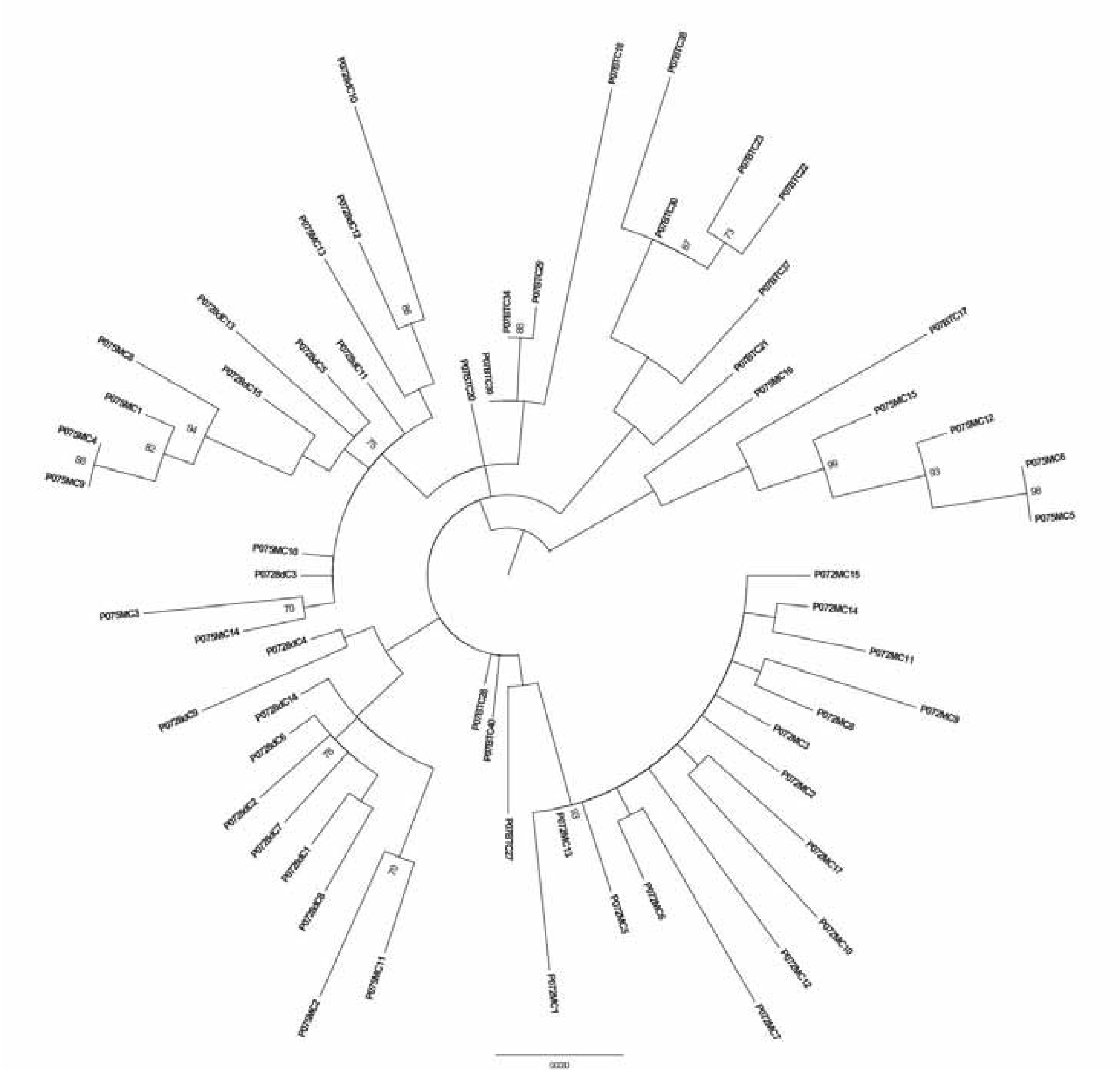


Figure SR-1: P07 Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 1000 replicates. Values above 70% were considered significant. BT - Before treatment; w - weeks; d - days; m - months.

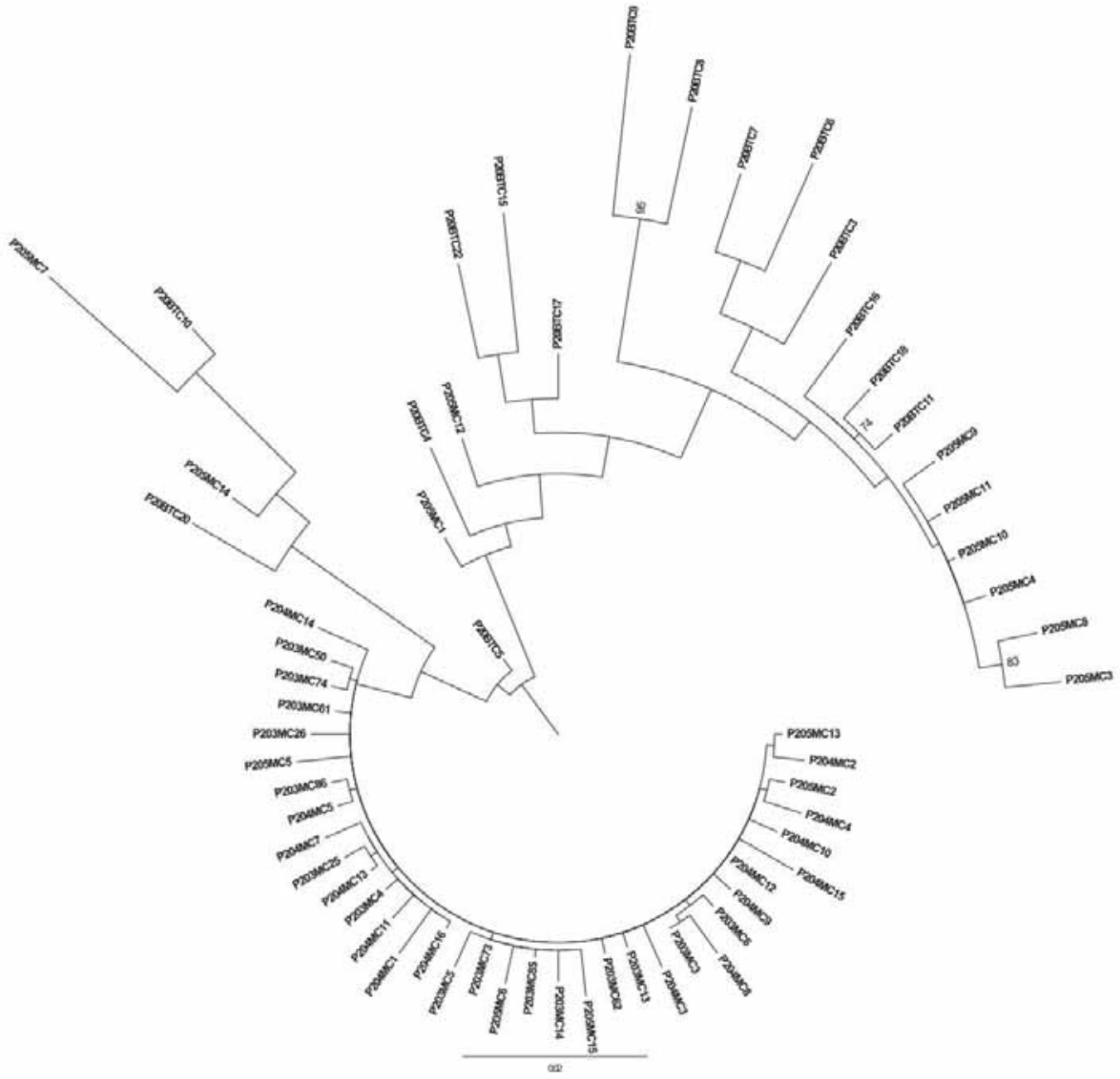


Figure SR-2: P20 Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 1000 replicates. Values above 70% were considered significant. BT - Before treatment; w - weeks; d - days; m - months.

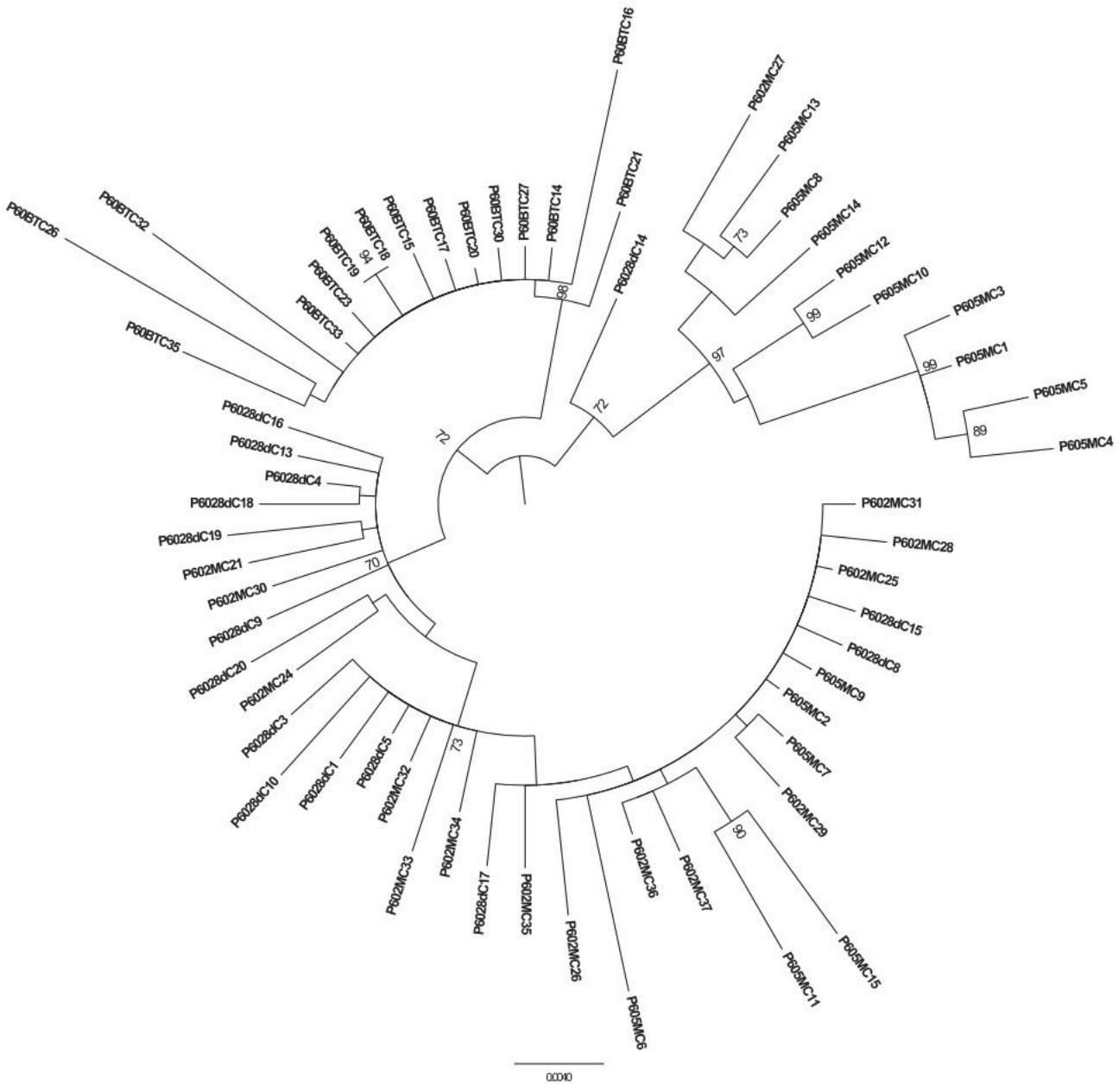


Figure SR-3: P60 Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 1000 replicates. Values above 70% were considered significant. BT - Before treatment; w - weeks; d - days; m - months.

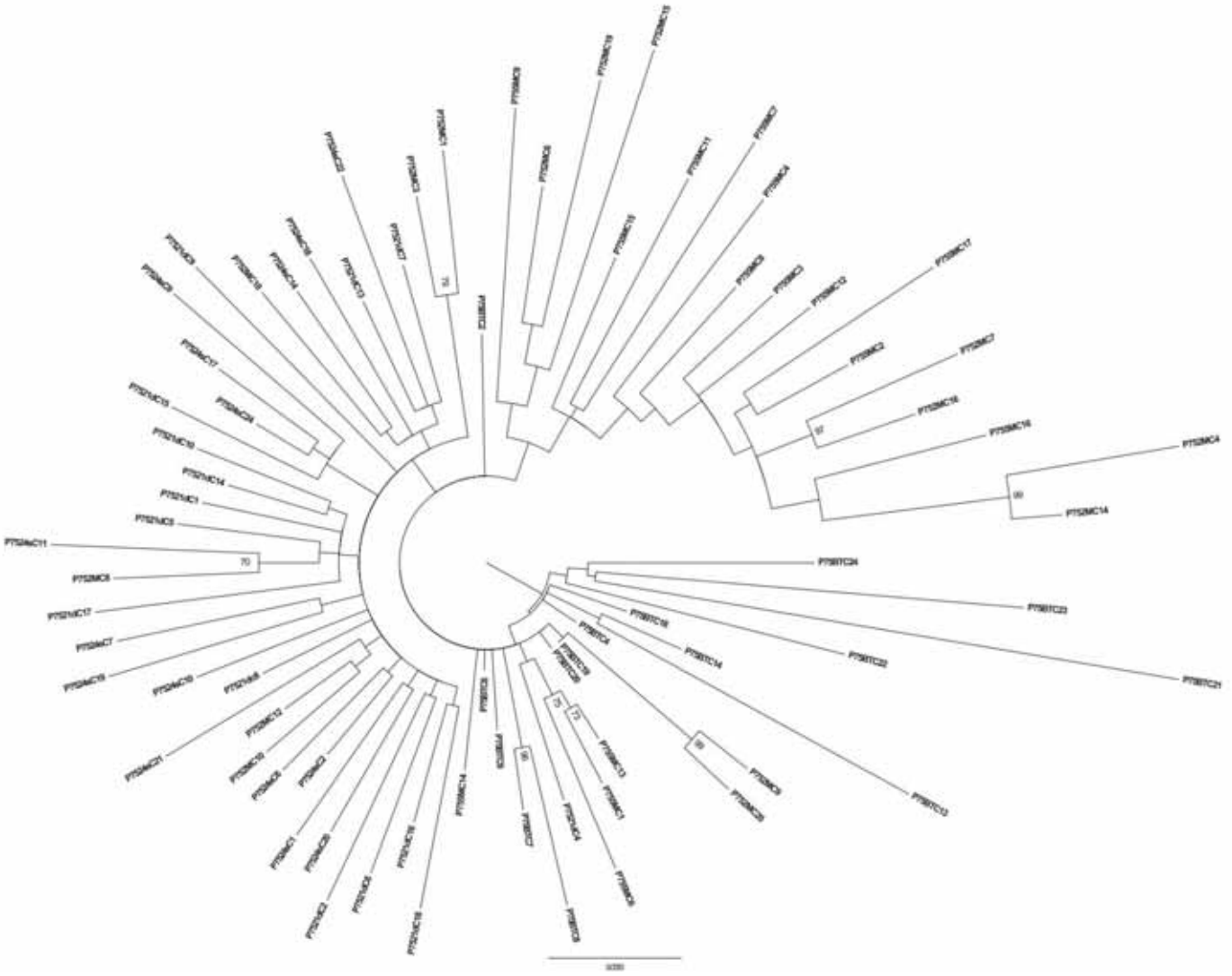


Figure SR-4: P75 Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 1000 replicates. Values above 70% were considered significant. BT - Before treatment; w - weeks; d - days; m - months.

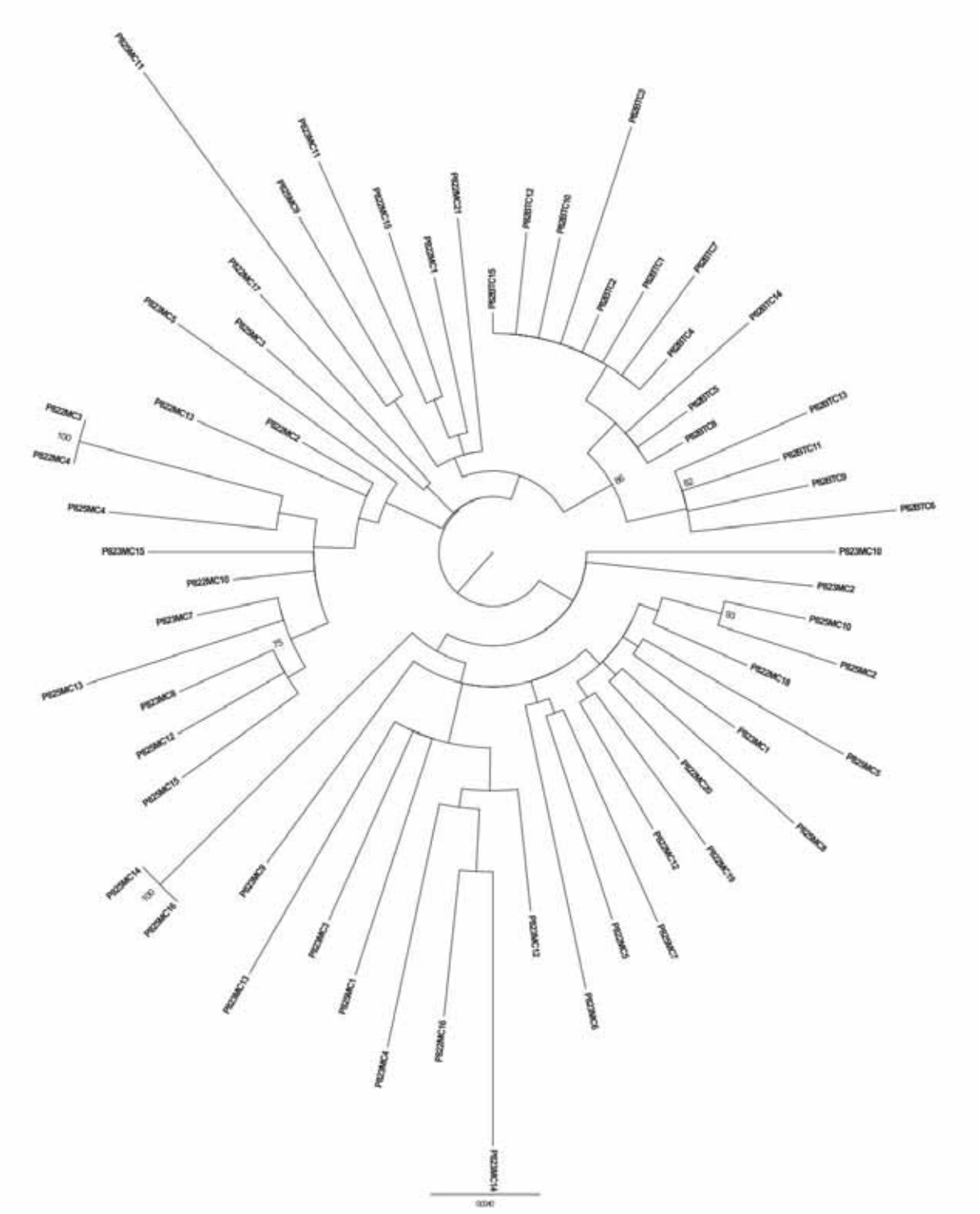


Figure SR-5: P82 Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 1000 replicates. Values above 70% were considered significant. BT - Before treatment; w - weeks; d - days; m - months.

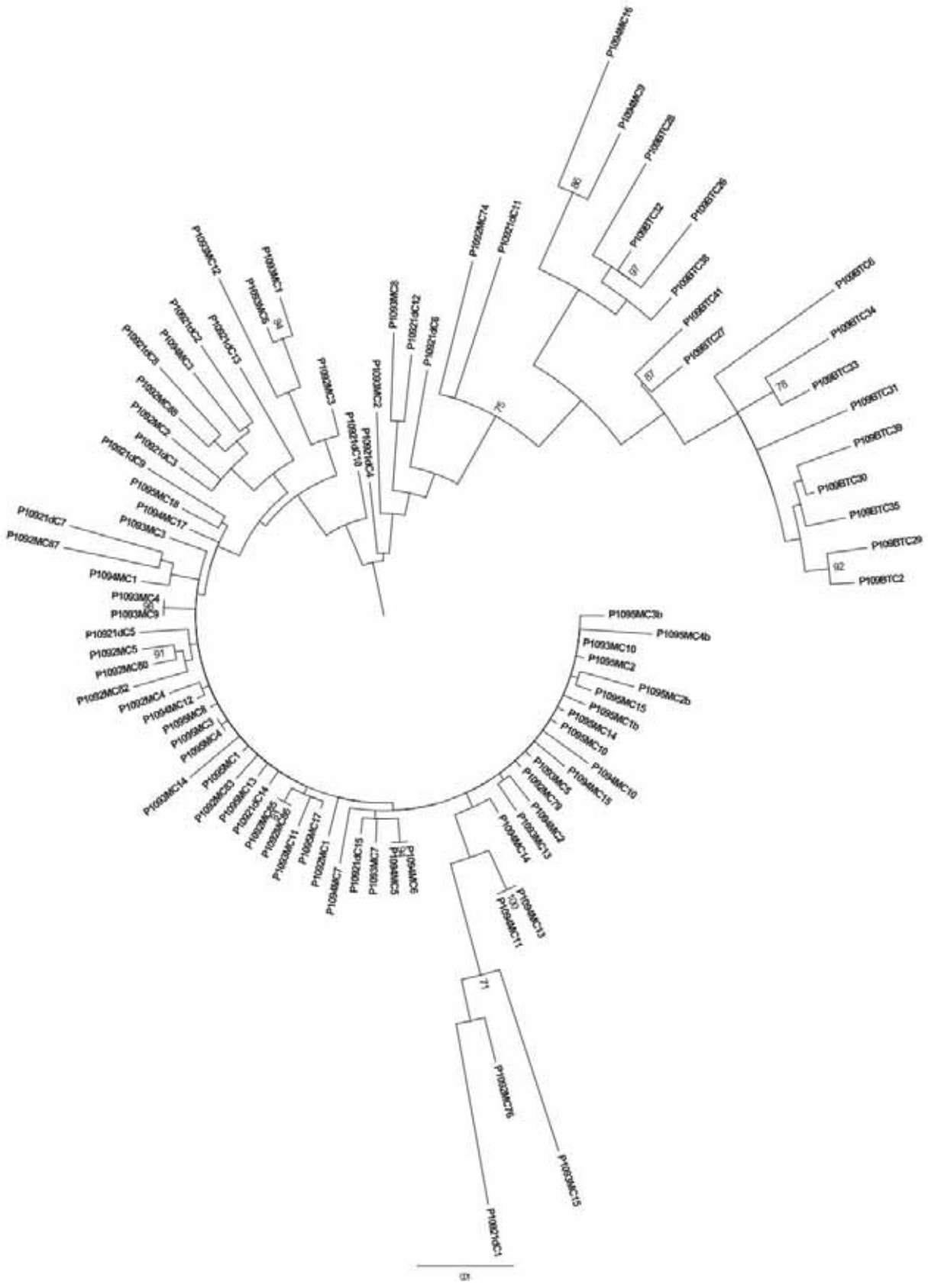


Figure SR-6 P109 Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 1000 replicates. Values above 70% were considered significant. BT - Before treatment; w - weeks; d - days; m - months.

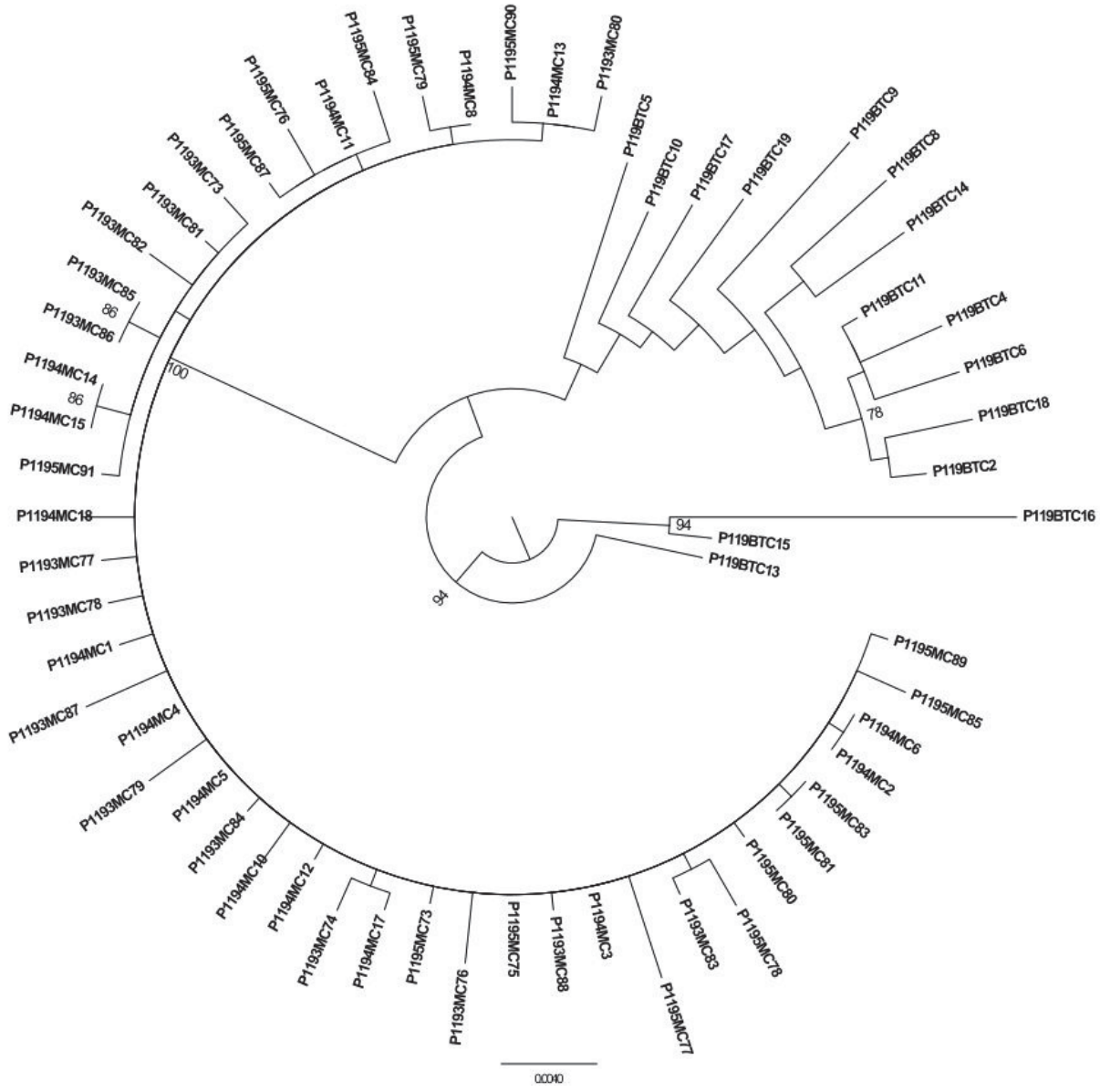


Figure SR-7: P119 Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 1000 replicates. Values above 70% were considered significant. BT - Before treatment; w - weeks; d - days; m - months.

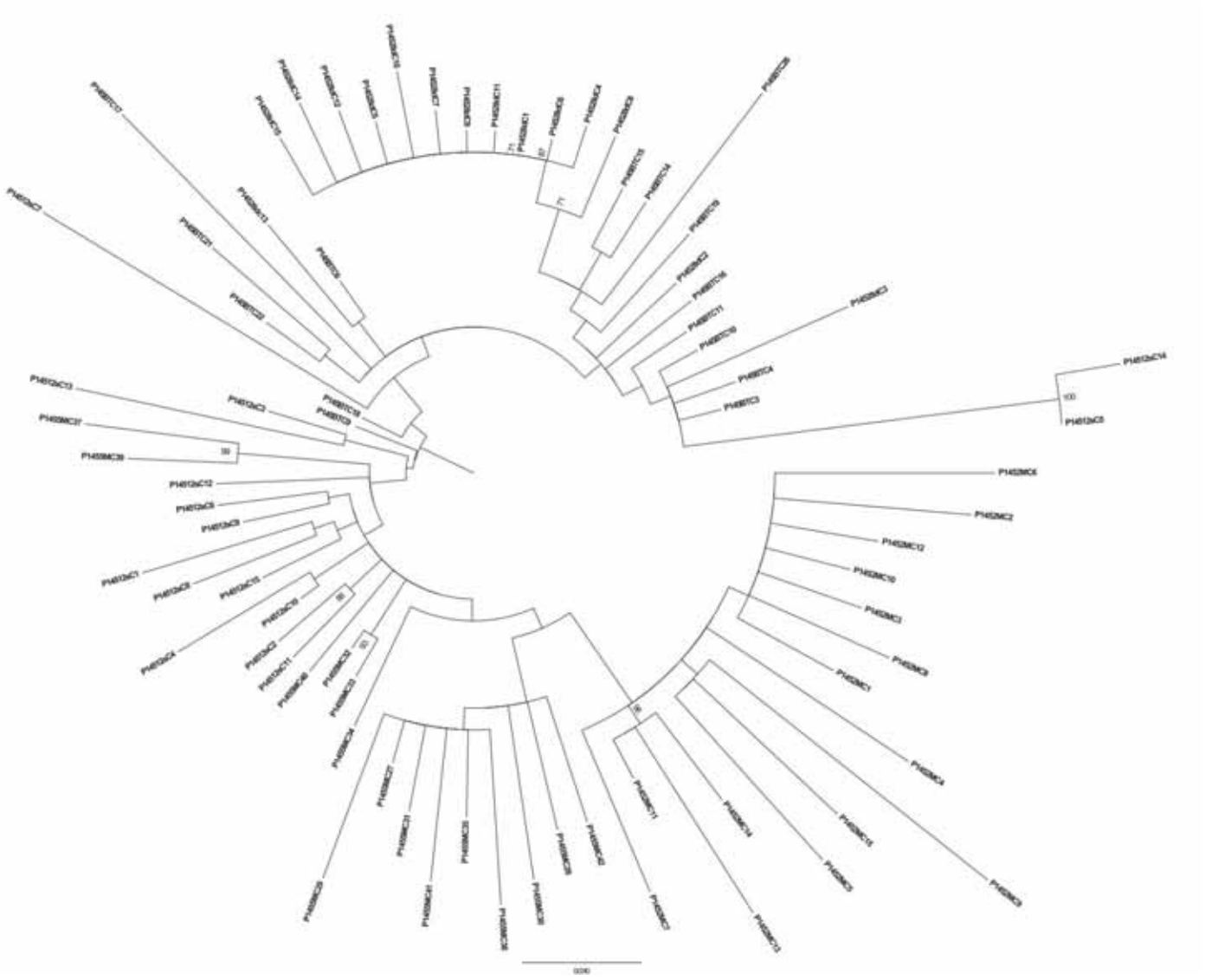


Figure SR-8: P145 Unrooted Maximum-Likelihood phylogenetic tree was constructed with HKY85 substitution model including Gamma distribution parameter (HKY+G). Bootstrap was performed with 1000 replicates. Values above 70% were considered significant. BT - Before treatment; w - weeks; d - days; m - months.

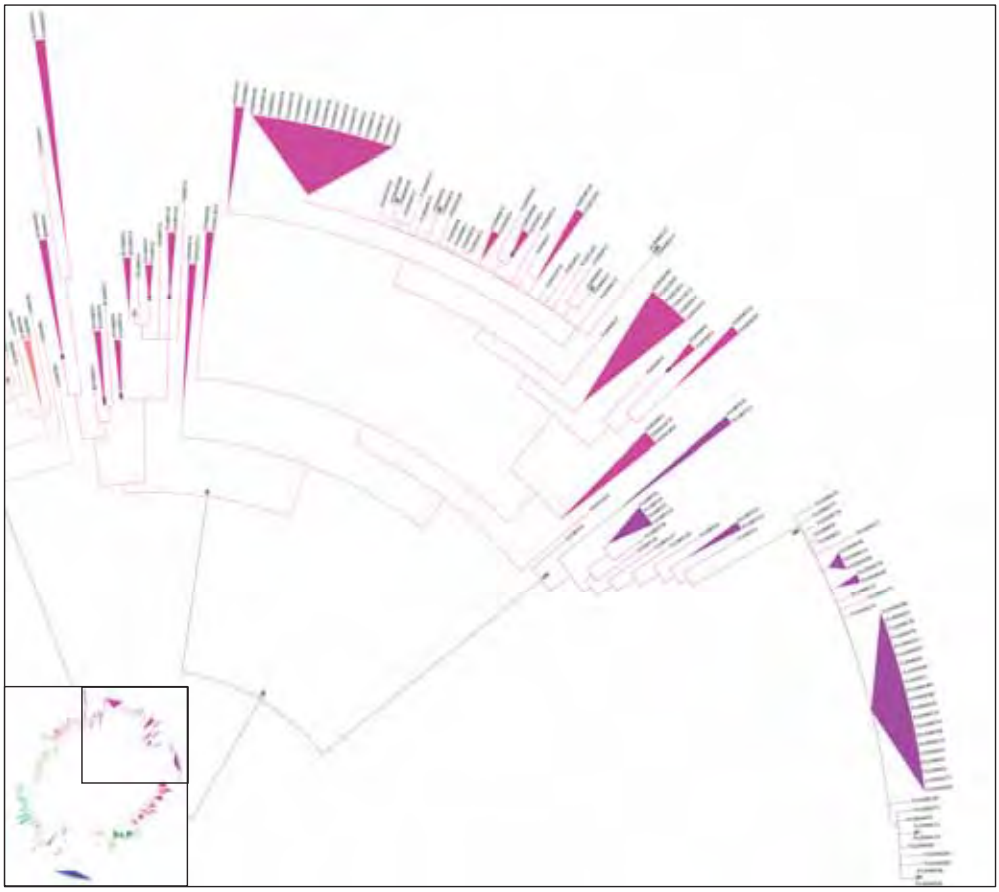


Figure SR-9: Amplification of the branch from patients P109 and P119 of the phylogenetic tree presented in Figure 8. Amplified region is represented on the left lower side.

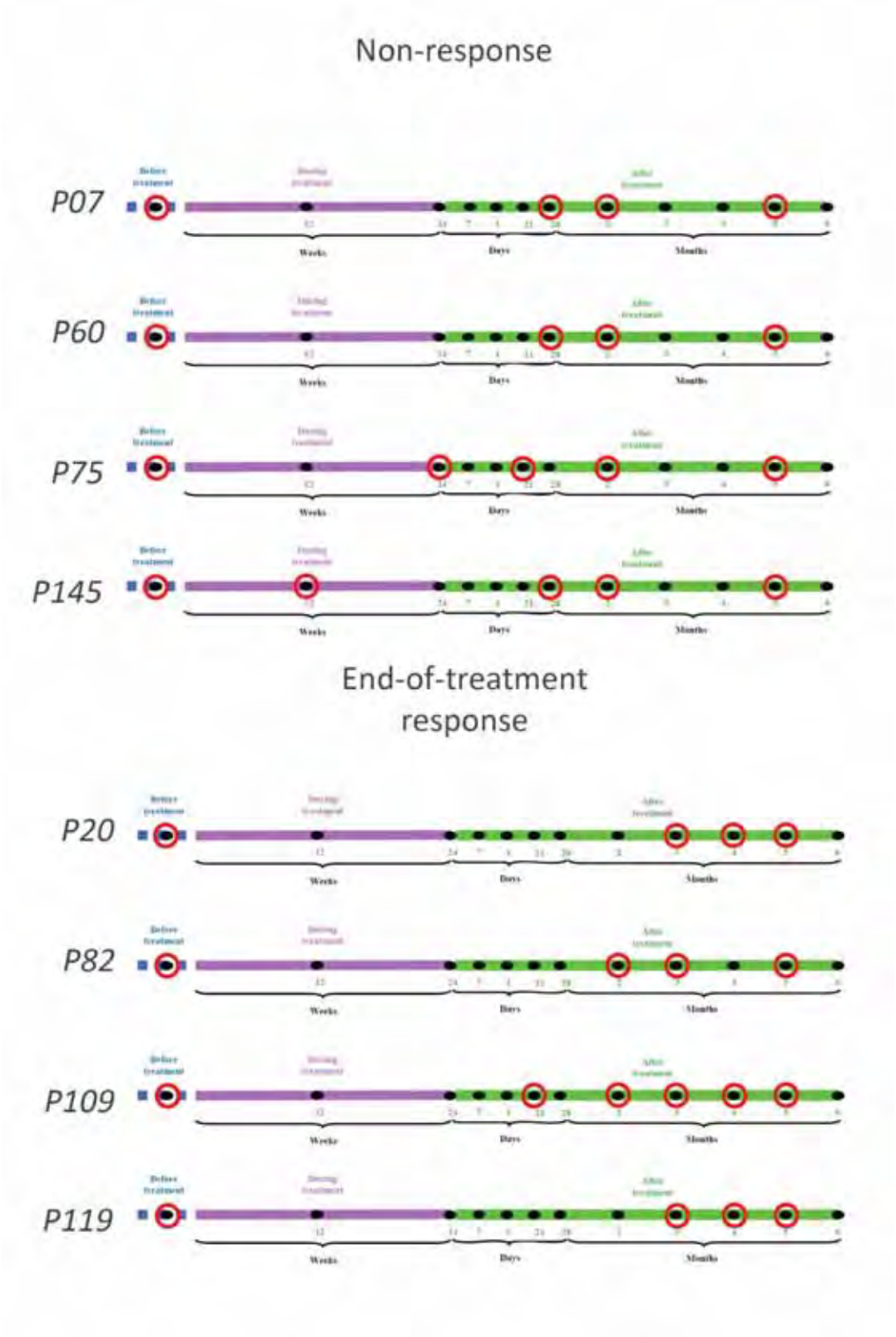


Figure SM-1: Graphic representation of the time-points used for treatment follow-up. The red circles represent the time-points when the sequences were used for the analysis from this study. Before treatment sequences from P07, P20, P60, P75, P109, P119, P145 were from a previous study detailed in Bittar *et al* (2010).

Non response (NR)

A.

P07		BT	28d	2M
	BT			
	28d	0.6628		
	2M	0.0000	0.0000	
	5M	0.0000	0.0000	0.0000

B.

P60		BT	28d	2M
	BT			
	28d	0.0000		
	2M	0.0000	0.3618	
	5M	0.0000	0.0000	0.0000

C.

P75		BT	24s	21d	2M
	BT				
	24s	0.1870			
	21d	0.6558	0.0000		
	2M	0.0000	0.0000	0.0000	
	5M	0.0068	0.0003	0.0000	0.0003

D.

P145		BT	12s	28d	2M
	BT				
	12s	0.0000			
	28d	0.0000	0.0000		
	2M	0.0000	0.0196	0.0000	
	5M	0.0005	0.0027	0.0000	0.2595

End-of-treatment response (ETR)

E.

P20		BT	3M	4M
	BT			
	3M	0.0000		
	4M	0.0000	0.0000	
	5M	0.4333	0.0000	0.0000

F.

P82		BT	2M	3M
	BT			
	2M	0.0000		
	3M	0.0000	0.0000	
	5M	0.0000	0.0000	0.0199

G.

P109		BT	21d	2M	3M	4M
	BT					
	21d	0.0001				
	2M	0.0025	0.0000			
	3M	0.3221	0.0001	0.1184		
	4M	0.0001	0.0000	0.5671	0.3030	
	5M	0.0000	0.0000	0.0000	0,0000	0.0000

H.

P119		BT	3M	4M
	BT			
	3M	0.0000		
	4M	0.0000	0.0000	
	5M	0.0000	0.0030	0.0000

Table ST-1: Mann-Whitney statistical test on distances between each time-point. A. P07; B. P60; C P75; D. P145; E. P20; F. P82; G. P109; H. P119. In red differences that were not significant (significance $p<0.005$).

Set	Reaction	Name	Sequence (5' – 3')	Nucleotide position in NZL1 (GenBank D17763G)
1*	PCR	*H.NS5AP-F*	GAGCGGTACAGTGGATGAAC	6089 to 6108
		H.NS5AP-R	CCTCCTTTAATGCAGTCTTG	7821 to 7840
	Nested-PCR	*H.NS5AN-F*	CGCATTGCTGAGTTCTCTAAC	6206 to 6226
		H.NS5AN-R	CAACAAGGAGTTGCTGAGTG	7703 to 7722
2*	PCR	*H.NS5AP-F2*	GGTACAGTGGATGAACAGG	6093 to 6111
		H.NS5AP-R2	ACGACGTTGAATAGACTAGG	7734 to 7753
	Nested-PCR	*H.NS5AN-F2*	CTCTAACTGTCACAAGTCTGC	6206 to 6226
		H.NS5AN-R2	CAGCACTACATGGTGTTATC	7659 to 7678
3	PCR	*H.NS5AP-F4*	TTGCCCGCCATACTATCT	6004 to 6021
		HNS5AP-R4	TTTACCTCCTKTAATRCA	7827 to 7844
	Nested-PCR	*H.NS5AN-F4*	GTVCAGTGGATGAACAG	6094 to 6110
		H.NS5AN-R4	AGGTAACCTTCYTCTGAC	7772 to 7789

Table ST-2: Primers used on PCR and NESTED-PCR reactions.

* published in Bittar *et al* (2010).

REFERÊNCIAS BIBLIOGRÁFICAS DA INTRODUÇÃO

REFERÊNCIAS BIBLIOGRÁFICAS DA INTRODUÇÃO

Aach, R. D., C. E. Stevens*, et al.* Hepatitis C virus infection in post-transfusion hepatitis. An analysis with first- and second-generation assays. N Engl J Med, v.325, n.19, Nov 7, p.1325-9. 1991.

Alter, H. J., C. Conry-Cantilena*, et al.* Hepatitis C in asymptomatic blood donors. Hepatology, v.26, n.3 Suppl 1, Sep, p.29S-33S. 1997.

Alter, H. J., R. H. Purcell*, et al.* Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B hepatitis. N Engl J Med, v.321, n.22, Nov 30, p.1494-500. 1989.

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

_____. HCV routes of transmission: what goes around comes around. Semin Liver Dis, v.31, n.4, Nov, p.340-6. 2011.

Arase, Y., K. Ikeda*, et al.* Efficacy and changes of the nonstructural 5A GENE by prolonged interferon therapy for patients with hepatitis C virus genotype 1b and a high level of serum HCV-RNA. Intern Med, v.38, n.6, Jun, p.461-6. 1999.

Bartenschlager, R., M. Frese*, et al.* Novel insights into hepatitis C virus replication and persistence. Adv Virus Res, v.63, p.71-180. 2004.

Bartosch, B. e F. L. Cosset. Cell entry of hepatitis C virus. Virology, v.348, n.1, Apr 25, p.1-12. 2006.

Bartsch, H. e J. Nair. Oxidative stress and lipid peroxidation-derived DNA-lesions in inflammation driven carcinogenesis. Cancer Detect Prev, v.28, n.6, p.385-91. 2004.

Berg, T., A. Mas Marques*, et al.* Mutations in the E2-PePHD and NS5A region of hepatitis C virus type 1 and the dynamics of hepatitis C viremia decline during interferon alfa treatment. Hepatology, v.32, n.6, Dec, p.1386-95. 2000.

Blight, K. J., A. A. Kolykhalov*, et al.* Efficient initiation of HCV RNA replication in cell culture. Science, v.290, n.5498, Dec 8, p.1972-4. 2000.

Bowen, D. G. e C. M. Walker. The origin of quasispecies: cause or consequence of chronic hepatitis C viral infection? J Hepatol, v.42, n.3, Mar, p.408-17. 2005.

Brasil. Protocolo Clínico e Diretrizes Terapêuticas - Hepatite Viral Crônica C e coinfecções: Ministério da Saúde 2011.

Bukh, J., R. H. Purcell*, et al.* At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. Proc Natl Acad Sci U S A, v.90, n.17, Sep 1, p.8234-8. 1993.

Busch, M. P., S. A. Glynn*, et al.* A new strategy for estimating risks of transfusion-transmitted viral infections based on rates of detection of recently infected donors. Transfusion, v.45, n.2, Feb, p.254-64. 2005.

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

Carrere-Kremer, S., C. Montpellier, *et al.* Regulation of hepatitis C virus polyprotein processing by signal peptidase involves structural determinants at the p7 sequence junctions. J Biol Chem, v.279, n.40, Oct 1, p.41384-92. 2004.

Casiraghi, M. A., M. De Paschale, *et al.* Long-term outcome (35 years) of hepatitis C after acquisition of infection through mini transfusions of blood given at birth. Hepatology, v.39, n.1, Jan, p.90-6. 2004.

Chayama, K., A. Tsubota, *et al.* Effect of lymphoblastoid alfa-interferon in patients with chronic hepatitis C having different genotypic subtype of hepatitis C virus. Gastroenterol Jpn, v.28 Suppl 5, May, p.45-7. 1993.

_____. 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, Mar, p.745-9. 1997.

Chevaliez, S. e J. M. Pawlotsky. Hepatitis C virus: virology, diagnosis and management of antiviral therapy. World J Gastroenterol, v.13, n.17, May 7, p.2461-6. 2007.

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

Chung, R. T., A. Monto, *et al.* Mutations in the NS5A region do not predict interferon-responsiveness in american patients infected with genotype 1b hepatitis C virus. J Med Virol, v.58, n.4, Aug, p.353-8. 1999.

Cocquerel, L., C. Voisset, *et al.* Hepatitis C virus entry: potential receptors and their biological functions. J Gen Virol, v.87, n.Pt 5, May, p.1075-84. 2006.

Coffin, J. M. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science, v.267, n.5197, Jan 27, p.483-9. 1995.

Cornberg, M., H. A. Razavi, *et al.* A systematic review of hepatitis C virus epidemiology in Europe, Canada and Israel. Liver Int, v.31 Suppl 2, Jul, p.30-60. 2011.

Corvino, S. M., R. M. S. Henriques, *et al.* Distribuição dos genótipos do HCV em pacientes das regiões de Botucatu, Bauru, Assis, SP, Brasil. Rev Inst Adolfo Lutz, v.65(2), p.137--140. 2006.

Deng, L., M. Nagano-Fujii, *et al.* NS3 protein of Hepatitis C virus associates with the tumour suppressor p53 and inhibits its function in an NS3 sequence-dependent manner. J Gen Virol, v.87, n.Pt 6, Jun, p.1703-13. 2006.

Domingo, E., C. Escarmis, *et al.* Basic concepts in RNA virus evolution. Faseb J, v.10, n.8, Jun, p.859-64. 1996.

Domingo, E. e J. J. Holland. RNA virus mutations and fitness for survival. Annu Rev Microbiol, v.51, p.151-78. 1997.

Dubuisson, J. Hepatitis C virus proteins. World J Gastroenterol, v.13, n.17, May 7, p.2406-15. 2007.

Dubuisson, J., F. Penin, *et al.* Interaction of hepatitis C virus proteins with host cell membranes and lipids. Trends Cell Biol, v.12, n.11, Nov, p.517-23. 2002.

Duverlie, G., H. Khorsi, *et al.* Sequence analysis of the NS5A protein of European hepatitis C virus 1b isolates and relation to interferon sensitivity. J Gen Virol, v.79 (Pt 6), Jun, p.1373-81. 1998.

Egger, D., B. Wolk, *et al.* Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. J Virol, v.76, n.12, Jun, p.5974-84. 2002.

Eigen, M. Viral quasispecies. Sci Am, v.269, n.1, Jul, p.42-9. 1993.

Eigen, M. e C. Biebricher. Sequence space and quasispecies distribution. Boca Raton, FL: CRC Press, v.3. 1988. 211–245 p. (RNA Genetics)

Enomoto, N., I. Sakuma, *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, Jul, p.224-30. 1995.

_____. 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, Jan 11, p.77-81. 1996.

Feinstone, S. M., A. Z. Kapikian, *et al.* Transfusion-associated hepatitis not due to viral hepatitis type A or B. N Engl J Med, v.292, n.15, Apr 10, p.767-70. 1975.

Forns, X. e J. Bukh. The molecular biology of hepatitis C virus. Genotypes and quasispecies. Clin Liver Dis, v.3, n.4, Nov, p.693-716, vii. 1999.

Fried, M. W., M. L. Shiffman, *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med, v.347, n.13, Sep 26, p.975-82. 2002.

Gale, M. J., Jr., M. J. Korth, *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, Apr 14, p.217-27. 1997.

Girard, S., P. Shalhoub, *et al.* An altered cellular response to interferon and up-regulation of interleukin-8 induced by the hepatitis C viral protein NS5A uncovered by microarray analysis. Virology, v.295, n.2, Apr 10, p.272-83. 2002.

Griffin, S. D., L. P. Beales, *et al.* The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. FEBS Lett, v.535, n.1-3, Jan 30, p.34-8. 2003.

Griffin, S. D., R. Harvey, *et al.* A conserved basic loop in hepatitis C virus p7 protein is required for amantadine-sensitive ion channel activity in mammalian cells but is dispensable for localization to mitochondria. J Gen Virol, v.85, n.Pt 2, Feb, p.451-61. 2004.

Hauri, A. M., G. L. Armstrong, *et al.* The global burden of disease attributable to contaminated injections given in health care settings. Int J STD AIDS, v.15, n.1, Jan, p.7-16. 2004.

Helle, F., C. Wychowski, *et al.* Cyanovirin-N inhibits hepatitis C virus entry by binding to envelope protein glycans. J Biol Chem, v.281, n.35, Sep 1, p.25177-83. 2006.

Hissar, S. S., A. Goyal, *et al.* Hepatitis C virus genotype 3 predominates in North and Central India and is associated with significant histopathologic liver disease. J Med Virol, v.78, n.4, Apr, p.452-8. 2006.

Hladik, W., P. Kataaha, *et al.* Prevalence and screening costs of hepatitis C virus among Ugandan blood donors. Trop Med Int Health, v.11, n.6, Jun, p.951-4. 2006.

Hoofnagle, J. H., K. D. Mullen, *et al.* Treatment of chronic non-A,non-B hepatitis with recombinant human alpha interferon. A preliminary report. N Engl J Med, v.315, n.25, Dec 18, p.1575-8. 1986.

Hourioux, C., R. Patient, *et al.* The genotype 3-specific hepatitis C virus core protein residue phenylalanine 164 increases steatosis in an in vitro cellular model. Gut, v.56, n.9, Sep, p.1302-8. 2007.

Kalinina, O., H. Norder, *et al.* A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg. J Virol, v.76, n.8, Apr, p.4034-43. 2002.

Kanai, K., M. Kako, *et al.* HCV genotypes in chronic hepatitis C and response to interferon. Lancet, v.339, n.8808, Jun 20, p.1543. 1992.

Kato, N. Molecular virology of hepatitis C virus. Acta Med Okayama, v.55, n.3, Jun, p.133-59. 2001.

Kenny-Walsh, E. Clinical outcomes after hepatitis C infection from contaminated anti-D immune globulin. Irish Hepatology Research Group. N Engl J Med, v.340, n.16, Apr 22, p.1228-33. 1999.

Kershenobich, D., H. A. Razavi, *et al.* Trends and projections of hepatitis C virus epidemiology in Latin America. Liver Int, v.31 Suppl 2, Jul, p.18-29.

Kim, J. L., K. A. Morgenstern, *et al.* Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. Cell, v.87, n.2, Oct 18, p.343-55. 1996.

Kimura, M. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. Genet Res, v.11, n.3, Jun, p.247-69. 1968.

_____. The Neutral Theory of Molecular Evolution. Cambridge: Cambridge University Press. 1983

Krugman, S., J. P. Giles, *et al.* Infectious hepatitis. Evidence for two distinctive clinical, epidemiological, and immunological types of infection. Jama, v.200, n.5, May 1, p.365-73. 1967.

Kurosaki, M., N. Enomoto, *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, Mar, p.750-3. 1997.

Lavanchy, D. Evolving epidemiology of hepatitis C virus. Clin Microbiol Infect, v.17, n.2, Feb, p.107-15.

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

Levrero, M. Viral hepatitis and liver cancer: the case of hepatitis C. Oncogene, v.25, n.27, Jun 26, p.3834-47. 2006.

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

Lin, C., B. D. Lindenbach, *et al.* Processing in the hepatitis C virus E2-NS2 region: identification of p7 and two distinct E2-specific products with different C termini. J Virol, v.68, n.8, Aug, p.5063-73. 1994.

Locasciulli, A., M. Testa, *et al.* Prevalence and natural history of hepatitis C infection in patients cured of childhood leukemia. Blood, v.90, n.11, Dec 1, p.4628-33. 1997.

Lohmann, V., F. Korner, *et al.* Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science, v.285, n.5424, Jul 2, p.110-3. 1999.

Love, R. A., H. E. Parge, *et al.* The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. Cell, v.87, n.2, Oct 18, p.331-42. 1996.

Lyra, A. C., X. Fan, *et al.* Molecular biology and clinical implication of hepatitis C virus. Braz J Med Biol Res, v.37, n.5, May, p.691-5. 2004.

Manns, M. P., J. G. Mchutchison, *et al.* Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet, v.358, n.9286, Sep 22, p.958-65. 2001.

Martell, M., J. I. Esteban, *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, May, p.3225-9. 1992.

Mclauchlan, J. Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. J Viral Hepat, v.7, n.1, Jan, p.2-14. 2000.

Mello, I. M. V. G. C. Evidências moleculares da transmissão horizontal do vírus da hepatite C (VHC) entre cônjuges Faculdade de Medicina da Universidade de São Paulo Universidade de São Paulo - USP, São Paulo, 2006.

Mihm, S., A. Fayyazi, *et al.* Analysis of histopathological manifestations of chronic hepatitis C virus infection with respect to virus genotype. Hepatology, v.25, n.3, Mar, p.735-9. 1997.

Mizushima, H., M. Hijikata, *et al.* Two hepatitis C virus glycoprotein E2 products with different C termini. J Virol, v.68, n.10, Oct, p.6215-22. 1994.

Moradpour, D. e H. E. Blum. A primer on the molecular virology of hepatitis C. Liver Int, v.24, n.6, Dec, p.519-25. 2004.

Murphy, D., J. Chamberland, *et al.* A new genotype of hepatitis C virus orginating from central Africa. Hepatology, v.46, p.623A. 2007.

Murphy, M. D., H. R. Rosen, *et al.* Analysis of sequence configurations of the ISDR, PKR-binding domain, and V3 region as predictors of response to induction interferon-alpha and ribavirin therapy in chronic hepatitis C infection. Dig Dis Sci, v.47, n.6, Jun, p.1195-205. 2002.

Nakamoto, Y., L. G. Guidotti, *et al.* Immune pathogenesis of hepatocellular carcinoma. J Exp Med, v.188, n.2, Jul 20, p.341-50. 1998.

Nakamura, I., Y. Tanaka, *et al.* Clarification of interspousal hepatitis C virus infection in acute hepatitis C patients by molecular evolutionary analyses: Consideration on sexual and non-sexual transmission between spouses. Hepatol Res, v.41, n.9, Sep, p.838-45. 2011.

Nakano, I., Y. Fukuda, *et al.* Why is the interferon sensitivity-determining region (ISDR) system useful in Japan? J Hepatol, v.30, n.6, Jun, p.1014-22. 1999.

Nakayama, H., Y. Sugai, *et al.* Molecular investigation of interspousal transmission of hepatitis C virus in two Japanese patients who acquired acute hepatitis C after 40 or 42 years of marriage. J Med Virol, v.75, n.2, Feb, p.258-66. 2005.

Nousbaum, J., S. J. Polyak, *et al.* Prospective characterization of full-length hepatitis C virus NS5A quasispecies during induction and combination antiviral therapy. J Virol, v.74, n.19, Oct, p.9028-38. 2000.

Okuda, M., K. Li, *et al.* Mitochondrial injury, oxidative stress, and antioxidant gene expression are induced by hepatitis C virus core protein. Gastroenterology, v.122, n.2, Feb, p.366-75. 2002.

Pascu, M., P. Martus, *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, Sep, p.1345-51. 2004.

Pavlovic, D., D. C. Neville, *et al.* The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. Proc Natl Acad Sci U S A, v.100, n.10, May 13, p.6104-8. 2003.

Pawlotsky, J. M. e G. Germanidis. The non-structural 5A protein of hepatitis C virus. J Viral Hepat, v.6, n.5, Sep, p.343-56. 1999.

Pietschmann, T., A. Kaul, *et al.* Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. Proc Natl Acad Sci U S A, v.103, n.19, May 9, p.7408-13. 2006.

Polyak, S. J. Hepatitis C virus--cell interactions and their role in pathogenesis. Clin Liver Dis, v.7, n.1, Feb, p.67-88. 2003.

Polyak, S. J., K. S. Khabar, *et al.* Hepatitis C virus nonstructural 5A protein induces interleukin-8, leading to partial inhibition of the interferon-induced antiviral response. J Virol, v.75, n.13, Jul, p.6095-106. 2001.

Poordad, F. Big changes are coming in hepatitis C. Curr Gastroenterol Rep, v.13, n.1, Feb, p.72-7. 2010.

Qiang, G., L. Yang, *et al.* Recombinant adenoviruses expressing steatosis-associated hepatitis C virus genotype 3 core protein produce intracellular lipid accumulation in cultured and primary hepatocytes. Virus Res, v.139, n.1, Jan, p.127-30. 2009.

Reyes, G. R. The nonstructural NS5A protein of hepatitis C virus: an expanding, multifunctional role in enhancing hepatitis C virus pathogenesis. J Biomed Sci, v.9, n.3, May-Jun, p.187-97. 2002.

Rubbia-Brandt, L., G. Leandro, *et al.* Liver steatosis in chronic hepatitis C: a morphological sign suggesting infection with HCV genotype 3. Histopathology, v.39, n.2, Aug, p.119-24. 2001.

Rubbia-Brandt, L., R. Quadri, *et al.* Hepatocyte steatosis is a cytopathic effect of hepatitis C virus genotype 3. J Hepatol, v.33, n.1, Jul, p.106-15. 2000.

Sakai, A., M. S. Claire, *et al.* The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. Proc Natl Acad Sci U S A, v.100, n.20, Sep 30, p.11646-51. 2003.

Santolini, E., G. Migliaccio, *et al.* Biosynthesis and biochemical properties of the hepatitis C virus core protein. J Virol, v.68, n.6, Jun, p.3631-41. 1994.

Sarrazin, C., T. Berg, *et al.* Mutations in the protein kinase-binding domain of the NS5A protein in patients infected with hepatitis C virus type 1a are associated with treatment response. J Infect Dis, v.181, n.2, Feb, p.432-41. 2000.

Seeff, L. B. Strategies for assessing the long-term consequences of Hepatitis C virus infection. Minerva Gastroenterol Dietol, v.46, n.4, Dec, p.207-16. 2000.

_____. The history of the "natural history" of hepatitis C (1968-2009). Liver Int, v.29 Suppl 1, Jan, p.89-99. 2009.

Sharma, P., V. Balan, *et al.* Hepatic steatosis in hepatitis C virus genotype 3 infection: does it correlate with body mass index, fibrosis, and HCV risk factors? Dig Dis Sci, v.49, n.1, Jan, p.25-9. 2004.

Shepard, C. W., L. Finelli, *et al.* Global epidemiology of hepatitis C virus infection. Lancet Infect Dis, v.5, n.9, Sep, p.558-67. 2005.

Shirota, Y., H. Luo, *et al.* Hepatitis C virus (HCV) NS5A binds RNA-dependent RNA polymerase (RdRP) NS5B and modulates RNA-dependent RNA polymerase activity. J Biol Chem, v.277, n.13, Mar 29, p.11149-55. 2002.

Sievert, W., I. Altraif, *et al.* A systematic review of hepatitis C virus epidemiology in Asia, Australia and Egypt. Liver Int, v.31 Suppl 2, Jul, p.61-80. 2011.

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

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

Simmonds, P., E. C. Holmes, *et al.* Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. J Gen Virol, v.74 (Pt 11), Nov, p.2391-9. 1993.

Smith, D. B., S. Pathirana, *et al.* The origin of hepatitis C virus genotypes. J Gen Virol, v.78 (Pt 2), Feb, p.321-8. 1997.

Squadrito, G., M. E. Orlando, *et al.* Long-term response to interferon alpha is unrelated to "interferon sensitivity determining region" variability in patients with chronic hepatitis C virus-1b infection. J Hepatol, v.30, n.6, Jun, p.1023-7. 1999.

Tahan, V., C. Karaca, *et al.* Sexual transmission of HCV between spouses. Am J Gastroenterol, v.100, n.4, Apr, p.821-4. 2005.

Thomas, D. L., J. Astemborski*, et al.* The natural history of hepatitis C virus infection: host, viral, and environmental factors. Jama, v.284, n.4, Jul 26, p.450-6. 2000.

Tsai, W. L. e R. T. Chung. Viral hepatocarcinogenesis. Oncogene, v.29, n.16, Apr 22, p.2309-24. 2010.

Urbanus, A. T., T. J. Van De Laar*, et al.* Hepatitis C virus infections among HIV-infected men who have sex with men: an expanding epidemic. Aids, v.23, n.12, Jul 31, p.F1-7. 2009.

Vachon, M. L. e D. T. Dieterich. The era of direct-acting antivirals has begun: the beginning of the end for HCV? Semin Liver Dis, v.31, n.4, Nov, p.399-409. 2011.

Vogt, M., T. Lang*, et al.* Prevalence and clinical outcome of hepatitis C infection in children who underwent cardiac surgery before the implementation of blood-donor screening. N Engl J Med, v.341, n.12, Sep 16, p.866-70. 1999.

Wakita, T., T. Pietschmann*, et al.* Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. Nat Med, v.11, n.7, Jul, p.791-6. 2005.

Who. World Health Organization 2000.

Yoshioka, K., S. Kakumu*, et al.* Detection of hepatitis C virus by polymerase chain reaction and response to interferon-alpha therapy: relationship to genotypes of hepatitis C virus. Hepatology, v.16, n.2, Aug, p.293-9. 1992.

Zeuzem, S., J. H. Lee*, et al.* Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon alfa. Hepatology, v.25, n.3, Mar, p.740-4. 1997.

ANEXO

1. COMPROVANTE DE SUBMISSÃO EM REVISTA CIENTÍFICA

Gmail by Google

Cintia Bittar <cibittar@gmail.com>

6821133706651362 On Hepatitis C virus evolution: the interaction between virus and host towards treatment outcome.

BioMed Central Editorial **Mon, Jan 30, 2012 at 12:58 PM**
<editorial@biomedcentral.com >
To: Dr Isabel Maria de Carvalho-Mello <imvgcmello@butantan.gov.br >
Cc: Cintia Bittar <cibittar@gmail.com >, Ana Carolina G Jardim <caroljardim@gmail.com >, Lilian H T Yamasaki <lilian.hty@gmail.com >, Claudia M A Carareto <carareto@ibilce.unesp.br >, João Renato R Pinho <jrrpinho@usp.br >, Philippe Lemey <philippe.lemey@rega.kuleuven.be >, Isabel M V G de Carvalho-Mello <imvgcmello@butantan.gov.br >, Paula Rahal <rahalp@yahoo.com.br >

Article title: On Hepatitis C virus evolution: the interaction between virus and host towards treatment outcome.
MS ID : 6821133706651362
Authors : Cintia Bittar, Ana Carolina G Jardim, Lilian H T Yamasaki, Claudia M A Carareto, João Renato R Pinho, Philippe Lemey, Isabel M V G de Carvalho-Mello and Paula Rahal
Journal : BMC Evolutionary Biology

Dear Dr de Carvalho-Mello

Thank you for submitting your article. This acknowledgement and any queries below are for the submitting author. This e-mail has also been copied to each author on the paper. Please bear in mind that all queries regarding the paper should be made through the submitting author.

A pdf file has been generated from your submitted manuscript and figures. We would be most grateful if you could check this file and let us know if any aspect is missing or incorrect. Any additional files you uploaded will also be sent in their original format for review.

http://www.biomedcentral.com/imedia/6821133706651362_article.pdf

(2969K)

For your records, please find below link(s) to the correspondence you uploaded with this submission. Please note there may be a short delay in creating this file.

If the PDF does not contain the comments which you uploaded, please upload the cover letter again, click "Continue" at the bottom of the page, and then proceed with the manuscript submission again. If the letter will not upload, please send a copy to editorial@biomedcentral.com.

The submitting author can check on the status of the manuscript at any time by logging into 'My BioMed Central' (http://www.biomedcentral.com/my).

In the meantime, if you have any queries about the manuscript you may contact us on editorial@biomedcentral.com. We would also welcome feedback about the online submission process.

Best wishes,

The BioMed Central Editorial Team

Tel: +44 (0) 20 3192 2013
e-mail: editorial@biomedcentral.com
Web: http://www.biomedcentral.com/

Autorizo a reprodução xerográfica para fins de pesquisa.

São José do Rio Preto, 10/03/2012

_____

Assinatuta