HEPATITIS C VIRUS COMPARTMENTALISATION: UNRAVELLING THE GENETIC COMPLEXITY

by

DITTE CHRISTIANE EMMA LINDEMANN HEDEGAARD

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

School of Immunity and Infection

College of Medicine and Dentistry

University of Birmingham

June 2014

ABSTRACT

Hepatitis C virus (HCV) is a global health problem with over 150 million individuals infected worldwide. Many of these patients will develop end-stage liver diseases, such as cirrhosis and hepatocellular carcinoma (HCC), and will require liver transplantation.

HCV exists as a heterogeneous population in infected individuals, however, the processes which maintain this genetic complexity are unknown. Recent observations suggest that HCV transmits between hepatocytes via a cell-to-cell route of infection, supporting a “demic” model of evolution where HCV diversity arises from independent evolution in small isolated hepatic populations.

To investigate the distribution of HCV within the liver we sampled eight segments of the liver explant from 22 HCV infected subjects undergoing liver transplant and measured viral RNA burden and sequence diversity. Comparable HCV RNA levels were observed across all 8 samples from a single liver, however, between patients we observed a 100-fold range in the hepatic viral load that was independent of hepatic expression of anti-viral and pro-viral interferon stimulated genes (ISGs). Sequence analysis of the viral envelope E1E2 region, obtained from PCR generated single molecules or ultra-deep sequencing approaches, showed minimal evidence of genetic compartmentalisation between hepatic sites or between the liver and plasma.

Modelling the HCV population structure in infected patients will have a major impact on our understanding of how HCV escapes host immune responses and anti-viral therapies.

DEDICATION

I dedicate this thesis to my loving parents Tove and Niels, who have always been there to support me when times have been difficult.

"*Mor og Far, jeg kunne aldrig have klaret det her uden jer. Jeg elsker jer over alt i verden*".

And to my brothers, who are the best older- and younger-brother a girl can wish for.

Finally, I would also like to dedicate this thesis to my close friends in Birmingham. Thank you for always making me smile and for supporting me during the past 3.5 years in Birmingham. My amazing Sniff, Wifee, Tunisian Baklava and English Muffin you know who you are……..

ACKNOWLEDGEMENT

I would like to thank my supervisors Prof Jane McKeating, Prof Deirdre Kelly and Dr Peter Balfe for their supervision, support and guidance throughout my studies. I especially like to mention Dr Peter Balfe, who has been my guru and mentor in sequencing and single molecule amplification in the laboratory. I also like to thank past and present members in the Viral Hepatitis group for their technical support and scientific discussions during the past years. It has been a joy working with all of you; Colin Howard, Zania Stamataki, Dave Mason, Tessa Lawrence, Megan Roberts, Lizzy Benedikz, Sam Lissauer, Suki Galsinh, Amy Barnes, Ian Rowe, Michelle Farquhar, Garrick Wilson, Nick Frampton, Reina Lim, Ke Hu, Mandy Diskar, Luke Meredith and Nicola Fletcher. Furthermore I would like to thank the staff at the Children's Hospital Hepatology Department for providing children biopsy samples.

I also owe a special thanks to Dr Ian Rowe and Dr Gary Reynolds from the Liver Research Unit, UK, for collecting patient samples for my project, to Annika Wilhelm from the Liver Research Unit, for providing RNA extractions from varies liver diseases, to Dr Damien Tully, Prof Todd Allen and Dr Karen Power at the Ragon Institute of MGH, MIT and Harvard, Boston, USA, for performing the next-generation sequencing and helping with the analysis, to Dr Alex Tarr at University of Nottingham, UK, for helping with the HCV genotype 1a amplification, Prof Stuart Ray and Dr Ashwin Balagopal at John's Hopkins University in Baltimore, USA, for inviting me to visit their lab and to all the patients who donated samples to this project.

ABBREVIATIONS

Aa	Amino acids
Ab	Antibody
AIH	Auto-immune hepatitis
ALD	Alcoholic liver disease
ALT	Alanine aminotransferase
Au	Australian antigen
CHV	Canine hepacivirus
CLDN1	Claudin 1
DAA	Direct acting anti-viral
DDX3	DEAD box helicase 3, X-linked
DMSO	Dimethyl sulfoxide
dn	Non-synonymous
ds	Synonymous
E1/2	Envelope protein 1 / 2
ER	Endoplasmic reticulum
FDA	Food and drug administration
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GC	Guanine-cytosine
Gt	Genotype
HBsAg	Hepatitis B virus antigen
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C virus
HCVcc	HCV cell culture
HCVpp	HCV pseudo particle
HTATIP2	HIV-tat interactive protein 2

IFI27	α-inducible protein (IFI) 27
IFN	Interferon
IRF	IFN regulatory factor
ISRE	Interferon-stimulated response element
ISG	Interferon stimulated gene
ISG15	Interferon stimulated gene 15
IRES	Internal ribosome entry site
IU	International unit
JFH	Japanese fulminant hepatitis
MTCT	Mother-to-child-transmission
MxA	IFN-regulated resistance GTP-binding protein
nAb	Neutralising antibody
nA/nB/nC	non-A/non-B/non-C
NASH	Nonalcoholic steatohepatitis
NF-kB	Kappa-light-chain-enhancer of activated B cells
NGS	Next-generation sequencing
NPC1L1	Niemann-Pick C1-like 1 cholesterol absorption receptor
NS	Non-structural
OCLN	Occludin
ORF	Open reading frame
PAMPs	Pathogen-associated molecular patterns
PCR	Polymerase chain reaction
PNG	Potential N-linked glycosylation
Pt	Patient
qPCR	Quantitative PCR
qRT-PCR	Real-Time qPCR
Rd-Rp	RNA-dependent RNA polymerase

RIG-I	Retinoic acid inducible gene-I
RPM	Revolutions per minute
RSAD2	Radical S-adenosyl methionine domain containing 2
SR-BI	Scavenger receptor B type I
sE2	Soluble E2
SGA	Single genome amplification
SSCP	Single strand conformational polymorphism
SVR	Sustained viral response
TfR1	Transferrin receptor 1
TLR	Toll-like receptor
Tm	Melting temperature
UTR	Un-translated region
VLDL	Very-low-density-lipid

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

INTRODUCTION

1.1 A brief introduction to viral hepatitis

Epidemic jaundice was first mentioned in the Babylonian Talmud in the 5th century B.C. Since then epidemic jaundice has been recognised as an important military disease, especially linked to times of war, without anyone knowing the cause. In the 1930's several doctors noticed an increase in the number of soldiers with hepatitis following yellow fever immunisation with glycerinated human lymph-node preparation. Often these incidents were blamed on other drugs being administrated and not on the human serum inoculum. However, accumulating transmission and epidemiology studies during the Second World War lead to the discovery of two different types of hepatitis caused by the ultra-filterable agents hepatitis type A (infectious) and type B (serum) (Krugman, 1976; MacCallum, 1972; Reuben, 2002). Several studies documented that hepatitis type A is transmitted through the faecal/oral transmission route and that hepatitis B is transferred via blood and blood products, making the later the source of hepatitis following immunisation (MacCallum and Bradley, 1944; Neefe et al., 1944).

In 1965, Blumberg and Alter set out to investigate the immune-reactivity of plasma proteins, without any particular interest in viral hepatitis. During their experiments they discovered a unique protein in the serum of Australian aborigines, which was soon after named "Australian antigen" (Au) (Blumberg

et al., 1965). A few years later it was reported that the Au had a strong link to acute hepatitis and that the antigen in fact was the surface protein of what became known as the hepatitis B virus (HBsAg) (Dane et al., 1970; London et al., 1969; Prince, 1968). Screening of blood and blood products decreased the incidence of post-transfusion and immunisation associated hepatitis dramatically, however about 10% of patient having received HBsAg-negative blood-products were still developing hepatitis, suggesting that the individuals had contracted a non-A, non-B (nA/nB) viral hepatitis (Alter et al., 1975a; Alter et al., 1975b; Feinstone et al., 1975).

1.2 Hepatitis C virus discovery

Extensive research was carried out from the 1950's throughout the 1980's to identify the causative behind nA/nB viral hepatitis, but it wasn't until the late 1980's that Houghton and colleagues at Chiron and Bradley at the Centre for Disease Control and Prevention made a breakthrough (Choo et al., 1989). These researchers performed a "blind" immune-screening of a cDNA expression library previously generated from total plasma DNA and RNA collected from nA/nB viral hepatitis infected chimpanzees with high viral titer (Bradley and Maynard, 1986). The library was screened against serum deriving from a patient with chronic nA/nB viral hepatitis with an unusually high serum alanine aminotransferase (ALT) level and from thousands of cDNA clones they identified and sequenced a single immuno-reactive clone, labelled as 5-1-1 (Choo et al., 1989). This 5-1-1 clone and subsequent overlapping clones encoded a positive single-stranded RNA molecule of about 10,000 nucleotides that was closely related to the viral genomes of the *Flavivridae*

family members (Choo et al., 1991). The virus was named hepatitis C virus (HCV), following the alphabetical labelling system introduced by MacCallum in 1944 (MacCallum and Bradley, 1944) and was classified within a separate genus (*Hepacivirus*) of the *Flavivridae* family.

The origin of HCV is unclear; it could either have originated from a single emergence in humans followed by diversification or by several cross-species transmissions from a currently unidentified source. Recent identification of several HCV-like viruses in different species of wild rodents and bats strongly support the latter hypothesis (Drexler et al., 2013; Kapoor et al., 2013). In addition, a HCV homolog, later named non-primate hepacivirus (NPHV), has been identified in the respiratory tract of domesticated dogs (also called canine hepacivirus (CHV)) (Burbelo et al., 2012; Kapoor et al., 2011) and in plasma from horses (Lyons et al., 2012). Detailed phylogenetic investigation of the human and canine strains suggest that the HCV and CHV clades diverged within the past 500 - 1000 years, well after the domestication of canines. Close contact with dogs (and potentially horses (Lyons et al., 2012)) may therefore have played an important role in the emergence of the human *hepacivirus*, HCV (Kapoor et al., 2011).

1.3 HCV epidemiology

According to WHO 2013, approximately 150 million people are chronically infected with HCV worldwide and there are about 3 - 4 million new cases every year. The sero-prevalence varies for different geographical regions and the

latest estimates predict that approximately 4% of the populations in Central, East, South and Southeast Asia, North Africa and Middle East are infected, whereas 2% of the North America, Western and Central Europe populations are infected (Mohd Hanafiah et al., 2013; Sharma and Feld, 2014). The virus is commonly transmitted directly through blood or blood products and is often associated with injectable therapy and intra-venous drug-use (Chisari, 2005). Another, less considered, route is mother-to-child transmission (MTCT), which occurs in approximately 5% of cases and is the leading route for HCV infection in children (Abdel-Hady et al., 2011; Yeung and Roberts, 2010).

HCV is divided into seven genotypes (1, 2, 3, 4, 5, 6 and 7), which differ by approximately 30 - 35% over the complete viral genome, corresponding to between 22 – 31 amino acids. Each of the six genotypes can further be divided into more closely related subtypes with 20 - 25% genetic differences (Bostan and Mahmood, 2010; Irshad et al., 2010; Simmonds, 1995, 2004). The geographic distribution varies for the genotypes; 1a, 1b and 3a are mainly found in Europe and North America, whereas genotypes 2, 4, 5 and 6 are in Central, East, South and Southeast Asia, North Africa and the Middle East (Bruggmann, 2014; Irshad et al., 2010; McOmish et al., 1994; Simmonds, 1995, 2004).

Most patients are unaware they have been infected and can be asymptomatic for decades until signs of liver damage appear (Sharma and Feld, 2014). Of the people infected about 70% will develop chronic infection, despite the

presence of a humoral and cellular immune response against the virus (Bowen and Walker, 2005a) (see section 1.6). Once chronic infection is established approximately 20% will develop cirrhosis and 2.5% hepatocellular carcinoma (HCC). Both of these conditions are linked with end-stage liver disease and these patients may require liver transplantation (Afdhal, 2004; Hoofnagle, 2002; Sharma and Feld, 2014). Today, HCV related diseases are the leading indication for liver transplantation in the Western world. Following transplantation the new liver graft becomes re-infected by circulating virions in 100% of the cases and the patient has an increased risk of requiring a second liver transplantation. HCV is therefore the biggest burden on the liver transplantation lists (Brown, 2005; Charlton et al., 2004).

1.4 HCV in children

Viral hepatitis is estimated to be the most important cause of liver diseases in children and affects between 0.05% and 5% of children worldwide (Schwimmer and Balistreri, 2000). Since the introduction of blood and blood product screening for HCV in the early 1990s, MTCT has become the leading course for paediatric HCV infection in developed countries (Mok et al., 2005). Dependent on the geographic area, about 0.1 to 8% of pregnant women are infected with HCV (Abdel-Hady et al., 2011; Arshad et al., 2011) and between 5 to 10% of these will transmit the virus to their children (Resti et al., 1998; Yeung and Roberts, 2010). Mothers co-infected with HIV have a 5-10% higher risk of transmitting HCV (Marine-Barjoan et al., 2007; Thomas et al., 1998), but apart from HIV co-infection, associations between transmission rate and host / viral factors has proven controversial. Some studies claim an increased

risk of MTCT when the maternal viral loads are above 10^6 HCV RNA copies per ml (Lin et al., 1994b; Thomas et al., 1998), while others have shown that there is a broad overlap in viral loads between transmitting and non-transmitting mothers (Gibb et al., 2000; Indolfi and Resti, 2009; Steininger et al., 2003). Furthermore, contradictory results have been found for vaginal delivery versus caesarean section (Boxall et al., 2007; Gibb et al., 2000; Inui et al., 2002; Zanetti et al., 1995) and there is currently no guideline for the mode of delivery for HCV positive mothers. In addition, small-scale studies have suggested that genotype 3 and 1 are more frequently transmitted (Gibb et al., 2000; Zuccotti et al., 1995), however investigation of larger cohorts of infected children have suggested that the genotype is irrelevant for the transmission rate (Abdel-Hady et al., 2011).

Spontaneous clearance rate in children is strongly dependent on the mode of infection. Children infected vertically have a clearance rate comparable with adults (25 – 30%), whereas children infected following blood transfusion have 35% – 45% chance of spontaneously clearing the virus (Abdel-Hady et al., 2011; Bortolotti et al., 2008; Iorio et al., 2005). Those children who progress to chronic infection will usually have a milder disease progression than adults, indicated by low serum ALT levels and minimum or mild histological evidence for liver fibrosis (Jonas, 2002; Kage et al., 1997; Murray et al., 2005). This is further supported by the observation that only 0.09% of all liver transplantations performed on children in North America are the result of HCV-related liver damage (Mack et al., 2012).

1.5 Treatment and outcome for HCV infection:

At the time of writing the standard-of-care HCV therapy consists of pegylated Interferon-α (PEG-IFN-α) and ribavirin for up to 48 weeks. This combined treatment will lead to a cure for more than half the patients, but the individual treatment outcome may depend on the infecting virus genotype; patients infected with genotype 1 will clear the virus in 40% to 64% of the cases, whereas those with genotype 2 and 3 infections have a 76% to 86% clearance rate following treatment (Davis et al., 2003; Di Bisceglie and Hoofnagle, 2002; Fried et al., 2002; Lawitz et al., 2013; McHutchison et al., 2009).

In 2009, two independent studies reported an association between HCV genotype 1 treatment-response and polymorphisms on chromosome 19, at locations rs12979860 and rs8099917 (Ge et al., 2009; Tanaka et al., 2009). Interestingly, these sites are in close proximity to the cytokine IFN-γ gene (Interleukin (IL) 28-β) and frequently referred to as the IL28-β genotype. A cytosine at position rs12979860 predicts a higher chance (76%) of achieving sustained viral response (SVR) following anti-viral therapy for 24 weeks, compared to a non-cytosine (35%) (Tanaka et al., 2009). A thymine at position rs8099917 results in SVR for 80% of the patients compared to 35% for the patients with non-thymine genotypes (Ge et al., 2009). The exact mechanism of action for the IL28-β genotypes are not known, however reports have shown that the favourable IL28-β may be associated with higher IL28-β expression (Dill et al., 2011). Although the IL28-β genotype may be a reliable predictor for genotype 1 infected patients, it is less reliable for genotype 3a (Gupta et al., 2014; Mangia et al., 2013) and has no association to treatment response for

genotype 2 patients (Rauch et al., 2010). Researchers have now turned their attention to the intra-hepatic specific gene expression, especially the intra-hepatic interferon stimulated genes (ISGs) as better predictors of treatment response.

ISGs are usually expressed in large quantities just a couple of days after experimental HCV infection of chimpanzees (Bigger et al., 2001) and acutely infected humans (Dill et al., 2012), demonstrating that HCV induces an innate immune response (see section 1.7). About half of chronically infected patients continue to have a high hepatic ISG expression during chronic infection, whereas others have no detectable expression (Chen et al., 2005; Dill et al., 2011; Sarasin-Filipowicz et al., 2008). Administrating PEG-IFNα to patients with high levels of hepatic ISG mRNA will not further increase the expression levels and as a consequence these patients are less likely to respond to treatment than patients with low pre-treatment ISG levels (Chen et al., 2005; Dill et al., 2011; Honda et al., 2010; Sarasin-Filipowicz et al., 2008). Recently Dill *et al.* reported that the combined expression pattern of ISG15, α-inducible protein (IFI) 27 and radical S-adenosyl methionine domain containing (RSAD) 2, together with the HIV-tat interactive protein (HTATIP) 2; a ubiquitously tumour suppressor gene, can accurately predict the treatment outcome to an error rate of 15% (Dill et al., 2011). These 4-gene classifiers were also reported to be a better predictor than the IL28-β genotype (Dill et al., 2011).

To this date, it is not fully understood why some patients have a higher ISG expression then others. Low intra-hepatic ISG expression has previously been associated with the favourable IL28-β genotype (Honda et al., 2010; Urban et al., 2010). In contrast, Dill *et al.* observed an association between the favourable IL28-β genotypes and high intra-hepatic IL28-β mRNA level, which may prime ISG expression (Dill et al., 2011). Other studies have reported correlations between high ISG expression and HCV genotype 1 (Broering et al., 2010; Sarasin-Filipowicz et al., 2008), which may explain why genotype 1 patients are less likely to respond to treatment.

Of note, the cellular origin of ISG expression may be important in predicting treatment outcome. Immunohistochemical staining for ISG15 and IFN-regulated resistance GTP-binding protein (MxA) in needle biopsies prior to treatment has revealed a link between immune-positive Kupffer cells (liver recident macrophages) and treatment response, whereas expression by hepatocytes has been associated with non-responders (Chen et al., 2010; McGilvray et al., 2012).

Researchers are now more than ever trying to find a more successful cure against HCV infection and a new generation of directly acting anti-viral (DAA) treatments are under development and in late stage clinical trials (Asselah and Marcellin, 2011; Scheel and Rice, 2013). These new DAA drugs are designed to directly inhibit the viral encoded NS3-NS4A protease, the multi-functional role of NS5A in viral replication and assembly and the NS5B RNA polymerase

activity (see section 1.12). At the time of writing the NS3-NS4A protease inhibitors Telaprevir, Boceprevir (U.S. Food and Drug Administration (FDA) approved in 2011) and Simeprevir (FDA approved in 2013), together with the NS5B polymerase inhibitor Sofosbuvir (FDA approved in 2013), are the only clinically approved DAAs. The new drugs are currently given in combination with ribavirin and PEG-IFN-α. Triple combination with the first generation DAAs, Telaprevir and Boceprevir resulted in SVR for up to 75% of patients infected with genotype 1 compared to the PEG-IFN/ribavirin therapy (Jacobson et al., 2011; Poordad et al., 2011). Nevertheless, the second generation of DAAs have reported even higher rates of SVR for the difficult to treat genotype 1 patients; triple therapy with Simeprevir showed SVR for 91% of patients following 12 weeks of treatment and combination therapy with Sofosbuvir has been reported to clear the virus for 89% of genotype 1 infected patients (Lawitz and Gane, 2013).

Although the new generation DAAs have promising effect in the clinic, HCV escape mutations have been detected for the three NS3-NS4A protease inhibitors *in vivo* and *in vitro* (Fried et al., 2013; Imhof and Simmonds, 2011; Lenz et al., 2013; Wu et al., 2013). HCV escape mutants have also been detected for the NS5B polymerase inhibitor Sofosbuvir *In vivo* (Tong et al., 2014).

One strategy to avoid HCV escape mutants is to design drugs against host factors instead of viral. One of the drugs that are currently being evaluated in

the clinic is the small molecule SR-BI antagonist ITX5061, which has shown to inhibit HCV entry *in vitro* (Syder et al., 2011). Recent reports have found that this drug is well tolerated, but that administrating the antagonist alone does not have a significant reduction in viral load (Sulkowski et al., 2013). Instead ITX5061 may be a strong candidate in the liver transplantation setting to prevent reinfection of the allograft (Rowe, unpublished data).

The standard-of-care treatment for children between 3 and 17 years of age is PEG-IFNα-2a/b in combination with ribavirin (Wirth, 2012). The treatment will continue for 24 to 48 weeks and results in SVR in more than 90% of the genotype 2 and 3 cases and between 44 and 57% of genotype 1a infected children (Baker et al., 2007; Ghany et al., 2009; Jara et al., 2008; Schwarz et al., 2011; Wirth et al., 2005; Wirth et al., 2010). The use of different dose concentrations and treatment times makes it difficult to directly compare adult and child response rates. However, despite overlapping values, HCV genotype 2 and 3 infected children do tend to have a higher response rate than adults, whereas the SVR rate is comparable between genotype 1 infected children and adults (Davis et al., 2003; Di Bisceglie and Hoofnagle, 2002; Fried et al., 2002; Lawitz et al., 2013; McHutchison et al., 2009). Predicting the response rate for children is still a largely unexplored area and so far only the pre-treatment serum/plasma viral load seems to be a reliable predictor (low viraemia being an indicator of a higher chance of SVR (Sokal et al., 2010; Wirth et al., 2010)). Interestingly the IL28-β genotype, which is predictive for genotype 1 infected adults, does not have any predictive value for children (Domagalski et al., 2013; Komatsu et al., 2013). The notion that intra-hepatic

ISG expression prior to treatment can be predictive for adults makes it an attractive predictor for paediatric patients as well, however to date no one has investigated this area and our study will therefore be the first to explore hepatic ISG expression in HCV infected children.

1.6 Immunity

During the initial infection the plasma viral load will usually stay below the level of detection for 7 to 21 days. Over the following 8.2 to 21 days the level of detectable plasma HCV RNA accelerate exponentially, with a HCV RNA doubling-time of 7.4 to 17 hours, resulting in a viral peak of 10^5 - 10^7 International Units (IU)/ml blood (Busch, 2001; Glynn et al., 2005; Ribeiro et al., 2012). After the initial viremic burst the viral titer stays constant for the following 4-8 weeks and then decline with the onset of an adaptive cellular immune response (Farci et al., 1991; Heim, 2013a; Major et al., 2004; Thimme et al., 2001). Towards the end of the high viral titer phase, anti-HCV antibodies can be detected in the serum, however this sero-conversion is not required to clear the acute infection (Farci et al., 1991; Thimme et al., 2001) (Figure 1.1). A clear example of this is seen for immune-compromised patients who spontaneously resolve HCV infection (Christie et al., 1997). Furthermore, presence of antibodies (Ab) against the hypervariable region 1 (HVR1) in the 5'UTR of E2 (section 1.9.2) and anti-E2 sero-conversion has previously been linked with persistent infection in chimpanzees and humans (Major et al., 2004; Prince et al., 1999). However, the function of the detected Abs was not evaluated in these studies and it is therefore impossible to rule out that they

potentially were non-functional. Other studies have found that early induction of cross-reactive anti-E2 nAbs is associated with viral clearance in humans (Dowd et al., 2009; Osburn et al., 2014; Pestka et al., 2007). These data suggest that that the mere presence of Abs does not eliminate the virus, but that the width and affinity of the Ab spectrum may be important.

During the humoral immune response, nAbs against the HCV antigens are constantly being produced. However the high viral mutation rate means that the virus may evolve escape mutants that are not recognised by nAbs, leading to viral persistence (Dowd et al., 2009; Kato et al., 1994; von Hahn et al., 2007). The presence of nAbs may form a selective pressure on the HCV population complexity and indirectly influence the evolutionary path (Farci et al., 2000; Ray et al., 1999) (see section 1.14).

An early strong adaptive immune response is important for successful clearance of the HCV infection. Studies have show that patients with a broadly directed, vigorous and sustained activation of the CD8+ T cells, together with a strong intra-hepatic IFN-γ expression tend to have a higher clearance rate (Billerbeck et al., 2013; Gruner et al., 2000; Lechner et al., 2000; Thimme et al., 2001). However, the CD8+ T cells require the CD4+ T cells for the initial HCV-priming and a broad-spectrum CD4+ T cell population is associated with an increased chance of clearing the virus as well (Day et al., 2002; Grakoui et al., 2003; Schulze zur Wiesch et al., 2005). The initial detection of HCV specific CD8+ and CD4+ T cells is strongly linked to a decline in viral load within the

first 4-8 weeks of infection, and to a significant increase in liver-associated serum ALT, which is associated with immune-mediated liver injury (Figure 1.1). The serum ALT level will usually decline after a few months and for some patients the level will return to the normal range (Hoofnagle, 2002).

For patients with persistent infection the number and function of responding CD8+ T will often decline. This reduction in T cells has previously been linked with an increase in inhibitor receptors such as programmed death-1 (PD-1) (Cox et al., 2005; Sumida et al., 2013). The increase in PD-1 may be the result of CD8+ T cell exhaustion, which is often the feature of infections with early and high replication rate and persistent high viral burden (Moskophidis et al., 1993; Wherry et al., 2003). T cell exhaustion is believed to increase the risk of viral persistence and may in part explain why 70% of infected individuals do not clear the virus (McMahan et al., 2010). In addition, identification of epitope escape HCV variants may also play a part in the progression to chronic infection (Bowen and Walker, 2005b; Neumann-Haefelin et al., 2008).

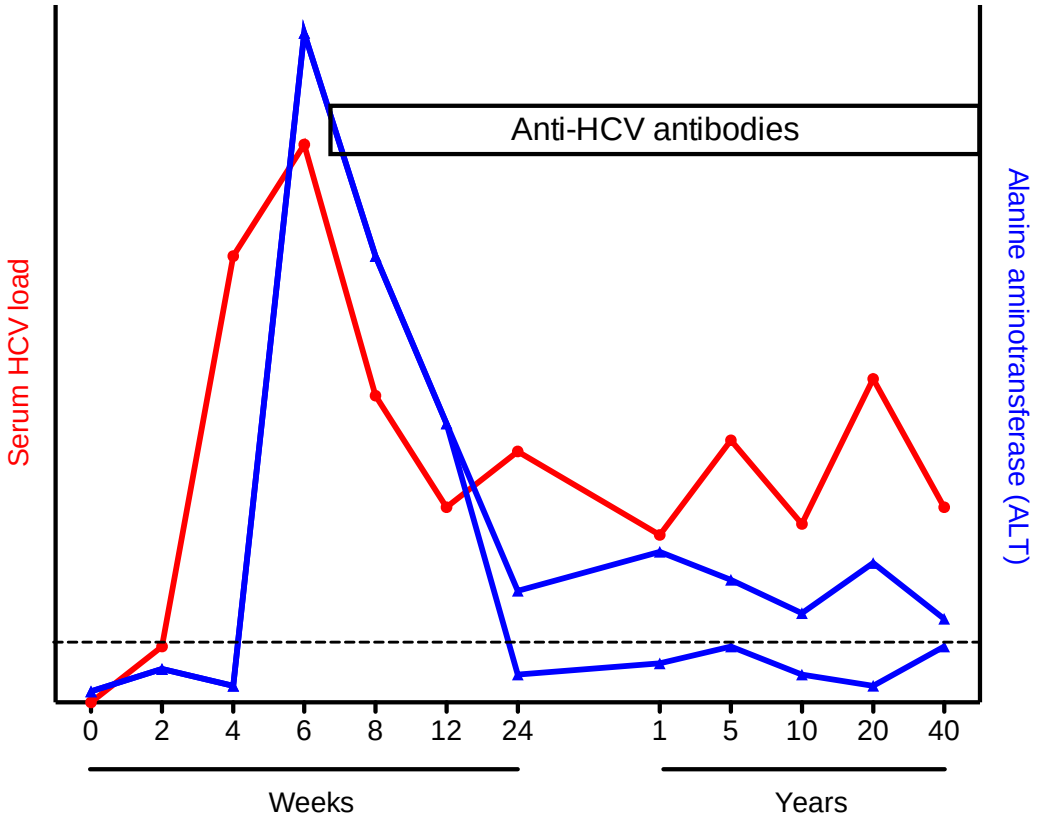


Figure 1.1 Natural course of HCV chronic infection.

In the early stage of HCV infection the serum viral load (red) is un-detectable. After 7-21 days the HCV serum level accelerate up to 10^7 IU/ml, but declines with the onset of the cellular and humoral immune response (anti-HCV antibodies are indicated with the transparent box). During the chronic infection the viral kinetics remain relatively stable. Detection of serum alanine aminotransferase (ALT, blue) is associated with the high hepatic infiltrate of CD8+ T cells. Following the initial burst of serum ALT the level will decline and in some cases return to normal (lower blue line), whilst for other patients the ALT level will remain elevated (top blue line). The dashed line indicates the detectable level of serum HCV RNA and the upper limited of normal serum ALT. The graph is made from modified data in Heim et al. 2013 (Heim, 2013b) and Hoofnagle et al. 2002 (Hoofnagle, 2002).

1.7 Interferon and Interferon stimulated genes

The type I interferons (IFN-α and IFN-β) are usually detected during the early stages of infection regardless of infection outcome (Bigger et al., 2001; Thimme et al., 2001) and have been found to correlate with viral load during the acute infection of experimentally inoculated chimpanzees (Su et al., 2002). The main source of IFN-α and IFN-β is unknown. However a recent study have reported ISG mRNAs in association with HCV RNA positive cells *in vivo* (Wieland et al., 2013), suggesting that the hepatocytes may be the main site of type I IFN expression. Alternatively it has been suggested that the type I IFN may derive from plasmacytoid dendritic cells (pDCs) located in close proximity to infected hepatocytes (Dreux et al., 2012).

Activation of type I and type III IFN expression in infected hepatocytes is a result of cellular viral sensors reacting to the HCV pathogen-associated molecular patterns (PAMPs). The two major pathways for IFN production are the retinoic acid inducible gene-I (RIG-I) and Toll-like receptor (TLR) cascades. Both pathways leads to the downstream activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB) and the IFN regulatory factor (IRF) 3 and IRF7. The activated transcription enhancers translocate to the nucleus and bind the interferon-stimulated response element (ISRE) in the IFN promotor to stimulate type I and III IFN expression (Heim, 2013a, b; Rosen, 2013; Sharma and Feld, 2014; Thimme et al., 2012). More detailed descriptions of the pathways can be found in Figure 1.2. The secreted type I and III IFN bind to their respective receptors on the membrane surfaces of

infected cell and naïve neighbouring cells, leading to the activation of the Jak-STAT pathway and downstream ISG expression (Figure 1.2) (Akira and Takeda, 2004; Heim, 2013a).

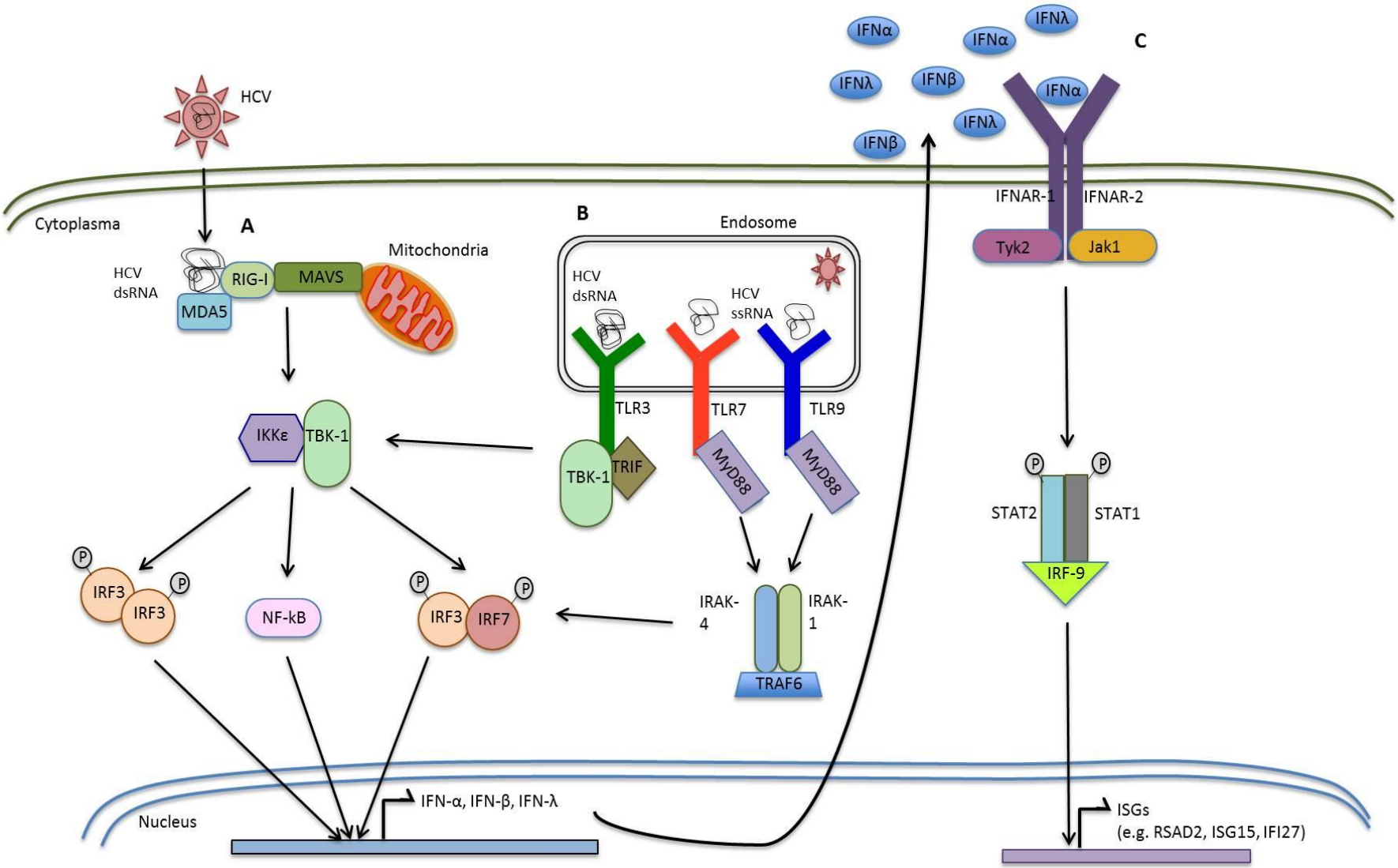

HCV
Cytoplasma
A
HCV
dsRNA
RIG-I
MAVS
MDA5
Mitochondria
IKKε
TBK-1
IRF3
IRF3
NF-kB
IRF3
IRF7
B
Endosome
HCV
dsRNA
HCV
ssRNA
TLR3
TLR7
TLR9
TBK-1
TRIF
MyD88
MyD88
IRAK-
4
IRAK-
1
TRAF6
C
IFNα
IFNα
IFNλ
IFNλ
IFNβ
IFNα
IFNλ
IFNβ
IFNAR-1
IFNAR-2
Tyk2
Jak1
STAT2
STAT1
IRF-9
Nucleus
IFN-α, IFN-β, IFN-λ
ISGs
(e.g. RSAD2, ISG15, IFI27)

Figure 1.2 Activation of the innate immune response.

A. During the RIG-I pathway the cytosolic HCV double stranded (ds) RNA is recognised by the RNA helicases RIG-I and melanoma differentiation-associated protein 5 (Mda5). These activated helicases will bind and activate the mitochondrial anti-viral signalling molecule (MAVS) on the mitochondria, which in turn leads to the activation of the IκB kinase (IKK)-related kinases (IKKε) and TANK-binding kinase-1 (TBK-1). This activation will result in the phosphorylation of IFN regulatory factor (IRF) 3 and IRF7, leading to homo-and hetero-dimerization and the activation of the transcription regulator nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB). These transcription factors will then translocate to the nucleus and stimulate the expressions of type I IFN-α, -β and type III IFN-λ. **B.** In the TLR pathway the TLR3, TLR7 and TLR9 endosome residential proteins recognise the intruding virus by the presence of viral dsRNA or single stranded (ss) RNA. Following stimulation, TLR3 initiates a signalling cascade through the TRIF proteins and the kinase TBK1, which translocate and form an active complex with IKKε. TLR7 and TLR9 signal through the MyD88 protein, which in turn leads to the association of TNF receptor associated factor 6 (TRAF6) and the Interleukin-1 receptor-associated kinases 4 and 1 (IRAK4/1). Activation of the IRAK4/1-TRAF6 complex results in the phosphorylation of IRF3 and 7, which lead to dimerization and activation of the IFN transcription. **C.** The secreted IFN-α, -β and -λ will bind their receptive receptors on the cell surface, here shown for IFN-α. Following ligand binding the Janus Kinase (Jak) 1 and tyrosine kinase (Tyk) 2 will initiate phosphorylation of the intra-cellular IFN-α receptor (IFNAR) domains, which in turn recruits the signal transducers and activators of transcription (STATs). The phosphorylated STAT1 and STAT2 will form the IFN-stimulated gene factor 3 (ISGF3) complex together with IRF9, which translocate to the nucleus and initiates the expression of interferon stimulated genes (ISGs), e.g. ISG15, IFI27 and RSAD2 (Akira and Takeda, 2004; Heim, 2013a, b; Rosen, 2013; Sharma and Feld, 2014; Thimme et al., 2012).

The ISG proteins can have either anti-viral or pro-viral properties (Schoggins and Rice, 2011). Among the anti-viral effectors are RSAD2 (also called Viperin) and IFI27. RSAD2 is located on the endoplasmic reticulum (ER) and associates with the HCV replication complex on the lipid droplets (see section 1.12.2). Here RSAD2 is capable of inhibiting the HCV replication by interacting with the viral protein NS5A (Helbig et al., 2011; Hinson and Cresswell, 2009). Another example of an anti-viral protein is IFI27, which has been shown to limit HCV replication *in vitro*. The exact method of anti-viral action remains unknown, but the evidence suggests that IFI27 may regulate HCV replication indirectly by being involved in cellular anti-viral responses (Itsui et al., 2006).

One of the most intensely studied ISGs is ISG15, however scientists are still discussing its function during viral infection. ISG15 functions as ubiquitin-like protein modifier and conjugates its target proteins through the action of the E1 activating protein Ube1L (Yuan and Krug, 2001), the E2 conjugating enzyme UbcH8 (Yuan and Krug, 2001) and E3 ligase HERC5 (Wong et al., 2006). ISG can be conjugated (ISGylation) to more than 150 cellular proteins and is involved in various cellular functions including signalling transduction and immune/stress responses (Kim and Zhang, 2003; Zhao et al., 2005). The effect of ISGylation during viral infection has proven to be controversial. Several studies have reported an anti-viral effect for DNA and RNA viruses. For example ISGylation of the HIV Gag protein *in vitro* will inhibit its interaction with host Tsg101 protein and block viral release (Okumura et al., 2006). Moreover, ISG15 knock-out mice have been found to be more susceptible to

influenza Sinbis virus and herpes simplex virus 1 infections (Lenschow et al., 2007). Controversially, ISG15 seems to have the opposite effect during HCV infection. For instance, over-expression of ISG15 has been reported to promote replication of HCV strain J6/JFH in hepatoma cell lines, whereas silencing led to decreased viral replication (Broering et al., 2010; Chua et al., 2009), suggesting a positive regulatory role in the viral life cycle. In addition, intra-hepatic ISG15 has been reported up-regulated in HCV infected patients who do not respond to treatment compared to those who do (Broering et al., 2010; Chen et al., 2005; Dill et al., 2011; Urban et al., 2010). The exact mechanism for this pro-viral effect is unknown, but evidence suggests that ISG15 may control the RIG-I activity through ISGylation, (Arnaud et al., 2011). Conjugation with ISG15 may inhibit the HCV dsRNA inducted activation of RIG-I, leading to an reduction in IFNβ expression and an increase in the accumulation of the HCV structural protein NS3 (Arnaud et al., 2011).

Acute HCV infection leads to an activation of IFN and subsequently ISG expression (Bigger et al., 2001; Su et al., 2002; Thimme et al., 2001), however in 70% of cases the virus is not cleared and persists. This paradox can partly be explained by the ability of the viral encoded NS3/4A to manipulate the RIG-I and TLR pathways in infected cells. This HCV protease is capably of blocking the phosphorylation of IRF3 (Foy et al., 2003) and can cleave and release the mitochondrial antiviral-signalling protein (MAVS) from the mitochondria (Bellecave et al., 2010; Meylan et al., 2005). Both actions lead to inactivation of the target protein and inhibition of the RIG-I pathway. In addition NS3/4A

has been found to cleave the adaptor protein TRIF, leading to abolishment of the TLR3-pathway in infected cells (Li et al., 2005). It is not known to what extent HCV is capable of inhibiting the PAMP sensing pathways, however recent studies using new methods for ISG mRNA detection reveals a correlation between HCV RNA and ISG mRNA expression, suggesting that that the infected cells are the major source of ISG expression and that HCV inhibition of the sensing pathways therefore only have a minor role (Wieland et al., 2013).

1.8 HCV genome

The viral genome is a 9.6 kilobase (kb) long uncapped positive single RNA strand. The 5'- and 3'-untranslated regions (UTRs) of the genome consist of control elements required for translation and replication. The 5'-UTR contains an internal ribosome entry site (IRES), which binds the 40S ribosomal subunit and initiates translation of the virus RNA-strand in a cap-independent manner (Penin et al., 2004). The 3'UTR include a poly(U/UC) tract, which is believed to be important for viral replication and translation of the viral genome (Friebe and Bartenschlager, 2002; Friebe et al., 2005; Tuplin et al., 2012). The two UTRs are separated by a single uninterrupted open reading frame (ORF), which encodes a single polyprotein of approximately 3,011 amino acids (aa). After co- and post-translational modifications, this polyprotein is processed into 3 structural (core, E1, E2) and 7 non-structural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B) (reviewed in (Ashfaq et al., 2011; Lindenbach and

Rice, 2005). A cartoon of the HCV genome, polyprotein and cleaved components can be viewed in Figure 1.3.

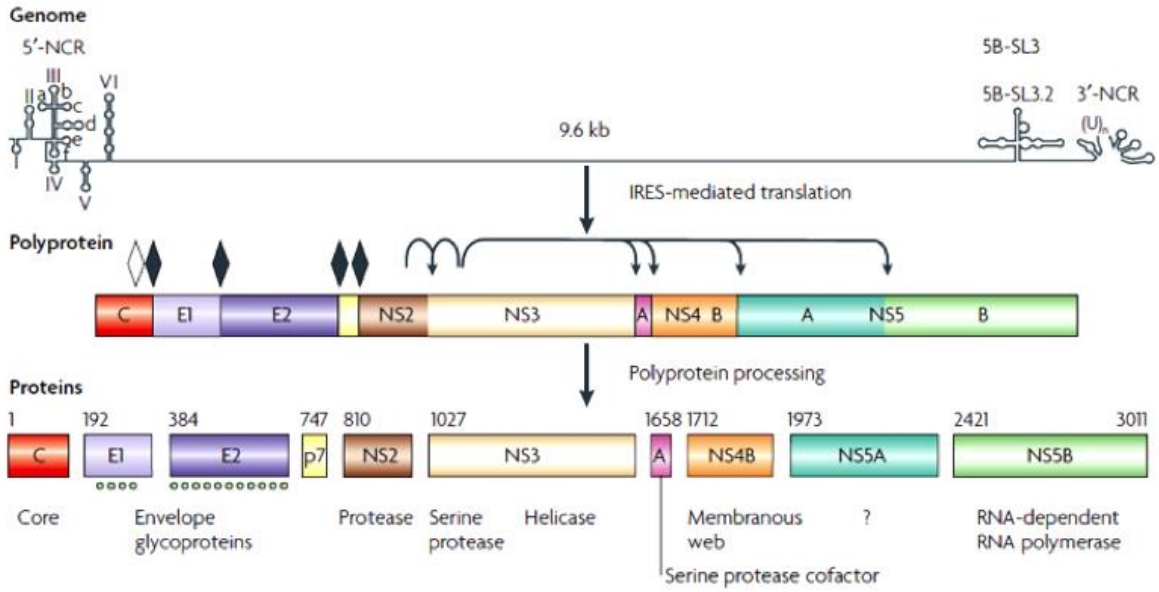


Figure 1.3 HCV genome and proteins.

The 9.6kb long ssRNA genome contains several secondary structures in both the 5' and 3'-non-coding regions (NCR) or UTR. Following transcription and translation of the HCV polyprotein, individual proteins become separated by host encoded ER signalling peptidases and peptide peptidases (rhombi), and by viral encoded *cis*- or *trans*-acting proteases (arrows). The key functions of the 3 structural genes (core, E1, E2) and 6 of the 7 non-structural proteins (p7, NS2, NS3, NS4A, NS4Band NS5B) are listed underneath their sequence location. The exact mechanism of NS5A is not known (indicated by a question mark). It has been suggested that the protein is involved in viral assemble by interaction with core and the ssRNA genome. Green dots underneath the E1 and E2 proteins indicate glycosylation sites (Kim and Chang, 2013; Moradpour et al., 2007; Ogawa et al., 2009; Suzuki et al., 2007). Illustration from (Moradpour et al., 2007).

1.9 The HCV particle and the viral structural proteins

HCV particles comprise the structural proteins core, E1 and E2 (Figure 1.4). The 191aa core protein forms the icosahedral capsid around the viral genome. This highly conserved non-mature protein can be divided into 3 domains according to their hydrophobicity. Domain I (aa 1-117) is hydrophilic, and thought to bind and stabilise the RNA genome, domain II (aa 118-171) mediates the interaction with the lipid droplets in the viral replication complex (section 1.12.2) and finally domain III (aa 172-191) correspond to the signal sequence for E1, which leads to E1 translocation to the ER lumen (McLauchlan, 2000; Rouille et al., 2006; Santolini et al., 1994). Following insertion of E1 into the ER membrane, signal peptidases will release the signalling sequence from the envelope protein. The signalling sequence is further processed by a signal peptide peptidase located in the ER membrane (Pene et al., 2009) which leads to the release of the mature core from the signalling sequence (McLauchlan et al., 2002). A high-resolution structure of core and its topology in the capsid remains to be discovered.

The nucleocapsid is surrounded by a cell-derived lipid envelope bilayer. The viral type-I transmembrane proteins E1 and E2 are associated with the envelope where they form non-covalent hetero-dimers, which stabilise the viral particle and interact with viral entry factors on permissive cells (Bartenschlager et al., 2011). Density gradient analyses of infected patient serum have shown that the HCV particles may have variable densities and that only the low-density particles are infectious (Bradley et al., 1991; Hijikata et al., 1993). It

was later found that the infectious fraction contains high content of triglycerol, and the apolipoproteins (Apo) B and E, which leads to the current model that the HCV particle resembles the very-low-density lipid-proteins (VLDL) (Andre et al., 2002; Hishiki et al., 2010).

It has previously been reported that the E2-p7 cleavage by the ER signalling peptidases is incomplete, resulting in both E2 and E2-p7 species (Lin et al., 1994a). Later investigation of HCV pseudo-particles (HCVpp) (see section 1.10) generated from E1E2p7 transient expressing 293Ts (instead of only E1E2), revealed that p7 may become incorporated into the HCV envelope, suggesting that p7 may function as a structural protein as well. The incorporation, however, had no effect on the infectivity of the HCVpp particle (Flint et al., 2004). The role of p7 is still not fully understood and most reports suggests that p7 may have both structural and non-structural features during the HCV life cycle (Khaliq et al., 2011). A description of the non-structural feature can be found in section 1.12.3.

The viral envelope proteins E1 and E2 are type-I transmembrane proteins, which consist of three domains; (*i*) the transmembrane domain, which is responsible of anchoring the proteins to the viral particle and to the ER during protein maturation, (*ii*) the membrane proximal domain, involved in dimerization and (*iii*) the ecto-domain which interacts with entry factors on the

cell surface (Ashfaq et al., 2011; Dubuisson, 2007; Suzuki et al., 2007). A detailed description of the envelope proteins can be found below.

1.9.1 Envelope protein 1 (E1)

The E1 protein is encoded by amino residues 192 – 383 (length 181aa) and contains five sites for the addition of asparagine-linked glycans (N-linked glycosylation) labelled E1N1 to E1N5 (Meunier et al., 1999). Site-directed mutation studies of the glycosylation sites in the HCVpp envelope proteins have revealed that E1N1 and E1N4 are important for dimerization of E1E2 in the viral particle and that E1N1, E1N2 and E1N4 are required for successful HCVpp entry (Goffard et al., 2005). Unlike observations for E2, glycosylation of E1 does not seem to have an impact on the secondary structure of the protein (Lorent et al., 2008). The glycosylation sites may instead have an effect on the di-sulphide bonds and interactions with E2 in the particle. Furthermore, E1 has proven to be far less antigenic than E2, with only 10 - 50% of HCV chronically infected patients having reactivity against *E.coli* expressed E1, compared to 70%-73% for E2 (Hussy et al., 1997; Lorenzo et al., 2000).

1.9.2 Envelope protein 2 (E2)

The E2 protein is encoded by amino residues 384 -746 (length 362aa). Over the years, E2 has been the main focus of many research studies because of its high antigenicity and its direct interaction with entry factors on the cell surfaces (see section1.11). However, it was not until recently that the crystal

structure of the E2 core domain was published (Kong et al., 2013). In this study Kong *et al.* strategically truncated and/or replaced regions of E2 before co-crystallisation with a high affinity antibody, to achieve an E2 crystal structure (Kong et al., 2013). The crystal structure revealed a general globular shape of E2 with multiple loops, short helices and beta-sheets. Furthermore, the E2 surface featured a highly glycosylated face, a neutralising face, containing strain-restricted and cross-reactive epitopes, and several regions, which are believed to be concealed in the viral protein (Ball et al., 2014; Fauvelle et al., 2014). The globular E2 structure reported by Kong *et al*. contradicts the previously suggested model of the E2 having a class II membrane fusion protein structure (Keck et al., 2005; Krey et al., 2010).

The E2 ecto-domain encodes three highly variable regions, termed hypervariable region 1 (HVR1, residues 384-410) hypervariable region 2 (HVR2, residues 460-485) and the intergenotopic variable region (igVR, residues 570 - 580). Additional HVRs have been identified exclusively in genotype 3a viruses; HVR495 and HVR575 (Humphreys et al., 2009). The most variable of these is HVR1, which appears to be highly tolerant to aa substitutions (Cocquerel et al., 2000), however there seems to be a strong selective pressure to maintain a length of 27aa (Casino et al., 1999). This region is highly antigenic and contains key neutralising epitopes (Bankwitz et al., 2010; Farci et al., 1996; Kato et al., 1994; Manzin et al., 1998; Ray et al., 1999; von Hahn et al., 2007). Despite the antigenic properties of HVR1, studies have shown that HCV cell culture particles (section 1.10) lacking HVR1

are more prone to neutralisation by both E2-specific monoclonal antibodies and patient serum (Bankwitz et al., 2010), which suggests that HVR1 may shield neutralising epitopes outside HVR1 in the E2 configuration. Further studies have shown that only one out of five anti-HVR1 antibodies were able to neutralize HCVpp entry, whereas antibodies against E2 aa regions outside HVR1 (412-423aa, 432-443aa and 436-447aa) completely neutralized infection (Hsu et al., 2003a). These observations suggest that the Ab epitopes identified in HVR1 may not lead to neutralisation. Instead it has been suggested that HVR1 may function as a immune decoy, which lead the attention away from the functional nAb sites in E2 (Ray et al., 1999).

The HVR1 and part of HVR2 are located in the receptor-binding domain (RBD: residues 384 – 661 aa), however neither region has been reported to directly bind receptor CD81, an essential receptor for HCV entry. Instead, deletion of the HVR1 region may result in increased binding of soluble E2 (sE2) to cell lines expressing CD81. Furthermore, studies have shown that deleting or replacing the HVR1 region in the HCVpp lead to increased affinity to CD81 (Bankwitz et al., 2010; Roccasecca et al., 2003) and neutralisation by anti-CD81 antibodies (Bankwitz et al., 2010), suggesting that HVR1 may act to shield the CD81 binding region in E2 (Roccasecca et al., 2003). In support of this, mutagenesis analyses have revealed that the direct interaction between CD81 and E2 residues are found outside HVR1 and HVR2 at positions W420, W437, L438, L441, F442, Y527, W529, G530, D535 and at a region spanning

613 – 618aa (Drummer et al., 2006; Owsianka et al., 2006; Roccasecca et al., 2003).

Although it has been reported that HVR1-deleted HCV particles may have increased recognition of CD81, deletion of the region has also been reported to decrease the viral entry (Bankwitz et al., 2010; McCaffrey et al., 2011; Prentoe et al., 2014). Interestingly, viral particles generated from HCV genomes lacking HVR1 have increased particle density, suggesting alteration of the lipid composition in the viral particle (Bankwitz et al., 2010). Recent, reports have further shown that HVR1-deleted HCV cell culture particles (section 1.10) have decreased dependency of low density lipoprotein receptor (LDLr) and SR-BI during entry (Bankwitz et al., 2010; McCaffrey et al., 2011; Prentoe et al., 2014), however the receptor independency were not associated with variable incorporation of apolipoprotein E (ApoE) into the envelope. It is therefore thought that HVR1 instead directly or indirectly binds to LDLr and SR-BI during the early stages of viral entry (Prentoe et al., 2014).

The HCV E2 envelope protein contains eleven sites for N-linked glycosylation (E2N1-E2N11). Of these, ten are highly conserved between genotypes (Goffard et al., 2005). Site-directed mutagenesis studies in the HCVpp systems have revealed that E2N6, E2N8 and E2N10 are involved in viral assembly (Goffard et al., 2005; Helle et al., 2007). Goffard and colleagues also reported that the E2N8 and E2N10 are involved in binding CD81 and

mediating virus entry (Goffard et al., 2005). Interestingly Helle *et al.* found that mutation of the E2N1, E2N6 and E2N11 significantly increased the sensitivity of virus neutralisation by anti-E2 mono- and polyclonal antibodies and increased the affinity for CD81 both in the HCVpp and HCVcc system (Helle et al., 2007; Helle et al., 2010). The presence of glycans at these 3 sites is believed to result in conformational changes that will shield important neutralising sites and the CD81 binding site (Helle et al., 2007; Helle et al., 2010). Similar observations have been made for the HIV gp120 protein, where the heavy glycosylation of the envelope protein provides a “glycan shield” making the protein completely immunological silent (Wei et al., 2003).

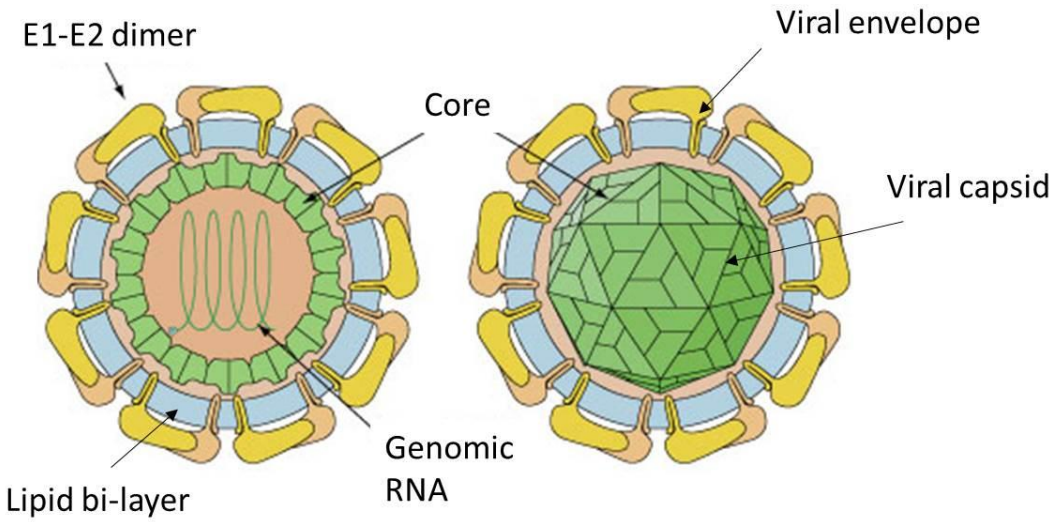


Figure 1.4 HCV particle.

The viral particle consists of the viral genome, the capsid and the envelope. The icosahedral capsid is made up of the viral encoded core protein, which binds and stabilises the RNA genome. The nucleocapsid is surrounded by a cell-derived lipid bi-layer (also called the envelope) into which the envelope proteins E1 and E2 are incorporated. These proteins form non-covalent hetero-dimers that stabilise the viral particle and interact with viral entry factors on the permissive cells (Bartenschlager et al., 2011). Illustration from *http://viralzone.expasy.org*

1.10 Model systems for studying HCV life cycle

Since the discovery of HCV in 1989 different *in vitro* models have been developed for studying the HCV life cycle. The first *in vitro* system of significance was the HCV replicon system, which was developed in 1999 by Lohmann and his colleagues (Lohmann et al., 1999). In the original HCV replicon the structural genes were replaced by a neomycin phosphotransferase gene followed by an IRES sequence deriving from the encephalomyocarditis virus (EMCV). This IRES lead to the synthesis of HCV non-structural proteins required for HCV replication, whereas the HCV IRES mediated the translation of the selective marker. In this setting only cells supporting HCV replication would express sufficient amount of neomycin phosphotransferase to become resistant against G418 treatment and thereby allow the studies of the HCV RNA replication (Lohmann and Bartenschlager, 2014; Lohmann et al., 1999).

In 2003 the HCV pseudo-particle (HCVpp) system was developed, which provided an important tool for studying viral entry (Bartosch et al., 2003; Drummer et al., 2003; Hsu et al., 2003b). The HCVpp consist of an envelope made up of HCV E1 and E2 glycoproteins, which surrounds a capsid consisting of retro-viral core proteins. The capsid contains a packed lenti-viral genome with a luciferase gene insertion. Following successful entry, the HCVpp genetic material becomes integrated into the host cell genome, wherefrom the reporter gene is expressed. The level of entry is evaluated indirectly by quantifying the activity of the luciferase in the cell lysate. The

HCVpp envelope is the only part of the particle that has derived from an authentic HCV particle and the HCVpp system can therefore only be used for studying HCV entry (Bartosch et al., 2003; Drummer et al., 2003; Hsu et al., 2003b).

In 2005 another milestone was reached in HCV research, when an authentic HCV particle was reported to replicate and release infectious particles within a cell culture system (HCVcc) (Kato and Wakita, 2005; Lindenbach et al., 2005; Zhong et al., 2005). The strain was generated by transfecting HCV genotype 2a clonal RNA from a Japanese patient with fulminant hepatitis-1 into producer cell lines and the strain is therefore termed Japanese fulminant hepatitis-1 (JFH1) (Lindenbach et al., 2005). The HCVcc system has been further developed and chimeric JFH1 strains have been created by replacing the core, E1, E2, p7 and NS2 of JFH1 with genotype specific sequences for genotype 1a, 1b, 3a, 4a and 5a (Gottwein et al., 2007; Jensen et al., 2008; Pietschmann et al., 2006; Scheel et al., 2008; Yi et al., 2007). The advantage of the HCVcc system is that it allows the full HCV life cycle to be studied in hepatocyte derived cell lines and cultured primary hepatocytes (Ploss et al., 2010).

1.11 HCV entry factors and the internalisation process

The complex process of HCV entry into hepatocytes is thought to be initiated by the interaction of low-affinity attachment factors, such as heparin sulphate glycosaminoglycans and low-density lipoprotein receptors with the E2 protein,

ApoB and ApoE in the viral envelope (Barth et al., 2003; Hishiki et al., 2010; Owen et al., 2009). Internalisation of the HCV particle requires a series of interactions with the four essential HCV entry receptors and entry factors; The tetraspanin CD81 (Flint et al., 1999; Pileri et al., 1998), scavenger receptor class B type I (SR-BI) (Scarselli et al., 2002) and the tight-junction proteins Claudin-1 (CLDN1) (Evans et al., 2007) and Occludin (OCLN) (Ploss et al., 2009) (Figure 1.5). A description of each entry factor can be found below.

1.11.1 CD81

CD81 is a member of the tetraspanin superfamily and is ubiquitously expressed on most nucleated cells. The protein contains 4 transmembrane domains, which create 2 extracellular loops defined as EC1 and EC2 (Hemler, 2005). CD81 was the first HCV entry factor to be identified and was discovered by screening a human cDNA expression library with soluble E2 (sE2) (Pileri et al., 1998). Further proof of the role of CD81 as an entry factor came when the non-permissive hepatoblastoma cell line (HepG2) was found to support infection following the exogenous expression of CD81 (McKeating et al., 2004; Zhang et al., 2004). In addition, HCVpp entry can be inhibited by anti-CD81 monoclonal antibodies, by si-RNA mediated CD81 silencing (Zhang et al., 2004) and by soluble forms of the CD81 EC2 domain (Bartosch et al., 2003; Hsu et al., 2003b).

1.11.2 Scavenger receptor B type I

Scavenger receptor B type I (SR-BI) is primarily expressed in liver and steroidogenic tissues, where it selectively takes up the cholesterol by binding high cholesterol ligands such as high-density lipoprotein (Acton et al., 1996). SR-BI was identified as a putative HCV receptor as sE2 was shown to bind the CD81-negative HepG2 (Scarselli et al., 2002). Furthermore, siRNA silencing of SR-BI reduces HCVcc infection *in vitro* (Zeisel et al., 2007), while over-expression of SR-BI enhances virus infection (Grove et al., 2007). E2 HVR1 is thought to interact with SR-BI since the deletion of the HVR impaired sE2-SR-BI binding (Scarselli et al., 2002). In addition, deletion of HVR1 has been reported to decrease SR-BI dependency of the virus, but increase the affinity for E2-CD81 interaction (Bankwitz et al., 2010; McCaffrey et al., 2011; Prentoe et al., 2014). This observation suggests that HVR1 may interact directly or indirectly with SR-BI, leading to conformational change of E2 exposing the CD81 binding site (Bankwitz et al., 2010).

1.11.3 Claudin-1

Claudin-1 (CLDN1) plays a key role in the formation of cellular tight junctions (Krause et al., 2008). The entry factor was identified as a HCV entry co-receptor when introducing the cDNA library generated from a highly permissive huh-7.5 hepatoma cell line into the non-permissive 293T cell line (Evans et al., 2007). Residues required for HCV entry have been mapped to the first extracellular loop of CLDN1, particularly to the highly conserved CLDN motif W_{30}-GLW_{51}-C_{54}-C_{64} (Cukierman et al., 2009). Fluorescent resonance

energy transfer (FRET) studies from the McKeating laboratory have shown that CLDN1 associates with CD81 to form a complex on the cell surface (Harris et al., 2010; Harris et al., 2008). This association appears essential for endocytosis of the receptor complex and for later fusion with Rab5a expressing endosomes (Farquhar et al., 2012).

1.11.4 Occludin

Occludin (OCLN) is an additional tight junction protein, which is required for HCV entry. Ploss *et al.,* confirmed OCLN as a HCV receptor via screening of a huh7.5 cDNA library in mouse embryonic fibroblasts. Briefly, transduction of human OCLN into non-permissive murine fibroblasts transduced to express human CD81, SRBI and CLDN1 enabled HCVpp entry. Furthermore, silencing OCLN would abolish the entry of both HCVpp and HCVcc (Ploss et al., 2009). The mechanism of HCV-OCLN interaction remains uncertain. Although OCLN has been reported to interact with the E2 glycoprotein, it is unknown whether this interaction is direct, or indirect via CD81/CLND1 (Meredith et al., 2012).

1.11.5 Additional entry receptors

Recently, the transferrin receptor (TfR) 1 (Martin and Uprichard, 2013), Niemann-Pick C1-like 1 cholesterol absorption receptor (NPC1L1) (Sainz et al., 2012), the epidermal growth factor receptor (EGFR) and ephrin receptor A2 (EphA2) (Lupberger et al., 2011) have been reported as novel HCV entry factors.

The most recently identified entry factor is TfR1. The receptor was initially found down-regulated in huh7 cells infected with HCVcc, however further characterisation of the receptor protein revealed that HCVcc and HCVpp infection could be inhibited by TfR1 silencing and in the presence of anti-TfR1 antibodies. Furthermore, adding the antibodies after HCV attachment would inhibit viral entry, suggesting that TrF1 acts at the post-binding step (Martin and Uprichard, 2013). The NPC1L1 is a cholesterol sensing receptor expressed on the apical surface of hepatocytes. Blockage of the receptor by the FDA-approved NPC1L1 antagonist ezetimbe (used to treat hypercholesterolemia) potently prevented HCV uptake in a cholesterol-dependent manner (Sainz et al., 2012). The EGFR and EphA2 were identified as HCV entry factors by RNAi kinase screening. Inhibiting the receptor kinase activity resulted in impaired HCV infection *in vitro* and *in vivo*. In addition, the receptors are believed to have a regulatory role in the formation of CD81-CLDN1 co-receptor complexes after the initial attachment of the virus (Lupberger et al., 2011).

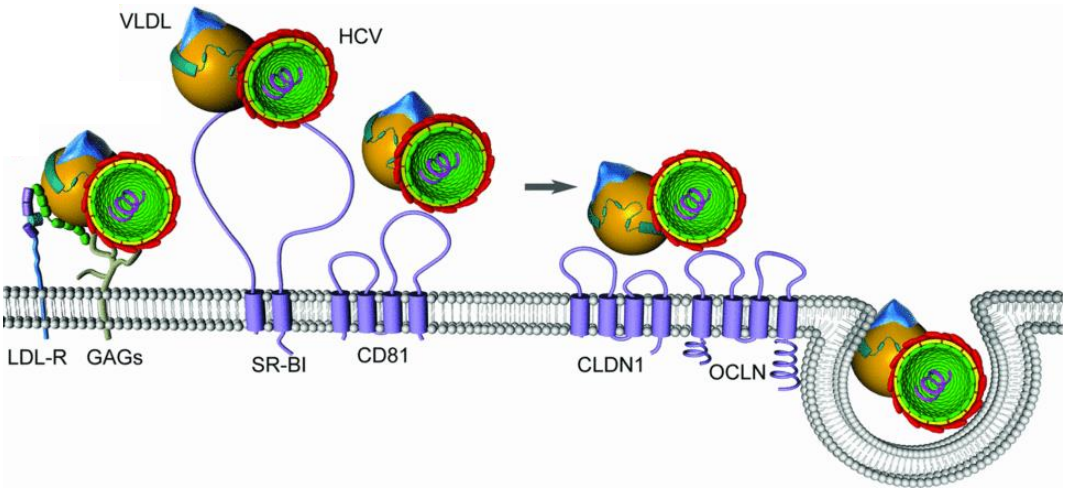


Figure 1.5: HCV entry factors.

The initial binding of the viral particle is believed to be initiated by low-affinity interactions between the surface proteins heparin sulphate glycosaminoglycans (GAGs) and low-density lipoprotein receptors (LDL-R), and the particle components E2 and ApoB and ApoE (here represented in the form of very-low density lipoprotein (VLDL) (Barth et al., 2003; Hishiki et al., 2010; Owen et al., 2009). Following the initial attachment, the viral envelope interacts with SR-BI, CD81, CLDN1 and OCLN, leading to clathrin-mediated endocytosis (Blanchard et al., 2006; Meertens et al., 2006; Popescu and Dubuisson, 2010). Illustration from (Popescu and Dubuisson, 2010).

HCV internalisation

The HCV particle enters the cell via clathrin-mediated endocytosis driven by dynamin (Blanchard et al., 2006; Meertens et al., 2006). By using inhibitors of vascular acidification, researchers demonstrated that the release into the cytoplasm of the HCVpp (Hsu et al., 2003b) or HCVcc (Tscherne et al., 2006) nucleic acid contents is highly pH-dependent. In other enveloped viruses, such as influenza, lowering the pH in the endosome will lead to conformational changes of the glycoproteins, which expose fusion peptides, also called fusion loops. These domains then trigger the fusion process between the endosomal membrane and the glycoproteins, leading to the cytosolic release of the viral genome (Thorley et al., 2010). Identification of 3 classical fusion loop motifs at positions 270 to 284aa, 416 to 430aa and 600 to 620aa in the E1 and E2 proteins, suggests that the HCV genome may be released by fusion (Lavillette et al., 2007). The biological importance of these regions was confirmed by generating mutant HCVpps by site-directed mutagenesis and testing them in liposome/HCVpp fusion assays (Lavillette et al., 2007). The initial model of the E2 ecto-domain structure confirmed the presence of fusion loops and predicted that the loops were positioned close to anti-parallel beta-sheets in the N-terminus and thereby resembles the flavivirus and alphavirus encoded class II fusion proteins (Krey et al., 2010). However, the recently published E2 crystal structure show no signs of classical class II fusion protein folds and thereby differ markedly from the previous predicted model (Khan et al., 2014; Kong et al., 2013). In addition, the new E2 model does not undergo significant conformational changes when exposed to low pH (Khan et al., 2014). It therefore remains unknown how the viral genome is released into the cytosol.

1.12 HCV non-structural proteins and their role in viral replication, assembly and release

The roles of the various HCV non-structural proteins are described below in the context of the intracellular viral life cycle. This life cycle can be viewed in Figure 1.6

1.12.1 Translation and polyprotein maturation

The translation occurs by the rough ER and yields a single polyprotein precursor of about 3000 amino acids in length. At the ER surface the structural proteins and the non-structural p7 become cleaved by host signal peptidases. Studies on the E1 and E2 maturation have suggested that the C-terminal trans-membrane domains of E1 and E2 form hairpin structures, which go through the ER membrane twice. Upon separation of these proteins the C-terminal will translocate to the cytoplasm in order to form the type I membrane topology of mature E1 and E2, leaving the proteins retained in ER (Dubuisson, 2007; Joyce and Tyrrell, 2010; Kim and Chang, 2013; Moradpour et al., 2007; Suzuki et al., 2007). The final glycosylation and di-sulphide rearrangement of the envelope proteins is believed to occur in the Golgi apparatus (Vieyres et al., 2010) (section 1.12.3). The NS2-NS3 junction is cleaved by a zinc-dependent cysteine proteinase activity found at the C-terminus of NS2 and the N-terminus of NS3 (Dubuisson, 2007; Joyce and Tyrrell, 2010; Kim and Chang, 2013; Moradpour et al., 2007; Suzuki et al., 2007). The cleavage of the remaining polyprotein is performed by a serine protease activity in the N-terminus of NS3

in association with the NS4A co-factor protein. In addition, to the serine protease, NS3 contains an RNA helicase/NTPase activity, which is required for eliminating RNA secondary structures, separation of the dsRNA replication intermediates and removal of nucleic acid binding proteins from the genome. RNA unwinding occurs in a 3' to 5' direction and requires energy from NTP hydrolysis. The NS4A protein functions as a co-factor for both the protease and RNA helicase/NTPase activity and facilitates the location of the highly hydrophobic NS3 by anchoring it to the ER membrane (Dubuisson, 2007; Joyce and Tyrrell, 2010; Kim and Chang, 2013; Moradpour et al., 2007; Suzuki et al., 2007).

1.12.2 Viral replication

HCV replication processes occur at sponge-like inclusions of the ER membrane, also termed the membranous web. The formation of this membranous web is believed to be initiated by the viral NS4B proteins, which may play an essential role in gathering the non-structural proteins together at the replication site, including the RNA-dependent RNA polymerase (RdRp) NS5B (Elazar et al., 2004). The HCV replication proceeds through NS5B synthesising a complementary negative strand RNA directly from the positive-stranded genome. The viral RdRp afterwards rely on this intermediate RNA template to generate the positive strand genome (Gu and Rice, 2013; Kim and Chang, 2013; Moradpour et al., 2007; Suzuki et al., 2007). NS5B initiates the synthesis of the negative and positive strand genomes in a primer dependent-manner, where structures in the 3'UTR are thought to create a *cis*-acting "copy

back loop", in which the secondary structures leads to the priming for the complementary strand. The NS5B RdRp is highly error-prone and will generate multiple mistakes during the replication process. The mistakes are incorporated as mutations in the viral genome and the positive-stranded replicons are often not 100% complementary to the template (Behrens et al., 1996; Elazar et al., 2004; Friebe et al., 2005; Tuplin et al., 2012) (see section 1.14). Interestingly, the polymerase activity of NS5B is able to copy both homogenous and heterogeneous RNA templates in a primer-dependent manner (Behrens et al., 1996), however is it not well understood how HCV template specificity is achieved. It is believed that the organisation of the NS5B domains may be important (Gu and Rice, 2013). Alternatively, precipitation studies reported that NS5A binds both negative and positive strand 3'UTR (Huang et al., 2005) and could potentially play a role in the delivery of HCV specific genomes to the NS5B polymerase.

The exact function of NS5A in the HCV lifecycle is largely unknown, however identification of both viral and host interaction partners strongly indicates that this serine phosphoprotein may have multiple functions. Besides binding both negative and positive strand RNA prior to replication (Huang et al., 2005), NS5A has been reported to interact with the viral core protein, leading to an increase in HCV particle production (Masaki et al., 2008). It has also been reported to bind host proteins, such as the cellular chaperon cyclophilin A, which functions in protein folding and trafficking (Lim and Hwang, 2011) and with the phosphatidylinositol 4-kinase III α, which is involved in the formation of

the phospholipid PIP_2 (Foster et al., 2011). These interactions are believed to be dependent on the phosphorylation status of NS5A, which can either be hypo- or hyper-phosphorylated (Shulla and Randall, 2012).

1.12.3 Assembly and release

In recent years researchers have started to uncover the mechanisms behind HCV assembly and release. Although more research is required, current data suggests that the newly synthesised RNA and the structural proteins are assembled on ER-derived lipid droplets in association with the original ER membranous web. During the assembly process the core becomes concentrated on the lipid droplets and becomes surround by membranous structures containing the non-structural proteins (Miyanari et al., 2007). Transfection studies using full-length viral genomes lacking core have shown a reduced recruitment of the replication complex and viral RNA to the lipid droplets, suggesting that core play an essential role in the initial assembly process (Miyanari et al., 2007). In addition, the lack of core will lead to a reduction in infectious HCV particle being released and an increase in the ratio of high-density non-infectious virus particles compared to the wild-type viral genome (Miyanari et al., 2007).

During the genome packing process, NS2 is believed to interact with the viral p7 protein, which stabilises the structure of NS2. The association of these two proteins are essential for the interaction of NS2 with the NS3/4A complex

(Lindenbach, 2013; Lindenbach and Rice, 2013; Stapleford and Lindenbach, 2011), which in turn leads to the trafficking of core protein from lipid droplets (Coller et al., 2012). The exact relationship between core and NS3-4A is unknown, but the fact that one is the capsid protein and the other an RNA-helicase suggests that NS3/4A assists in the packaging of the HCV genome into the nucleocapsid (Lindenbach, 2013; Lindenbach and Rice, 2013). In addition, NS2 has been shown to recruit the ER-resident E1E2 complex to the assembly site (Stapleford and Lindenbach, 2011). A further essential requirement for viral assembly is the interaction of NS5A with core on the lipid droplet. This interaction is believed to assist genome packaging and formation of the icosahedral capsid (Miyanari et al., 2007). In addition, NS5A is associated with the host Apo E protein, which is often incorporated into the envelope (Benga et al., 2010).

The full details on the HCV assembly mechanism is not yet clear, however it is believed that the HCV particle forms via budding into the ER, where the final maturation of the particle is closely linked with the VLDL secretion pathway. This prediction is supported by the high levels of immature intracellular HCV particles following treatment of infected huh7 cells with an inhibitor of the microsomal triglyceride transfer protein; a key regulator of VLDL precursor synthesis (Gastaminza et al., 2008). In the same study Gastaminza and colleagues also found evidence for translocation of the HCVcc particle from the ER to the Golgi apparatus (Gastaminza et al., 2008), where the final maturation of the particle is believed to occur, notably E2 glycosylation by the

Golgi glycosidases and glycosyltransferases (Vieyres et al., 2010). Following maturation in the Golgi the viral particles are sorted into endosome compartments, which are subsequently released at the plasma membrane (Coller et al., 2012). The viral p7 has previously been shown to form ion channels in artificial lipid bi-layers and is believed to be essential for preventing acidification of the endolysosomal compartments during the secretory pathway of the viral particles (Wozniak et al., 2010).

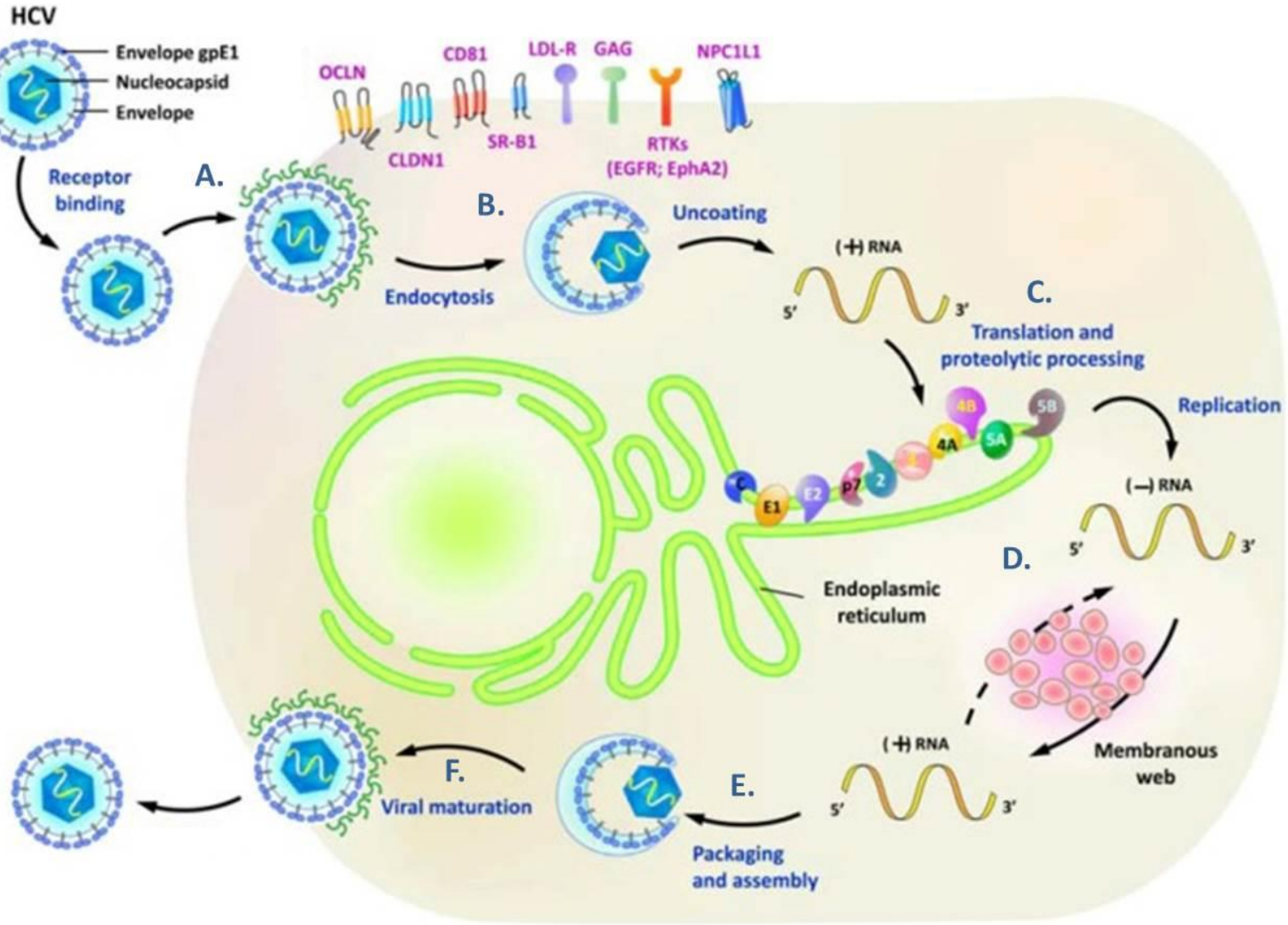

HCV
Envelope gpE1
Nucleocapsid
Envelope
OCLN
CD81
LDL-R
GAG
NPC1L1
CLDN1
SR-B1
RTKs
(EGFR; EphA2)
Receptor binding
A.
B.
Endocytosis
Uncoating
(+) RNA
5'
3'
C.
Translation and proteolytic processing
Replication
(−) RNA
D.
Endoplasmic reticulum
Membranous web
(+) RNA
E.
Packaging and assembly
F.
Viral maturation

Figure 1.6: HCV life cycle.

A. The viral particle interacts with surface receptors on the hepatocytes, which initiate clathrin-mediated endocytosis. Of note, only the entry factors SR-BI and CD81 are represented in the illustration, however claudin 1 and occludin are equally important for HCV internalisation. **B.** Following endocytosis the HCV particle un-coats in response to a reduction in endosomal pH level. **C.** The HCV genome is translated into a single polyprotein by the host ribosomal complex and transferred to the surface of the ER membrane. Here the polyprotein undergoes post-translational modifications leading to the processing of three structural and seven non-structural proteins. **The v**iral proteins remains directly linked to the ER membrane where non-structural proteins direct viral replication. **D.** Prior to replication NS4B initiates the formation of the membranous web where the RNA-dependent RNA polymerase, NS5B, syntheses a complementary negative strand RNA directly from the positive-stranded HCV genome. This intermediate RNA template is subsequently used by NS5B to generate numerous positive strand HCV genomes. **E.** The newly synthesised genomes are packaged into icosahedral capsid. The capsids bud into the ER lumen creating an ER-derived envelope surrounding the nucleocapsid. **F.** From the ER lumen the completed viral particles translocate to the Golgi, where the final maturation of the E1 and E2 glycoproteins occurs. The particles are sorted into endosomal compartments that are trafficked to the cell surface, resulting in viral release from the infected cell (Ashfaq et al., 2011; Dubuisson, 2007; Kim and Chang, 2013; Lindenbach, 2013; Moradpour et al., 2007; Ogawa et al., 2009; Suzuki et al., 2007). Illustration from (Ke and Chen, 2012).

1.13 HCV transmission routes between hepatocytes

The main site for HCV replication is the hepatocytes in the liver. By determining the viral kinetics researchers estimate that it takes approximately 6 hours for the hepatocyte to generate a viral particle starting from the negative HCV RNA-strand (Ribeiro et al., 2012). During the chronic infection approximately 1 x 10^{12} RNA copies are being produced every day and the half-life of the viral particle is about 3 – 5 hours in the blood (Layden et al., 1999; Neumann et al., 1998). Even though most (if not all) circulating HCV particles arise from infected hepatocytes, it has proven difficult to accurately determine the fraction of infected cells and the number of HCV RNA copies within each of them. Recent reports using highly sensitive *in situ* hybridisation and single cell laser capture micro-dissection suggest that between 1-54% of the hepatocytes are infected during chronic infection (Kandathil et al., 2013; Wieland et al., 2013). In addition, Kandathil and colleagues found that a single HCV positive cell contain 1 - 50IU HCV RNA molecules (Kandathil et al., 2013), a quantification that correlates well with previous estimations (Chang et al., 2003; Stiffler et al., 2009) and suggests significantly lower viral burden *in vivo* compared with *in vitro* studies.

The intra-hepatic replication of HCV will lead to the production of virions, which in turn may infect new hepatocytes. Most of the studies published to date have focused on the HCV particle being released into a peri-cellular (cell-free) environment before infecting naïve hepatocytes. However, recent data has demonstrated that HCV is also capable of direct cell-cell transmission between

adjacent cells in the presence of nAbs (Brimacombe et al., 2011; Meredith et al., 2013a; Timpe and McKeating, 2008; Witteveldt et al., 2009) (Figure 1.7). The fraction of virions being transmitted through cell-free and cell-cell routes is unknown, but *in vitro* data from Brimacombe *et al.* shows that all genotypes utilise both transmission pathways to differing degrees; genotype 3a transmission is dominated by cell-cell, whereas genotype 6a mainly transmits through a cell-free route (Brimacombe et al., 2011). The cell-cell transmission route has been reported dependent on SR-BI expression (Brimacombe et al., 2011; Meredith et al., 2013b; Xiao et al., 2014), NPC1L1 (Barretto et al., 2014), occludin and claudin-1 (Ciesek et al., 2011; Timpe et al., 2008; Xiao et al., 2014), but independent of CD81 (Timpe et al., 2008; Witteveldt et al., 2009) and TrF1 (Martin and Uprichard, 2013).

The exact mechanism for HCV cell-cell transmission is at the time of writing not fully understood. Studies on HIV have shown that infection induces activation of the actin cytoskeleton, leading to the formation of filopodia that may carry HIV-particles towards the recipient cell (Aggarwal et al., 2012). A similar route of transmission has been suggested for HCV, since hepatocytes form multiple cell contacts, however a recent publication have reported that filopodia express limited levels of the essential entry receptor SR-BI *in vitro*, making it unlikely that the virus is transmitted to naïve cells via these filamentous contacts (Meredith et al., 2013a). Reports of HCV cell-cell transmission being resistance to monoclonal and polyclonal anti-HCV Ig isolated from infected individuals (Brimacombe et al., 2011; Meredith et al.,

2013a; Timpe et al., 2008; Zhong et al., 2013) and to the DAAs Telaprevir, Boceprevir and Simeprevir (Xiao et al., 2014) suggests that HCV may transmit through synapses or ”pockets” between adjacent hepatoma cells. However, recent observations that heavy-chain antibodies (also referred to as nanobodies) raised in HCV E2 immunized alpacas (Tarr et al., 2013) and polyclonal anti-E1E2 sera from immunized guinea pigs (Chmielewska et al., 2014) can limit cell-cell transmission, suggest that the viral particle is not transmitted via a synopsis that is protected from Ab access.

In 2013 researchers reported that exosomes isolated from HCVcc transduced human hepatoma cell lines contained full-length HCV RNA and viral particle proteins and that these were capable of transmitting productive infection in a process that were partly resistant to IgGs isolated from chronically infected patients and rat antisera raised against CD81, SR-BI and CLDN1 (Ramakrishnaiah et al., 2013). The observation that the sub-genomic replicon cell line were capably of transmitting HCV RNA by exomes as well, suggests that the transmission is independent of E1E2 and explains why the antisera did not lead to full neutralisation of the exosomes containing full length HCV RNA (Cosset and Dreux, 2014; Ramakrishnaiah et al., 2013). The report of HCV transmission via exosomes stands in contrast to previous observations that nAb resistant transmission requires direct cell contact (Brimacombe et al., 2011).

The presence of a cell-cell transmission route for HCV *in vivo* has been difficult to prove, however recent studies using high-sensitive imaging techniques, such as two-photon microscopy and *in situ* hybridization with individually designed probes against the infecting HCV variant have demonstrated the presence of HCV antigen expressing foci in histologically stained needle biopsies from chronically infected patients (Liang et al., 2009; Wieland et al., 2013). In addition, single cell laser capture micro-dissection followed by HCV qRT-PCR, reported a similar focal distribution of HCV positive cells in hepatic biopsies (Kandathil et al., 2013; Stiffler et al., 2009). These observations strongly support the existence of cell-cell dissemination *in vivo* and highlight the importance of considering multiple transmission routes when designing new treatment therapies against HCV entry.

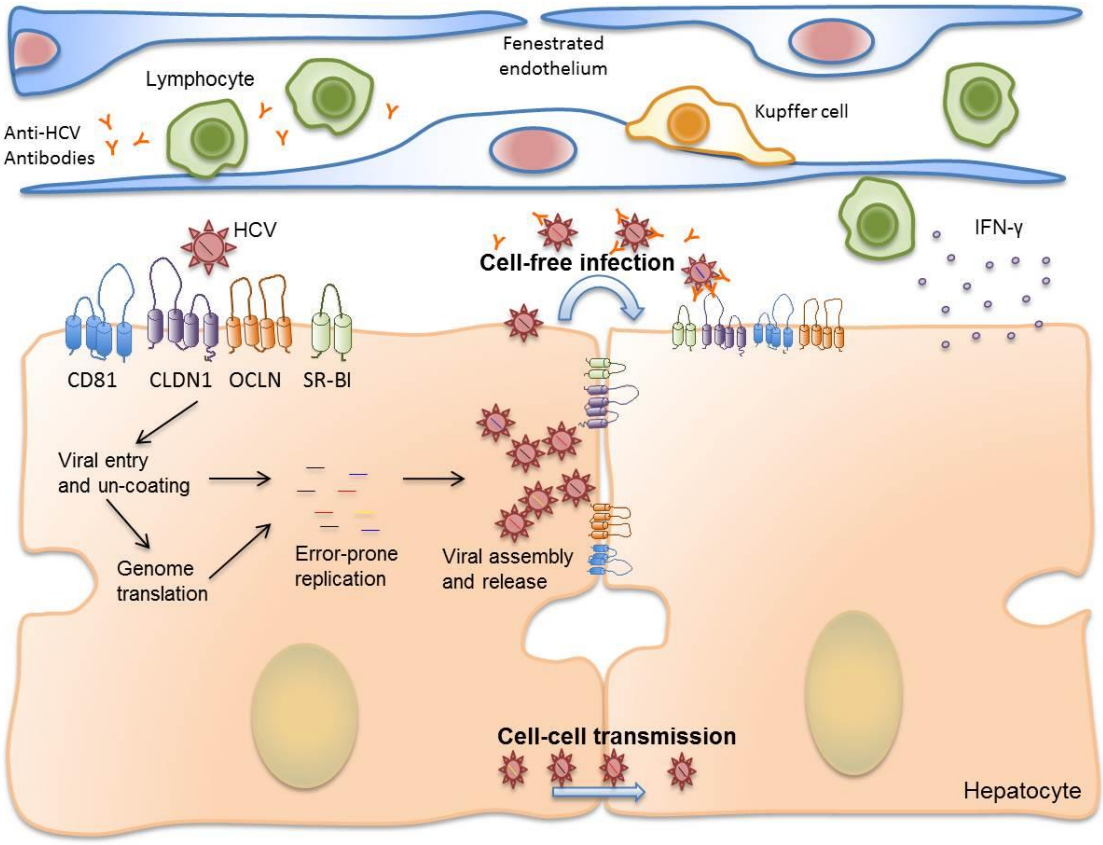


Figure 1.7: Modes of HCV transmission.

Following HCV entry, un-coating and genome translation, the RNA-dependent RNA polymerase (NS5B) initiates replication. The NS5B is highly error-prone and will generate genetically distinct, but closely related progeny strands (indicated with different colour strands) (see section 1.14). Following viral assembly the particles may disseminate through cell-free or cell-cell transmission. During cell-free transmission the virions are released to the extracellular environment. Here the virus will encounter neutralising anti-HCV antibodies (Y), which may prevent the virus from entering new target cells. (Brimacombe et al., 2011; Timpe et al., 2008; Witteveldt et al., 2009). (CLDN1: Claudin 1, OCLN: Occludin).

1.14 HCV quasispeices

Longitudinal studies have shown that HCV accumulates between 1.11 and 1.44 x 10^{-3} nucleotide substitutions per site per year over the whole genome (Gomez et al., 1999; Gray et al., 2011; Simmonds, 2004). This is the equivalent of 0.1% -10% of the daily virus production having at least one nucleotide change (Gregori et al., 2014; Guedj et al., 2013). As a result HCV sequence diversity evolves rapidly and the patient becomes infected with a group of closely related variants, rather than a clonal population. Each variant in the heterogeneous population is represented at a different quantity, also referred to as the major/dominant and minor variants, the major being the variant with the highest number of detectable copies.

In 1971, Eigen developed the quasispecies model, which has proven to be a useful tool for looking at general evolutionary modelling of error-prone self-replicative systems (Eigen, 1971). The term is often used to define a population of genetically variable sequences where natural selection acts on the group rather than the individual (Holmes, 2010a, b; Holmes and Grenfell, 2009). In 2006, Vignuzzi *et al.* set out to challenge the quasispeices hypothesis by generating laboratory poliovirus variants with high- (wild type) and low-fidelity (mutant) polymerases. In support of the quasispecises model they found that the wild type poliovirus polymerase lead to a lower viral diversity in inoculated animal models and a more attenuated pathogenic phenotype compared to the mutant (Vignuzzi et al., 2006). This observation has been described as proof that quasispecies variants function as an

interactive unit, which together contribute to the phenotype of the whole population. HCV is often described as a quasispecies population with similar properties as the poliovirus mutant. However, Holmes challenges the proof that HCV or any other RNA viruses form quasispecies *in vivo* (Holmes, 2010a). Instead, Holmes *et al.* suggests that all previous *in vivo* studies merely studied HCV variability and phenotype in transit, which is necessary for demonstrating quasispecies behaviour, but not sufficient. Most studies investigating natural selection, such as immune escape, drug resistance or co-receptor affinity for individual HCV variants, have failed to show natural selection on the whole group. According to Holmes *et al.* natural selection at the group level can only occur if the mutation rates are high enough to generate an effective "mutational coupling" between the variants, so that they evolve as a unit. However it is unlikely that the mutation rates of RNA viruses are high enough to form this "coupling", thereby reinforcing Holmes' contention that RNA viruses do not form quasispecies (Holmes, 2010a, b; Holmes and Grenfell, 2009; Jenkins et al., 2001).

In contrast to Holmes, other researchers argue that quasispecies dynamics do exist for error-prone asexually-replicating organisms (e.g. RNA viruses). The argument is that RNA viruses follow a concept of "mutation-selection balance", where the quasispecies dynamics rely on mutation-fitness and replication rate. In a situation where the mutation rate is low or zero, the homogenous replicator with a high replication rate will quickly take over the population, but this homogenous population will be less robust to external factors, which may

change the population structure (survival of the fittest model). However, if the mutation rate is high, the resulting heterogeneous population will have higher mutational robustness, where the population is less sensitive to external selection against individual variants (the concept of quasispecies population) (Sardanyes et al., 2008; Wilke, 2005; Wilke et al., 2001). This latter criterion is also called “survival of the flattest”, with a strong reference to Darwin’s survival of the fittest concept. In this model the heterogeneous population covers a wide range of fitness values, including some with very high individual fitness (Mas et al., 2010; Sardanyes et al., 2008; Wilke, 2005; Wilke et al., 2001). Most studies today, including those from the McKeating group, consider HCV as a quesispecies-forming virus that follows the concept of “survival of the flattest” during persistent infection. The term quasispecies will therefore be used in our study to describe the swarm of genetically distinct, but closely related HCV variants isolated from the study subjects.

1.15 HCV quasispecies and persistent infection

Analyses of the E2 HVR1 have previously shown proportionally higher accumulation of non-synonymous (*dns*) than synonymous (*ds*) substitutions per site over time (Farci et al., 2000; Liu et al., 2010; Manzin et al., 1998; Ray et al., 1999; Wang et al., 2007b). For protein-coding genes this suggests positive selection on the mutation rate. The observation that HCV establishes persistent infection in the presence of cellular and humoral responses strongly suggests that positive selection is being driven by the host immune response, leading to the evolution of HCV “escape” variants capable of evading the

immune response. This is further supported by the observations that (*i*) HVR1 sequences diverge around the time of sero-conversion (Bowen and Walker, 2005a), (*ii*) major HVR1 variants are not neutralised by autologous nAbs present in simultaneously collected serum samples, but are neutralised when testing against later collected sera (Kato et al., 1994; von Hahn et al., 2007) and (*iii*) that patients with hypogammaglobulinemia, (lack of a functional humoral immune system) and co-infected HIV patients have significantly lower nucleotide substitution rates for HVR1 compared to immuno-competent HCV infected individuals (Booth et al., 1998; Canobio et al., 2004).

Studies evaluating the HCV HVR1 *dn/ds* ratio in patients who subsequently resolve their infection and in those that progress to chronic diseases have shown contradicting results. Farci and colleagues found that resolving patients had higher HVR1 *dn/ds* ratios than progressing patients, favouring the model that a strong immune selection pressure towards the antigenic HVR1 is required to eliminate the virus (Farci et al., 2000). In contrast Ray *et al.* found high HVR1 *dn/ds* ratios linked to persistent infection and so argued that HVR1 may function as an immunological decoy, stimulating a strong immune response that is unsuccessful in clearing the infection. Instead HVR1 may distract the humoral immune response from the effective nAb epitopes (Ray et al., 1999) (section 1.9.2).

A potential association of HCV quasispecies complexity with hepatic diseases has also proven controversial. While some studies have reported that high HCV sequence divergence is associated with cirrhosis (Hayashi et al., 1997; Honda et al., 1994; Koizumi et al., 1995) and HCC (El-Shamy et al., 2013), others have found no significant differences in HVR1 diversity for patients with mild or severe liver disease (Duffy et al., 2002; Naito et al., 1995). Most studies on this topic have been cross-sectional and have not taken into account the fluctuation of the HCV genetic population over time, which may explain the reported disagreements.

More recent studies have looked at the longitudinal dynamics of HCV quasispecies during hepatic disease progression. Interestingly, most published papers have found that patients who, will later develop severe liver disease, have a lower level of sequence diversity than patients with mild progression (Curran et al., 2002; Li et al., 2011; Sullivan et al., 2007; Wang et al., 2007b). A high HVR1 genetic diversity prior to liver transplantation has also been linked to a milder disease progression following transplantation (Li et al., 2010). In addition, patients with mild disease progression accumulate a higher *dn/ds* ratio over time than those with more aggressive disease (Curran et al., 2002; Li et al., 2011; Wang et al., 2007b), suggesting that a high selective immune pressure will lead to milder disease progression. Interestingly, the observations listed above contrast with the immune-pathogenesis theory that strong HCV-specific immunity (leading to high selective pressure on the HCV

particles) is the major contributor to liver injury (Spengler and Nattermann, 2007).

1.16 Analyses of HCV quasispecies variability

Over the years several different approaches have been made to measure the genetic distance between the identified HCV sequences (*diversity*) and the number and distribution of each HCV genetic variant (*complexity*) within the quasispecies population. Two of the most popular methods, when HCV was first discovered, were single strand conformation polymorphism (SSCP) and Heteroduplex mobility analysis (Gomez et al., 1999; McCaughan et al., 2003). The results from SSCP analyses only provided information about the quasispecies complexity and not diversity (Peters et al., 1997). That said, SSCP is still widely used for analysing changes in the quasispecies population during PEG-IFN-α and ribavirin treatment (Caraballo Cortes et al., 2012; Salmeron et al., 2008). Unlike the SSCP approach, heteroduplex mobility analysis is capable of measuring both quasispecies complexity and diversity (McCaughan et al., 2003). However, heteroduplex mobility analysis has proven to be less sensitive at differentiating single nucleotide substitutions and detecting HCV minor sequences than molecular cloning followed by Sanger sequencing (Wilson et al., 1995). Today, heteroduplex mobility analysis is mainly used to distinguish between HCV genotypes and to detect super-infection (Li et al., 2008; White et al., 2000).

Sanger sequencing of PCR amplified cDNA has been the preferred route to investigate HCV quasispecies populations. Often the investigator will generate a bulk PCR product and perform direct sequencing to generate an initial consensus sequence of all the viral polymorphisms in the sample. Using the consensus sequence to assess viral heterogeneity offers an easy and low cost overview of the sequence polymorphisms in the sample (Rodrigo et al., 2001) However, many scientists are sceptical about the approach, claiming that the method is less sensitive to detect minor sequences and that it does not provide sufficient information about the population structure (Forns et al., 1999; Gomez et al., 1999; Rodrigo et al., 2001). Results presented in this thesis show that the approach is also less sensitive than single molecule amplification (SGA) and Next generation deep-sequencing (NGS) (see Chapter 6).

A more desirable, but labour intensive approach is to bulk PCR amplify desired regions of the viral genome, followed by cloning and Sanger sequencing of individual clones (Fishman and Branch, 2009; Forns et al., 1999; Gomez et al., 1999; Rodrigo et al., 2001). Although this approach provides an overview of the different variants it may be subject to several technical errors, leading to over- or under-estimation of the population variability and divergence. For instance, Forns *et al.* reported that transformation of the *DH5* competent *E. coli* would favour HCV sequences with mutations or frame-shift mutations, resulting in an under-estimation of the genetic diversity (Forns et al., 1997). Furthermore, over-estimation of the diversity may be a consequence of low-

fidelity polymerase creating artificial stop codons, frame-shifts and/or substitutions during PCR amplification. These errors can be minimised by using proof-reading high-fidelity polymerases, such as Phusion (Qin et al., 2011). Amplifying closely related sequences may result in a phenomena called "PCR bridging", where the PCR polymerase "jumps" from one template molecule to another, creating a chimeric sequence that does not exist in the true population (Liu et al., 2002). To eliminate PCR bridging, Simmonds *et al.* developed a highly sensitive PCR method that is able to detect a single molecule of target DNA (single molecule amplification, SGA) (Simmonds et al., 1990). Another advantage of amplifying from a single copy of target molecule is that the PCR product will be homogeneous and can be directly sequenced, saving time and limiting the risk of non-random cloning (Simmonds et al., 1990).

HCV genome heterogeneity occurs throughout the genome, however the mutation rates vary in different regions. The non-structural genes (p7, NS2, NS3, NS4, NS5A and NS5B) have an estimated mutation rate of 1 x 10^{-3} nucleotide substitutions/site/year, whereas the envelope proteins are estimated to have up to 2.7 x 10^{-3} nucleotide substitutions/site/year and core 0.28 x 10^{-3}. Unsurprisingly, the HVR1 has the highest rate of mutations ranging from 2.5 – 6.9 x 10^{-3} substitutions/site/year (Gray et al., 2011). As a consequence HVR1 has become the most extensively analysed region of the HCV genome for quasispecies population studies (Jacka et al., 2013). Sub-genomic sequence analysis offers a less expensive alternative to the full-

sequence analysis, but the disadvantage is the risk of creating false monophyletic groups of HVR1 identical variants, which potentially could have heterogeneity outside this region (Fishman and Branch, 2009; Forns et al., 1999; Gomez et al., 1999). A clear example of this is highlighted in the study reported in this thesis (see Chapter 6).

In recent years NGS approaches have been developed, enabling coverage of 96% and an accuracy of 99.96% for *Mycoplasma genitalium* (580kb); the first organism ever to be sequenced by NGS (Margulies et al., 2005). NGS has proven to be more reliable and more sensitive to detect minor viral variants than SGA detection followed by Sanger sequencing (Fonseca-Coronado et al., 2012; Gregori et al., 2014; Wang et al., 2007a) and the technique has become a popular tool for studying HCV and HIV population dynamics (Bull et al., 2011; Caraballo Cortes et al., 2013; Hoffmann et al., 2007; Wang et al., 2007a; Wang et al., 2010). One of the limitations with NGS is the manual filtering of reads; depending on where the threshold for *quality score* is set (e.g. calling an incorrect base), there may be a risk of filtering out true viral variants (*non-errors*) together with the sequencing errors (Cabanski et al., 2012; Ramirez et al., 2013).

1.17 Viral compartmentalisation

In this study viral compartmentalisation is defined as the existence of viral sub-populations in different tissues within a single host or in different

compartments of the same tissue, between which movement of viral variants is restricted (Beerli et al., 2001). The most widely analysed virus in terms of genetic compartmentalisation is HIV. Compartmentalised HIV sub-populations have been observed between PBMC and the CNS (i.e. brain, spinal cord and surrounding meninges) (Hughes et al., 1997; Lamers et al., 2011), the female genetic tract (Philpott et al., 2005), semen (Anderson et al., 2010), breast milk (Becquart et al., 2002) and most recently between PBMCs and the liver (Blackard et al., 2012; Penton and Blackard, 2013).

In addition, HIV-1 is reported to form intra-splenic compartmentalisation between the white pulps. In this study, Delassus *et al.* found that HIV proviral variable regions in the gp120 envelope protein (V1 an V2) had identical dominant sequences, but different minor sequences between splenic white pulps isolated from a patient with AIDS (Delassus et al., 1992). It is thought that the spatial HIV distribution in the spleen is the result of a few initial (founder) viruses being transferred to different periarteriolar lymphoid sheaths (PALS) in the spleen. The antigen-depended T cell activation in the PALS will afterwards result in an error-prone HIV replication, leading to an independent HIV evolution at the specific sites (Cheynier et al., 1992; Gratton et al., 2000). The spleen will eventually be infected with a large number of small HIV populations (meta-populations), which over time undergo genetic drift and turnover, leading to increased genetic distance between the sites (Frost et al., 2001).

1.18 HCV compartmentalisation between tissues

HCV genetic diversity has previously been reported between several tissue compartments. In 2004, Forton *et al.* who found statistical support for HCV genetic divergence between brain, serum and lymph node for one out of two patients, using cloning and Sanger sequencing methods (Forton et al., 2004). In addition, several studies have reported genetic compartmentalisation between PBMC and serum/plasma for different HCV genomic regions; NS5A (Blackard et al., 2012), HVR1 (Blackard et al., 2007; Navas et al., 1998; Ramirez et al., 2009) and the 5'UTR (Roque-Afonso et al., 2005; Vera-Otarola et al., 2009).

HCV compartmentalisation studies are often focused on either viral RNA quantity or genetic sequence divergence between serum and liver tissue. Comparative studies on the viral load between the two sites in a single host are often difficult as different approaches are used to isolate and measure viral RNA copies at each site. Serum and/or plasma loads are measured as either HCV copies or international unit (UI) per ml serum/plasma, whereas the liver burden is usually expressed as copies per amount of total hepatic RNA, making the two challenging to compare directly. Instead, studies have compared the ratio of the two parameters between multiple study objects. Whereas some have reported positive correlation between the serum/plasma and the liver viral RNA load for humans and chimpanzees (Bigger et al., 2004; Cabot et al., 2000; Descamps et al., 2012; Idrovo et al., 1996; Sakai et al., 1999; Vona et al., 2004) others have reported no correlation between the viral

burden in the two compartments (Ballardini et al., 1997; Fanning et al., 2001; Yuki et al., 2006). It should be mentioned that the majority of 'between' tissue compartmentalisation studies do not consider the disease stage of the study subjects. Of the studies mentioned above, only one has investigated the impact of fibrosis staging (Cabot et al., 2000). In this study, only patients with mild fibrosis showed correlation between the plasma and hepatic HCV RNA loads, however the correlation was lost for patients with cirrhosis (Cabot et al., 2000).

In addition to studying viral burden between the periphery and liver, reports have assessed the genetic diversity between these sites. Fan *et al.* showed that the liver has a higher number of non-identical sequences, which failed to correlate with the peripheral sequence set (Fan et al., 1999) and Li *et al.* has reported that 11 out of 14 patients showed statistical evidence for genetic divergence for HVR1 clones deriving from liver and plasma (Li et al., 2010). Others have found correlation in the level of sequence diversity and complexity between the compartments (Cabot et al., 2000; Sakai et al., 1999) and furthermore that 46 to 53% of patients have identical HVR1 dominant clones between the two compartments together with a unique set of minor (Cabot et al., 2000; Ramirez et al., 2009; Sakai et al., 1999). In this thesis an in depth investigation of the HCV genetic distribution between liver and plasma provides new insight to the topic (see Chapter 6).

The observation that the viral load and genetic pool differs between different compartments in one patient could potentially be the result of either excess hepatic replication, poor hepatic release of some HCV variants or a stronger immune selective pressure on systemic circulating virions (Cabot et al., 2000). Under high immune selective pressure, nAbs may define the clearance rate for some strains in the circulation, leading to an overrepresentation of escape mutants (Cabot et al., 2000; Fan et al., 1999; Ramirez et al., 2009). Another possibility is that extra-hepatic viral replication may contribute to the circulating population. This is supported by detection of HCV negative strands in PBMCs (Natarajan et al., 2010; Willems et al., 1994), lymph nodes (Pal et al., 2006) and in the CNS (Maggi et al., 1999). These extra-hepatic sites have been suggested to act as viral reservoirs and could be involved in the formation of viral variants capably of re-infecting the liver-graft after transplantation (Brown, 2005; Gray et al., 2011).

1.19 HCV intra-hepatic compartmentalisation

The liver is the main site of HCV replication, however several *in vivo* studies have shown that not all hepatocytes in the liver are infected. Techniques such as two-photon microscopy with antibodies specific for HCV have shown that 7-20% of hepatocytes express HCV antigens and are inferred to be infected (Liang et al., 2009). More recently, *in vivo* studies using single cell laser capture micro-dissection of individual hepatocytes followed by qRT-PCR (Kandathil et al., 2013) and *in situ* hybridisation with patient-specific HCV probes (Wieland et al., 2013) have estimated that between 1 and 50% of

hepatocytes are infected and that the number of infected cells correlates with the plasma viremia (Wieland et al., 2013). A further interesting observation from these studies is that infected hepatocytes are not randomly distributed across the liver, but exist in clusters of 5-10 infected cells (foci) (Kandathil et al., 2013; Liang et al., 2009; Wieland et al., 2013), similar to what has been observed *in vitro* (Brimacombe et al., 2011; Meredith et al., 2012).

The formation of foci supports a model where cell-cell infection is the dominant mode of HCV transmission *in vivo*, however it is not known what prevents the virus from expanding the infected foci beyond the 5-10 infected cell. Wieland *et al.* reported higher levels of IFI27 mRNA in hepatocytes expressing HCV antigens and neighbouring naïve cells, compared to more distant ones (Wieland et al., 2013). This suggests that the Jak-STAT pathway is initiated by endogenous IFN production, which in turn may limit viral transmission. In contrast, Kandathil and colleagues failed to observe a similar distribution of another ISG, the IFITM3 (Kandathil et al., 2013). Although the observation of non-random virus distribution is convincing, it should be mentioned that other studies have claimed uniform distribution of HCV across the liver. In 2004, Vona *et al.* found random distribution of HCV infected hepatocytes when using a fluorescein-conjugated serum-derived polyclonal Ab against viral NS5A on hepatic needle biopsies (Vona et al., 2004). However the antibodies used in this study are believed to have low sensitivity and the uniform detection of HCV antigens may be the result of high background staining. Nevertheless, a uniform distribution of infection would support previous reports on comparable

levels of HCV copies in needle biopsies taken from 2 – 7 individual sites in the liver (Fanning et al., 2001; Idrovo et al., 1996; Sakai et al., 1999; Terrault et al., 1997).

Information on HCV genetic intra-hepatic compartmentalisation is limited. The most extensive study to be published included 10 hepatic specimens taken from two explanted livers with end-stage liver disease. In that study heteroduplex mobility revealed that the 10 hepatic sites contained close to identical viral diversity, leading the authors to conclude that the viral diversity is uniform at end-stage liver disease, however the study did not include sequence analyses (Li et al., 2010).

As far as is known, only two reports have studied the HCV sequence diversity across distant sites of the liver. Sakai *et al.* studied 3 samples taken from unidentified sites in the livers of 8 patients with earlier stage liver disease and compared the HVR1 sequence diversity between those patients who progressed to advance disease and those that did not. The data showed correlation between the degree of sequence heterogeneity across the liver, but left out any phylogenetic or compartmentalisation testing for the HVR1 sequences obtained (Sakai et al., 1999). In a second study, Cabot *et al.* analysed liver biopsies taken from 3 different sites in one patient. In this study, comparison of the 5'UTR region and the E2-NS2 junction revealed identical dominant and sub-dominant sequences, together with a set of unique minor

sequences. Cabot and colleagues' concluded that HCV does not form genetic compartmentalisation, based on the fact that the intra-hepatic compartments shared major HCV sequences (Cabot et al., 1997).

The limitation of the studies described above is that most do not include statistical evidence of intra-hepatic compartmentalisation. As far known, the HCV genetic compartmentalisation study by Sobesky *et al.* is the only study to have included statistical testing (Mantel's test) of the isolated sequences from each hepatic site (Sobesky et al., 2007). In this study researchers found statistical evidence of compartmentalisation of HCV core sequences between micro-dissected sections from chronically infected individuals with HCC (Sobesky et al., 2007). In addition they reported evidence for compartmentalisation between non-tumorous and HCC tissue within the same biopsies, indicating that HCV may have evolved independently at different histological sites (Sobesky et al., 2007).

Performing single cell HCV sequencing would provide a new insight into HCV genetic compartmentalisation in the microenvironment. However amplifying and sequencing the quasispecies at a single cell level would require highly sensitive systems. As far known proviral HIV is the only viral genome to be sequenced in a single cell *in vivo* (Balfe et al., 1990; Josefsson et al., 2011; Simmonds et al., 1990) and single cell sequencing studies are yet to be performed for viral RNA genomes. Interestingly, a recent study has showed

that the NS5B sequences from single huh7 cells supporting sub-genomic JFH1 HCV replication, displayed *ds* and *dn* changes between cells in the same culture after 6 days incubation (McWilliam Leitch and McLauchlan, 2013). This observation suggests that HCV evolve independently in individual cells, leading to potential genetic compartmentalisation. It should be mentioned that the study by McWilliam Leitch *et al.* does not take into account viral transmission, since the replicon cells lack structural protein and do not assemble infectious particles. Repeating this study with a full-length infectious virus could potentially lead to less diversity between cells due to the presence of both cell-free and cell-to-cell transmissions.

2 PROJECT OPJECTIVES

The overall goal of this study was to gain a better understanding of the HCV quasispecies distribution within the liver and between the liver and plasma, and to evaluate potential markers, which may predict the intra-hepatic and plasma viral RNA burden.

The hypothesis was that the level of HCV RNA load and the diversity of quasispecies population would vary between different sites within the liver and between liver and plasma, and that intra-hepatic markers would reflect this variation.

The result chapters address the following objectives:

- To quantify the HCV RNA burden across 8 different sites in the livers of 22 patients with end-stage liver disease and to ascertain whether an association exists between hepatic and peripheral HCV RNA levels.

- To measure the expression levels of the 4-gene classifiers (ISG15, IFI27, RSAD2 and HTATIP2) across the 8 hepatic sites and compare the level with hepatic and plasma viral RNA loads and other clinical parameters, such as previous response to PEG-IFN-α + ribavirin treatment, disease diagnosis and infecting viral genotype.

- To investigate HCV sequence diversity across the 8 hepatic sites and plasma collected from patients with end-stage liver disease and to

determine whether genetic compartmentalisation exists between hepatic segments and between liver and plasma during ends-stage liver disease.

3 MATERIALS AND METHODS

3.1 Clinical samples

3.1.1 Adult samples

This project is part of a clinical trial for the SR-BI antagonist ITX5061 in a liver transplantation setting. The clinical trial was managed by Dr Ian Rowe (HCV Research Group, University of Birmingham). In total 23 patients were recruited to the study. The first 13 received placebo treatment and the remaining 10 patients received one dose of ITX5061 12 hours pre-transplantation and then daily for 7 days post-transplantation. Less than 24 hours after liver transplantation, pathologist Dr. Gary Reynolds (Liver Research Group, University of Birmingham) sampled 150mg specimens from each of the 8 Couinaud's segments in the liver explants. Parts of these were snap-frozen for later RNA extraction and the adjacent sections were fixed in formalin for histology. An overview of the 8 liver segments is illustrated in Figure 3.1A. The Couinaud's segments indicated on the patient 1 liver in Figure 3.1B.

In addition to the liver specimens, each of the transplanted patients donated 5ml plasma at the time of transplantation. The plasma viral load was quantified by the Heartlands Hospital, Birmingham, UK, using the Roche Cobas Amplicor 2.0 assay and expressed as International Unit (IU) / ml. To compare the HCV copies per sample, all IU were converted to HCV copy numbers by using the

guidelines on the Hepatitis C Support Project (HCSP) Fact Sheet for the Cobas Amplicor HCV monitor v2.0 (1 IU/ml = 2.7 HCV copies/ml).

Information about the patient age, indication for liver transplantation, previous response to IFN treatment, plasma viral load at time of transplantation and the time between the explanted liver was removed from the patient and processed in the laboratory can be found in Table 3.1. Further information regarding the trial is available at www.clinicaltrials.gov (NCT01292824).

Total hepatic RNA from six normal livers, six livers with alcoholic liver disease (ALD), six non-alcoholic steatohepatitis (NASH), six auto-immune hepatitis (AIH) and six non-A, non-B, non-C viral hepatitis (nA/nB/nC) were provided by Doctoral Researcher Annika Wilhelm (Liver Research Group, University of Birmingham).

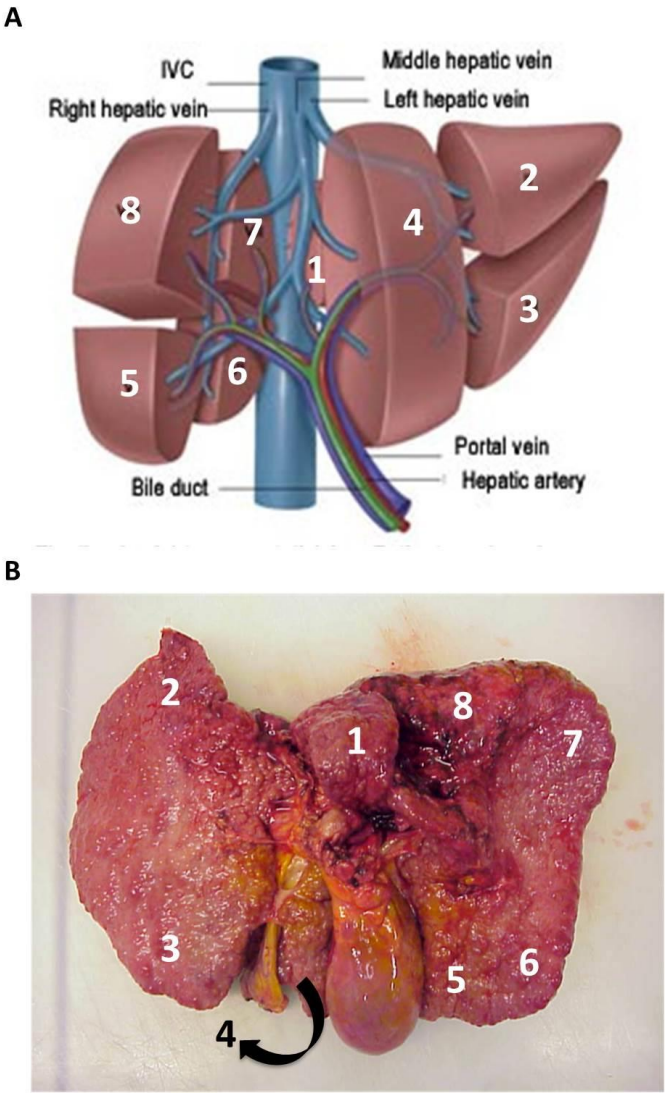


Figure 3.1 Anatomy of the liver.

A. Illustration of the Couinaud's segmental anatomy of the human liver. The main portal vein is divided into right and left branches, followed by further division into four vessels, which supply each of the 8 Couinaud segments with portal blood. In addition the segments have individual arterial blood supply and bile drainage. Segment 1 is the equivalent of the caudate lobe, which is separate from the other segments. Segments 2 and 3 are located on the left lateral liver section, segment 4 on the left medial sector, segment 6 and 7 on the right anterior sector and 5 and 8 on the right posterior sector (Dooley et al., 2011). In this study specimens were taken from the surface of each of the 8 sections (illustration from (Bornman and Beckingham, 2001)). **B**. Liver explant from patient 1. The numbers indicate where the specimens were taken. Sample 4 was dissected from the dorsal side.

Table 3.1: Adult patient information

No.	Age	Indication	Genotype	Viral load (HCV copies/ ml plasma)	IFN response	Time until processing
1	58	Liver failure	3	1.3×10^6	Untreated	16.00
2	39	Liver failure	1a	2.0×10^6	Untreated	13.83
3	62	HCC	1b	1.6×10^6	Non-response	2.50
4	57	HCC	3a	2.8×10^6	Non-response	20.25
5	48	Liver failure	3a	8.2×10^2	Relapse	4.00
6	56	Liver failure	1	2.9×10^6	Non-response	16.00
7	58	HCC	1a	5.2×10^5	Not tolerated	7.50
8	34	Liver failure	3a	1.3×10^5	Untreated	6.66
9	58	Liver failure	1	1.3×10^5	Relapse	15.25
10	52	Liver failure	4	1.3×10^5	Non-response	2.50
11	62	HCC	2	5.4×10^6	Non-response	3.50
13	49	Liver failure	1	3.5×10^5	Relapse	22.50
14	58	HCC	3	1.2×10^5	Relapse	12.00
15	39	Liver failure	1	8.4×10^4	Non-response	5.00
16	38	Liver failure	1	3.2×10^3	Non-response	2.33
17	58	Liver failure	3	8.1×10^5	Relapse	6.00
18	62	Liver failure	3	2.1×10^5	Not tolerated	5.50
19	47	Liver failure	1	3.2×10^7	Non-response	6.00
20	49	Liver failure	1a	3.0×10^6	-	3.25
21	54	Liver failure	1b	8.1×10^6	-	13.20
22	64	Liver failure	1a	1.1×10^6	-	21.00
23	56	Liver failure	3a	3.2×10^6	-	6.66

HCC: Hepatocellular carcinoma
Untreated: The patient has not received treatment for the infection.
Non-response: The plasma HCV RNA load remained unchanged during treatment.
Relapse: Initial response to treatment, however the plasma viral burden returned to pre-treatment levels either during or after the treatment had ended.
Not tolerated: The patient did not tolerate the therapy and the treatment had to be ended early due to server adverse effects.

3.1.2 Children samples

Frozen hepatic needle biopsies from 11 chronically infected children were provided by the Birmingham Children's Hospital Biobank. Information on the pediatric patients can be found in Table 3.2.

Table 3.2: Paediatric patient information

No.	Age	Genotype	Viral load (HCV copies/ ml plasma)	Length of treatment (weeks)	IFN response
1	4	1	1.3×10^6	49	Response
2	3	1	2.0×10^6	47	Response
3	3	1	1.6×10^6	23	Non-response
4	6	3a	2.8×10^6	23	Response
5	7	1a	8.2×10^2	46	Non-response
6	5	1	2.9×10^6	69	Response
7	3	1	5.2×10^5	24	Response
8	13	1a	1.3×10^5	60	Response
9	7	3a	1.3×10^5	62	Response
10	11	3a	1.3×10^5	23	Response
11	14	3a	5.4×10^6	23	Response

3.2 Histology

3.2.1 Sectioning

Hepatic specimens from patient 2, 7 and 9 were snap frozen in liquid nitrogen prior to cryo-sectioning (Leica Biosystems, Germany). The frozen tissue was embedded within Jung Tissue Freezing Medium (Leica Biosystems) and placed on a chuck in the cryostat (– 20°C). All 20µm sections were transferred to pre-chilled sterilised eppendorf tubes, which were placed back into liquid nitrogen. The 2 x 5µm sections for each patient were placed on glass-slides and allowed to air-dry before being fixed in 10% formal saline solution (Adam's Healthcare Ecolabs, UK) for 24 hours at room temperature. Following fixing, the tissue was stained with H&E using the standard protocol (Bancroft et al., 2013). All reagents for the H&E were provided by Leica Biosystems.

3.2.2 Microscope imaging

The 5µm sections from patients 2, 7 and 9 were scanned at 10 x magnification by the Leica DM6000 microscope (Leica Biosystems). The images were afterwards analysed using ImageJ 1.48b (Wayne Rasband, National Institute of Health, USA).

3.3 RNA extraction

3.3.1 Adult hepatic samples

Total RNA was extracted from the HCV infected hepatic specimens (weight ~150mg) using the RNeasy Midi-kit with the manufacturer's provided instructions for "Isolation of total RNA from animal cells" (Qiagen, Germany). Briefly, the snap frozen hepatic sample was placed in separate gentleMACS^{TM} M-tubes (Miltenyi Biotec, UK) together with 4ml of Buffer RLT (all buffers were provided with the kit) + 1% β-Mecaptoethanol (Sigma-Aldrich, USA). The M-tubes were placed on the gentleMACS^{TM} Dissociator (Miltenyi Biotec), which homogenised the liver tissue during the manufacture's default program RNA02.01. The homogenised lysate were added 4ml of 70% ethanol and transferred to RNeasy Midi-columns. The columns were centrifuged for 5 minutes at 12,500 *x g* and the flow-through discarded. Four ml of washing buffer RW1 was added to the column prior to centrifugation for 5 minutes at 12,500 *x g*. This was followed by two washes with 2.5ml buffer RPE followed by 2 minutes centrifugation at 12,500 *x g* between each wash step. The Rneasy Midi-columns were transferred to sterile15ml Falcon tubes (Fisher Scientific, UK) and 250µl of the provided RNase-free H_2O added. The membrane was allowed to soak for 5 minutes before the RNA was eluted during a 5-minute centrifugation at 12,500 *x g*. The eluted RNA was stored in 50µl aliquots at -80°C.

All samples from one adult liver were processed on the same day and only one liver processed per day, in order to decrease the risk of potential cross

contamination between patients. Plasma specimens were processed on separate days for the same reasons. Phylogenetic analyses of the plasma- and liver derived sequences showed no evidence of cross-contamination, validating this approach.

3.3.2 Children hepatic samples and adult liver sections

The paediatric hepatic samples (about 27mm^3) and the 12 sections of the adult liver specimens (about 0.72mm^3) were considerably smaller than the 8 hepatic adult specimens (about 256mm^3) and therefore required the RNeasy Micro-kit (Qiagen) for RNA extraction. The capacity of the RNeasy column is 300µl lysate, however the gentleMACS^{TM} M-tubes requires a minimum volume of 600µl for a complete homogenisation. The lysates were therefore split into 2 x 300µl after being homogenised on the gentleMACS^{TM} Dissociator (program: RNA02.01) in the presence of 600µl buffer RLT + 1% β-Mecaptoethanol + 4 ng/µl carrier RNA. The RNA extraction was done according to the manufacture's instructions for "Purification of total RNA from fibrous tissue". In short, 300µl of lysate were added to 300µl Protease K solution (0.33mg/ml) (Qiagen) and incubated at 55°C for 10 minutes followed by 3 minutes centrifugation at 20,000 *x g*. The supernatant was transferred to a new eppendorf tube and 500µl of 100% ethanol gently added. This mixture was transferred to the RNeasy Micro-column and the sample centrifuged for 30 seconds at 20,000 *x g*, followed by a 350µl buffer RW1 wash step. The flow-through was discarded and 80µl of DNAse I (0.34 Kunitz units/µl) added to the membrane. The column was incubated for 15 minutes at room temperature

before 350µl buffer RW1 was applied and the sample centrifuged for 30 seconds at 20,000 *x g*. The flow-through was discarded and 500µl of buffer RPE added prior to an additional centrifugation step. A volume of 500µl 80% ethanol was added to the RNeasy Micro-column, followed by a 7-minutes spin at 20,000 *x g*. The column was then transferred to a sterile eppendorf tube and the RNA eluted in 15µl of RNase-free H_2O by centrifugation at 20,000 *x g* for 2 minutes. The elutes from the duplicate RNA extractions were pooled and stored at -80°C.

3.3.3 Adult plasma samples

RNA was extracted from 500µl of plasma by the QIAamp MinElute virus spin kit (Qiagen, UK), using the provided protocol for "Purification of viral nucleic acids from plasma, serum or cell free body fluids". The viral RNA was extracted according to the manufacture's instruction without any alterations. The extracted viral RNA, was eluted in 30µl RNase-free water and stored at -80°C.

3.4 Optimising HCV detection

3.4.1 RNA heat denaturation

Investigation of the optimal temperature for dsRNA denaturation was performed with 1µg dsRNA ladder (New England Biolabs, UK) as the template. The dsRNA ladder was mixed with 1x SYBR-green jumpstart *Taq* ReadyMix (Sigma-Aldrich) and incubated at increasing temperatures going from 37 to

99°C over 30 minutes or decreasing temperatures going from 99 to 37°C over 1 hour, using the melting-curve setting on the Stratagene Mx3000P QPCR machine, (Stratagene, UK). The fluorescence intensity from the SYBR-green oligos binding dsRNA were analysed by the qPCR software provided by the manufacture (Stratagene Mx3000P QPCR machine system).

The compounds betain, trehalos dihydrate and dimethyl sulfoxide (DMSO) have previously shown to reduce secondary RNA structure during qRT-PCR and it was therefore decided to test these within our melting curve assay. The optimal effect of DMSO is usually found within a narrow range of concentrations (Masoud et al., 1992) and we therefore compared concentrations of 5 and 10% DMSO in the assay. The final concentrations of the remaining compounds were based on previous experience in the McKeating lab (Peter Balfe, unpublished data).

The assays were set up with the addition of either 10% betain (Sigma-Aldrich), 1% D-(+) trehalos dihydrate (Sigma-Aldrich), 5 or 10% DMSO (Fisher Scientific, UK) to the 1x SYBR-green jumpstart *Taq* ReadyMix and 1µg dsRNA ladder. The melting curve program on the Stratagene qPCR machine was set to reach 99°C within 30 minutes and the SYBR-green fluorescence was measured for each second. As a result of these data, HCV quantification in the hepatic RNA extractions was performed in the presence of 5% DMSO (Fisher Scientific) (see section 3.5).

The dsRNA denaturation rate was determined by incubating the dsRNA ladder at 85°C for 30 minutes on the Stratagene qPCR machine. The fluorescence intensity was measured every second. Heating up the samples to 85°C resulted in a rapid 1.25-log reduction in fluorescence intensity within the first few seconds. By 15 seconds all of the samples had reached plateaus of minimum fluorescence (Figure 3.2, here shown for the first 30 seconds). The fluorescence did not decrease further during the 30-minute incubation. These data suggest that the RNA remained single stranded for an extensive period and it was therefore decided to incubate RNA samples with 5% DMSO for 2 minutes at 85°C immediately prior to target-specific HCV cDNA synthesis (see section 3.6.1).

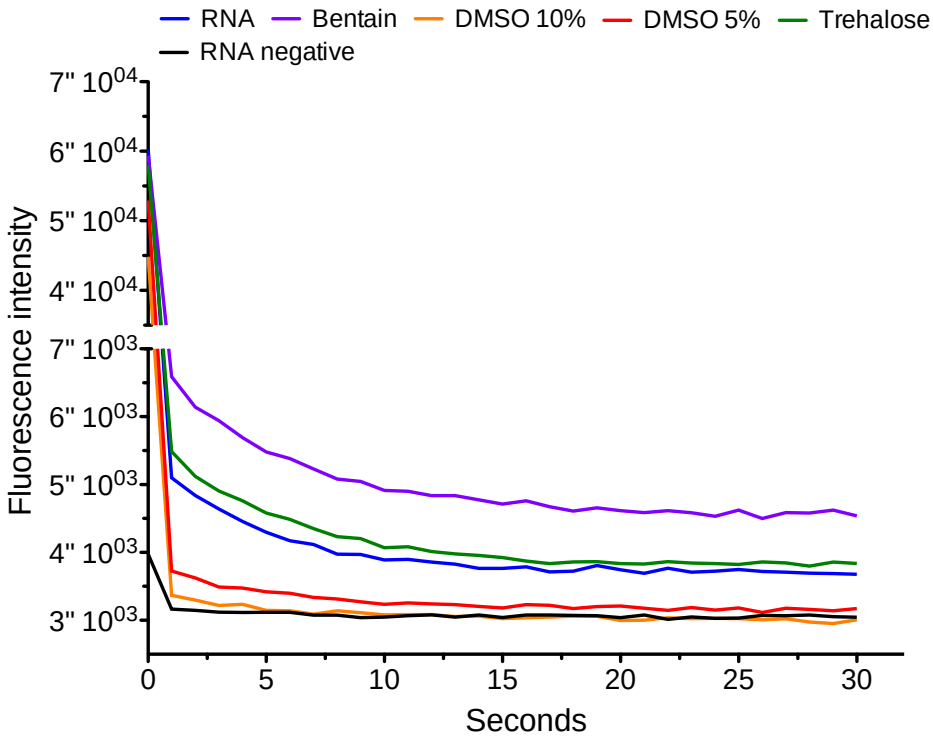


Figure 3.2 Heat-treatment of dsRNA.

One µg of the dsRNA ladder (New England Biolabs, UK) was used as template for the melting curve assay. The amount of dsRNA was indirectly detected by quantifying the SYBR-green (Sigma-Aldrich) bound to the dsRNA. Different compounds were mixed with the RNA-SYBR green mixtures and the samples incubated at 85°C for 30 minutes and the SYBR-green fluorescence measured every second. The fluorescence intensity was unchanged after 15 seconds of heat incubation.

3.4.2 HCV cDNA efficiency

HCV RNA was synthesised from the full-length genotype 1a vector, H77 (pcDNA3.1 Directional TOPO expression vector (Life Technologies, UK)). Firstly the vector was restriction digested with 5U/µl XbaI in the provided buffer (New England Biolabs Inc, UK), followed by RNA transcription of 2000ng vector DNA using the RiboMAX Large Scale RNA production system –T7 according to the manufacture's instructions (Promega). The RNA product was purified using manufacture's "RNA clean-up" protocol provided with the RNeasy Kit (Qiagen) and the RNA concentration measured by the GeneQuant Spectrophotometer (Amersham Biosciences). For the spiking assay increasing amounts of total liver RNA (18.75ng; 37.5ng; 75ng; 150ng; 375ng; 750ng; 1500ng; 2250ng and 3000ng) from a HCV negative ALD liver were spiked into a constant amount of HCV RNA (8.3×10^5 copies) followed by cDNA synthesis in the presence 10nmol v3463a-1a primer (Table 3.3) (SuperScript® VILO™ cDNA synthesis kit according to the manufacture's recommendations (Life Technologies)). The HCV and GAPDH content were measured by qPCR (see section 3.5.1).

3.4.3 Detection of minor HCV RNA in vitro

Using the same protocol as for the HCV spiking experiment (see section 3.4.1), RNA was generated from 1600ng of the patient-derived P1-B6.H3-1.6 and P5-B3.F3-B9 clones, followed by RNA purification with the RNeasy kit. The RNA concentration was quantified by spectrophotometry and decreasing amounts of P4-B3.F3-B9 (0.5log dilutions starting at 5000ng) was spiked into a constant

solution of 5000ng P1-B6.H3-1.6 RNA and 1500ng total ALD RNA. HCV specific cDNA synthesis (primer: v3471a-3a (10nmol)) was carried out using the Superscript III Reverse Transcriptase kit according to the manufacturer's instructions (see also section 3.6.1).

This cDNA was used as the template in bulk amplifications (see section 3.6.2) and the PCR products purified using the QIAquick PCR purification kit according to the manufacturer's recommendations (Qiagen). Consensus sequencing of the purified product was done by Sanger sequencing using the v1293s primer (see section 3.8).

3.5 Real time quantification PCR (qRT-PCR)

Absolute and relative quantifications of target RNA were measured in the presence of the CellsDirect One-Step qRT-PCR reagents (Life Technologies, UK); 1 x Reaction Mix, 0.225 x SuperScript® III RT/Platinum® Taq Mix together with 1 x human GAPDH (VIC) reference primer-set (Applied Biosystems) and 5% DMSO (Fisher Scientific). The total volume for the qRT-PCR was 15µl per well in 96 well plates and 10µl for 384 well plates. The content of target RNA was quantified by either the ABI 7500 Real Time PCR system (96 well plate) or the ABI 7900HT Real Time PCR system (384 well plate) (Applied Biosciences). All measurements were performed as duplicates in two technical repeats.

3.5.1 Quantification of HCV RNA

For the detection of HCV RNA, 0.015µl HCV 5'UTR specific primer/probe mixture (FAM) (PrimerDesign, UK) were added per 1µl CellDirect PCR reaction mixture (see section 3.5). Quantification of the HCV copies in adult specimens was performed for 0.5µl template in 384 well plates and 96 well plates. The viral burden in the children samples was evaluated in 96 well plates using 1.5µl of template.

A dilution series of HCV positive control standards were included on each qRT-PCR and qPCR plate (10^0 to 10^7 copies/µl, PrimerDesign) to generate a standard curve from which the absolute HCV copy numbers could be calculated. The following thermo-cycling protocol was used for HCV qRT-PCRs:

30 min	50°C	
5 min	95°C	
15 sec	95°C	⎫ 50 cycles
60 sec	60°C	⎭

The results were analyzed by manually setting the threshold values at the exponential phase of the FAM (HCV content) and VIC (GAPDH content) sigmoidal amplification curves. The HCV quantity for the 8 hepatic specimens were normalised to the concentration of total liver RNA measured by spectrophotometry. The HCV copy numbers in the children needle biopsies

and the adult 20µm hepatic sections were normalised to the internal GAPDH expression (2^Δ Ct).

3.5.2 Evaluation of hepatic expression

The expression patterns of *DDX3* (Hs00606179_m1), *albumin* (Hs00910225_m1), *ISG15* (Hs00192713_m1), *IFI27* (Hs00271467_m1), RSAD2 (Hs00369813_m1) and *HTATIP2* (Hs01091727_m1) were evaluated by target-specific qRT-PCR. The probes were obtained from Life Technologies TaqMan® Gene Expression Assay library, all with FAM-dye. The probes were diluted 1 in 3 with the CellDirect qRT-PCR reaction mix. DDX3, ISG15, IFI27, RSAD2 and HTATIP2 mRNA levels were measured in 0.5µl template from the 8 hepatic specimens. The gene expressions were quantified in duplicate in two independent qRT-PCR assays.

The 20µm sections from patients 2, 7 and 9 were evaluated for albumin, ISG15, IFI27, RSAD2 and HTATIP2 expressions in 96 well plates. A volume of 0.02µl total liver RNA were used for quantifying albumin mRNA and 0.5µl used for the quantification of the remaining gene mRNAs. The expressions of ISG15, IFI27, RSAD2 and HTATIP2 were evaluated in 1.5µl of the paediatric hepatic RNA extractions.

The relative expressions of the target genes were evaluated by normalising the Ct-values with the internal GAPDH expression (2^Δ Ct).

3.5.3 Evaluation of housekeeping genes

Evaluation of the housekeeping gene's mRNA levels was performed with the PerfectProbe geNorm 12 gene kit (PrimerDesign). The probes were prepared according to the manufacturer's instructions, and used together with the previously described protocol for CellsDirect One-Step qRT-PCR kit in a 96 well plate. The volume of hepatic specimens was 0.25µl / measurement.

Threshold values were set manually in the exponential phase of the FAM (housekeeping gene content) and VIC (GAPDH content) sigmoidal amplification curves. Calculation of the 2^Δ Ct was performed by using the Ct-value for each gene to normalise the Ct levels for the remaining 11 gene measurements for the sample template RNA.

3.5.4 Anti-viral gene-array

Biopsy 1 total hepatic RNA extraction collected from patient 4 (5.3ng RNA), patient 5 (16.9ng RNA), patient 6 (14.3ng RNA) and patient 13 (14.2ng RNA) were used as templates for cDNA synthesis using reagents from the RT^2 First Strand Kit (SABiosciences, USA) together with the manufacturer's instructions. The total volume of cDNA synthesis for each specimen was applied to 96 wells

on a 384 well Human antiviral Response PCR Array (SABiosciences, USA) and the gene expression measured according to the provided protocol (RT^2 Profiler PCR Array, using the Absolute Quantity RT-qPCR program). Data analyses were done in the web-based RT^2 Profiler PCR Array Data Analysis Program (www.sabiosciences.com/pcrarraydataanalysis.php, SABiosciences). Gene expressions of the candidate genes were normalised to the mean expression of the housekeeping genes; actin, B2M, GAPDH, HPRT1 and RPLP0 (2^∆ Ct).

Comparative analyses between patients with high and low hepatic HCV loads were performed by calculating the 2^∆∆ Ct levels for each candidate gene and the fold-difference displayed on heat-maps.

3.6 Generation of patient-derived HCV sequences

3.6.1 *cDNA synthesis*

Recent studies show that *cis*-interactions between core, E1, E2, p7 and NS2 from the same virus variant are important in the assembly of the infectious HCVcc particles (Stapleford and Lindenbach, 2011). We therefore aimed at generating the 3.4kb cDNA fragment from the RNA extractions using reverse primers binding at the N-terminal of the NS3 junction (see Table 3.3). For unknown reasons it was not possible to generate the 3.4kb fragment for HCV genotype 1a and instead the E1 and E2 fragment was obtained (1.7kb).

In addition, we found that the cDNA genesis efficiency varied between HCV genotype 1a and 3a and we therefore used different kits in order to obtain high cDNA yields from both genotypes. The “High capacity cDNA reverse transcription” (Applied Biosystems, ABI) were used for HCV genotype 3a variants and the SuperScript® reverse transcriptase kit (Life Technology) for genotype 1a. The cDNA syntheses were either used directly for nested PCR amplification or stored at -20°C.

Optimization for viral positive strand RNA availability (see section 3.4.1), resulted in the following cDNA synthesis protocols for the two kits:

- *High capacity cDNA reverse transcription kit* (Applied Biosystems, ABI): 7µl RNA were heat-treated for 2 minutes at 85°C before being added 10µM HCV specific reverse primer (v3471a-3a, see Table 3.3), 1 x Reverse transcriptase buffer, 1 x Reverse Transcriptase Mixture and 5% DMSO (Fisher Scientific) in a total volume of 20µl. The cDNA synthesis was carried out at 37°C for 2 hours, followed by heat-treatment at 95°C for 10 minutes to deactivate the reverse transcriptase.

- *SuperScript® III Reverse Transcriptase kit* (Life Technology): 11µl RNA were pre-heat treated for 2 minutes at 85°C before being added to 10nmol HCV specific reverse primers (v2616a-1a, see Table 3.3), 10mM dNTPs and 5% DMSO (Fisher Scientific) to a total volume of 13µl. The mixture was heated to 65°C for 5 minutes followed by a 1

minute incubation on ice. The template mixture was then added to 1 x First-Strand Buffer, 5mM DTT, 200 units of Superscript™ III Reverse Transcriptase and 40 units of RNaseOUT™ Recombinant Ribonuclease Inhibitor (Life Technologies) in a total volume of 20µl. The mixture was allowed to rest for 5 minutes at room temperature. The cDNA synthesis was carried out at 50°C for 2 hours followed by heat-treatment at 70°C for 15 minutes to inactivate the reaction.

3.6.2 Polymerase chain reaction (PCR)

The cDNA syntheses were used as templates for amplifying HCV cDNA sequences in the sample. Nested PCR reactions were set-up on ice using reagent for GC-rich templates (New England Biolabs): 1x GC buffer, 4% DMSO, 0.2µM HCV specific forward primer, 0.2µM HCV specific reverse primer, 40units/µl Phusion High-Fidelity DNA polymerase and 0.3µM dNTP mixture (Invitrogen). Amplifications were performed on the Mastercycler gradient thermocycler (Eppendorf), with the following program:

2 min	95°C	
15 sec	94°C	⎫
15 sec	60°C	⎬ 30 cycles
150 sec	72°C	⎭
∞	4°C	

One µl of the 1st round PCR product was used as template in the 2nd round PCR. The 2nd round PCR products were evaluated on 1% agarose gels (1% molecular grade agarose (Bioline, UK), 1x bionic buffer (Sigma-Aldrich), 1x SYRB Safe DNA gel stain (Invitrogen)) by gel electrophorese (100V). The sizes of the amplified fragments were verified by running 100ng of the 1kb Plus DNA ladder (Invitrogen) along side the PCR products on the gel. The amplified products were purified using the QIAquick PCR purification kit (Qiagen) according to the manufacturer's protocol before direct sequencing.

Table 3.3: List of primers

Application	Name	Direction	HCV Genotype / Patient no.	Sequences (5′ – 3′)
cDNA Synthesis	v3471a-3a	Reverse	3a / patient 1, 4, 5	CAA TAG TTC CAA GAA GGC CCC TAG TTT GCT G
1st round PCR	v2616a-1a	Reverse	1a / patient 2, 7, 9	GGG ATG CTG CAT TGA GTA
	v70s25	Forward	3a/patient 1, 4, 5	AGA AAG CGT CTA GCC ATG GCG TTA G
	v3471a-3a	Reverse	3a / patient 1, 4, 5	CAA TAG TTC CAA GAA GGC CCC TAG TTT GCT G
	v818s-1a	Forward	1a / patient 2, 7, 9	GGA CGG GGT AAA CTA TGC AAC AGG
2nd round PCR	v2616a-1a	Reverse	1a / patient 2, 7, 9	GGG ATG CTG CAT TGA GTA
	v166s-CAAC	Forward	3a / patient 1, 4, 5	CAC CGT GGT CTG CGG AAC CGG TGA GTA CAC CG
	v3443a-3a	Reverse	3a / patient 1, 4, 5	CCC CTA GTT TGC TGG GCG TAT GCT GTG ATC G
	v843s-1a	Forward	1a /patient 2, 7, 9	CAC CAT GGG TTG CTC TTT CTC TAT C
	v2582a-1a	Reverse	1a / patient 2, 7 9	TTA CGC CTC CGC TTG GGA TAT GAG TAA CAT CAT
PCR colony Screening	T7 (pcDNA3.1D)	Forward	1a, 3a / patient 1, 2, 4, 5, 7, 9	TAA TAC GAC TCA CTA TAG GG
	v2273s-pt4	Forward	3a / patient 4	ACG CCG CTT GCA ATT GGA C
	BGH (pcDNA3.1D)	Reverse	1a / patient 2, 7 9	TAG AAG GCA CAG TCG AGG
Sequencing	v1433a-1a	Reverse	1a /patient 2,9	CAT GGA GAA ATA CGC TAT GC
	v1711a-1a	Reverse	1a /patient 7,9	CTA CAG CTG GCC ATC CTC TCG
	v2273s-3a	Reverse	3a/ patient 1, 5	ACG CCG CCT GCA ACT GGA C
	v610s-3a	Forward	3a/ patient 1, 4, 5	GGC AGG GTG GCT CCT GTC
	v954s-3a	Forward	3a/ patient 1, 4, 5	AAC GAC TGT TCT AAT AGC AGT ATT G
	v1293s-pt1	Forward	3a/ patient 1	ATG GCT TGG GAT ATG ATG ATG
	v1679s-3a	Reverse	3a/ patient 1	GTT CAA CTC TAC TGG ATG TCC
	v2273s-3a	Reverse	3a/ patient 1, 5	ACG CCG CCT GCA ACT GGA C
	v1293s-pt4	Forward	3a / patient 4, 5	ATG GCT TGG GAC ATG ATG ATG
	v1679s-pt4	Forward	3a/ patient 4, 5	GTT CAA CTC TAC TGG GTG TCC
	v2273s-pt4	Forward	3a / patient 4	ACG CCG CTT GCA ATT GGA C
	v746a-3a	Reverse	3a/ patient 1, 4, 5	CCC CAT GAG GTC GGC AAA TCC
	v1364a-3a	Reverse	3a/ patient 1	GGC GAT TAT GTC GAA CAA GGT CTG G
	v1637a-3a	Reverse	3a/ patient 1, 4, 5	GGA ATC ATT ACA GTT CAG GGC AGT
	v2191a-3a	Reverse	3a/ patient 1, 4, 5	TGC CAA AGC CGG TAT GGG TAG
	v1364a-3a	Reverse	3a / patient 4, 5	AGC TAT TAT GTC GAA CAA GGT CTG G
	v1443a-3a	Reverse	3a/patient 5	ATG ATG ATA GCG ACC TTG GCC C
	v843s-1a	Forward	1a /patient 2, 7, 9	CAC CAT GGG TTG CTC TTT CTC TAT C
	v1293s-1a	Forward	1a /patient 2, 7, 9	ATG GCA TGG GAT ATG ATG ATG
	v1992s-1a	Forward	1a /patient 2, 7, 9	GGC TGC ACC TGG ATG AAT GC
	v2305s-1a	Forward	1a /patient 2, 7, 9	CGA CAG GGA CAG GTC CGA GC
	T7 (pcDNA3.1D)	Forward	1a /patient 2, 7, 9	TAA TAC GAC TCA CTA TAG GG
	v1433a-1a	Reverse	1a /patient 2, 7, 9	CAT GGA GAA ATA CGC TAT GC
	v1711a-1a	Reverse	1a /patient 2, 7, 9	CTA CAG CTG GCC ATC CTC TCG
	v1949a-1a	Reverse	1a /patient 2, 7, 9	GAC GTC CGT GTC ATT GCT ACC
	v2135a-1a	Reverse	1a / patient 2, 7, 9	CCA CAC CGA GAG TAT GTG GC
	v2372a-1a	Reverse	1a / patient 2, 7, 9	CAC GGA AGG ACC TGC CAG TG
	v2582a-1a	Reverse	1a / patient 2, 7 9	TTA CGC CTC CGC TTG GGA TAT GAG TAA CAT CAT
	BGH (pcDNA3.1D)	Reverse	1a / patient 2, 7 9	TAG AAG GCA CAG TCG AGG

3.6.3 End-stage limit dilution

Simmonds *et al.* previously described the technique for single genome amplification (SGA) (Simmonds et al., 1990). Here, modifications were made to the original protocol for estimating the quantity of the 3.4kb and 1.8kb HCV cDNA fragments in the hepatic and plasma samples. A volume of 1.5µl cDNA was added to 100µl of 1st round Phusion PCR mixture, which were afterwards split into 20µl aliquots. The fifth aliquot was added to 80µl of new Phusion PCR mixture (dilution 1/5) and the mixture divided into 20µl aliquots. This procedure was repeated until the quadruplicate samples had reached a dilution factor of 1/125. The schematic overview of the end-stage limit dilution assay can be viewed in Figure 3.3. After the 1st round PCR amplification, 1µl of each 1st round PCR products were transferred to 20µl of 2nd round Phusion PCR reaction mixture and re-amplified using the PCR program for the nested HCV amplicons (see section 3.6.2).

The nested PCR products were separated on 1% agarose (Bioline) gels and the number of negative reactions noted for each dilution. The fraction of negative reactions were used in the equation for the *null class of the Poisson distribution* to estimate the concentration of HCV cDNA copies that had been synthesised ($\mu = -\ln(F_0)$, where μ is the number of HCV copies and F_0 the fraction of negative reactions) (Figure 3.4). Calculation of the cDNA concentrations for samples collected from patient 1, 2, 4, 5, 7 and 9 can be found in Table 3.4.

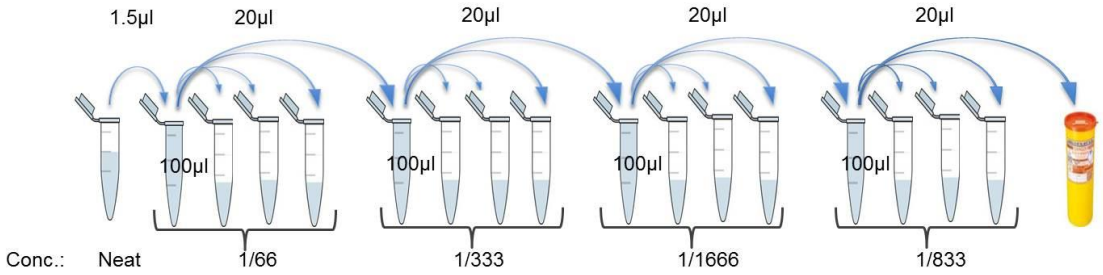
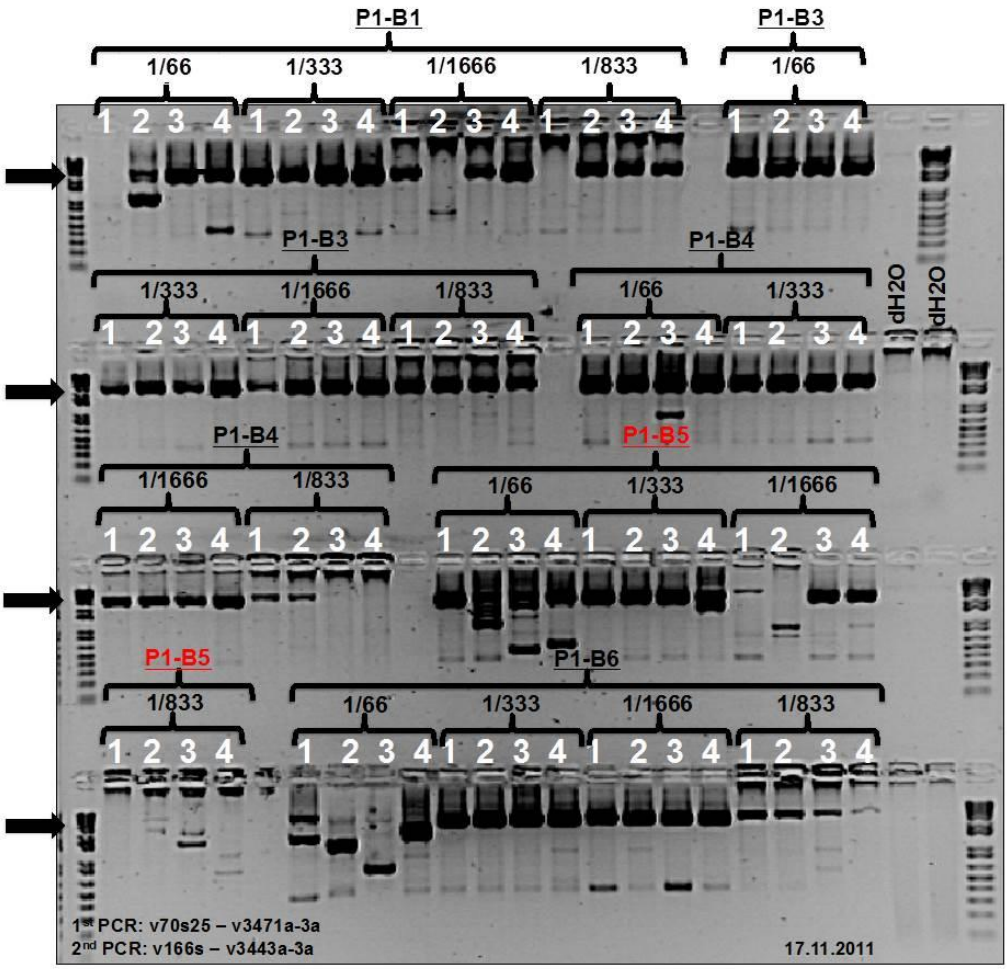


Figure 3.3 End-stage limit dilution PCR.

A limited dilution of the patient-derived HCV cDNA was made for the 1st round of PCR; 1.5µl of the template solution was transferred to 100µl 1st round PCR mixture. Volumes of 20µl PCR-template mixture were then split into 4 tubes and the remaining 20µl mixture transferred to 80µl fresh 1st round PCR mixture, making up 100µl (dilution 1:5). The procedure was repeated 3 times. The dilution step for the template cDNA can be viewed below the tubes.



Example P1-B5: One out of four PCR amplifications were negative at dilution 1/1666:

HCV copies at dilution 1/125: $\mu = -\ln(0.25)$ => $\mu = 1.38$ *copies*

Concentration of HCV cDNA copies in the undiluted sample:
μ/µl = (1.38 x dilution factor 1666)/1.5µl => *2299 HCV copies/µl cDNA synthesis*

Figure 3.4 Estimation of HCV cDNA copies.

End-point template dilution by nested PCR was set-up for selected biopsies. All dilutions were performed in quadruplicate. Here the PCR products are shown for patients 1 (3.4kb fragment, indicated with a black arrow). Underneath the calculation of cDNA copies is performed for patient 1, biopsy 5 (P1-B5, red), as an example of *the null class of Poisson distribution* ($\mu = -\ln(F_0$, where F_0 is the fraction of negative PCR reactions).

Table 3.4: Quantification of HCV cDNA copies synthesised (HCV copies/µl)									
Patient	Biopsy 1	Biopsy 2	Biopsy 3	Biopsy 4	Biopsy 5	Biopsy 6	Biopsy 7	Biopsy 8	Plasma
1	11437.8	5715.6	>11437*	11437.8	5715.6	11437.8	ND	ND	5715.6
2	285925.2	142962.6	11866.8	57182.4	>57156	59334	ND	ND	11437.8
4	2290.2	475.2	>2290.2*	>2290.2*	>2290.2*	NA	>2290.2*	2290.2*	11437.8
5	2290.2	2290.2	1141.8	>2290.2*	>2290.2*	>2290.2*	1141.8	475.2	NA
7	92.4	2376	NA	1141.8	92.4	2376	475.2	ND	2376
9	>57156	59334	>57156*	>57156*	2166.1	57182.4	ND	ND	45.54

NA: No Amplification, ND: Not Done,
**: All the quadruplicate samples were positive at the last dilution. Number of molecules is estimated as >(–ln(0.75) x dilution factor)*

3.6.4 Single Genome Amplification (SGA)

The number of HCV cDNA copies (Table 3.4) were used for calculating the volume of cDNA template required for generating 20% HCV positive amplifications in a 96 well plate (~20 HCV cDNA copies). The required amount of cDNA was mixed into 2000µl of 1st round Phusion PCR master mix before transferring 20µl aliquots to the 96-well plate. Nested PCR was performed according to the protocol stated above (section 3.6.2). Positive amplifications were sequenced across the highly variable 5'-region of E2 to reveal any ambiguities in the template. Only PCR products that generated monotypic chromatograms were carried forward for cloning (see section 3.7).

3.7 Molecular cloning

Plasmids carrying the 3.4kb genotype 3a fragment or the 1.8kb genotype 1a fragment were generated for the patient-derived SGA sequencens in order to

stably store the amplified HCV fragment and for future biological characterisation of their phenotypes.

3.7.1 Ligation

The purified PCR products were ligated into the pcDNA™3.1D/V5-His TOPO® vector using the reagents provided in the pcDNA3.1 Directional TOPO™ Expression Kit (Invitrogen). The ligation reaction was carried out in a total of 6µl including 1µl of salt solution, 1µl of the pcDNA™3.1D/V5-His TOPO® vector (15-20ng/µl) and 4µl of purified PCR product. The mixture was allowed to ligate for 10 minutes at room temperature.

3.7.2 Transformation

Chemically competent cells were thawed on ice for 5 minutes and kept on ice during the whole transformation procedure. For each transformation reaction 25µl of the One Shot® TOP10 Chemical Competent *E. coli* (Invitrogen) and the 6µl ligation reaction were carefully mixed and incubated on ice for 30 minutes. Subsequently, DNA uptake was enabled by heat shocking the *E. coli* for 45 seconds in a 42°C water bath. The cells were incubated on ice for 5 minutes before adding 75µl of the provided super optimal broth (S.O.C medium, Invitrogen). The cells were allowed to recover for 1 hour at 37°C with 300rpm before the 100µl cell suspension was spread on 2% Luria Broth Base (LB, Invitrogen) + 1% agarose (Bioline) plates supplemented with ampicillin (100ng/ml) (Sigma-Aldrich).

3.7.3 PCR colony screening

Overnight colonies of transformed One Shot® TOP10 Chemical Competent *E. coli* were PCR screened for the desired insert by using primers targeting the HCV insert and the pcDNA™3.1D/V5-His TOPO® plasmid backbone (see Table 3.3). The PCR master mixture was made from the reagents in the GoTaq® DNA Polymerase kit (Promega) and consisted of 1x Go Taq Buffer, 8% $MgCl_2$, 0.13µM forward and reverse primers and 0.75U/µl Go Taq Flexi polymerase, together with 0.1µM dNTPs (Invitrogen). Single colonies were picked up with sterile toothpicks and spread on to 2% LB plates supplemented with ampicillin (100ng/ml) (Sigma-Aldrich). After the spread plating the toothpicks were dipped into wells containing 15µl of GoTaq® PCR master mixture. The screening plates were incubated at 37°C overnight to allow colonies to form. The PCR reactions containing the matching colonies were screened using the following program:

5 min	95°C	
15 sec	94°C	⎫
30 sec	57°C	⎬ 15 cycles
2 min	72°C	⎭
∞	4°C	

The PCR products were separated on 1% agarose gels and positive PCR amplifications evaluated by the size of the amplified product.

3.7.4 Plasmid purification and restriction digestion

Two HCV positive colonies from each transformation were transferred from the PCR screening plate to 5ml of 2% LB medium, supplemented with ampicillin (100ng/ml) and the suspension incubated at 37°C with 300rpm overnight. The following day the plasmid DNA was isolated using the Qiaprep Spin Miniprep Kit (Qiagen) according to the manufacture's instructions. Briefly, the 5ml bacterial culture was pelleted by 5 minutes centrifugation at 12.500 *x g*. The supernatant was discarded, 250µl Suspension Buffer P1 (containing LyseBlue) added and the mixture vortexed thoroughly until the pelleted bacterial cells were completely re-suspended. A volume of 250µl ice-cold Lysis Buffer P2 were added and the tube inverted gently until the suspension had turned blue. The homogenously coloured lysate then had 350µl of Neutralising Buffer N3 added and the tube was inverted until the blue colour had disappeared. The lysates were transferred to 1.5ml sterile eppendorfs, centrifuged for 10 minutes at 20,000 *x g* and the supernatant transferred to the Qiaprep spin column. The spin column was centrifuged for 1 minute at 20,000 *x g*, the filtrate discarded and 750µl of Washing Buffer PE applied to the filter. The column was centrifuged again and the filtrate removed before a further centrifugation step for 1 minute at 20,000 *x g*. Subsequently, the Spin Column was transferred to a Collection Tube and the filter allowed soaking for 1 minute in 40µl of Elution Buffer EB. The plasmid DNA was eluted by spinning the column for 1 minute at 20,000 *x g*. The purified plasmids were quantified by spectrometry before storing at -20°C.

The correct size of the inserted fragment was evaluated by digesting in the multiple restriction sites on each side of the insert, followed by gel electrophoresis. The restriction digestions were performed on 1µg plasmid in the presence of 10 units of *XbaI* (Promega), 10 units of *BamHI* (Promega), 0.1µg/µl acetylated bovine serum albumin (BSA, Promega) together with the buffer recommended by the manufacture (1 x Buffer D), making up a total volume of 20µl. The digests were performed for 60 minutes at 37°C.

In total 120 HCV clones were generated from liver and plasma samples collected from patients 1, 2, 4, 5, 7 and 9 (Table 3.5). In average 20 clones were generated per patient. The naming system for the clones is based on the well they were identified in. The first 2 characters refer to the initial positive well during the end-point dilution PCR and the last 2 characters derive from the colony screening PCR. This way all clones can be traced back to the day of amplification and the initial E2 sequencing. A few clones are giving numeric first characters, these refers to the dilution of the cDNA template from which they where amplified.

Table 3.5: Patient-derived HCV clones verified by Sanger sequencing

Genotype	Patient	Biopsy 1	Biopsy 3	Biopsy 6	Biopsy 7	Plasma	Total
1a	2	G4-B2 G5-F11 F8-F5 A9-G1 A12-A5 D8-B4 H7-C12 F7-F5 E2-C8 F6-H8 H3-F11	ND	H4-D11 H10-F12 E6-G11 C10-C7	ND	B5-B11 G6-C8 G10-D3 C2-C11 All-D2 D12-E3 G11-D6	22
	7	C3-C11 C8-D4 C6-F3 F2-D11 F8-A5 G7-F5 F5-B1 F9-A8 D9-D2	ND	A2-E8 C7-B8 B7-C9 F11-E11 F9-B6	ND	E8-D4 B8-C12 1/125-4-B11 1/5-3-G4 G7-F12 A9-E3 E7-B1 E2-A10 A9-C6 B9-A4 A2-G5	25
	9	C8-G8 B9-C2 A10-F3 C7-D5 G6-H8	ND	B12-B2 C3-G10 D10-G1 G6-C12 C12-E4	ND	D7-D9 B2-G3 1/1-2-C5 1/1-3-B12 A5-H6 E8-D4 F11-E4	17
3a	1	C2.1-B7 C2.2-D1 B2-F7 B4-H2	ND	H3-1.6 D6-G10 D7-D12 B6-F4	ND	F6-E3 A5-G2 B10-F2 D11-A11 ¼-1-1B ¼-3-F3 ¼-4-A5 ¼-6-E8 1/8-1-B12	17
	4	B11-A4 C2-A6 C4-G11 C2-B7 F5-D2 E8-G2	E9-B2 E11-C3	NA	F3-E3 F9-H6 H3-C12 G7-E3 B6-C8	D7-E2 C3-F3 E6-G5 G7-C8 G12-A9 H8-H10 D10-B12	20
	5	C2-B1 C9-C11 C10-D9 C8-D11 F4-C2 G4-E3 D9-A12 G9-A10 G7-C6	F3-B9 F11-F2 G2-G2 F5-C3	1/125-4-D10 D7-E5 B7-G9 H6-H11 C12-E9 B2-H6	ND	NA	19

NA: No Amplification, *ND: Not Done,*

3.8 Sanger sequencing

All sequencing was performed by the Functional Genomics Lab at the School of Biosciences, University of Birmingham, UK, who used a 3730 DNA Analyzer (Life Technologies). For each sequencing assay 500ng template was used together with 1nmol HCV specific primer.

3.8.1 Sequencing sensitivity assay

The concentration of the patient 1 and patient 4 nested bulk PCR products were quantified by spectrophotometry. Patient 1 biopsy 8 and Patient 4 biopsy 4 had comparable DNA concentrations (560ng/µl) and were directly mixed at ratios of 0:10, 1:9, 2:8, 5:5, 8:2, 9:1 and 10:0. The mixtures were sequenced with the v1293s-3a primer (Table 3.3) to identify ambiguities within the HVR1.

3.9 454 ultra-deep sequencing

The 454 ultra-deep sequencing were performed and analysed by our collaborators, Damien Tully, Karen Power ad Todd Allen from the Ragon Institute of MGH, MIT and Harvard at Harvard University, USA. In brief, cDNA was generated using a one-step RT-PCR for the core, E1, E2, p7 and NS2 fragment. The amplified cDNA sequences were digested into smaller fragments (300-600bp) and ligated to adaptor sequences with individual barcodes (Nextera Tagmentation), allowing the HCV cDNA sequence to bind agarose beads whose surfaces carried oligonucleotides complementary to the ligated adaptors in the cDNA library. Each bead was associated with a single

fragment. The bead complexes were subsequently isolated in oil:water micelles containing PCR reagents. Thermal cycling (emulsion PCR) of the micelles resulted in the formation of thousands of clonal sequences on the surfaces of each bead, which were afterwards transferred to individual wells on a PicoTiterPlate where the sequencing reaction could be performed. During the sequencing reaction the Roche 454 genome sequencer introduces the nucleotides in a step-wise manner, allowing the DNA polymerase to incorporate one nucleotide at a time. Once incorporated, the nucleotide releases a pyrophosphate, which initiates a series of downstream reactions leading to the production of light. The emitted light is recorded by a charge-coupled device (CCD) camera and translated into the genetic code of the sequenced fragment. The nucleotide sequences were used for generating a *de novo* HCV consensus genome using the Assemble Viral 454 (Henn et al., 2012) and VICUNA *de novo* assembler (Yang et al., 2012) and the detection of intra-host HCV variations was assessed by aligning the individual reads against the *de novo* assembled consensus sequence. The validation of true polymorphisms was done by ReadClean 454 and V-phaser algorithms (Macalalad et al., 2012), which detect any processing errors such as homopolymers and carry-forward-incomplete-extensions (CAFIE) miscalls by considering the frequency of the polymorphisms compared to the coverage. Information about the number of reads and sequence coverage for each sample can be found in Table 3.6. The 454 ultra-deep sequencing process information was provided by Dr. Damien Tully and from Mardis *et al.* (Mardis, 2008) and Schuster *et al.* (Schuster, 2008).

Table 3.6: 454 ultra-deep sequencing coverage

Patient	Genotype	Sample	Number of reads	Total bases	Mean Sequence Length (± STD)	Mean Coverage (± STD)
1	3a	Plasma	2,681	1,004,840	375 ± 132	295 ±113
		Biopsy 1	1,631	578,576	355 ± 140	176 ± 63
		Biopsy 2	3,743	1,345,717	360 ± 133	420 ± 124
		Biopsy 3	3,549	1,285,335	362 ± 135	345 ±132
		Biopsy 4	3,676	1,355,148	369 ± 132	412 ± 144
		Biopsy 5	2,717	978,419	360 ± 133	297 ± 115
		Biopsy 6	4,911	1,792,511	365 ± 134	540 ± 184
		Biopsy 7	3,483	1,300,182	373 ± 133	400 ± 130
		Biopsy 8	2,594	948,527	366 ± 133	247 ± 123
2	1a	Plasma	5,756	2,094,910	363 ± 129	623 ± 222
		Biopsy 1	4,968	1,781,786	359 ± 131	539 ± 80
		Biopsy 2	5,549	1,979,135	357 ± 131	587 ± 226
		Biopsy 3	5,433	1,941,660	357 ± 131	579 ± 198
		Biopsy 4	6,814	2,438,294	358 ± 133	726 ± 261
		Biopsy 5	5,296	1,902,800	359 ± 132	566 ± 222
		Biopsy 6	6,514	2,342,614	360 ± 133	696 ± 242
		Biopsy 7	9,328	3,276,485	351 ± 135	973 ± 364
		Biopsy 8	8,708	3,083,990	354 ± 130	915 ± 312

3.10 Sequence analysis

3.10.1 Phylogenetic analysis

The *CLC Main Workbench 6* (CLC Bio) software was used for analysing the genetic distances for the HCV variants. Phylogenetic Neighbour joining modelling was performed with 1000 bootstraps and the Hasegawa-Kishino-Yano (HKY) substitution model with a transition/transversion ratio of 2.0. Phylogenetic tree were generated in FigTree version 1.3.1. (Andrew Rambaut, Institute of Evolutionary Biology, University of Edinburgh).

The Slatkin-Maddison tree-based statistical model was used for estimating the minimum number of migration events required to predict the observed spatial distribution of sequences between sites (Beerli et al., 2001). The Slatkin-Maddison modelling were performed with the HyPhy 2.220130723beta(MP) software, created by S. Pond, A. Poon and S. Frost, Antiviral Research Centre, University of California, USA, and S. Muse, Statistical Genetics, North Carolina State University, USA. The analyses were performed with the maximum number of unique permutations available. *P*-values below 0.05 were considered statistically significant evidence that the sequences were compartmentalised (Beerli et al., 2001).

3.10.2 Shannon entropy

Shannon Entropy is the most widely used informational theory to quantitatively measure uncertainties in a data set. The entropy measure can be used as a strategy to calculate sequence variability for a single column of nucleotides in a sequence alignment (one-way entropy) or between two sequence alignments (two-way entropy). The model considers both nucleotide frequencies (e.g. 50% G and 50% C will have a higher entropy than 80% G and 20% C) and the number of nucleotide possibilities (e.g.. 80% T, 10% A and 10% C will have a higher entropy than 90% T and 10% A) (Sherwin, 2010). This gives each position an individual entropy score from 0 (no variability) to 1 (maximum variability).

One-way and two-way entropies were found for each nucleotide in the SGA alignments using the Shannon Entropy software on the http://hcv.lanl.gov website.

3.10.3 Synonymous and non-synonymous changes

The number of synonymous (*ds*) and non-synonymous (*dn*) changes were found for each codon in the SGA-derived HCV sequence alignments using the SNAP v1.1.1 program on the

http://hcv.lanl.gov/content/sequence/HCV/ToolsOutline.html website.

3.10.4 BLOcks SUbstitution Matrix (BLOSUM)

In nature amino acid substitutions do not occur randomly, some are more likely to happen than others. For instance, tryptophan (W) will only on very rare occasions mutate, whereas histidine (H) more often changes to arginine (R). The BLOSUM will weigh these substitutions according to the likelihood of them occurring and give a similarity score between the analysed sequences. As a result BLOSUM matrices are not based on explicit evolutionary modelling, but instead on counting the relative frequency and the substitution probabilities between protein families. The BLOSUM62 matrixes for the SGA-derived HCV sequences were generated in the ClustalW2 program

(https://www.ebi.ac.uk/Tools/msa/clustalw2, European Bioinformatics Institute, UK).

3.11 Statistics

Most statistical analyses were performed using the Prism 5.0 software (Graphpath, USA) and the results expressed as mean ± standard deviation, unless otherwise stated in the figure legends. Tests were performed using the two-tailed paired or unpaired t-test with a confidence interval of 95% and corrected for multiple comparison by the Bonferroni method. The Boniferroni calculation for multiple comparisons was also used in the post-test for the one-way and two-way ANOVA. Calculations of the parametric Pearson's correlation were applied to all correlation analyses.

The multiple component linear regression modelling was performed in the R commander 3.0.2 software (The R Foundation for Statistical Computing 2012, USA) and scatter plot matrixes generated and exported.

4 HCV RNA BURDEN IN THE CHRONICALLY INFECTED LIVER AND PLASMA.

Previous studies have shown that HCVcc can disseminate to naïve cells via cell-free and cell-cell routes *in vitro*. Cell-cell transmission results in direct viral transfer to adjacent cells without an extracellular stage, leading to the formation of infected foci in the presence of nAbs (Brimacombe et al., 2011; Timpe et al., 2008). Confocal imaging of liver biopsies has demonstrated similar foci of HCV antigen presenting cells consistent with the hypothesis that cell-to-cell route of virus dissemination occurs *in vivo* (Liang et al., 2009; Wieland et al., 2013). HCV cell-cell transmission in the infected liver may support a distribution model where a founder virus infect the hepatocyte followed by the dissemination of the virus progeny to adjacent cells, forming defined foci across the liver. Formation of HCV positive foci in the liver would imply that not all hepatocytes are permissive to HCV infection and that the virus may form replication hot spots in distinct areas of the liver.

Clinical monitoring of HCV replication is usually performed with RNA isolated from plasma and only rarely performed on liver biopsies. As a result most studies on viral burden have been conducted on plasma samples. Of those studies which have investigated liver biopsies, most have only included one single needle biopsy (Ballardini et al., 1997; Cabot et al., 2000; Descamps et al., 2012; Idrovo et al., 1996; Sugano et al., 1995) and only in a few cases

have the investigators looked at more than 2 sites in the liver (Cabot et al., 1997; Sakai et al., 1999; Terrault et al., 1997). For this study we investigated the distribution of HCV RNA copies across 8 anatomical segments in 22 subjects undergoing liver transplantation for HCV associated end-stage liver disease (Table 4.1) and compared these hepatic viral burdens with the plasma load.

Table 4.1: Adult patient overview*	
Patients (n = 22)	
Age (year)	35 – 64 (Median 57)
Plasma viral load (HCV copies/ml)	274 – 1.1 x 10^7 (Median 2.2 x 10^5)
HCV Genotype:	
Genotype 1a/b	9
Genotype 2a	1
Genotype 3a	6
Genotype 4a	1
Unknown	5
Indication of transplantation:	
Cirrhosis	17
Hepatocellular Carcinoma (HCC)	5
Previous treatment with IFN:	
Non-responders	8
Treatment naïve	3
Relapsers	5
Not tolerated	2
Unknown	4
Average time until sampling processing post-transplantation (hours)	2.3 – 22.5 (median 6.7)

* A more detailed characterisation of each patient can be found in section 3.1.1.

In this thesis, the 8 hepatic specimens will be referred to as biopsies and in some cases be shortened to B1, B2, … B8 and the patients they were collected from will be referred to as P1, P2, … P23.

4.1 Optimising detection of HCV in human specimens

Although HCV is a single stranded RNA virus, double stranded RNA (dsRNA) structures, such as hairpin loops, occur both in the coding and non-coding regions (Tuplin et al., 2012). Moreover, RNA duplexes of negative and positive strands are generated during the HCV replication.

Quantification of the HCV RNA requires that a specific HCV probe bind to its ssRNA target (in this study a 88nt conserved region in the 5'UTR). Any dsRNA structures in the HCV genome may inhibit this binding and underestimate the amount of HCV RNA present. Previous studies on dsRNA reovirus indicate that high temperatures increase the formation of viral ssDNA (Gomatos and Tamm, 1963a, b). In addition, researchers at the Mount Sinai Medical Centre, NY, USA, observed an increase in PCR detectable HCV RNA (up to 2-logs) following 106°C heat-treatment of total liver RNA from chronically infected individuals (unpublished data, presented at the Annual HCV meeting, 2011)..

The one-step HCV qRT-PCR assay used in this thesis requires several different temperature steps ranging from 50 to 95°C (see section 3.5). It was therefore decided to evaluate dsRNA denaturation during these different temperatures. A melting-curve was made using 1µg of a commercial dsRNA ladder together with SYBR-green fluorescence probes. As the temperature of the mixture increased from 37 to 99°C the SYBR-green fluorescence declined, indicating denaturation of dsRNA (Figure 4.1A). At 52°C the fluorescent intensity reached 50% of the maximum, indicating that half the dsRNA had denatured (Tm_{50} = 52°C). A plateau of minimum fluorescence was reached at 85°C (~3700 fluorescent intensity), suggesting complete dsRNA denaturation at this temperature. Decreasing the temperature for the same samples to 37°C resulted in limited RNA re-annealing, and at 37°C 26% of the original dsRNA formations were reformed (Figure 4.1A). These data show that maximum denaturation occurs within the temperature range of the one-step HCV qRT-PCR assay and that only limited re-denaturation occurs at the 50°C cDNA synthesis step.

Several compounds have been reported to improve the efficiency of the enzymes used in cDNA synthesis, PCR amplification and qRT-PCR quantification by reducing secondary RNA (or DNA) structures in the template. These compounds include the disaccharide trehalose (Spiess et al., 2004), betaine (Schwinefus et al., 2013) and DMSO (Mamedov et al., 2008). The 3 compounds were tested in our melting curve assay with the aim to optimise the dsRNA denaturation assay further. Each compound was added to a

SYBR-green solution containing 1µg of a dsRNA ladder, which were then incubated at increasing temperatures going from 25 to 99°C over 30 minutes (Figure 4.1B). Samples containing DMSO showed a reduction in fluorescent intensity of approximately 44% at 25°C compared to the control sample contained dsRNA. In comparison betaine and trehalose had 4% and 6% lower levels of fluorescence intensity at 25°C, respectively. This observation may suggest that the compounds result in dsRNA denaturation even at room temperature. Alternatively the compound could have a negative effect on the SYBR, resulting in reduced fluorescent independent on denaturation.

Increasing the temperature towards 99°C showed that samples containing DMSO reached the minimum intensity plateau (approximately 3100 fluorescent intensity) at a lower temperature (85-90°C) than any of the other solutions (94 to 98°C) and also had a lower Tm_{50} (Figure 4.1B). These melting assays led us to investigate the effect of DMSO on the quantification of HCV RNA by qRT-PCR. Since the 5 and 10% DMSO only showed minor difference in Tm_{50}, we chose 5% DMSO to limit any potential toxicity effects on the enzyme activity.

HCV specific qRT-PCR master mixes were made with or without 5% DMSO together with total hepatic RNA isolated from patient 1 (P1-B1 and P1-B5) and patient 5 (P5-B1 and P5-B7). The assay included total RNA from a HCV negative liver with alcoholic liver disease (ALD) and the full-length HCV H77

RNA genome, which had been transcribed from a vector (Figure 4.1C). The presence of 5% DMSO resulted in a higher mean HCV RNA detection for the patient derived samples (patient 1: 1.9-fold, patient 5: 2.6-fold). However, the difference was only statistical significant for P1-B5 (two-tailed unpaired t-test $p = 0.03$). The laboratory strain H77 was included as a negative control for the presence of positive-negative HCV RNA strand formations. The lack of increased H77 RNA detection in the presence of DMSO indicates that the H77 genome, and potentially also patient-derived HCV, may not contain any dsRNA folds in the amplified region (Figure 4.1C). This supports our conclusion that the small changes in viral load seen for the hepatic RNA extractions may be the result of RNA duplex denaturation.

Although the effect of 5% DMSO in the qRT-PCR only had a small effect on the amount of HCV RNA detected, it did not have any toxic effect and it was decided to use 5% DMSO in all subsequent measurements of HCV RNA.

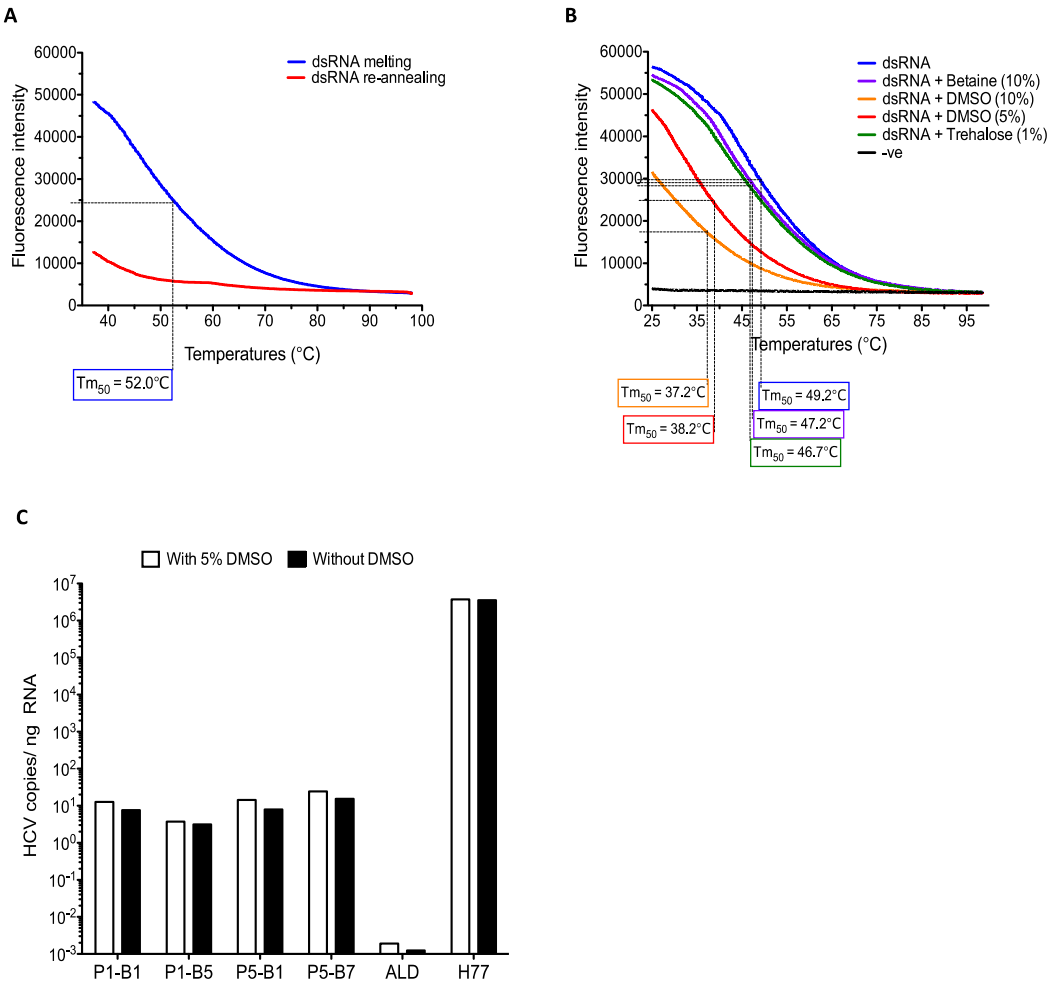


Figure 4.1 Optimisation of ssRNA detection.

One µg of a commercially obtained dsRNA ladder (Promega) was used as template for the melting curve assay. The amount of dsRNA was detected by the intercalation of SYBR-green fluorescent dye into dsRNA. The Stratagene qPCR machine was set to record SYBR-green fluorescence every second over 30 minutes. **A.** The dsRNA ladder was incubated at increasing temperatures over a 30-minute period (blue line). Following heat denaturation the template was re-cooled towards 37°C (-1°C/minute, red line), allowing the RNA to re-anneal. **B.** Different additives were tested for their ability to accelerate dsRNA denaturation during heat-treatment. The dsRNA templates were either mixed with betaine (10%, purple line), DMSO (10%, orange line), DMSO (5%, red line), trehalose (1%, green line) or water (blue line) before being heated from 37 to 99°C over a 30-minutes period. H_2O with SYBR-green master mix were included as an RNA negative control (black line). The coloured boxes indicate the melting temperature when 50% denaturation occurred (Tm_{50}). **C.** HCV specific qRT-PCRs with or without 5% DMSO were performed with total hepatic RNA from patient

1 (P1), biopsy 1 (B1) and 5 (B5) and patient 5 (P5), B1 and B7 in duplicates (data represented as a mean). Total hepatic RNA was included from alcoholic liver disease (ALD) together with the full-length HCV H77 RNA genome. The absolute quantification of HCV copies were interpolated from an internal standard curve (Primer Design Ltd.) in the assay (data not shown) and the viral load normalised to the total amount of RNA in the samples.

4.2 Identification of stable housekeeping genes in HCV related end-stage liver disease

Investigation of HCV between specimens requires a stable reference for normalisation. In most of this chapter, the amount of total liver RNA (ng) was used as a reference, but in a few cases an internal mRNA reference had to be used. Waxman and Wurmback have previously shown that HCV positive patients with HCC exhibit a large variation in the expression of housekeeping genes compared to patents with early stage fibrosis or cirrhosis (Waxman and Wurmbach, 2007). It was therefore important to establish an optimal reference gene to enable HCV normalisation between samples obtained from our patient cohort.

The expression of 12 common reference genes was measured with the GeNorm assay. For this experiment, biopsy 1 from three subjects with HCC (patients 4, 7 and 14) and two with cirrhosis (patient 2 and 17) were chosen as templates. The expression of the 12 candidate genes in the GeNorm panel was measured by target-specific qRT-PCR. For the five hepatic samples evaluated, only candidate genes with Ct-values between 15 and 25 were considered to have reliable expression patterns. Initial analysis of the raw Ct-values indicated high expression of 18-sRNA (Ct-value <15) and low expressions of RPL13A, CYC1 and TOP1 (Ct-value > 25). These genes were therefore excluded from further investigation (data not shown). Eight genes, ATCB, GAPDH, UBC, B2M, YWHAZ, EIFA2, SDHA and ATP58 remained in the analysis. The expression of each of these 8 genes was afterwards used as

internal references for normalising the expression levels for the other 11 genes included in the array (Figure 4.2). Normalising all genes with the same internal reference revealed large variation in the 2^Δ Ct values. However, comparing the values between patient samples normalised with the same reference gene did not reveal any significant differences in expression levels (One-way ANOVA for each reference gene, p = 0.35). In addition, two-way ANOVA of the entire data set did not reveal any systemic variance of the housekeeping gene expressions for the different hepatic templates (p = 0.74). Separate analyses of the variability between normalising with different reference genes and between the expression patterns for different patients, revealed that 1.21% of the systemic variation (p = 0.56) was due to normalisation with different reference genes. However including different patient templates accounts for 2.74% of the variation (p <0.05). This observation is likely to be the result of larger variation seen in the gene expressions for patient 17 (Figure 4.2). Excluding patient 17 from the data set makes the observed variation between patient templates non-significant (p = 0.6).

The low variability for normalising with different reference genes means that any one of the 8 housekeeping genes could be used as a reference for normalising HCV RNA in the 22 explanted livers. Given this result, we decided to employ the widely used GAPDH as the reference gene for normalising RNA quantities in samples where the total amount of liver RNA cannot be determined.

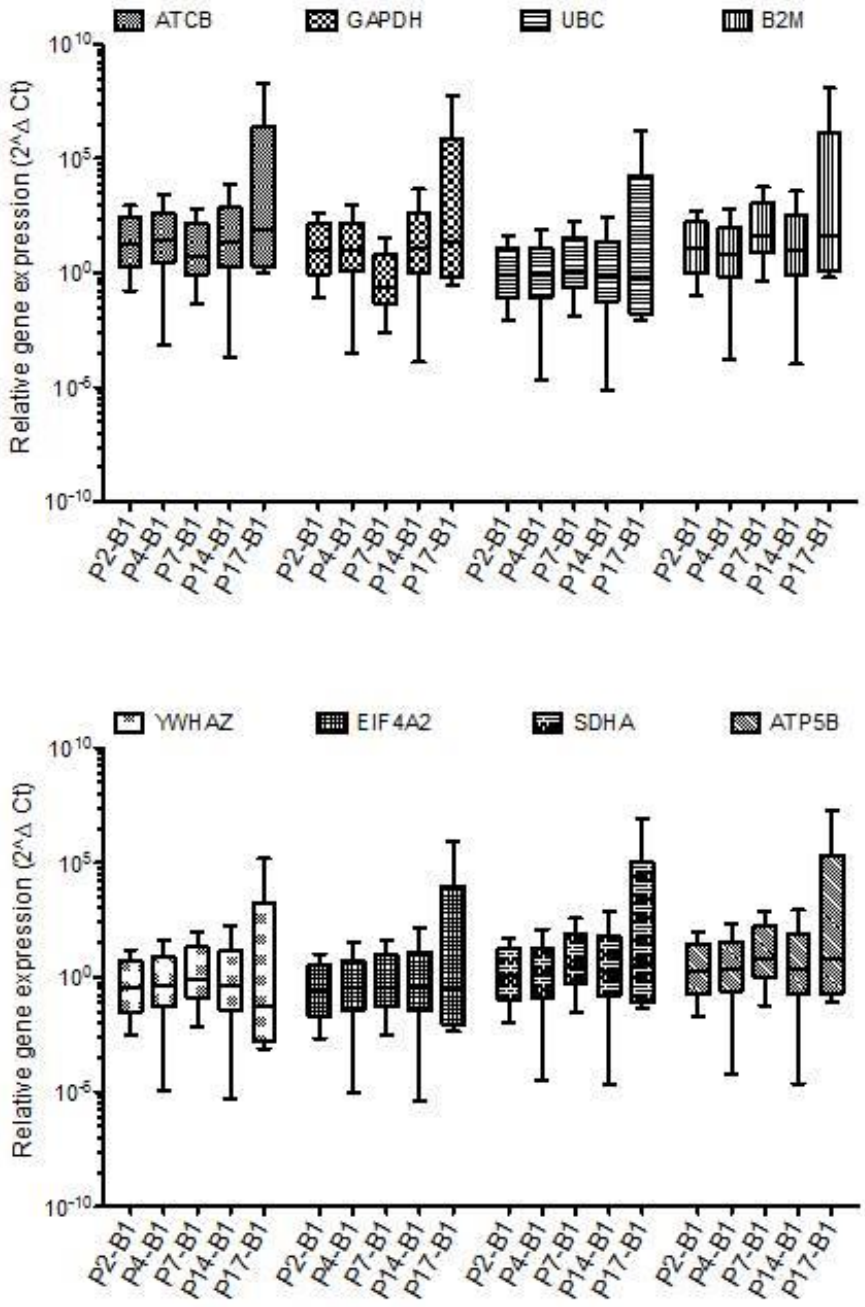


Figure 4.2 Evaluation of reference genes in end-stage liver disease.

The GeNorm panel of 12 different housekeeping genes were used for evaluating the most stable reference genes among HCV positive patients with end-stage liver disease. Biopsy 1 from two patients with cirrhosis (patients 2 and 17) and three patients with HCC (patients 4, 7, 14) were used as template. Between 63 and 270ng total liver RNA was used for each evaluation. The expression levels were measured

by qRT-PCR with target specific probes. Four of the 12 genes had Ct values either above 25 or below 15 and were not included in subsequent analysis. The raw Ct values for each of the remaining 8 genes were used in turn to normalise the remaining 11 expression levels (Δ Ct) and the fold difference calculated (2^Δ Ct). The data is presented as median (horizontal line), upper and lower quartile and the maximum and minimum values. One-way ANOVA between gene expressions normalised with the same reference gene did not reveal any significant differences between the patients (in all cases $p > 0.05$). In addition, the data set did not show any systemic variance by two-way ANOVA ($p = 0.74$).

4.3 HCV burden across the liver explant

To this date, it is not known whether the HCV chronically infected liver is homogenously infected or if infected cells exists in pockets of isolated populations. In order to investigate hepatic HCV distribution, multiple biopsies were taken from 22 HCV infected explanted livers. Twenty-one of the patients had 8 biopsies taken and one (patient 10) had 4 biopsies collected because of a large hydatid cyst in the left lobe due to an *Echinococcus* infestation. Patient 12 samples are not included in the analysis as the time between the removal of the explant and sample processing exceeded 24 hours.

As biopsy sampling is highly invasive it would not be ethically possible to take 8 samples from a functioning liver, justifying the use of samples collected from explanted HCV positive livers following liver transplantation. All 22 patients gave informed consent for the donation of their explant liver and for blood sampling taken at the time of transplantation. The patients included in this study were mainly infected with genotype 1 (n=9), followed by 3a (n=6), 2a and 4a (n=1 and n=1), which agree with the local geographical distribution of these HCV genotypes (Mohd Hanafiah et al., 2013; Sharma and Feld, 2014; Simmonds, 2004). The higher number of patients with cirrhosis compared to hepatocellular carcinoma is in agreement with the previous observation that approximately 20% of chronic infected patients will proceed to develop cirrhosis and 2.5% will develop HCC (Hoofnagle, 2002). An overview of the patient characteristics can be found in Table 4.1.

Samples were taken form the 8 anatomical segments in the liver less than 24 hours after it was removed from the transplant recipient (see section 3.1.1). Total liver RNA was extracted from the hepatic biopsies and the viral load measured by HCV specific qRT-PCR. Each of the 8 biopsies from the 22 included livers was evaluated in duplicate by two experimental repeats (n=2). In Figure 4.3A the results are shown for two randomly selected patients (patient 1 and 2). The repeated measurements for each biopsy varied by less then 0.5 log and are therefore within the expected variation of the assay (0.5-log). This pattern was also observed for the remaining 22 patients when measuring duplicate samples in two experimental repeats.. These data justify the use of a mean viral load for each biopsy in subsequent analyses. The mean biopsy loads for the remaining patients are shown in Figure 4.3B. Each of the 22 livers had comparable levels of HCV RNA between the 8 biopsies, with a variance ranging from 0.2-log to 0.75-log, however the HCV RNA levels between patients varied up to 2-logs (Figure 4.3B). Comparison of the different hepatic anatomical sites showed no correlation between the viral burden and the anatomical origin of the sample. (Figure 4.3C; one-way ANOVA; $p = 0.44$), suggesting that no single site within the liver hosts a higher viral burden than others.

Patient 14 had a mean viral load of 1.4 HCV copies / ng total liver RNA, which is below the estimated assay's threshold (10 copies / sample, McKeating group protocol applied to Primer Design reagents in accordance with manufacturer's instructions). This patient is not considered HCV negative as

viral copies were detected in the plasma at the time of transplantation (3.9 x 10^4 HCV copies /ml). Likewise, it is unlikely that the low detection is due to errors in the RNA extraction, since all 8 hepatic extractions contained high levels of total RNA and gave typical signals with the GAPDH endogenous control (data not shown). The apparent low level HCV RNA in patient 14 may either be real or a case of poor qRT-PCR primer affinity.

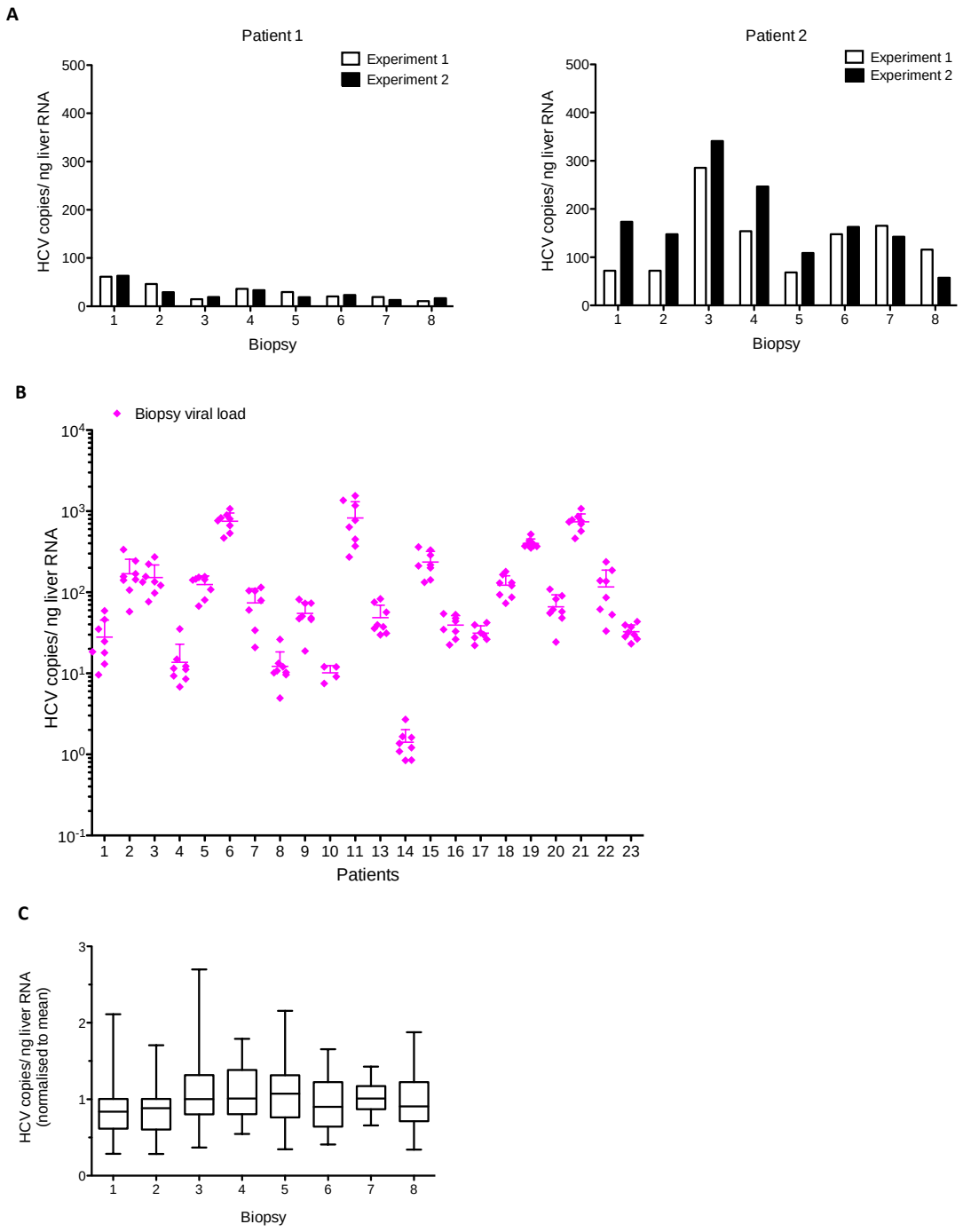


Figure 4.3 HCV RNA expression in the liver explant.

Biopsies were taken from the 8 anatomic segments in 22 explanted livers and the total hepatic RNA extracted. The amount of HCV RNA per biopsy was measured for duplicate samples in two experimental repeats (n=2) and the absolute value calculated from an internal standard curve. The HCV copy number was normalised to

the total amount of liver RNA in the particular sample. **A.** The data are here shown for patient 1 and 2. The plotted bars represent the mean measurements with standard deviations. **B.** The mean HCV copies / ng total liver RNA for the 22 patients included in the study represented as a scatterplot. Each pink diamond represents one biopsy and the mean ± standard deviations are shown for each set of 8 biopsies. **C.** The HCV copies / ng liver RNA for each biopsy was normalised to the mean load for livers. The normalised HCV loads for the 22 patients are stratified according to the hepatic site the biopsy originated from. The data is presented as median (horizontal line), upper and lower quartile and the maximum and minimum spreads. The data set were analysed by the one-way ANOVA, which showed no statistical evidence of variation across the samples ($p = 0.98$).

Linear regression analyses were done for the hepatic viral load versus the patient's age and the time between transplantation and sample processing. Pearson's correlation coefficient (r^2) revealed no association between the viral load and patient's age (r^2 = 0.02, p = 0.051) (Figure 4.4A) and between viral load and processing time (r^2 = 0.0007, p = 0.74) (Figure 4.4B). The hepatic viral loads were also compared between the HCV genotypes 1 and 3 Genotypes 2 and 4 were only represented by one patient each and were therefore not included in the analysis. The hepatic viral load in genotype 1 infected patients were significantly higher than for genotype 3 (Mann-Whitney p = 0.006), consistent with previous publications (Rong et al., 2012; Soriano et al., 2008).(Figure 4.4C). Furthermore we did not find an association between viral load and indication for transplantation (cirrhosis versus HCC, Mann-Whitney p = 0.87, Figure 4.4D) or between viral load and the patient's known response to PEG-IFN response (One-way ANOVA: p = 0.2) (Figure 4.4E).

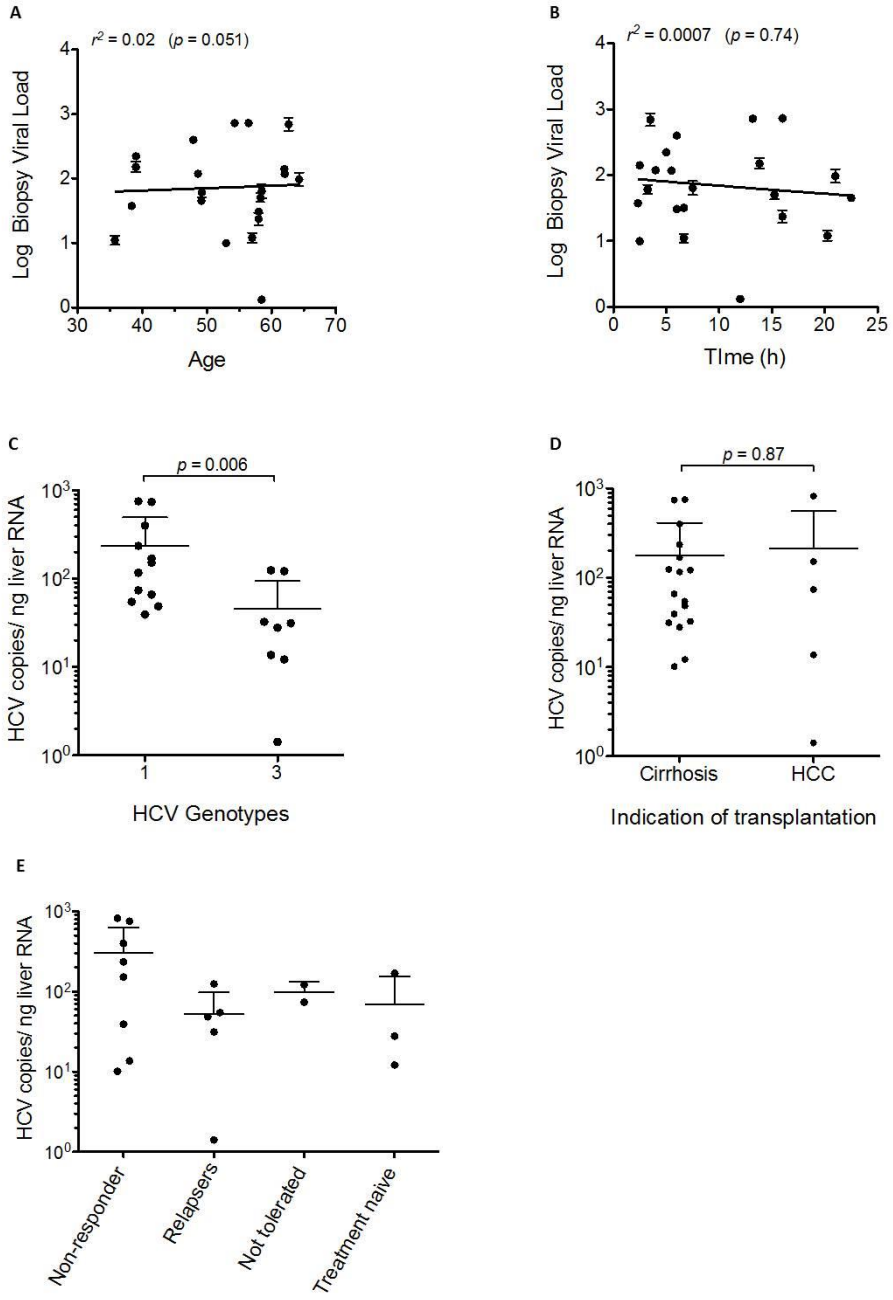


Figure 4.4 Association between hepatic RNA burden and clinical features.

A. Linear regression of hepatic viral burden and the age of the patient. **B.** Linear regression of hepatic viral burden and the time between the liver was removal and biopsies processed. Each patient is represented by the mean viral load (black circle) ± standard deviations and the line of best fit between the data points is indicated in the graph. Pearson's correlation coefficient (r^2) was calculated for each of the data sets (**A** + **B**). **C.** HCV viral loads were compared between patients infected with genotype 1 and 3. **D.** Comparison of viral loads between patients diagnosed with

cirrhosis or hepatocellular carcinoma (HCC). Statistical analysis using Mann-Whitney were applied to the data in **C** + **D**. **E.** Viral load comparisons between patients with known PEG-IFN response. One-way ANOVA revealed comparable viral load between patients who did not respond to treatment, those who relapsed, those who did not tolerate the treatment and those who did not received treatment (p = 0.20). Each black circle represents the mean hepatic load for one patient. The mean and standard deviations are shown for each group.

4.4 HCV RNA burden in hepatic sections

During chronic infection, HCV-induced apoptosis of the hepatocytes triggers a pro-fibrogenic activation of hepatic stellate cells. These activated cells generate fibrotic tissue, which takes over the normal hepatic architecture (Wang et al., 2013). In end-stage liver disease the natural hepatic structure is usually lost and the hepatocytes are isolated in islets (nodules) surrounded by fibrotic tissue. Since hepatocytes are the major reservoir supporting HCV replication, differences in viral populations may exist between these healthy nodules, leading to hepatic compartmentalisation of the HCV RNA at the microscopic level.

Hepatic samples ($1cm^3$) collected from patients 2, 7 and 9 (genotype 1a) were sectioned into 12 x 20µm slices. Lysis of the twelve sections for RNA extraction meant we could not directly estimate the exact number of hepatocytes in each section. We therefore decided to estimate the number of hepatocytes in 5µm sections taken directly before and after the twelve sections. These sections were stained by Hematoxylin and Eosin (H&E) and the fraction of hepatocytes estimated by scanning the slides and manually highlighting islets of hepatocyte nodules. The frequency of hepatocytes were comparable for the first and last 5µm section for all three patients; patient 2 showing 57.6% and 60.4%; patient 7, 71.9% and 73.5%; and patient 9 66.8% and 69.7% for the first vs. last sections, respectively (Figure 4.5). We infer that the number of hepatocytes is comparable across the 12 x 20µm sections.

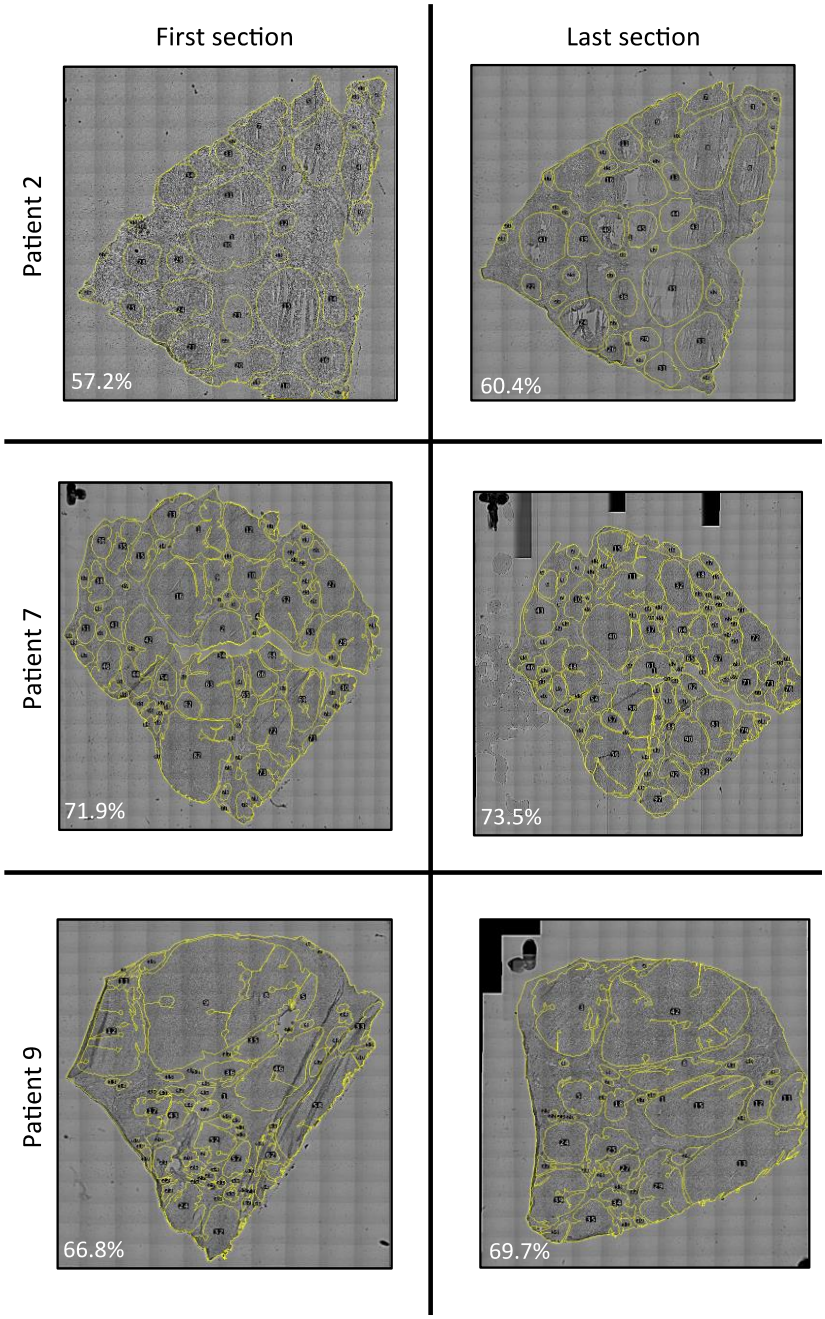


Figure 4.5 Quantification of hepatocytes *in vivo.*

Five µm sections were taken directly before and after a series of 12 x 20µm sections. The 5µm sections were histologically stained with H&E, scanned at a magnification of 10.000 x and analysed by ImageJ. The outline of the whole section and the individual nodules are highlighted with yellow lines and each unit has a number. The highlighted areas were used to calculate the fraction of healthy hepatocytes in each section. The estimated percentage of healthy hepatocytes is shown in the bottom left-hand corner for each image.

The HCV RNA load was quantified in the total hepatic RNA extractions from each section. Since the total hepatic RNA extractions were performed with exogenous RNA as a “carrier” to precipitate the relatively low levels of RNA, the viral loads could not be normalised to the total amount of liver RNA. For these samples normalisation to endogenous GAPDH was used, since previous studies showed stable GAPDH expression between hepatic samples from 5 patients included in this study (Figure 4.2). The quantities of HCV RNA varied between the sections with patient 2 showing up to 100-fold variation, patient 7 approximately 10-fold and patient 9 demonstrated 30-fold variation in the HCV RNA burden (Figure 4.6). One-way ANOVA demonstrated that each patient had significant heterogeneity in viral RNA loads across the 12 sections ($p < 0.001$). Interestingly, we found that the HCV RNA levels had continuous inclines and declines over several adjacent sections, instead of fluctuating levels (Figure 4.6). This suggests that the infected hepatocytes are not in 20µm sections within the liver. Instead, some of the islets of infected cells may spread over two or more adjacent sections resulting in smoothing of the HCV detection curve.

The presence of hepatocytes in each section was indirectly confirmed by measuring the albumin mRNA present (*in vivo*, albumin is almost exclusively produced by hepatocytes (Hoofnagle, 2002)). All the investigated hepatic sections proved positive for albumin mRNA (Figure 4.6B). Patient 7 and 9 showed significant variation in albumin levels by one-way ANOVA (patient 7: $p = 0.02$ and patient 9: $p < 0.0001$), this either indicates variation in the number

of hepatocytes in each section or that the hepatocytes in the individual sections show varying levels of albumin expression. Patient 2 did not show statistical variation in albumin expression across the twelve sections (one-way ANOVA: $p = 0.83$) (Figure 4.6).

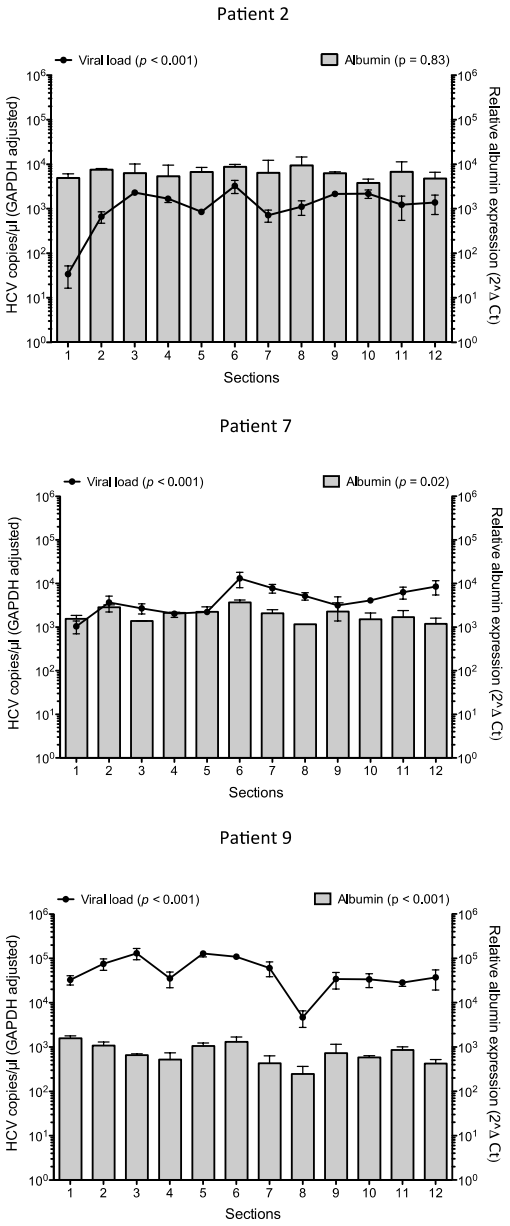


Figure 4.6 HCV RNA and albumin expression in hepatic sections.

Total liver RNA was extracted from 12 x 20µm adjacent liver sections obtained from patients 2, 7 and 9. The viral load for each section was determined by HCV specific qRT-PCR and the absolute quantity calculated from an internal standard. Relative amounts of albumin mRNA estimated by target specific qRT-PCR. All HCV and albumin estimates were normalised to the endogenous GAPDH expression. The data is represented as the mean of triplicate measurements with standard deviation. One-way ANOVA was performed for the HCV RNA levels and albumin expressions across the 12 sections. The probability (*p*) can be found in the graphs.

4.5 Comparison of hepatic and plasma viral burden

The association between the hepatic and plasma viral burden for the same patient is a controversial topic. Some investigators claim to have found a strong correlation between these parameters (Descamps et al., 2012; Idrovo et al., 1996), whereas others found none (Ballardini et al., 1997; Cabot et al., 2000). To our knowledge no study have looked at the association between plasma and liver viral load in end-stage liver disease and we decided to throw some light on the subject by comparing the viral load in the plasma and liver at the time of transplantation. The viral load in the last plasma samples obtained before anaesthesia (1-2 hours before the explant was removed) was quantified at the Heartlands Hospital, Birmingham, UK.

The viral load at the time of transplantation for the 22 patients varied between 823 and 3.2 x 10^7 HCV copies/ ml plasma. This large variation was independent of the patient's age (Figure 4.7A; r^2 = 0.005, p = 0.77), the response to previous PEG-IFN treatment (Figure 4.7B; One-way ANOVA p = 0.57), the indication of liver transplantation (Figure 4.7C; p = 0.27) and the infecting genotype (Figure 4.7D; p = 0.51).

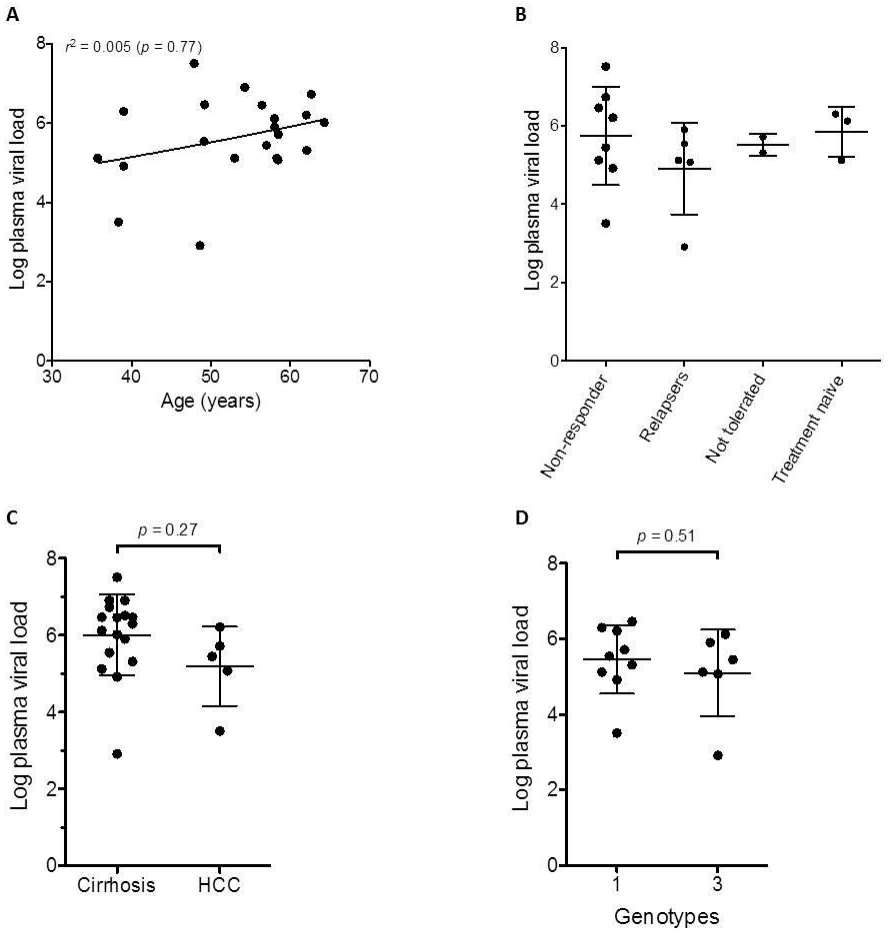


Figure 4.7 Comparative analysis of plasma HCV RNA.

Plasma viral loads were measured by the Roche Cobas amplicor 2.0 and the estimated HCV load (IU) converted to HCV copies /ml. **A.** Linear regression analysis between the plasma viral burden and the age of the patient. A black circle represents each patient and the linear regression (line of best fit) is indicated. Spearman's correlation coefficient (r^2) can be found at the top of the graph. **B.** Plasma viral loads were compared between patients who did not respond to PEG-IFN treatment, those who had relapsed, those who did not tolerate the treatment and those patients who had not been treated with PEG-IFN previously. One-way ANOVA revealed no sign of significant variation between the HCV RNA levels (p = 0.57) **C.** Comparison of plasma viral loads between patients diagnosed with cirrhosis and hepatocellular carcinoma (HCC). The statistical analysis was performed by two-tailed unpaired t-test. **D.** The HCV viral loads in genotype 1 and 3. The statistical analysis was performed by two-tailed Mann-Withney. The mean plasma loads for the different groups are represented with a mean ± standard deviation (**B** + **C** + **D**).

Comparison of hepatic HCV copies/ ng total liver RNA with the plasma HCV copies/ml revealed significant differences between the loads, with the liver showing an average of 184.1 ± 256.4 HCV RNA copies/ng total RNA for each patient and the plasma showing $2.9 \times 10^6 \pm 6.89 \times 10^6$ HCV RNA copies/ml. Investigation of the distribution of plasma and liver HCV RNA levels for the patient cohort revealed that the two data sets showed non-Gaussian distribution of the loads (Kurtosis ≠ 0). Instead, HCV RNA quantified in the plasma showed a flat distribution (Kurtosis = -0.07) and in the liver a peaked distribution (Kurtosis = 2.24) (Figure 4.8B). Further investigation of the symmetry in the data sets showed that both the plasma and liver HCV RNA loads were non-symmetrical (skewness ≠ 0), however the distribution of HCV RNA quantities were highly similar in the two compartments, with plasma revealing a skewness of 1.27 and liver a skewness of 1.77 (Figure 4.8B).

The smallest difference between HCV RNA quantities in the plasma and liver were found for patient 5 (0.5-log) and the largest for patient 23 (4-logs) (Figure 4.8A). Analysis of the viral load ratios for the patients, indicated a limited, albeit significant, correlation between hepatic and plasma viral load estimates (Figure 4.8C; $r^2 = 0.49$, $p < 0.02$).

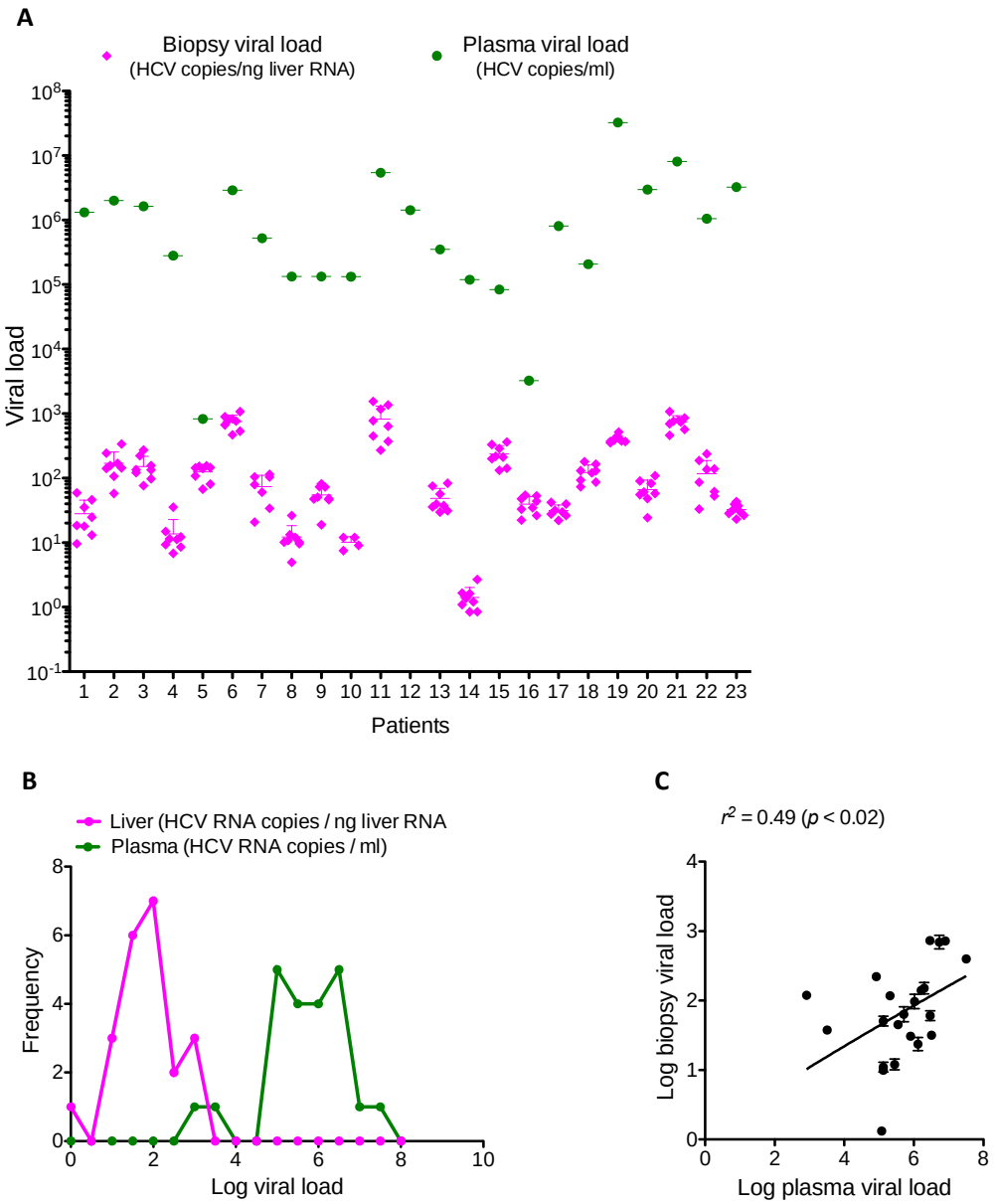


Figure 4.8 Comparative analysis between hepatic and plasma viral load.

A. The plasma and liver viral loads are shown for each of the 22 patients. The plasma HCV copies/m is represented with green circles and the mean hepatic HCV copies /ng total liver RNA per biopsy by pink diamonds. The mean and standard deviation are indicated for the hepatic viral loads. **B.** Histogram of the liver and plasma HCV RNA loads. **C.** Linear regression analysis of plasma and hepatic viral burden. Each black circle represents the mean hepatic and plasma load mean ± standard deviations for each patient. The linear regression (best fit) line is indicated. Spearman's coefficient is indicated for the data set.

4.6 Discussion

The intrinsic limitation of studying liver biopsies is their potential bias in terms of representing the whole liver. In this study, we investigated the spatial pattern of HCV RNA expression in the liver by sampling 150mg pieces from each of the 8 anatomical hepatic segments in chronically infected explanted livers. HCV RNA levels were quantified and compared to vireamic loads, infecting HCV genotype, indication for transplantation and previous response to PEG-IFN treatment.

The levels of HCV RNA were comparable across all 8 liver biopsies from one liver, suggesting that the liver is uniformly infected during end-stage liver disease. This observation supports previous publications stating that HCV RNA is comparable between multiple hepatic sites (2 – 7 sites) (Fanning et al., 2001; Idrovo et al., 1996; Sakai et al., 1999; Terrault et al., 1997).

Recent studies have reported that only a limited number of hepatocytes are infected and that HCV antigen or RNA expressing cells exists in clusters surrounded by naïve cells (Kandathil et al., 2013; Liang et al., 2009; Wieland et al., 2013). It is therefore possible that variation in HCV levels exist at a microscopic level, instead of a macroscopic. These microscopic differences may be overlooked when sampling larger sections of the liver. Characterising viral RNA expression across twelve adjacent 20µm liver sections revealed significant differences in HCV RNA levels. However, we did not observe any

sections that failed to express HCV RNA, contradicting previous reports by Liang *et al.,* Kandathil *et al.* and Wieland *et al.*.The H&E stained sections (5µm) immediately before and after the twelve 20µm sections, analysed for viral RNA, demonstrated high numbers of hepatocytes (57.2 – 73.5%, Figure 4.5). Given the low levels of HCV RNA detected in each of the 20µm sections from patients 2, 7 and 9, it is unlikely that all hepatocytes are infected. Instead, it is more reasonable to model that only a limited number of hepatocytes are infected with HCV and that each section contained a fraction of infected cells, explaining why all sections contain HCV RNA. Quantifying viral RNA within isolated nodules containing healthy hepatocytes or single hepatocytes would provide a deeper understanding of the distribution of HCV infected cells and the potential formation of compartments supporting HCV infection at a microscopic level.

For 5 of the 22 investigated patients HCV RNA levels varied with more than 0.5-log across the liver, exceeding the variability of the assay. According to a previous report, variable HCV RNA levels between hepatic biopsies sampled from the same liver may represent the time between sample collection and RNA processing (Terrault et al., 1997). In the Terrault *et al.* study, researchers were comparing two sets of samples obtained from the same liver explant with a 30-minute interval between collections. The sampling process in our study takes approximately 1 hour, which potentially could result in significant load differences between first and the last conducted sample. However, when comparing the viral RNA levels for individual biopsies, we did not see any

significant differences in viral load for the hepatic specimen, suggesting that the variability seen for these 5 patients is true. The lack of association between HCV RNA levels and the specimen collection time for one liver, furthermore suggests that RNA in the explanted liver is stable for at least 1 hour and that the HCV RNA degradation occurs simultaneously across the liver within the first 24-hours post-transplantation.

The mean intra-hepatic load measured for our cohort differed up to 2-logs between patients. These differences were independent of patient age, indication for transplantation, infecting HCV genotype or previous response to treatment. It is unlikely that the variation in viral load would result from varying numbers of hepatocytes between patients, since the estimated percentage of hepatocytes for patients 2, 7 and 9 microscopic sections did not reflect the hepatic viral load for these patients. However, it is likely that patients with high hepatic viral loads have a higher fraction of HCV infected cells, an observation previously made by Wieland *et al.* (Wieland et al., 2013) or that these patients have higher viral load per infected hepatocyte. Histological staining for HCV antigens in the collected biopsies would have provided further information on the frequency of infected cells, however staining for HCV antigens *ex vivo* has proven difficult and requires highly sensitive antibodies (Liang et al., 2009) or patient-specific mRNA probes (Wieland et al., 2013) that currently are not available in our laboratory. The reasons for variable hepatic viral loads between patients are not known, however it likely that differences in hepatic immune pressure may play a role. Investigation of the potential association

between host innate immune response in the liver and hepatic HCV RNA levels can be found in Chapter 5.

Measurements of plasma HCV RNA at the time of transplantation were provided by the clinic. Comparing the mean intra-hepatic and plasma load-ratio between patients showed a minimal correlation (r^2 = 0.13, p < 0.001), demonstrating that the plasma viral RNA level is a poor predictor of the hepatic burden in end-stage liver disease. Our study contrasts with previously reports of high correlation between liver and plasma viral burdens (Cabot et al., 2000; Descamps et al., 2012; Idrovo et al., 1996; Sakai et al., 1999; Vona et al., 2004), however most of these did not consider the stage of hepatic inflammation or fibrosis of subjects under study. The detection of correlation between plasma and liver viral burden for patients with mild fibrosis and the lack of correlation for patients with cirrhosis (Cabot et al., 2000) highlights the importance of matching patient disease stage when performing cross-patient comparisons. Other studies have failed to observe a correlation between hepatic and plasma load (Ballardini et al., 1997; Fanning et al., 2001; Sugano et al., 1995). All of these studies were retrospective and utilised frozen samples, which may have made it more difficult to obtain complete medical records. However, it is possible that patients in Ballardini *et al.*, Fanning *et al.* and Sugano *et al.'s* studies also had late chronic infection or end-stage liver disease, which would explain our comparable results.

The variable HCV RNA levels between plasma and liver observed in this study may be the result of extra-hepatic viral replication followed by the release into circulation. Previous studies have detected negative strand HCV RNA in tissues such as PBMCs and CNS (Natarajan et al., 2010; Ramirez et al., 2009; Willems et al., 1994), suggesting viral replication at other anatomical compartments. However, HCV negative strand qRT-PCR probes are considered to have low strand-specificity and these results should be interpreted with care (Blackard et al., 2006). Another reason for the poor correlation could be explained by the presence of viral variants with differing modes of virus transmission. Some variants may have a higher rate of cell-cell transmission than others and may be retained in the liver. High numbers of liver resident viruses could lead to a high intra-hepatic load and low plasma HCV RNA levels. Different modes of dissemination have previously been reported for laboratory prototypes of all the HCV genotypes (Brimacombe et al., 2011).

During the optimisation of the qRT-PCR protocol to quantify HCV RNA, we observed up to 2.6-fold increase in detectable HCV RNA when adding 5% DMSO to the RNA template and pre-incubating the mixture at 95°C prior to the quantification assay. Surprisingly this was significantly less than the 2-log change presented at the Annual HCV meeting in 2011). The discrepancy between our studies may be the result of different types of starting material (the presented study was conducted on hepatic needle biopsies collected from patients with earlier stages of liver fibrosis) or an inhibitory presence of DMSO

during the PCR. DMSO is widely used for its ability to lower the Tm of RNA and DNA double-stranded structures, but it may also increase the renaturation rate (Markarian et al., 2006). We did not test potential RNA re-annealing in the presence of denaturation enhancers, but did observe limited re-annealing of the dsRNA ladder in the absence of the chemical compounds and consider it unlikely that 5% DMSO would have a significant effect on this process. A future experiment could investigate whether DMSO accelerates the renaturation of double-stranded structures during the 50°C cDNA synthesis stage and 60°C annealing step of the qRT-PCR. As far known the RNA denaturation data presented at the Annual HCV meeting 2011 has not been published and it is possible that the result were not reproducible.

During our investigation of optimal dsRNA denaturation we observed a decrease in detected SYBR green for the dsRNA ladder in the presence of DMSO. It is possible that DMSO may affect the dsRNA structure at room temperature (25°C), however it should be mention that the effect could have resulted from DMSO interfering with the SYBR green. As the experiments stands today, we did not test the effect of DMSO on SYBR green fluorescence. Decreasing the temperature for the RNA ladder showed decreased re-annealing, which are likely to be the result of the ladder containing different length of the same nucleotide sequence. Mismatched re-annealing of sequences with different length may explain the lower detection of SYBR during decreasing temperatures.

The HCV genome has a high guanine-cytosine (GC) content of 58% (Tan et al., 2012) and will have a higher Tm_{50} than other sequences of the same length with lower GC content. The GC-level in the dsRNA ladder (NEB N0363) is 51.1% and heat-treating the ladder will show a lower Tm than the HCV genome. It is therefore possible that we would have obtained a larger increase in HCV detection if RNA extracts were incubated at 106°C, as suggested by Branch, instead of the 95°C.

Heat-treating liver RNA samples led to a small increase in the detection of HCV RNA, but not synthetic HCV strain H77 genomic RNA. The HCV H77 RNA was included as a control for potential secondary structures in the genome that could affect the 5'UTR availability for the probe. Failure to detect any change in the quantification of HCV H77 RNA following DMSO and heat treatment suggests that either the H77 genome does not form secondary structures in the 5'UTR probing site or that potential secondary structures doe not interfere with the reverse transcription of amplification. Considering the close relationship between HCV sequences, it is unlikely that the detection of patient-derived HCV genomes are affected by secondary structures in the 5'UTR either. If this is the case, the increase in detectable HCV genomes for the patient-derived total RNA extractions from patient 1 and 5 may be the result from denaturation of the negative / positive stranded duplex RNA, formed during viral replication. A recent study has reported that patients with end-stage liver disease express negative to positive RNA at a ratio of 1:3 to 1:340 (Lin et al., 2009). The authors reported that the amount of positive and

negative HCV RNA strands correlated between different sites in the liver, but not with the detection of serum positive strand HCV RNA. (Lin et al., 2009; Yuki et al., 2006). Patients being infected for different lengths of time may explain the large range in ratios reported by Lin et al.. Assuming that the increase in HCV RNA detection following heat and DMSO treatment is due to RNA duplex denaturation, the increased value would represent the number of negative strands in the sample. The negative : positive ratio for the two patient biopsies would be inferred to 1:1.5 and 1:5.4 for patient 1 and 1 : 1.3 and 1:1.7 for patient 5. The comparable ratio of negative and positive strands between samples collected from the hepatic lobes suggests that HCV does not form replication hot-spots during end-stage liver disease. Still, more samples needs to be evaluated to confirm the observation. In addition, qRT-PCR of the HCV negative strand in the hepatic samples is required to confirm the potential detection of HCV negative strands following heat and DMSO treatment.

To our knowledge, this is the first time that intra-hepatic HCV RNA levels across multiple sites have been studied in a cohort of patients with comparable liver disease staging. Investigating 8 different anatomical sites in the liver demonstrates that HCV RNA levels are comparable across macroscopic sites in the liver, suggesting that HCV does not form compartments with variable levels of HCV RNA. However small, but significant, differences in HCV RNA exist at the microscopic level. Comparison of plasma and liver HCV RNA levels between patients clearly indicates that a plasma sample does not

always represent the hepatic viral burden. Ideally, monitoring HCV burden in infected patients should include plasma and hepatic measurements.

5 EXPRESSION OF HEPATIC HOST FACTORS IN CHRONICALLY HCV INFECTED ADULTS AND CHILDREN

The infecting HCV genotype and viral load prior to PEG-IFN treatment are widely used as predictors of treatment response (Berg et al., 2003). However recent studies have suggested that the expression pattern of hepatic ISGs prior to treatment also can predict treatment response in patients infected with HCV genotype 1, 2, 3 and 4 (Chen et al., 2005; Dill et al., 2011; McGilvray et al., 2012; Sarasin-Filipowicz et al., 2008).

Dill and colleagues recently reported a predictive set of host genes including 3 ISGs (ISG15, IFI27, RSAD2), and a ubiquitously expressed tumour suppressor gene (HTATIP2) (Dill et al., 2011). Detection of high levels of intra-hepatic ISG expression suggests endogenous interferon signalling, which is thought to dampening the cellular response to therapeutic IFNs and the patient is therefore less likely to achieve SVR (Dill et al., 2011; Heim, 2013b; Honda et al., 2010; Qashqari et al., 2013). However, minimum information is available on whether the hepatic ISG expression varies across the liver. Information about the hepatic distribution of ISG expression may increase their accuracy as predictors for PEG-IFN response. In addition, evidence of a potential association between intra-hepatic ISG expression and viral burden are limited. This chapter will focus on studying the spatial ISG distribution across the

infected liver to gain a better understanding of the relationship between HCV replication and the ISG expression.

For this study, the expression patterns of intra-hepatic ISG15, IFI27, RSAD2 and HTATIP2 were evaluated for the 8 hepatic specimens collected from 22 chronically HCV infected adult patients with end-stage liver disease. For comparison, expression levels were evaluated in adult patients with end-stage liver disease due to NonAlcoholic SteatoHepatitis (NASH), Alcoholic Liver Disease (ALD), Auto-Immune Hepatitis (AIH) and non-A, non-B, non-C viral hepatitis (nA/nB/nC). Donor livers from deceased individuals were included as normal controls. Pre-treatment needle biopsies from 11 chronically infected paediatric patients were also included to evaluate any age-dependent expression of the 4 classifiers.

5.1 ISG expression in HCV related end-stage liver disease

The previously described RNA extractions from the 8 segments in 22 HCV infected livers were subjected to gene expression profiling by qRT-PCR using primers specific for ISG15, IFI27, RSAD2 and HTATIP2. The relative quantities for each target gene in the biopsies were evaluated in duplicate samples in two experimental repeats (n=2). (Figure 5.1; here represented by patients 1 and 2). The low variation observed for these repeated measurements (< 0.5log) for each biopsy justify the use of mean expression levels for the individual ISGs in the following analysis. Linear regression

analysis between the gene expressions and the collection time of the hepatic specimen revealed no correlation (Pearson's correlation coefficient: ISG15: r^2 = 0.007, p = 0.78, IFI27: r^2 = 0.003, p = 0.79, RSAD2: r^2 = 0.015, p = 0.30 and HTATIP2: r^2 = 0.004, p = 0.75, data not shown).

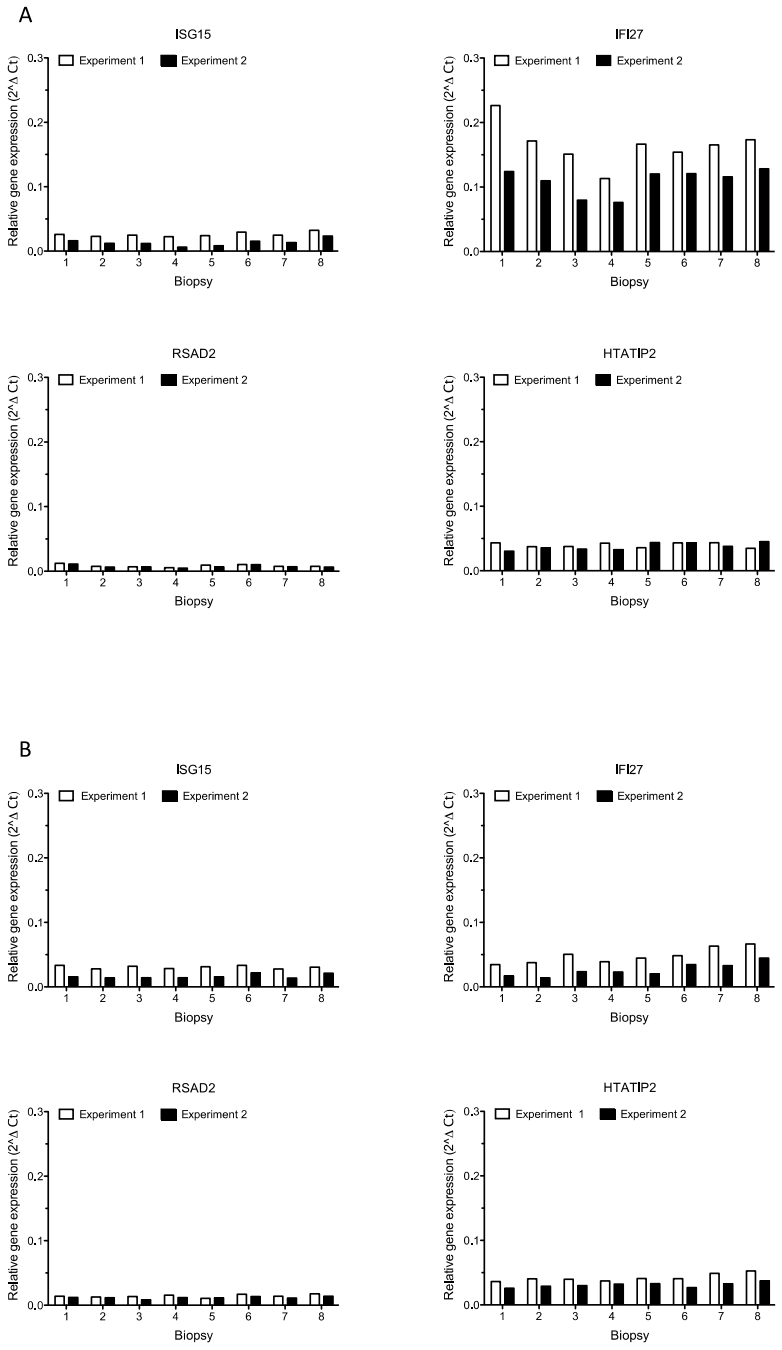


Figure 5.1 Intra-hepatic expression of the 4-gene classifiers for two representative patients.

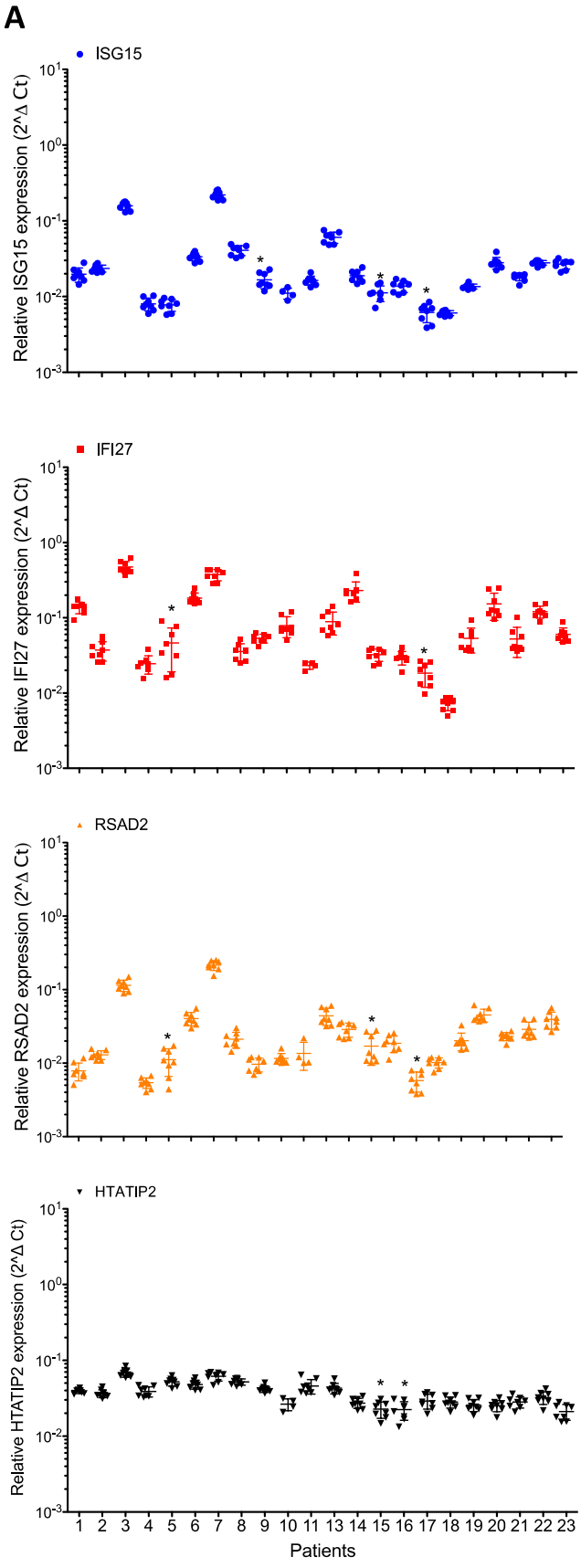
The amount ISG15, IFI27, RSAD2 and HTATIP2 mRNA were quantified by target specific qRT-PCR. The internal GAPDH level was used for normalisation. The data is presented as means ± standard deviations for quadruplicate values. Patients 1 (**A**) and patient 2 (**B**) are representative of the high consistency between duplicate samples measured in two experimental repeats (n=2).

The mean expression levels for each of the 4 classifiers across the 22 HCV positive livers are shown in Figure 5.2. The expression levels for most individual gene were comparable across all 8 biopsies, with less then ± 15% deviation from the mean. However, a few patients showed a greater level of diversity with up to ± 58% deviation from the mean (patients 5, 9, 15 and 17) (Figure 5.2A). These atypical observations are independent of HCV genotype, indication of transplantation, previous respond to PEG-IFN treatment and plasma and hepatic viral loads (data not shown).

Comparative analysis of the gene expression pattern within one liver showed that the level of relative expression significantly differed for the 4 genes with IFI27 usually showing the highest expression level (mean expression level = 0.11 ± 0.12), followed by HTATIP2 (mean expression level = 0.04 ± 0.01), ISG15 (mean expression level = 0.04 ± 0.05) and RSAD2 (mean expression level = 0.03 ± 0.05) (Figure 5.2B). Calculation of the Pearson's correlation coefficient between the individual gene expressions showed that ISG15, IFI27 and RSAD2 are highly correlated with one another (mean $r^2 = 0.89$, $p < 0.001$), whereas the correlation with HTATIP2 is somewhat lower (mean $r^2 = 0.63$, $p < 0.01$) (Table 5.1).

Cross-sectional analysis of the mean gene expressions revealed significant differences in hepatic ISG expression between the patients. The relative expression differed up to 34-fold for ISG15, 62-fold for IFI27 and

approximately 43-fold for RSAD2. In comparison HTATIP2 only differed by 3.5-fold among patients (Figure 5.2A and B).


A
ISG15
Relative ISG15 expression (2^Δ Ct)
IFI27
Relative IFI27 expression (2^Δ Ct)
RSAD2
Relative RSAD2 expression (2^Δ Ct)
HTATIP2
Relative HTATIP2 expression (2^Δ Ct)
1 2 3 4 5 6 7 8 9 10 11 13 14 15 16 17 18 19 20 21 22 23
Patients

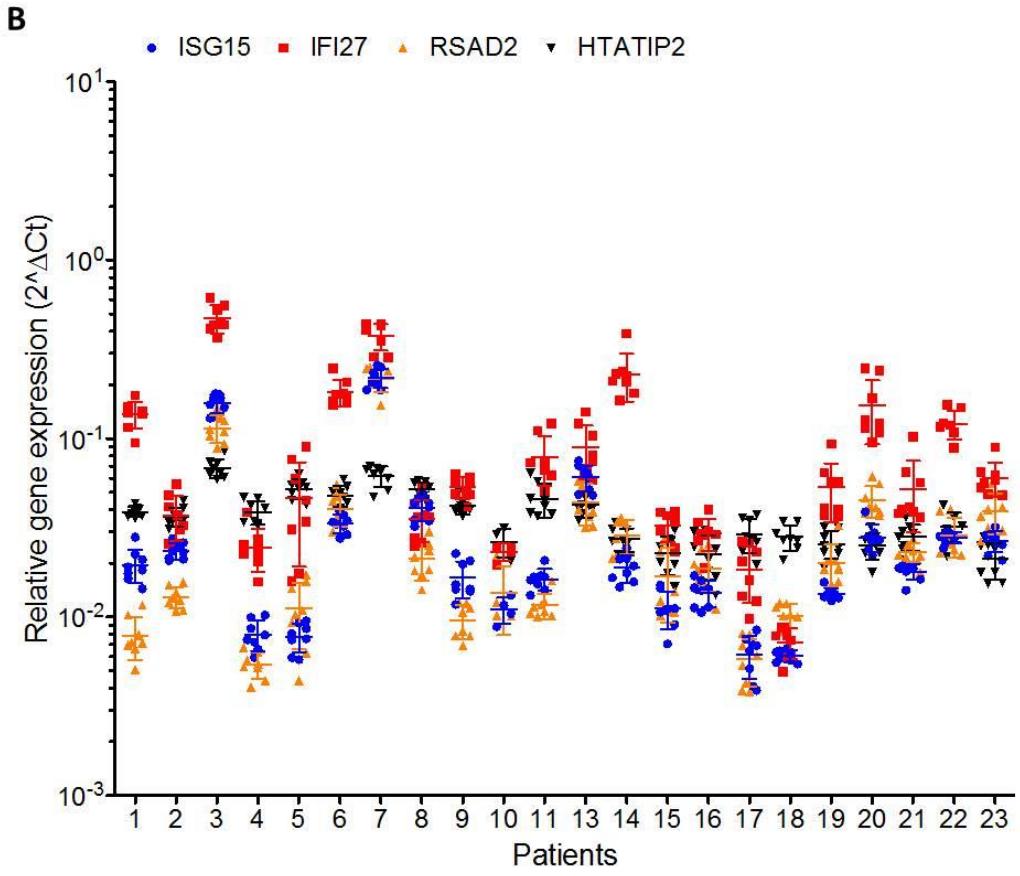


Figure 5.2 Intra-hepatic expression levels of the 4-gene classifiers for 22 chronically infected patients.

A. The expression levels for the 4 classifiers were quantified for the 8 hepatic specimens collected from 22 HCV positive liver explants and normalised to GAPDH. Each point in the graph represent the mean of duplicate samples measured in two independent qRT-PCR runs. The horizontal lines indicate the mean of all biopsies in one liver ± standard deviation. Hepatic samples with a deviation above 30% from the mean are indicated with asterisks. **B.** The combined expression patterns for the 4 classifiers in each patient.

Table 5.1: Pearson's correlation matrix				
	ISG15	IFI27	RSAD2	HTATIP2
ISG15	1.00	0.85^{***}	0.97^{***}	0.69^{***}
IFI27	0.85^{***}	1.00	0.81^{***}	0.63^{**}
RSAD2	0.97^{***}	0.81^{***}	1.00	0.57^{**}
HTATIP2	0.69^{***}	0.63^{**}	0.57^{**}	1.00

Correlation coefficient p-value: **: $p < 0.01$, ***: $p < 0.001$

Most HCV infected patients with end-stage liver disease has received PEG-IFN based treatment for the viral infection at some stage, but the treatment has either failed to eliminate the virus or had to be terminated early due to adverse effects in the patient. In this study the PEG-Interferon treatment status is known for 18 of the 22 patients (Table 3.1) and it was decided to investigate whether patients with end-stage liver disease who had either never received PEG-IFN treatment (n = 3), not responded to the administrated PEG-IFN treatment (n = 8), relapsed following initial successful treatment (n = 5) or had the treatment terminated prematurely (n = 2) would show different patterns of expression for the 4 gene classifiers (Figure 5.3A). Stratifying the gene expressions for the four treatment response groups revealed that those patients that were treatment naïve and those that had relapsed from previous treatment had lower expression of ISG15, IFI27 and RSAD2, but the observation was not statistically significant when comparing to the non-

responders (unpaired two-tailed t-test: $p > 0.26$). The high variation observed for ISG15 and IFI27 expressions for patients who did not tolerate the treatment is likely to be the consequence of the small sample number (Figure 5.3A).

Previous studies have reported that patients infected with genotype 1 and 4 show significantly higher hepatic expression of ISG15 and IFI27 compared to patients infected with genotype 2 and 3 (Broering et al., 2010; Sarasin-Filipowicz et al., 2008). Analysing the expression patterns of the 4-gene classifiers for patients infected with genotype 1 (n = 9) and 3 (n = 6) showed that patients with genotype 1 tended to have a higher expression of the three ISGs, however the observation were not statistical significant (unpaired two-tailed t-test: ISG15: $p = 0.14$, IFI27: $p = 0.33$, RSAD2: $p = 0.06$) (Figure 5.3C).

The expression patterns for the 4 gene classifiers were furthermore stratified for patients who had been diagnosed with HCC (n = 5) and cirrhosis (n = 17) prior to the transplantation. The results showed that patients diagnosed with HCC tended to have a higher expression of all the investigated genes, however in none of the cases were the observation statistical significant (unpaired two-tailed t-test: ISG15: $p = 0.34$, IFI27: $p = 0.06$, RSAD2: $p = 0.58$ and HTATIP2: $p = 0.16$) (Figure 5.3B).

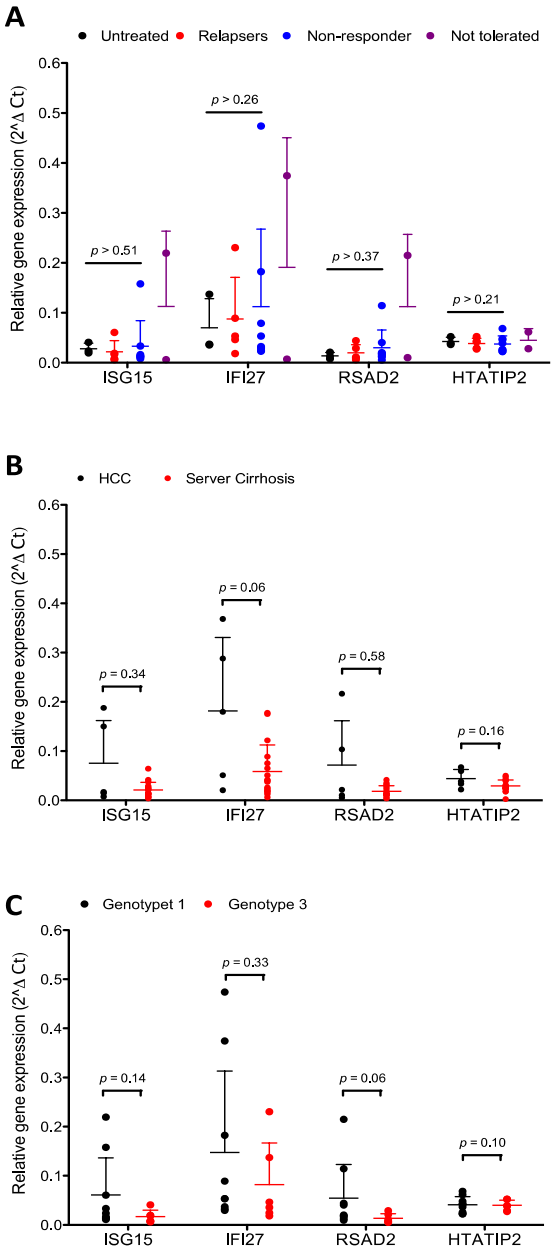


Figure 5.3 Influence of host and viral factors on the expression patterns for the 4-gene classifiers.

The hepatic expression patterns of the 4-gene classifiers were divided into groups according to; **A.** previous response to PEG-IFN treatment, **B.** the diagnosis prior to liver transplantation (HCC: Hepatocellular carcinoma) and **C.** the viral genotypes 1 and 3 (Genotype 2 and 4 were not included since these consisted of only 1 patient each). All data are presented as means with standard deviations for each patient group and are based on duplicate measurements in two independent qRT-PCRs. Statistical analyses for each data set was performed by unpaired two-tailed t-test.

5.2 ISG expressions in non-HCV related end-stage liver diseases and in donor livers

Most current research studying hepatic ISG expression has been performed with samples from HCV or HBV infected subjects. It is therefore not known whether patients with comparable end-stage liver disease would have similar levels of hepatic ISG expressions. To address this question we measured the relative levels of ISG15, IFI27, RSAD2 and HTATIP2 in patients with end-stage ALD, NASH, AIH and nA/nB/nC viral hepatitis, together with hepatic tissues donated from diseased healthy individuals.

The hepatic specimens from the above mentioned non-HCV associated liver diseases were all diagnosed with end-stage liver disease staging by Pathologist Dr. Gary Reynolds (data not shown), nevertheless different liver diseases may have variable expression of housekeeping genes. It was therefore decided to identify the housekeeping gene with the least variable expression between the ALD, NASH, AIH, nA/nB/nC viral hepatitis liver specimens and the healthy donor liver by the GeNorm panel of housekeeping genes (see section 4.2), before comparing the normalised relative expressions between patient groups.

Normalising the gene expressions for each disease group with the same reference gene ($2^{\Delta Ct}$) revealed no significant variation in the relative expression levels between the different groups (One-way ANOVA; $p > 0.1$).

However, comparing the normalised values by two-way ANOVA showed that there is a systemic difference in the expression of housekeeping genes for the different diseases ($p < 0.01$). This variation mainly reflects the different reference genes applied to the template (9.9% of the variation, $p < 0.05$) and not variation in gene expression between disease groups (0.47% of the variation, $p = 0.055$) (Figure 5.4). The expression levels of the housekeeping genes are therefore considered independent of the type of end-stage liver disease tissue, justifying our use of GAPDH for normalising the expression of the 4 gene classifiers.

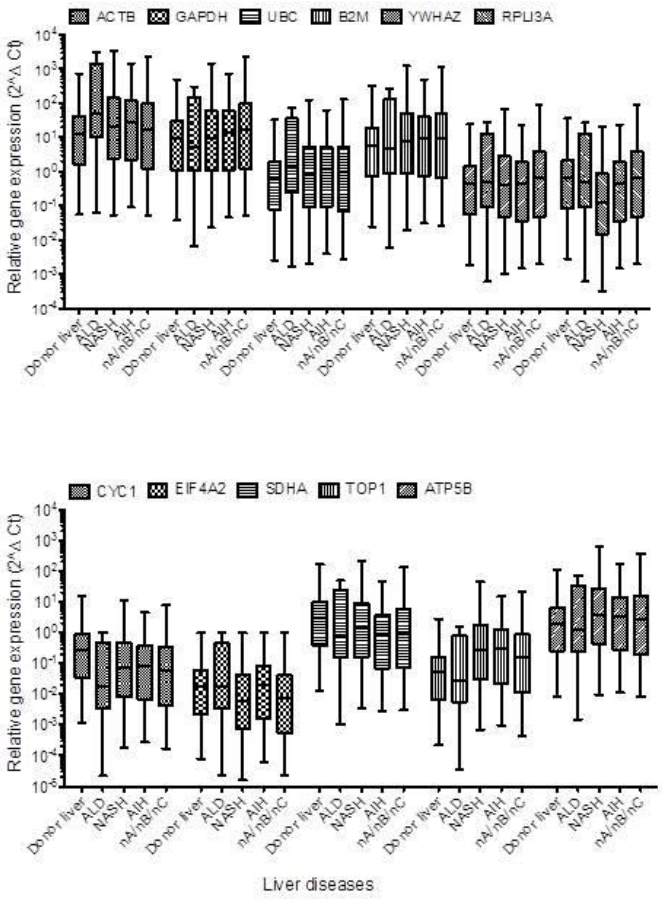


Figure 5.4 Evaluation of reference genes for different end-stage liver diseases and donor livers.

The GeNorm panel of 11 different housekeeping genes were used for evaluating the stability of the reference gene between single samples from livers with alcoholic liver disease (ALD), non-alcoholic steatohepatitis (NASH), auto-immune hepatitis (AIH), non-A, non-B, non-C viral hepatitis (nA/nB/nC). Livers from healthy donors were included as normal controls. The expression levels were measured by qRT-PCR. The raw Ct values for each housekeeping gene were used to normalise each of the other 10 expressions and the result expressed as 2^Δ Ct. The data is presented as median (horizontal line), upper and lower quartile and the maximum and minimum spreads. One-way ANOVA between gene expressions normalised with the same reference gene did not reveal any significant differences between the disease groups (in all cases $p > 0.1$). However, two-way ANOVA revealed systemic variance between different RNA templates and the different reference genes ($p < 0.01$).

Total hepatic RNA was extracted from single specimens collected from 6 patients in each disease group and from 6 donated livers. The relative expression levels of the classifiers were quantified in duplicate samples by two independent qRT-PCRs. Minimal variation was observed between the repeated measurement for each sample (< 0.5log, data not shown) and it was therefore decided to use the mean expression for each hepatic specimen in the following analyses.

The individual gene expressions for ISG15, IFI27 and RSAD2 varied between 2.5- and 5.2-fold for the samples in the same disease group, whereas the gene expression of HTATIP2 showed lower variability (2.1- to 3.2-fold, dependent on the disease group) (Figure 5.5). Comparing the mean and the range of the individual gene expression between the groups revealed comparable levels of variation in the expressions (Table 5.2). Interestingly, the hepatic specimens from the healthy donor livers showed a significantly greater range of the three measured ISG expressions compared to hepatic samples from non-HCV associated end-stage liver disease, with ISG15 showing 110% variation from the mean expression, IFI27 showing 102% and RSAD2 100% variation from the mean expression. Statistical analyses of the gene expression levels between the disease groups and the donor liver revealed that the level of RSAD2 significantly differ between the groups (One-way ANOVA: $p = 0.046$), however the observation is likely to be the result of the large range in expression levels for the donor liver. None of the remaining

classifiers differed significantly between the groups (One-way ANOVA: $p > 0.1$).

Table 5.2: Mean hepatic gene expression for end-stage liver diseases and donor livers

Hepatic tissue	Mean ISG15 expression (±STDV*)	Mean IFI27 expression (±STDV)	Mean RSAD2 expression (±STDV)	Mean HTATIP2 expression (±STDV)
HCV infected	0.036 (± 0.052)	0.105 (± 0.119)	0.034 (± 0.047)	0.037 (± 0.013)
ALD	0.013 (± 0.005)	0.016 (± 0.010)	0.022 (± 0.012)	0.039 (± 0.011)
NASH	0.010 (± 0.005)	0.009 (± 0.006)	0.013 (± 0.006)	0.043 (± 0.013)
AIH	0.015 (± 0.006)	0.019 (± 0.011)	0.015 (± 0.006)	0.030 (± 0.011)
nA/nB/nC viral hepatitis	0.024 (± 0.009)	0.031 (± 0.021)	0.028 (± 0.014)	0.046 (± 0.019)
Donor liver	0.012 (± 0.015)	0.016 (± 0.017)	0.012 (± 0.012)	0.045 (± 0.018)

*STDV: Standard deviation

The expression of the 4-gene classifiers was compared between the hepatic samples collected from patients with HCV-related liver disease, non-HCV associated liver disease and from donor livers (Figure 5.5). The HCV infected livers show higher median expression for all ISGs compared to the other disease groups. However the difference were only statistical significant for the ISG15 expression in samples collected from NASH (Mann-Whitney test: $p = 0.03$) and for the IFI27 expression in samples from ALD, NASH, AIH, nA/nB/nC viral hepatitis and the donor liver (Mann-Whitney test: $0.0006 > p < 0.04$) (Figure 5.5).

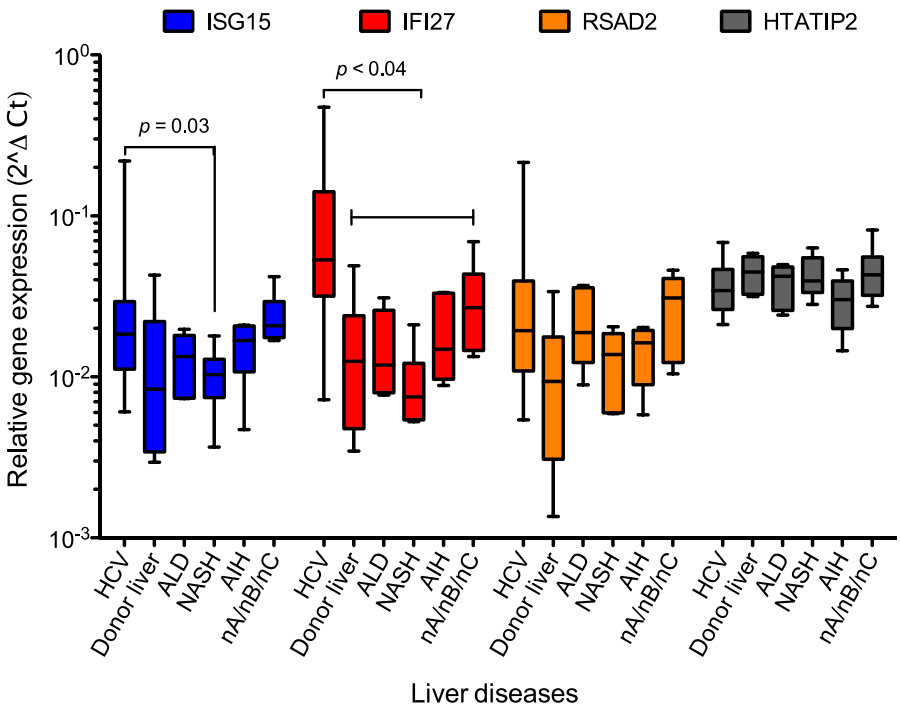


Figure 5.5 Intra-hepatic expression patterns for the 4-gene classifiers for end-stage liver diseases and donor livers.

Total RNA was extracted from 6 specimens representing the disease group; ALD, NASH, AIH and nA/nB/nC viral hepatitis. RNA extractions from 6 donor livers were also included. Here the expression levels are shown as means from duplicate measurement in two independent qRT-PCRs. All 4 classifiers are normalised to the internal GAPDH mRNA expression. The data is presented as median (horizontal bar), upper and lower quartile and the maximum and minimum spreads. Statistical analysis was performed by Mann-Whitney and the statistical values (*p*) below 0.05 are noted in the graph.

5.3 Association between ISG expression levels and the hepatic and plasma HCV RNA loads at end-stage liver disease

In section 4.3 we found that chronically infected patients express comparable levels of HCV RNA across the liver, but that these levels vary by up to 100-fold between patients. These observations are independent of infecting viral genotype, plasma load, patient age and indication for transplantation. Similar variation was seen for the expression patterns of the 4-gene classifiers on the same sample set (Figure 5.6A). To investigate whether the expression patterns of these 4-classifiers correlate with the hepatic viral loads we compared the two data sets.

Multiple linear regression modelling was performed to evaluate the correlations between the hepatic viral loads and the measured gene expression (Figure 5.6B). As shown in in Table 5.1 the expression of the 4-gene classifiers are strongly inter-correlated with one other, however single-variant analysis of the viral load against the individual gene expression show limited correlation (-0.095 to 0.087, $p > 0.67$). Calculation of the multiple correlation coefficient (r^2) revealed that the combination of the 4-gene expression patterns do not predict the hepatic viral load ($r^2 = 0.1473$, $p = 0.58$). The F statistic is calculated as the ratio of expression variance within each gene groups and between the groups of gene expressions. The null hypothesis, that all correlation coefficients are 0, can only be rejected if the F ratio is significantly greater than 1. For this dataset the F ratio was 0.73, indicating that the correlation coefficients between the viral load and the combined gene expressions are not different

from 0. As previously mentioned, the HTATIP2 expression remained constant for all patients, which is likely to be a reflection of the fact that HTATIP2 is not an ISG. The multiple linear regression modelling was therefore repeated without HTATIP2 and the result demonstrated no association between the hepatic HCV RNA load and the expression of ISG15, IFI27 and RSAD2 (r^2 = 0.0394, p = 0.86, F-ratio = 0.25).

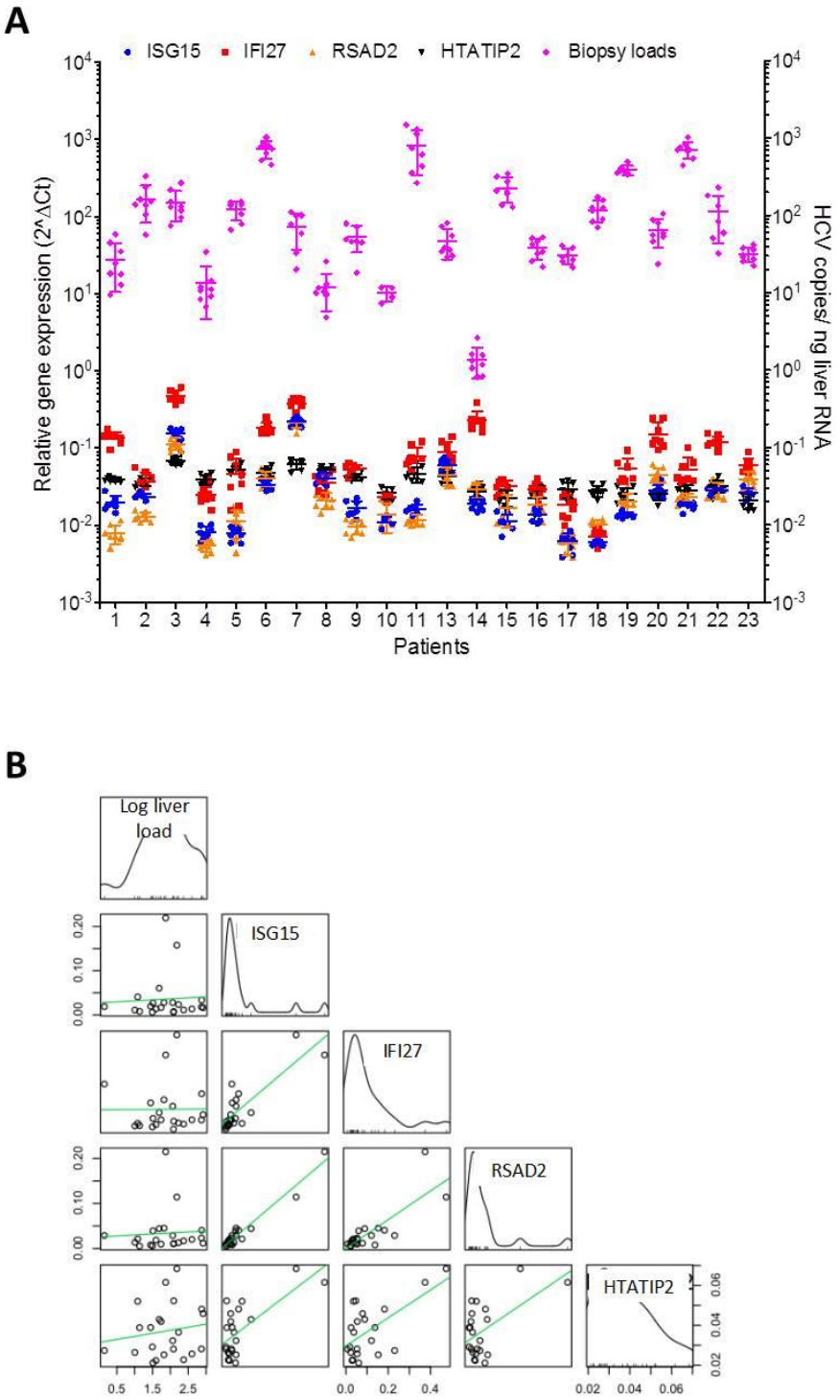


Figure 5.6 Association between the expression of the 4-gene classifier and hepatic HCV RNA.

A. The expression patterns for the 4 classifiers in the liver specimens collected from HCV infected patients. Left Y-axis; gene expressions normalised to GAPDH, Right Y axis HCV copies/ng total liver RNA. Each point in the graph represents the mean of

duplicate samples for two independent qRT-PCRs. The horizontal lines indicate the mean of all biopsies in one liver ± standard deviation. **B**. Scatter plot matrix of all components. The density plot for each gene expressions and viral load are seen in the diagonal boxes. On each side of the diagonal are the scatter plot (black circles) and the linear correlation line (green) for each pairwise comparison. The x- and y-axes are labelled relative gene expression (2^delta Ct). The matrix is exported from the software program R-commander.

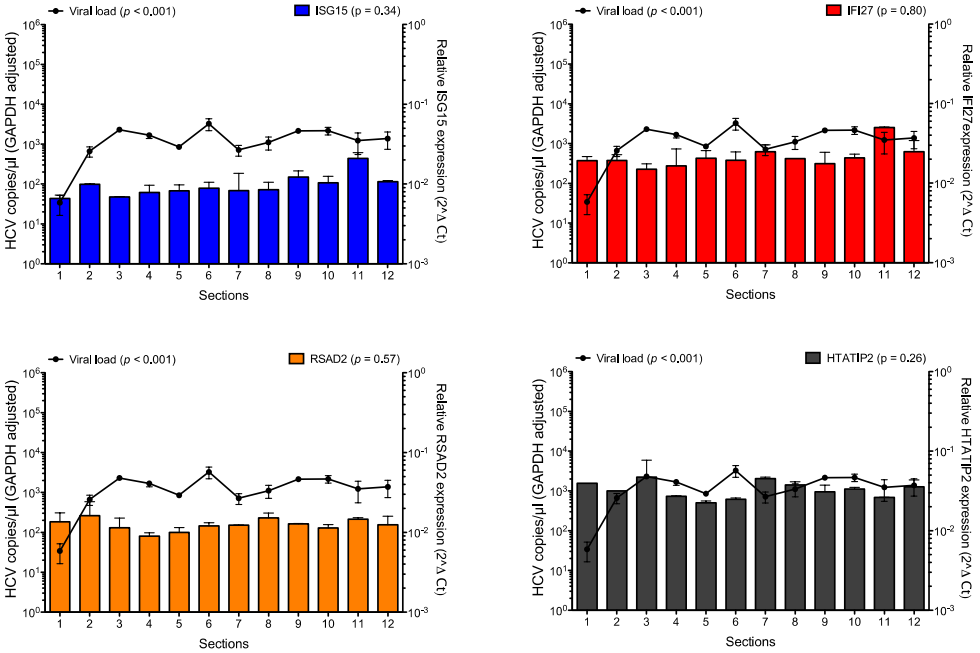
In the analysis above we showed that the hepatic HCV RNA load and the 4-gene classifiers expression pattern do not correlate, however the hepatic ISG expressions could instead have an effect on the level of released virus from the infected hepatocytes. A potential association between ISG15, IFI27, RSAD2 and HTATIP2 would explain the relative poor correlation between plasma and hepatic viral load observed for our patient cohort in section 4.5.

Linear regression modelling revealed a negative correlation between the plasma viral load and the individual gene expressions, but in none of the cases did the correlation significantly differ from 0 (Pearson's correlation coefficients (r^2): ISG15 = -0.11, IFI27 = -0.10, RSAD2 = -0.07 and HTATIP2 = -0.26. p > 0.05) (data not shown). The multiple correlation coefficient between the plasma viral burden and the 4-gene classifiers were 0.077, however the correlation was not statistical significant (p = 0.84, F-ratio = 0.35) (data not shown). Repeating the multiple linear regression modelling after excluding the expression of HTATIP2 resulted in a lower non-significant multiple correlation (r^2 = 0.0411, p = 0.85, F-factor = 0.26) (data not shown).

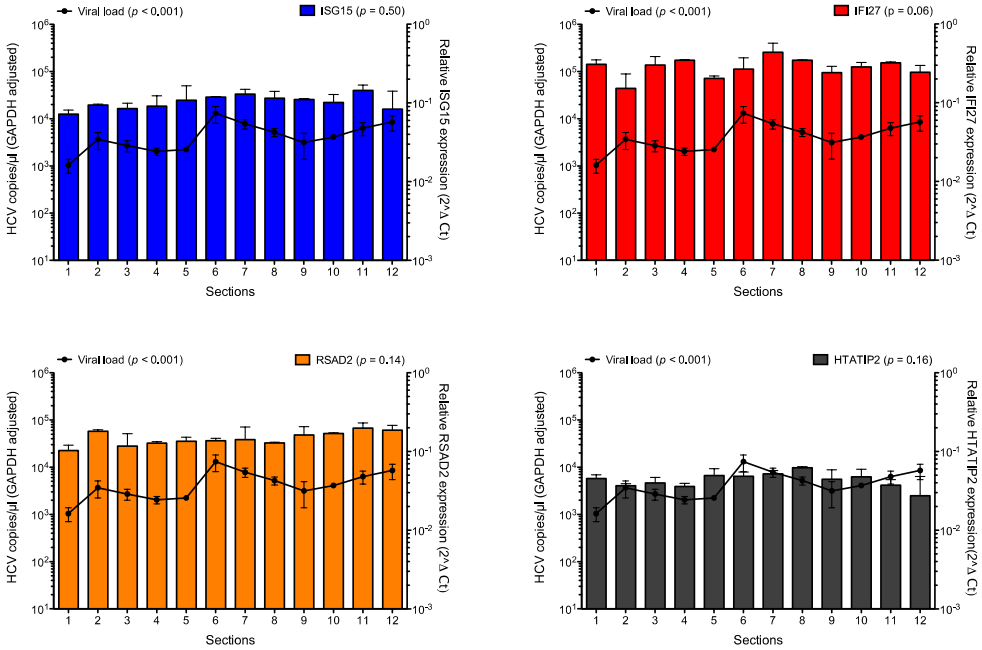
We have previously showed that although the patients do not have any significant variation in viral loads across 8 anatomical segments in the liver, they do show differences in viral burden in defined tissue sections. This observation led us to investigate the expression patterns for the 4 classifiers in adjacent liver sections (12 x 20µm thick) from patients 2, 7 and 9.

The relative gene expression profiles for ISG15, IFI27, RSAD2 and HTATIP2 were highly comparable between the 12 sections obtained from the three patients (one-way ANOVA: $p > 0.1$), however the viral load differed significantly among the same sections (one-way ANOVA: $p < 0.001$) (Figure 5.7). This discrepancy between the HCV RNA levels and the ISG expression patterns supported our earlier conclusion that the expression levels of the 4 classifiers do not correlate with HCV RNA levels.

Patient 2


Viral load (p < 0.001)
ISG15 (p = 0.34)
HCV copies/µl (GAPDH adjusted)
Relative ISG15 expression (2^Δ Ct)
Sections
Viral load (p < 0.001)
IFI27 (p = 0.80)
HCV copies/µl (GAPDH adjusted)
Relative IFI27expression (2^Δ Ct)
Sections
Viral load (p < 0.001)
RSAD2 (p = 0.57)
HCV copies/µl (GAPDH adjusted)
Relative RSAD2 expression (2^Δ Ct)
Sections
Viral load (p < 0.001)
HTATIP2 (p = 0.26)
HCV copies/µl (GAPDH adjusted)
Relative HTATIP2 expression (2^Δ Ct)
Sections

Patient 7


Viral load ($p < 0.001$)
ISG15 ($p = 0.50$)
HCV copies/µl (GAPDH adjusted)
Relative ISG15 expression (2^Δ Ct)
Sections
Viral load ($p < 0.001$)
IFI27 ($p = 0.06$)
HCV copies/µl (GAPDH adjusted)
Relative IFI27 expression (2^Δ Ct)
Sections
Viral load ($p < 0.001$)
RSAD2 ($p = 0.14$)
HCV copies/µl (GAPDH adjusted)
Relative RSAD2 expression (2^Δ Ct)
Sections
Viral load ($p < 0.001$)
HTATIP2 ($p = 0.16$)
HCV copies/µl (GAPDH adjusted)
Relative HTATIP2 expression(2^Δ Ct)
Sections

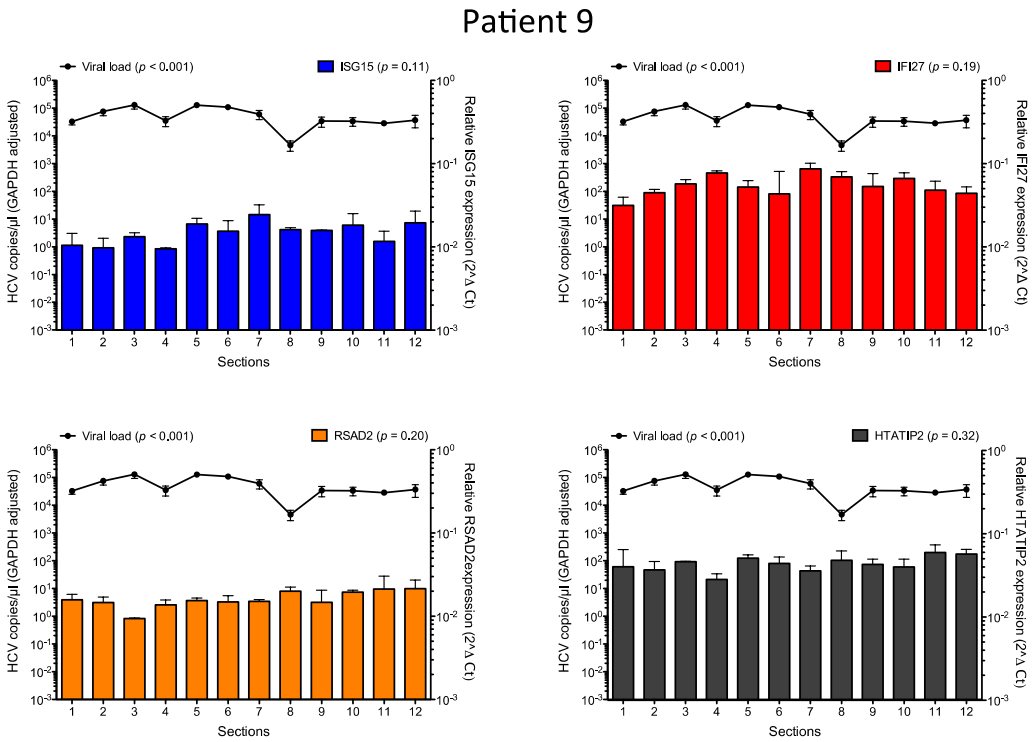


Figure 5.7 Expression pattern of the 4-gene classifiers within hepatic sections.

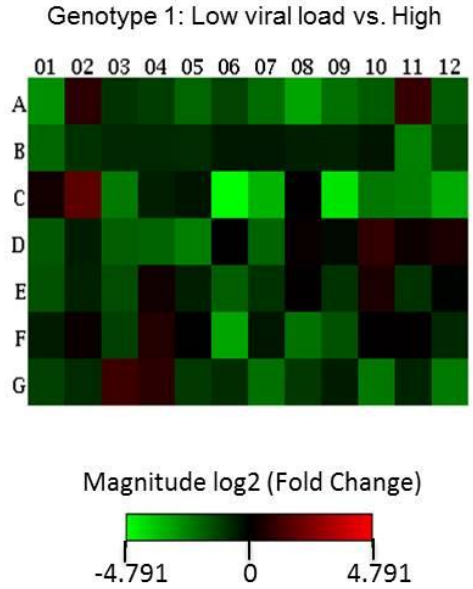
The HCV copies /µl and the expression levels of the 4 classifiers (normalised to GAPDH) were quantified in 12 adjacent 20µm sections from patients 2, 7 and 9. The samples were measured in duplicates in one qRT-PCR run. Significant differences were found for the viral loads in all patients (left Y-axis, p <0.001), but not for the classifiers' expression levels (right Y-axis, p > 0.1) when applying the one-way ANOVA to the data sets.

5.4 Anti-viral gene expression profile in HCV-related end-stage liver disease

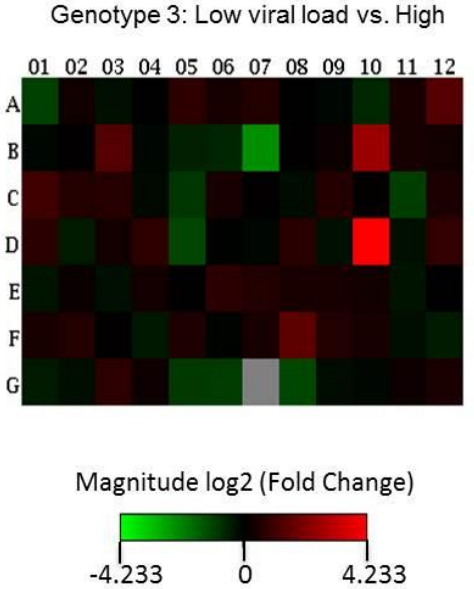
In previous sections we found that ISG15, IFI27, RSAD2 and HTATIP2 do not correlate with the viral RNA loads in the plasma or in the liver. To identify new potential candidate genes that may predict host pathways regulating HCV replication, we screened the expression of 84 different anti-viral genes using a SABiosciences anti-viral PCR array. Biopsy 1 from four patients was selected for these arrays, representing HCV genotype 3a (patients 4 and 5) and 1a (patients 6 and 13), which had low (patient 4 and 13) or high (patient 5 and 6) intra-hepatic viral loads compared to the mean hepatic HCV RNA level in the patient cohort. The samples were evaluated twice and on each occasion the templates passed the human genomic DNA contamination (HGDC) test, the reverse transcription control (RTC) test and the positive PCR control (PPC) test. Heat-maps were generated from the fold-difference in expression levels between the tissue containing low and high HCV RNA loads. Comparing hepatic expression levels for patients infected with the same HCV genotype demonstrated that genotype 1a patient 6 had significantly higher expression level for most genes compared to patient 13 (Figure 5.8A). The genotype 3a patients showed more comparable expression patterns across the array, with only a few exceptions of above 2-fold differences, one being ISG15 (Figure 5.8B, well D10). The expression levels of the 84 anti-viral genes were normalised to the combined expressions of housekeeping genes ACTB, B2M, GAPDH, HPRT1 and RPLPO and used for cluster analysis (Figure 5.9). The cluster analysis of the gene expressions in the individual specimens verified that the genotype 3a patients 4 and 5 are closely related in regards to hepatic

anti-viral gene expression and are closely followed by patient 13 (genotype 1a). Approximately 60 to 75% of the detected gene expressions for patients 4, 5 and 13 were below the level of the housekeeping genes (Figure 5.9, green), whereas patient 6 showed significantly higher expression of most genes, thereby forming an outlier in the cluster analysis (Figure 5.9). The close association of the genotype 3a patients with low and high intra-hepatic viral loads (patient 4 and 5), leads us to conclude that the expression of the evaluated genes are independent of hepatic HCV RNA load. The high difference in expressions between patient 6 and 13 may be specific to genotype 1a patients.

A



B



C

	1	2	3	4	5	6	7	8	9	10	11	12
A	AIM2	APOBEC3G	ATG5	AZI2	CARD9	CASP1	CASP10	CASP8	CCL3	CCL5	CD40	CD80
B	CD86	CHUK	CTSB	CTSL1	CTSS	CXCL10	CXCL11	CXCL9	CYLD	DAK	DDX3X	DDX58
C	DHX58	FADD	FOS	HSP90AA1	IFIH1	IFNA1	IFNA2	IFNAR1	IFNB1	IKBKB	IL12A	IL12B
D	IL15	IL18	IL1B	IL6	IL8	IRAK1	IRF3	IRF5	IRF7	ISG15	JUN	MAP2K1
E	MAP2K3	MAP3K1	MAP3K7	MAPK1	MAPK14	MAPK3	MAPK8	MAVS	MEFV	MX1	MYD88	NFKB1
F	NFKBIA	NLRP3	NOD2	OAS2	PIN1	PSTPIP1	PYCARD	PYDC1	RELA	RIPK1	SPP1	STAT1
G	SUGT1	TBK1	TICAM1	TLR3	TLR7	TLR8	TLR9	TNF	TRADD	TRAF3	TRAF6	TRIM25
H	ACTB	B2M	GAPDH	HPRT1	RPLP0	HGDC	RTC	RTC	RTC	PPC	PPC	PPC

Figure 5.8 Heat-maps of anti-viral gene expressions between patients with high and low HCV RNA burdens.

Total hepatic RNA extractions from samples collected from patients with high (patient 5 and 6) and low (patient 4 and 13) HCV RNA loads were used as template for the 2-step RT^2 profiler SABiosciences human antiviral micro-array (384 wells). The gene profiles were measured twice for each template. Representative results are here shown for one array. **A.** Comparison of expression levels for the genotype 1a patients 13 (low viral load) versus patient 6 (high viral load). The colours represent the magnitude of log2 (fold changes). Green represent decreased expressions, red increased. **B.** Comparison of expression levels for genotype 3a patients 4 (low viral load) versus patient 5 (high viral load). **C.** Overview of the investigated genes on the 96 well-plate. Well A1 to G12 are anti-viral genes; well H1 to H5 are housekeeping genes; well H6 is human genomic DNA control (HGDC), H7 to H9 are reverse transcription controls (PPC) and well H10 – H12 are positive PCR controls (PPC).

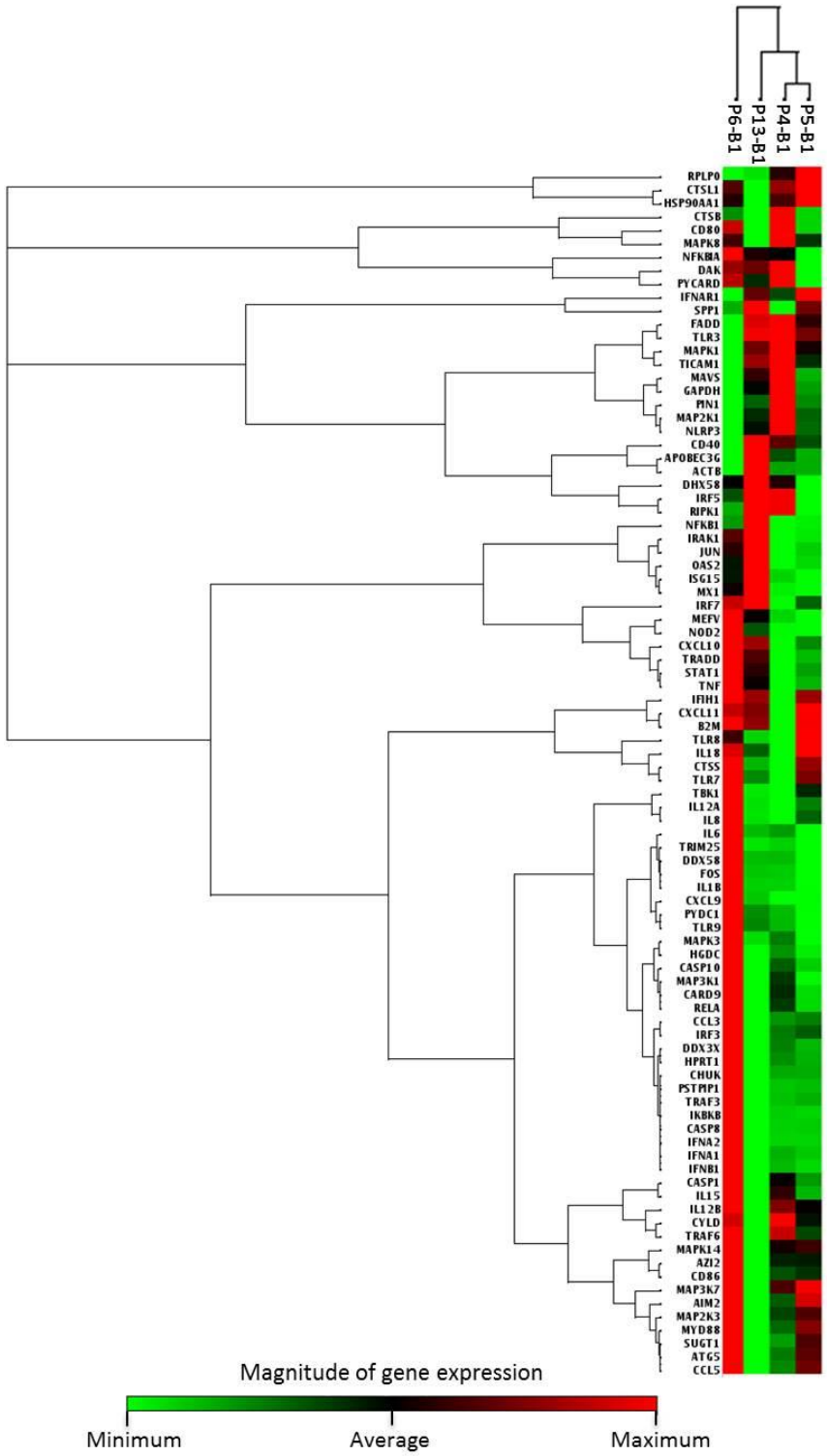


Figure 5.9 Dendrogram of anti-viral expressions for patients with high and low HCV RNA burdens.

Cluster analyses were performed for the anti-viral expression profiles for patients infected with HCV genotype 1a (patients 6 and 13) and genotype 3a (patients 4 and 5). All antiviral gene expressions were normalised to the internal reference genes

ACTB, B2M, GAPDH, HPRT1 and RPLPO. Association between the patients' hepatic gene expressions is shown as the X-axis and between individual gene expressions as the Y-axis. Two independent anti-viral arrays were evaluated. Representative results are here shown for one array.

Comparing the micro-array data between tissues collected from patients with low and high hepatic HCV RNA levels revealed that the regulated genes were involved in the RIG-I, TLR and nucleotide-binding oligomerization domain (NOD-like) receptor pathways. Unexpectedly, the raw Ct-values for IFN-α1, IFN-α2, IFN-β1, IL12α or IL12β were all above 30, suggesting low expression of these RIG-I and TLR responsive genes (Table 5.3).

Table 5.3: Anti-viral gene-array*

Genotype	Low viral load vs. High viral load	Pathways	Up-regulated genes in the tissue with low HCV RNA loads compared to high. (> 2-fold)	Down- regulated genes in the tissue with low HCV RNA loads compared to high. (> 2-fold)	Undetectable (>30 Ct)
1a	Patient 13 vs. Patient 6	RIG-I receptor and chaperon	-	TRIM25	-
		RIG-I downstream signaling	FADD	AZ12, CASP10, CASP8, DDX3, RIG-I, RELA, TNF, IRF3, TRAF3, IKBKB, MAP3K7	-
		RIG-I signaling responsive gene	-	IL8,	-
		Toll-Like receptor and chaperon	-	TLR7, TLR9	IFNα-1/2, IFNβ-1, IL12-α/β
		Toll-Like downstream signaling	TICAMP1	FOS, IKBKB, IRF3, MAP3K7, MAPK3, RELA, TRAF3,	-
		Toll-Like signaling responsive genes	CD40	CCL3, CCL5, CD80,CD86, IL6, IL15.	-
		NOD-Like signaling	-	AIM2, CARD9, CASP1, PSTPIP1, SUGT1.	IFNα-1/2, IFNβ-1, IL12-α/B
				IL1B	NOD2, PYDC1
		NOD-Like signaling responsive genes	-	IL15	-
1a	Patient 4 vs. Patient 5	RIG-I receptor and chaperon	DAK, RIG-I	-	-
		RIG-I downstream signaling	-	TNF	-
		RIG-I signaling responsive gene	-	IL8	IFNα-1/2 IFNγ-1, IL12-α/β, TLR9
		Toll-Like receptor and chaperon	CTSB	-	-
		Toll-Like downstream signaling	-	TNF	-
		Toll-Like signaling responsive genes	CD80	CXCL11	IFNα-1/2, IFNβ-1, IL12-α/β
		NOD-Like signaling	PYDC1	AIM2	-
		Interferon-Stimulated gene	ISG15	-	-

* Comparable gene expressions between templates are not included

Interestingly the DDX3 gene was found to be 5-fold down-regulated in patient 13 (low viral load) compared to patient 6 (high viral load). This Dead-Box helicase has previously been demonstrated to facilitate HCV replication *in vitro* (Angus et al., 2010; Ariumi et al., 2007) such that low expression may limit viral replication. To investigate this hypothesis, DDX3-specific qRT-PCRs were performed on all liver biopsies from the 22 HCV infected patients and on the 6 biopsies from donor, ALD, NASH, AIH and nA/nB/nC viral hepatitis livers. All 8 biopsies showed comparable levels of DDX3 for the same patient (Figure 5.10A +B) and the expression was highly comparable between 21 of the 22 HCV positive livers (mean relative expression of 0.28 over GAPDH) (Figure 5.10C). Comparing these values to the hepatic viral loads measured in the same patient cohort revealed a weak correlation (r^2 = 0.017, p = 0.94). Furthermore, comparing the plasma viral load with the expression levels for all 22 patients showed no correlation (r^2 = 0.22, p = 0.4).

Comparative analyses of DDX3 expression between samples collected from subjects with different liver diseases and from the donor livers demonstrate comparable levels of mean DDX3 expression between subject groups, however high variations were observed within most of the data sets (Figure 5.10C). Hepatic samples from NASH, AIH and donor livers revealed an outlier for each group with up to 2.2-fold higher expression levels than the mean expression (Figure 5.10C). These observations suggest that variation in DDX3 expressions occurs naturally in the population and is not associated with a specific liver disease.

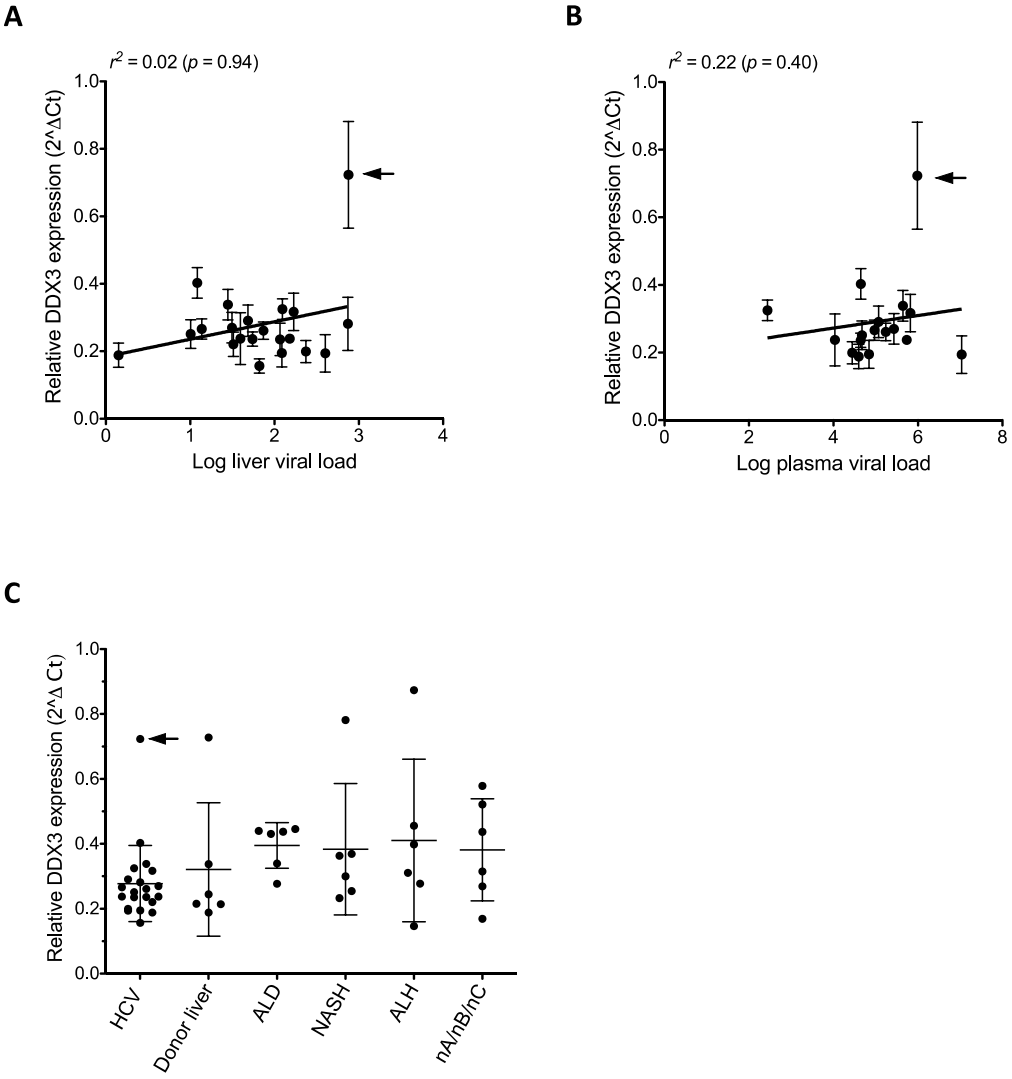


Figure 5.10 Association between hepatic DDX3 expression and HCV RNA loads. DDX3-specific qRT-PCR was performed for HCV positive liver samples and the additional non-HCV hepatic specimens. The expression levels were normalised to GAPDH. Linear regression analyses were performed for the DDX3 expressions versus **A.** mean hepatic viral loads (Spearman's correlation coefficient: r^2 = 0.0.02, p = 0.94) or **B.** the plasma viral loads (Spearman's correlation coefficient: r^2 = 0.22, p > 0.40). **C.** DDX3 expressions for all studied disease groups. Each point represent the mean DDX3 expression for one patient/donor and the horizontal lines the overall mean of all study subjects with the standard deviation. Patient 6 is indicated with an arrow.

5.5 Hepatic ISG expression in HCV-infected paediatric patients

Between 0.05 and 5% of children are believed to be infected with HCV worldwide (Le Campion et al., 2012). These children are likely to respond better to PEG-Interferon-α + ribavirin treatment than adults and have a higher chance of achieving SVR independent of their IL28-β genotype (Domagalski et al., 2013; Komatsu et al., 2013). The reasons for these different treatment outcomes for adults and children are unknown, but could potentially be related to age-dependent hepatic response to interferon. Knowledge of the paediatric hepatic immune response to HCV is limited and it is not known if the ISG expression profiles can predict treatment response.

To assess this, the expressions of the 4-gene classifiers were evaluated in single needle biopsies obtained from 11 chronically infected children with genotype 1a/1b (n = 7) or 3a (n = 4). The biopsies were taken prior to the treatment with PEG-Interferon-α and ribavirin, which led to SVR for 9 of the 11 children. Information about the pediatric patient cohort can be viewed in Table 5.4.

Table 5.4: Paediatric patient overview	
Number of patients	11
Age (year)	3.5 – 14.1 (Median 6.4)
Plasma viral load (copies/ml plasma)	$1.7 \times 10^4 – 2.4 \times 10^6$ (Median 1.9×10^5)
HCV Genotype:	
Genotype 1a/b	7
Genotype 3a	4
Length of Peg-IFN treatment (days)	164- 483 (Median 324)
Outcome of Peg-IFN treatment:	
Non-responders	2
Responders	9

Total liver RNA was extracted from the biopsies for quantifying viral load and gene expressions by qRT-PCR. The quantity of GAPDH-adjusted HCV RNA copies demonstrated up to 4-log variation between the children samples (mean: 98 HCV copies/µl total RNA extract ± 166.8) (Figure 5.11A) The estimated quantity of HCV RNA in the sample from child 2 (conc. 0.056 HCV RNA copies/µl RNA extract) is well below the estimated assay threshold of 10 copies / µl sample (McKeating group protocol). However we know that this child is HCV positive since his/her plasma burden was 1.7×10^4 copies / ml at

the time of biopsy sampling and the low hepatic load could be due to technical errors in the RNA extraction.

Comparison analysis between the hepatic HCV RNA burden in the needle biopsy and the plasma load measured at the same time the biopsy was taken (excluding child 2), revealed inverted correlation between the loads in the two compartments, however the association were non-significant (r^2 = -0.58, p = 0.08) (Figure 5.13B). It is unknown why infected children would have an inverted association between HCV RNA loads, but low release of virus from the liver or extra-hepatic HCV expression in children can potentially explain the association.

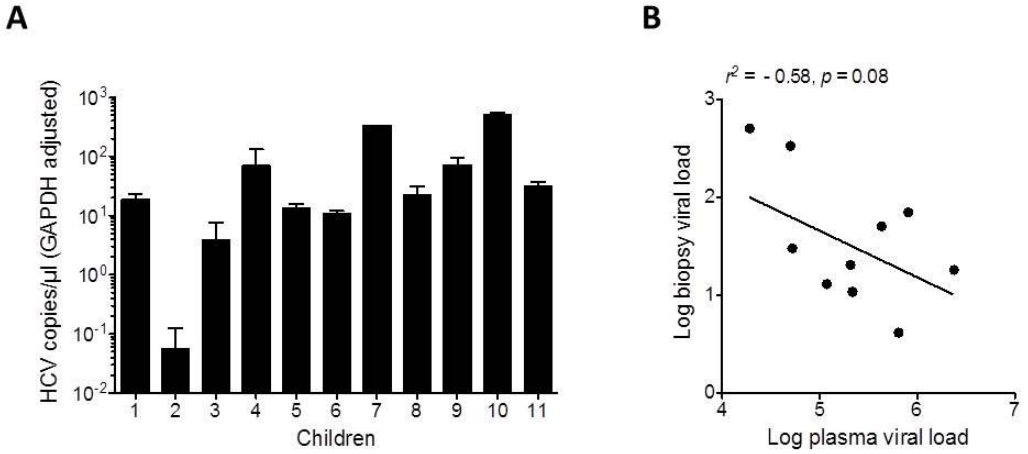
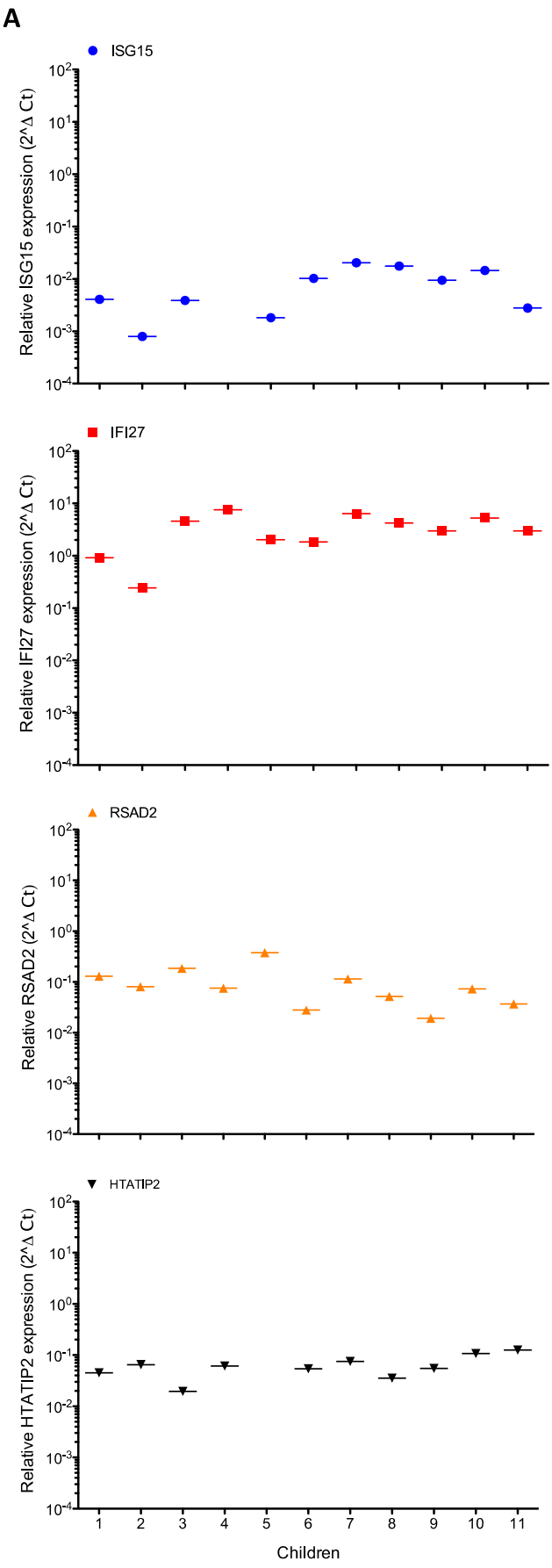


Figure 5.11 Paediatric intra-hepatic HCV RNA loads and the correlation with plasma load.

A. Total liver RNA was extracted from needle biopsies collected from 11 HCV positive children and the HCV RNA levels quantified by target-specific qRT-PCR. The viral loads are normalised to the internal GAPDH expression level. The data is presented as means plus standard deviations for duplicate samples in two independent measurements. **B.** Linear regression curve between paediatric plasma and mean hepatic viral load (Pearson's correlation coefficient: r^2 = -0.58, p = 08).

The expression levels of the 4-gene classifiers were quantified using the same protocol as for the adult samples. Like the adults, the IFI27 expression levels in the children samples were significantly higher than the other 3 gene classifiers. The data set also revealed that the expression levels were variable between subjects with the ISG15 expression differing by 28.6-fold, IFI27 by 37.5-fold and RSAD2 up to 19.5-fold. Unexpectedly, variations were also observed for HTATIP2 (up to 6.5-fold) (Figure 5.12A). Plotting the ISG15, IFI27, RSAD2 and HTATIP2 expression levels together with the plasma and hepatic viral loads for each child provides a clearer overview of the patient-specific RNA levels (Figure 5.12B). The ISG15 expression level in the sample from child 5 and HTATIP2 from child 5 were not quantified due to technical failures.


A
ISG15
Relative ISG15 expression (2^Δ Ct)
IFI27
Relative IFI27 expression (2^Δ Ct)
RSAD2
Relative RSAD2 (2^Δ Ct)
HTATIP2
Relative HTATIP2 expression (2^Δ Ct)
1 2 3 4 5 6 7 8 9 10 11
Children

B

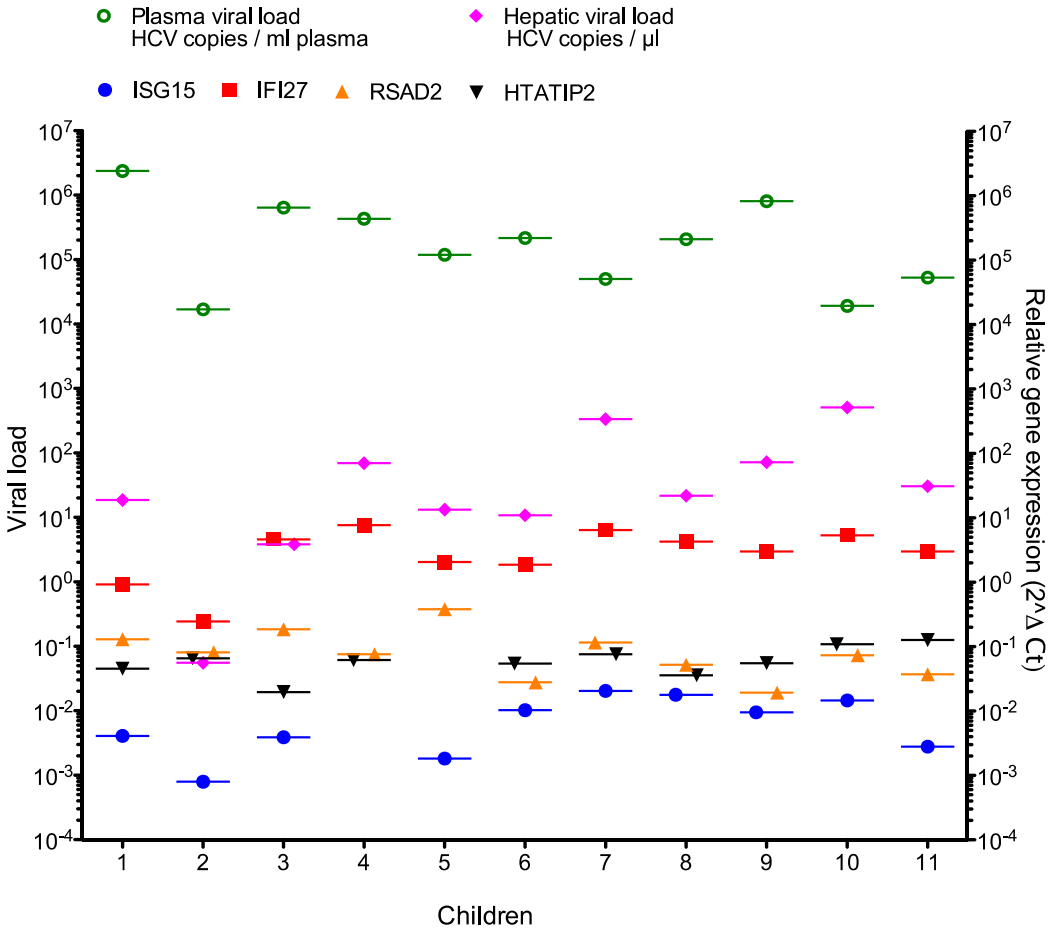


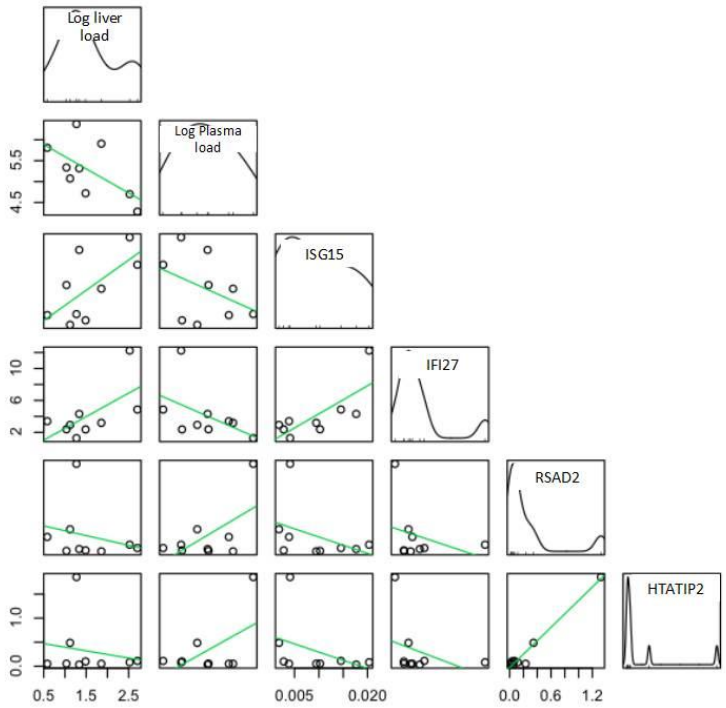
Figure 5.12 Hepatic expression profile for the 4-gene classifiers in paediatric patients.

Total liver RNA was extracted from needle biopsies of 11 HCV positive children and HCV viral load quantified by target-specific qRT-PCR. All mRNA values are normalised to GAPDH. **A.** Expression pattern for the individual classifiers for 11 children. Each point represents the expression in one biopsy. **B.** Re-plotting the individual gene expressions (right Y-axis) in one graph together with hepatic and plasma HCV RNA loads (left Y-axis) for each child. Green hollow circle: plasma load (HCV copies/ ml plasma), pink diamond: biopsy load (GAPDH adjusted HCV copies / µl), blue circle: ISG15, red square: IFI27, orange triangle: RSAD2, black inverted triangle: HTATIP2.

Linear regression analysis was performed for the plasma and hepatic viral RNA loads measured for the paediatric patients and the expression levels of the 4-gene classifiers (Figure 5.15A + B). Amongst the classifiers, correlations were noted between ISG15 and IFI27 (r^2 = 0.78, p < 0.01) and between RSAD2 and HTATIP2 (r^2 = 0.99, p < 0.001). Unlike our earlier biopsy data collected from adults, we noted an association between the liver viral load and IFI27 (0.64, p < 0.05), but not for any of the other classifiers. However, this value alone was not sufficient to make the combined expression pattern for the classifiers predictive of the hepatic viral load (multiple correlation coefficient: r^2 = 0.49, p = 0.41, F-factor = 0.41). Excluding the non-ISG gene, HTATIP2, from the analysis did not change the result significantly (r^2 = 0.45, p = 0.28, F-factor = 1.6).

The plasma viral loads in the children showed a strong correlation with both the RSAD2 (0.90, p < 0.001) and HTATIP2 (0.88, p < 0.001) expressions (Figure 5.13A + B). As a result the overall multiple correlation coefficient indicates an association between the plasma load and the combined gene expression of the classifier (r^2 = 0.84, p = 0.03, F-factor = 6.95). Excluding HTATIP2 makes the data more significant (r^2 = 0.83, p = 0.009, F-factor = 10.05).

A



B

Correlation matrix for viral loads and the 4 classifiers expressions

	Liver load	Plasma load	ISG15	IFI27	RSAD2	HTATIP2
Liver Load	1.00	0.07^{NS}	0.63^{NS}	0.64^{*}	-0.20^{NS}	-0.19^{NS}
Plasma load	0.07^{NS}	1.00	-0.24^{NS}	-0.29^{NS}	0.90^{***}	0.88^{***}
ISG15	0.63^{NS}	-0.24^{NS}	1.00	0.78^{**}	-0.31^{NS}	-0.32^{NS}
IFI27	0.64^{*}	-0.29^{NS}	0.78^{**}	1.00	-0.24^{NS}	-0.29^{NS}
RSAD2	-0.20^{NS}	0.90^{***}	-0.31^{NS}	-0.24^{NS}	1.00	0.99^{***}
HTATIP2	-0.19^{NS}	0.88^{***}	-0.32^{NS}	-0.29^{NS}	0.99^{***}	1.00

*: $p < 0.05$, **: $p < 0.01$ ***: $p < 0.001$, NS: Not Significant

Figure 5.13 Association between the 4-gene classifiers and HCV RNA loads for paediatric patients.

A. Scatter plot matrix of all components. The density plot for each gene expressions and viral loads are seen in the diagonal boxes. On each side of the diagonal are the pairwise comparisons represented as scatter plots (black circles) together with the linear correlation line (green). The graph is exported from the analyses using the R commander software. **B.** Pearson's correlation coefficient calculated in a pairwise manner for the viral loads and the 4 classifiers' gene expressions. Statistical significance was evaluated by the two-way unpaired t-test (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, NS: Not significant.

Dill and colleagues previously reported that adult patients with high expressions of ISG15, IFI27 and RSAD2 prior to PEG-IFN treatment are less likely to achieve SVR compared to those with lower expression (Dill et al., 2011). It is not know if the expression of these ISG predicts treatment outcome in paediatric patients. In this study 9 out of 11 children had responded to the treatment. Association analysis between gene expression and treatment outcome indicated higher ISG15 and IFI27 expression and lower levels of RSAD2 for the 9 children that responded compared to the 2 non-responders. However, the limited number of non-responders, prevented any statistical interpretation of this data (Figure 5.14A).

Comparative analysis of the gene expression levels between children and adults revealed that the two age groups vary in their relative hepatic expression patterns quantities. The children generally had lower expression of ISG15 and higher expression of IFI27, RSAD2 and HTATIP2, with IFI27, RSAD2 and HTATIP2 being statistical significant (unpaired two-tailed t-test $p < 0.5$) (Figure 5.14B). The observation that the treatment responsive children have greater gene expressions for 2 out of 3 ISGs contradicts the prediction by Dill and his colleagues that in adults high ISG expression decreases the likelihood of responding to treatment.

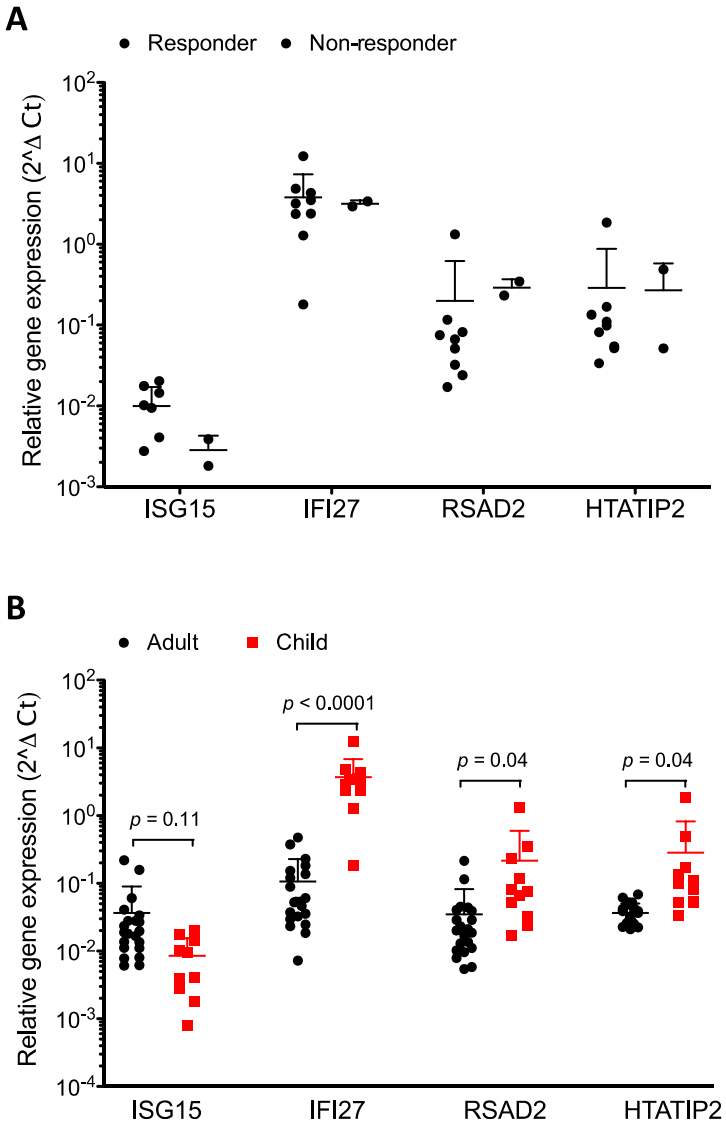


Figure 5.14 Association between the 4-gene classifiers and treatment response for paediatric patients.

A. The expression profiles for the 4-gene classifiers were divided into groups according to the treatment outcome for the particular paediatric patient. The data is represented by the means plus standard deviation for each patient group. **B.** Comparative analysis of the 4 classifiers' gene expressions between HCV chronically infected adults and children. The horizontal lines represent the means and standard deviation. Statistical significant analyses were performed with the unpaired two-tailed t-test.

5.6 Discussion

In 2011, Dill *et al.* reported that the level of hepatic ISG15, IFI27, RSAD2 and HTATIP2 expression in liver biopsies from chronically HCV infected patients collected prior to PEG-IFN treatment could predict treatment outcome. The predictive value of these 4-gene classifiers was independent of the infecting HCV genotype or IL28-β status (Dill et al., 2011). High expression of these transcripts in the liver is consistent with endogenous activation of IFN signalling that dampens the cellular response to IFN-based therapies and as a result patients with high expression of the 4-gene classifiers have a lower chance of responding to treatment.

Previous studies characterising hepatic ISG expression have utilised single biopsies from the chronically infected liver and it is unknown if the ISG expression varies across the liver. In this study we investigated the expression pattern of ISG15, IFI27, RSAD2 and HTATIP2 across multiple sites in the explant liver from subjects with HCV-associated end-stage liver disease and compared expression patterns to disease matched samples of non-viral aetiology including ALD, NASH, AIH and nA/nB/nC viral hepatitis. In addition, we had the opportunity to study hepatic ISG expression in biopsies from HCV infected children with early stage liver disease.

We investigated the potential association between hepatic ISG expression and HCV replication to gain a better understanding on how these parameters may

influence each other. The potential models include: (*i*) high levels of HCV replication would stimulate RIG-I and TLR pathways and consequently high ISG expression; (*ii*) high levels of HCV replication lead to an increased cleavage of MAVS and deactivation of the RIG-I pathway leading to low ISG expression; or (*iii*) high local production of IFN would lead to high ISG expression that controls viral replication and limits HCV RNA levels.

The expression level of ISG15, IFI27, RSAD2 and HTATIP2 were measured in 8 hepatic specimens collected from the explant liver of 22 subjects with end-stage liver disease. Liver biopsies showed similar expression levels of HTATIP2 across the liver and 18 of 22 (82%) patients also showed comparable levels of ISG15, IFI27 and RSAD2. This observation suggests a uniform stimulation of the hepatic innate and adaptive immune response by the production of type I, II and III IFNs from infected hepatocytes and from liver resident and infiltrating immune-cells. Dill *et al.* used single needle biopsies from early stage chronically infected patients without previously showing that a single specimen would represent the expression pattern in the whole liver (Dill et al., 2011). The observation from our patient cohort indicates that a single liver biopsy is representative for the hepatic expression level of the 4-gene classifiers, during end-stage liver disease in most cases.

Four of the 22 patients showed variable expression of the three ISGs across the liver (patients 5, 9, 15 and 17). This variability is not linked to a specific genotype or to differences in sample processing time. HTATIP2 hepatic mRNA

levels for these patients did not vary significantly, leading us to believe that the variability in ISG expressions across the liver for these patients is real. One possible explanation for the differences in ISG mRNA levels is that the three investigated ISGs respond to different stimuli; IFI27 and ISG15 expressions are stimulated by Type I IFNs (Dill et al., 2012) and RSAD2 by Type I, II and III (Fitzgerald, 2011). Although patients with chronic HCV infection show high levels of infiltrating T cells in the liver, it is possible that only a minority of these are HCV-specific (Bowen and Walker, 2005a). This would suggest that some livers might have compartments of high Type II IFN concentrations in association with the HCV specific lymphocytes, leading to a variable expression of Type II ISG such as RSAD2. Furthermore, the liver may in some instances have an uneven distribution of infected hepatocytes, leading to areas of high Type I and III IFN production and expression of ISG15 and IFI27.

Recently Wieland *et al.* reported that variable levels of IFI27 expression *in vivo* associated with the non-random distribution of infected hepatocytes (Wieland et al., 2013). However other studies failed to report a correlation between the frequency of infected hepatocytes and expression of the ISGs MxA, IFI1, IFI27 (Stiffler et al., 2009) and IFITM3 (Kandathil et al., 2013). The contradicting observations in the Wieland *et al.* and Kandathil *et al.* studies are likely to result from either different technical approaches being used, the study cohorts having different levels of fibrosis or that IFITM3 and IFI27 are stimulated by different types of interferon, i.e. IFITM3 expression is stimulated by Type I and II IFNs (Kandathil et al., 2013) and IFI27 by Type I IFN (Dill et al., 2012). We

were unable to determine the frequency of infected hepatocytes in biopsies from any patients in our study, due to technical difficulties in staining liver biopsies for HCV antigens and it is therefore not possible to reject or support the hypothesis of variable numbers of infected hepatocytes leading to differences in ISG expression for patient 5, 9, 15 and 17.

Comparative analysis of the expression level for the 4-gene classifiers demonstrated up to 65-fold variation in hepatic ISG15, IFI27 and RSAD2 expression between the explanted livers. We noted that the expression levels of the ISGs correlated, suggesting comparable expression within one liver, but that the level of stimulation may vary between patients.

In chapter 4 we found comparable levels of HCV RNA between different sites in the liver, however, the viral burden varied significantly between study subjects. These observations led us to investigate a potential association between the ISG expression and hepatic viral burden. To our surprise, we failed to detect any relationship between hepatic ISG expression levels and HCV RNA levels at macroscopic and microscopic levels. Furthermore, comparing the ISG expression levels with the plasma viral burden revealed no correlation, indicating that the hepatic ISG expression does not affect the level of HCV particles release to the periphery.

Our observations agree with previous reports stating that hepatic (Bellecave et al., 2010; Kandathil et al., 2013; Stiffler et al., 2009) and plasma HCV RNA levels (Bellecave et al., 2010; Urban et al., 2010) do not correlate with hepatic ISG expression during chronic infection. However, they stand in contrast to a previous study, which reported a negative correlation between serum HCV RNA and the transcript level of 15 different ISGs, when performing multivariate analysis for patients who responded to treatment and those who had relapsed (Honda et al., 2010). In that study, biopsies were taken prior to PEG-IFN and ribavirin treatment, but it is not known when the serum HCV RNA was measured. Comparing hepatic ISG levels and serum viral burden at different time points during IFN-based treatment may lead to misleading conclusions.

The expression pattern of the 4-gene classifiers was further stratified for clinical and viral parameters, such as HCV-associated liver disease (cirrhosis or HCC), prior response to IFN treatment and infecting HCV genotype. We did not find any association between these clinical parameters and ISG expression that could explain why these patients had failed to respond to IFN treatment. Neither did we find an association between ISG expression pattern and the infecting HCV genotype. Previous reports have shown that patients infected with HCV genotype 1 have higher hepatic ISG15 and RSAD2 expression compared to those infected with genotype 3 viruses (Bellecave et al., 2010; Broering et al., 2010). This is supposedly due to HCV genotype 1 exposing poor NS3/NS4 cleavage of MAVS compared to other HCV genotypes (Bellecave et al., 2010), leading to higher IFN-α/β expression.

Investigating the level of cleaved MAVS in our patient cohort would provide new insight into the HCV regulation of endogenous IFN and ISG expressions during end-stage liver disease.

In this study we investigated the expression of the 4-gene classifiers in hepatic tissue from patients diagnosed with NASH, ALD, AIH and nA/nB/nC viral hepatitis end-stage liver disease and from healthy liver donors. Comparing ISG15, IFI27 and RSAD2 expression levels within each disease group revealed up to 5.2-fold variations between samples. In contrast, the HCV infected patients showed variations up to 65-fold. The difference in spread between the HCV and the ALD, NASH, NIH and nA/nB/nC viral hepatitis cohorts are likely to reflect the variable number of samples in each disease group. Unexpectedly, hepatic samples donated from healthy individuals varied up to 24-fold in the ISG levels, leading us to question whether the donor livers are as healthy as first assumed. The term “donor livers” in this study, covers livers donated for transplantation by healthy donors. These livers were not used for transplantations because of their high fat content and instead send to the pathology department. A fatty liver will in 20-30% of the cases lead to liver inflammation (Tilg and Moschen, 2010) and it is therefore likely that some of these donated livers already had an activated immune response. The ideal control for all *in vivo* liver studies would be specimens from disease-free livers with no sign of hepatic inflammation, however there is no clinical justification to conduct needle-biopsies from healthy individuals or to take samples from livers suitable for transplantation.

Comparing ISG expression in hepatic tissue from HCV infected end-stage liver disease patients with tissue from donor liver and end-stage NASH, ALD, AIH and nA/nB/nC viral hepatitis livers revealed significant differences in IFI27 expression level, suggesting, that HCV infected patients have a unique expression pattern of IFI27 compared to other advanced liver disease and donor livers. In addition, significant differences were found for ISG15 in hepatic tissue collected from NASH patients, however the difference was minimal and may result of low sample numbers.

The mean expression of the 4-gene classifiers was comparable between tissues collected from the non-viral associated end-stage liver diseases, independent of aetiology. This observation suggests that these end-stage liver diseases have activated IFN signalling pathways. Searching the literature we found that all three non-viral associated liver diseases can be linked to active IFN production. The inflammatory condition in NASH is believed to be partly caused by free fatty acids acting on TLRs in the liver (Tilg and Moschen, 2010) and for ALD the heavy alcohol consumption may lead to increased gut epithelial permeability, which in turn result in elevated endotoxemia and an increased activation of hepatic TLRs (Wang et al., 2012). Although we did not observe high RSAD2 expression compared to ISG15 and IFI27 in the samples collected from AIH livers, this condition is believed to be driven by self-antigenic peptides being presented by the major histocompatibility molecule to

cellular members of the adaptive immune response, leading to high production of type II IFN (Liberal et al., 2013).

Although children tend to have a higher response rate than adults to PEG-IFN-α (Abdel-Hady et al., 2011; Wirth, 2012), it is not known whether this correlates with a lower hepatic ISG expression pre-treatment. To address this issue we studied the expression of the 4-gene classifiers in hepatic needle biopsies collected from children and adolescents chronically infected with HCV. Comparing hepatic and plasma viral burden with the expression of the 4-gene classifiers for the children samples revealed that IFI27 significantly correlated with hepatic HCV RNA load, whereas RSAD2 and HTATIP2 correlated with the plasma load. The current knowledge about the RSAD2 and HTATIP2 expression pattern and protein function does not reveal a reason why these genes are found to correlate with the plasma instead of hepatic load. Since HTATIP2 is a ubiquitously expressed tumour suppressor gene, we did not expect the expression level to be correlated with the plasma viral load and therefore suspect that the detected association between high plasma HCV RNA levels and HTATIP2 is unrelated and the same may apply to RSAD2.

Multiple component analysis revealed that the combined expression pattern for the 4 classifiers correlated with plasma, but not hepatic HCV RNA levels for the children samples. However looking at the expression of individual genes we found that the positive correlation was mainly driven by the high RSAD2 and HTATIP2 partial correlation coefficients, such that ISG15 and IFI27 did not

show any correlation. These observations highlight the importance of evaluating the individual expression patterns for the genes as well as the combined. Additional biopsies from other HCV infected children are currently being evaluated for the expressions of the 4-gene classifiers and will form a more robust data set for the multiple component analysis.

Nine of the eleven children in this study responded to IFN treatment. Stratifying ISG expression data for the treatment response revealed limited association between the 4-gene classifiers and treatment outcome. However, the low number of samples from non-responders (n = 2) makes it difficult to assign any statistical significance to this observation. High response rates in adults have previously been associated with young age, short infection period, low fibrosis levels, low alcohol consumption and low body weight (Komatsu et al., 2013). Since children have all these characterisations it may be that they automatically are more likely to clear the virus following treatment than adults, independent of their ISG expression status. The medical records of the 11 evaluated children have not yet been reviewed, but visiting these may help identifying any potential host factors, such as liver fibrosis stage and body-weight, which may explain the high response rate for this cohort.

Comparing the hepatic expression pattern of the 4-gene classifiers between adults and children reveal that children biopsies show higher mean expression levels of IFI27, RSAD2 and HTATIP2 and a lower expression of ISG15. The particular expression pattern of high anti-viral IFI27 and RSAD2, together with

low pro-viral ISG15 may pose an alternative explanation for the higher clearance rate in children. This rationalisation is in contrast with current knowledge about treatment response in adults. The detection of high IFI27 and RSAD2 suggests that children, do not experience endogenous IFN refractoriness – a condition where high endogenous IFN expression desensitises IFN signalling cascades. Instead the PEG-IFN treatment may boost anti-viral ISG expression even higher for children, leading to an increased clearance rate. Although this explanation seems attractive, it should be noted that the samples from adults and children are not matched for liver disease in this study and that adults with end-stage-liver disease may have a very different expression pattern compared to early-stage infections. An indication of unmatched adult and children samples is seen for the linear regression data for the HCV RNA loads between plasma and liver, where an inverted correlation is seen for the children samples (Figure 5.11) compared to a positive correlation for the adults (Figure 4.8).

Observations from this study suggest a potential age-dependent responsiveness to IFN treatment. Previous studies have shown that age has a negative effect on the IFN-γ production in individuals infected with influenza-A, leading to a higher clearance rate for younger individuals than adults (Powers, 1993). This observation is believed to result from lower CD4+ and CD8+ T-cells counts in elderly individuals (Linton and Dorshkind, 2004). To date it is not known whether aging effect hepatic ISG expression during HCV infection. To investigate this longitudinal needle biopsies would be required over the

period of childhood, adolescent and adulthood. Considering that most children clear the virus it is not clinical practise to take later hepatic samples. Instead we are currently assessing the ISG expressions in chronically infected adults without end-stage liver disease, which are better matched the children samples. In addition, we are planning to study hepatic ISG expression in biopsies from uninfected children.

In the search for alternative genes that may associate with the viral replication, we found that DDX3 was up-regulated in the hepatic tissue collected from adult patient 6 (the patient with the highest mean hepatic viral burden in our adult cohort). The human DDX3 belongs to the family of DEAD-box RNA helicases, which are involved in a large variety of cellular RNA processing. Knockdown of this helicase results in reduced HCV replication *in vitro* through an as yet unidentified mechanism (Angus et al., 2010; Ariumi et al., 2007). The observation has lead to a model where high DDX3 expression may associate with HCV RNA levels. Investigating the DDX3 levels in the 8 hepatic sites collected from the remaining 21 HCV positive patients revealed no correlation with hepatic or plasma HCV RNA loads. Furthermore, DDX3 levels were comparable between the 21 HCV positive livers and the specimens collected from livers with non-HCV associated advanced liver diseases and donor livers, suggesting that this RNA helicase is constitutively expressed in the liver regardless of the hepatic condition. A previous study reported that the vaccinia virus encoded K7 protein can bind DDX3 helicase, leading to a reduction in Type I IFN expression (Schroder et al., 2008). HCV core has also been

reported to bind DDX3 (Angus et al., 2010; Ariumi et al., 2007), but it is unknown if this interaction would blunt Type I IFN expression. A potential down-regulation of innate immune response by a DDX3-core complex could explain the high viral burden in patient 6, however the hepatic ISG expression for patient 6 was not significantly lower compared to the other patients in the cohort. The anti-viral array revealed that TNF and IL-8 were up-regulated in samples from patients 6 and 5 (patients with high viral loads). HCV NS5A has previously reported to induce IL-8 expression, which may desensitise cells to IFN mediated signalling (Polyak et al., 2001). This gene would be a relevant candidate for future investigations.

In conclusion we found comparable expression of the previously published 4-gene classifiers across the liver of subjects with HCV-related end-stage liver disease. Although the expression pattern within a liver was constant the pattern of ISG expression varied significantly between subjects. The variable hepatic ISG expression levels between patients did not correlate with plasma and hepatic viral load, infecting HCV genotype, indication for liver transplantation and prior response to IFN treatment. Comparing the expression of the 4-gene classifiers with samples collected from subjects with non-HCV associated advanced liver disease revealed that the high level of IFI27 was unique to patients infected with HCV. In this study we also found that biopsies from HCV infected children express higher levels of anti-viral ISGs and a lower expression of the pro-viral ISG15 compared to liver specimens collected from adults with end-stage liver disease. This observation

suggests age-dependent responsiveness to IFN and may explain the higher responds rate to IFN therapy seen for paediatric patients.

6 PHYLOGENETIC ANALYSIS OF PATIENT-DERIVED HCV VARIANTS

HCV is capable of forming heterogeneous populations *in vivo*, which has led to intensive studies of its quasispecies genetic complexity within infected individuals and experimentally infected chimpanzees (Brown et al., 2005; Farci, 2011; Farci et al., 2000; Farci et al., 1996; Gomez et al., 1999; Logvinoff et al., 2004; Ray et al., 2005; Simmonds, 2004). Most of these studies have focused on the longitudinal evolution of the hyper-variable region 1 (HVR1), located at the 5'-end of the viral encoded E2 gene. Several samples are required for these types of studies and the preferred specimens have therefore been plasma and/or serum, which are obtained by far less invasive methods than liver biopsies.

In other studies a single liver biopsy has been included in the investigation of HCV genetic distribution between the plasma and the liver, but without any empirical evidence that one needle biopsy is predictive for the whole liver. To date only 2 studies have looked at the HCV quasispecies at different anatomical sites in the liver, however the results from these have been contradicting. In 1999 Sakai and colleagues reported evidence for hepatic compartmentalisation between non-cancerous hepatic tissue and HCC nodules (Sakai et al., 1999), whereas Cabot et al. reported limited support for compartmentalisation between three hepatic sites (Cabot et al., 1997).

In contrast to most other HCV quasispecies studies, our study focuses on the full-length region of the most variable domains of the HCV genome; E1 and E2 (1.7kb). In addition, the downstream p7 gene was included for HCV genotype 3a. As mentioned in the previous chapters samples were taken from different sites in the explanted liver, which enabled the direct investigation of intra-hepatic distributions of the genetic HCV population (phylo-anatomy). In the same study we compared hepatic viral sequences to those in the plasma at the time of liver transplantation.

This study provides a snapshot into the distribution of HCV quasispecies variants within 8 different sites of the liver and in the plasma for 6 individual patients.

6.1 Evaluating HCV cDNA synthesis in the presence of total hepatic RNA

The amount of HCV RNA make up a tiny fraction compared to the vast amount of total liver RNA in the RNA extractions. To assess whether the amount of total RNA would interfere with the synthesis of HCV cDNA, a spiking assay was performed. Increasing amounts of total RNA from a HCV negative liver explant diagnosed with ALD, were mixed with a constant amount of full-length H77 HCV RNA (genotype 1). The RNA mixture was used for cDNA synthesis and the quantity of HCV cDNA assessed by qPCR (Figure 6.1). This spiking assay revealed that the synthesis of HCV cDNA remained constant even in the presence of large amounts of total liver RNA. This indicates that total liver RNA does not affect HCV cDNA synthesis.

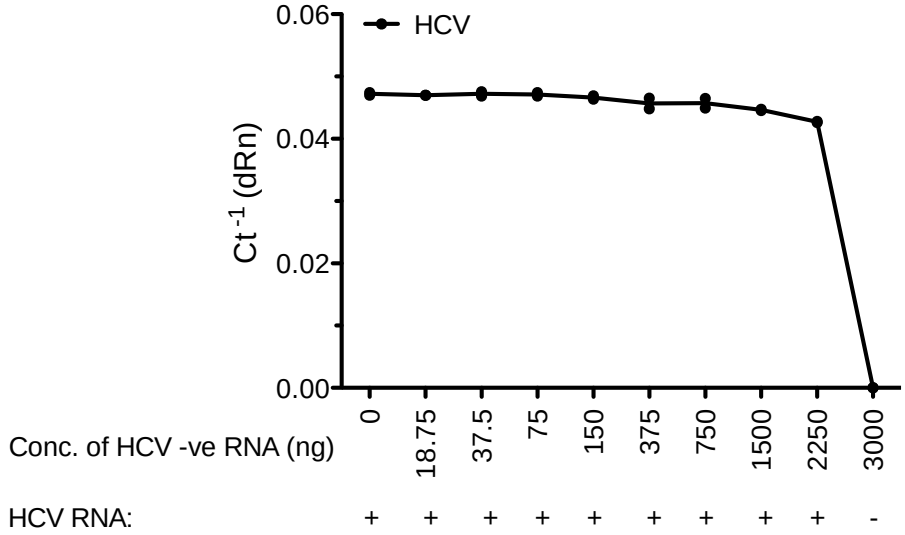


Figure 6.1 Detection of HCV RNA in the presence of total liver RNA.

Increasing amounts of total hepatic RNA from a HCV negative (-ve) ALD liver were spiked into a constant level of HCV RNA (8.3 x 10^5 copies). The concentration of the HCV –ve RNA can be seen in the x-axis. HCV specific and random cDNA was afterwards synthesised from the RNA mixtures and the HCV cDNA levels (black circles) detected by target specific qPCR. All relative quantifications were done in duplicate samples. The lines between the measurements represent the mean for the duplicate measurements.

6.2 Phylogenetic analysis of HCV consensus sequences

Total hepatic RNA from patients 1, 4 and 5 (genotype 3a) and from patients 2, 7 and 9 (genotype 1a) were used for reverse transcription and amplification of 5'UTR, core, E1, E2, p7 and NS2 (3.4kb) or E1, E2 (1.7kb), respectively. The initial aim was to amplify the core to NS2 fragment from both genotypes, which could be used to generate chimeric HCVcc clones (Lindenbach et al., 2005). However, after several unsuccessful attempts to amplify this fragment from genotype 1a, it was decided to assess the ease with which the shorter 1.7kb E1 E2 fragment could be amplified. The amplification of this shorter fragment was found to be reliable and we adopted this approach to recover HCV sequences from genotype 1 infected material.

In the first instance bulk amplification was performed for each total liver cDNA synthesis and the PCR products gel extracted on separate gels for later consensus sequencing of the desired fragments. The bulk amplifications of patient 1 and 2 can be seen in Figure 6.2A and are representative for the remaining four patients, except for the patient 4, biopsy 6 specimen which, for unknown reasons, did not contain RNA.

To ensure good detection of potential HCV variants in the bulk PCR products, a sequencing sensitivity assay was performed. Previous sequenced bulk amplification from P1-B8 and P4-B4 had showed no nucleotide ambiguities in their consensus sequence and these were therefore used for evaluating the

sequencing sensitivity. The PCR products from the P1-B8 and P4-B4 samples were mixed at different ratios and subjected to consensus sequencing of the 5'-end of E2. Evaluating the consensus sequences for each mixture revealed that the Sanger sequencing had detected patient-specific sequences down to a ratio of 1:9, indicating that the sequencing sensitivity of our assay is ≥ 90% (Figure 6.3B).

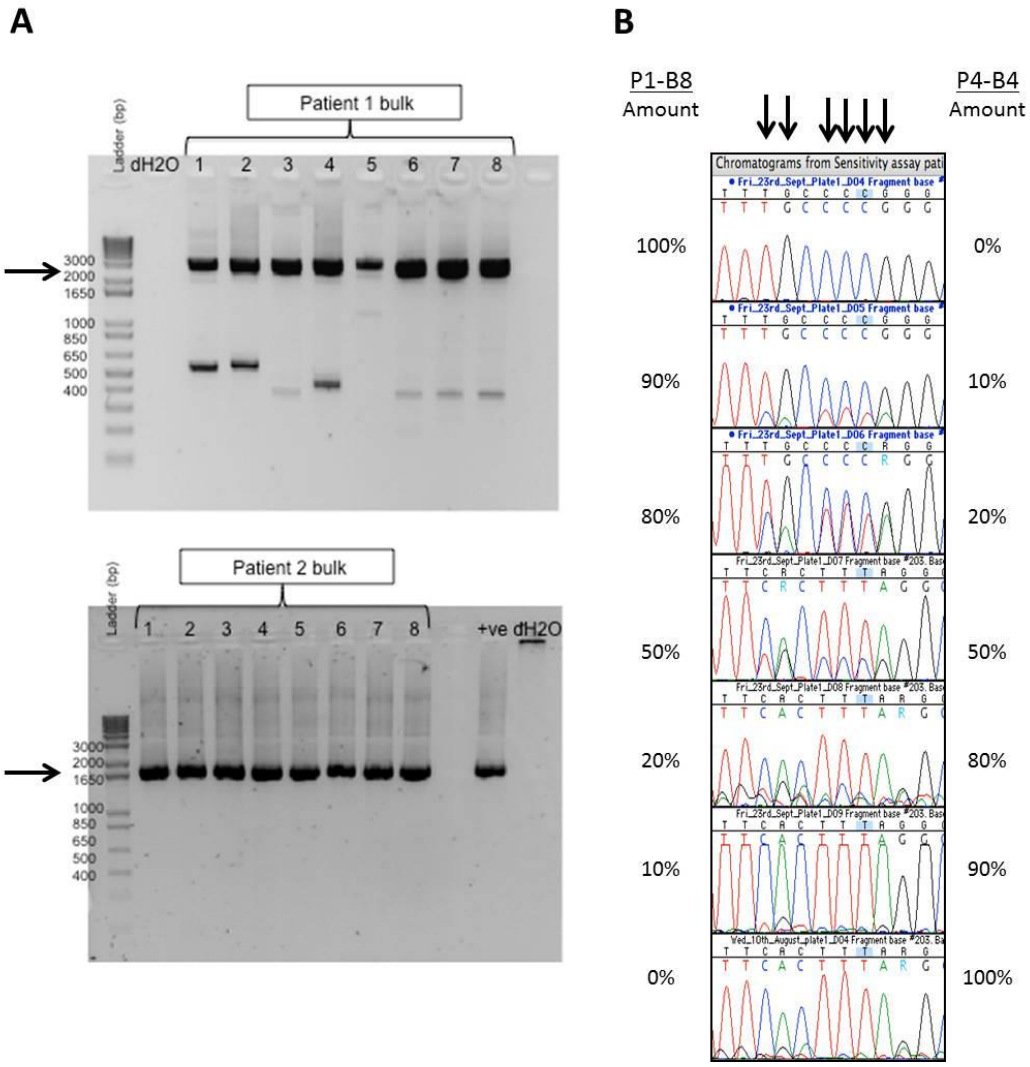


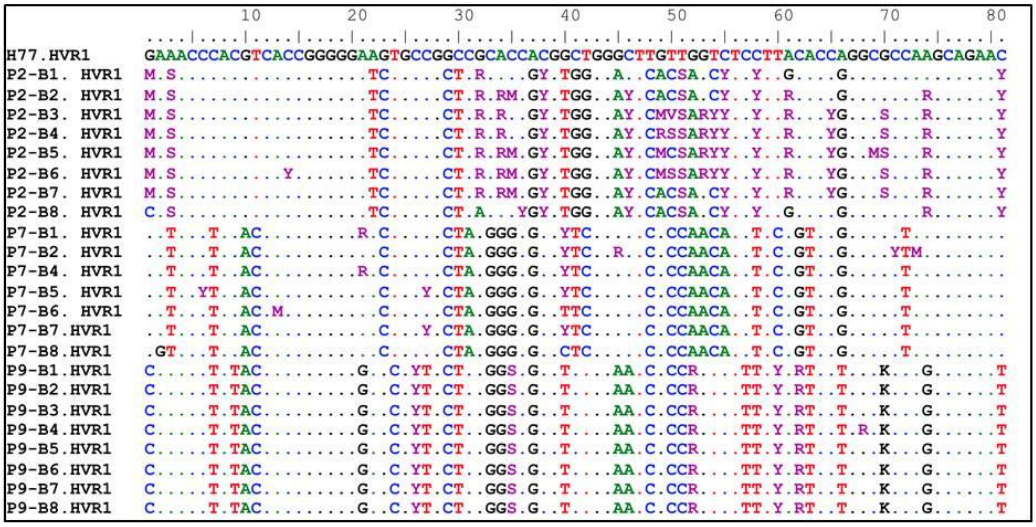
Figure 6.2 Bulk amplification of hepatic HCV sequences and evaluation of sequencing sensitivity.

A. Bulk amplification of the viral encoded genes 5'UTR, core, E1, E2, p7 and NS2 (3.4kb) from patient 1 hepatic tissue and E1 and E2 (1.7kb) from patient 2. The positive control (+ve) contained a genotype 1a template (H77) and the negative contained dH_2O instead of a cDNA template. **B.** Bulk PCR products from patient 1, biopsy 8 (P1-B8) and patient 4, biopsy 4 (P4-B4) were mixed at different ratios and the HVR1 sequenced by Sanger sequencing. The arrows indicate nucleotide ambiguities in the chromatograms.

The liver derived HCV bulk amplifications should in theory contain all the HCV variants found at the particular site and the HCV amplifications from each biopsy were therefore directly sequenced to get an estimate of the viral polymorphisms. The resulting consensus sequences contained relatively low numbers of ambiguities across the genotype 1; E1, E2 amplified region (average 112 difference ±39.59 in 1.7kb) and the genotype 3a; E1, E2, p7 region (average 64 difference ±17.34 in 1.8kb). Most of the changes were located in the E2 gene. An alignment of the consensus sequences for genotype 1 HVR1 can be seen in Figure 6.3A. Alignments of HVR1 and the genotype 3a specific HVR495 and HVR575 consensus sequences are shown in Figure 6.4A, B and C, respectively.

Nearest Neighbour sequence analyses of the genotype-specific full length consensus sequences demonstrate that HCV sequences isolated from one liver are highly comparable, with mean genetic distances between 0 and 8.0 x 10^{-4} (Figure 6.3B and Figure 6.4D). However, the genetic distances between patient-specific HCV sequences with the same genotype were significantly larger; mean branch length between genotype 1a sequences was 0.046 ±0.022 and between genotype 3a sequences the length was 0.057 ±0.028. These large distances between patient-derived HCV consensus sequences are supported by a bootstrap value of 1000.

A.



```
                    10        20        30        40        50        60        70        80
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
H77.HVR1            GAAACCCACGTCACCGGGGGAAGTGCCGGCCGCACCACGGCTGGGCTTGTTGGTCTCCTTACACCAGGCGCCAAGCAGAAC
P2-B1. HVR1         M.S.................TC.....CT.R....GY.TGG..A..CACSA.CY..Y..G....G.............Y
P2-B2. HVR1         M.S.................TC.....CT.R.RM.GY.TGG..AY.CACSA.CY..Y..R....G.......R......Y
P2-B3. HVR1         M.S.................TC.....CT.R.R..GY.TGG..AY.CMVSARYY..Y..R...YG...S...R......Y
P2-B4. HVR1         M.S.................TC.....CT.R.R..GY.TGG..AY.CRSSARYY..Y..R...YG...S...R......Y
P2-B5. HVR1         M.S.................TC.....CT.R.RM.GY.TGG..AY.CMCSARYY..Y..R...YG..MS...R......Y
P2-B6. HVR1         M.S.........Y.......TC.....CT.R.RM.GY.TGG..AY.CMSSARYY..Y..R...YG...S...R......Y
P2-B7. HVR1         M.S.................TC.....CT.R.RM.GY.TGG..AY.CACSA.CY..Y..R...YG...S...R......Y
P2-B8. HVR1         C.S.................TC.....CT.A...YGY.TGG..AY.CACSA.CY..Y..G....G.......R......Y
P7-B1. HVR1         ..T...T..AC.........R.C.....CTA.GGG.G..YTC.....C.CCAACA..T.C.GT..G.....T.........
P7-B2. HVR1         ..T...T..AC...........C.....CTA.GGG.G..YTC..R..C.CCAACA..T.C.GT..G....YTM........
P7-B4. HVR1         ..T...T..AC.........R.C.....CTA.GGG.G..YTC.....C.CCAACA..T.C.GT..G.....T.........
P7-B5. HVR1         ..T..YT..AC...........C...Y.CTA.GGG.G..YTC.....C.CCAACA..T.C.GT..G.....T.........
P7-B6. HVR1         ..T...T..AC.M.........C.....CTA.GGG.G..TTC.....C.CCAACA..T.C.GT..G.....T.........
P7-B7.HVR1          ..T...T..AC...........C...Y.CTA.GGG.G..YTC.....C.CCAACA..T.C.GT..G.....T.........
P7-B8.HVR1          .GT...T..AC...........C.....CTA.GGG.G..CTC.....C.CCAACA..T.C.GT..G.....T.........
P9-B1.HVR1          C.....T.TAC.........G..C.YT.CT..GGS.G..T....AA.C.CCR....TT.Y.RT..T...K...G......T
P9-B2.HVR1          C.....T.TAC.........G..C.YT.CT..GGS.G..T....AA.C.CCR....TT.Y.RT..T...K...G......T
P9-B3.HVR1          C.....T.TAC.........G..C.YT.CT..GGS.G..T....AA.C.CCR....TT.Y.RT..T...K...G......T
P9-B4.HVR1          C.....T.TAC.........G..C.YT.CT..GGS.G..T....AA.C.CCR....TT.Y.RT..T.R.K...G......T
P9-B5.HVR1          C.....T.TAC.........G..C.YT.CT..GGS.G..T....AA.C.CCR....TT.Y.RT..T...K...G......T
P9-B6.HVR1          C.....T.TAC.........G..C.YT.CT..GGS.G..T....AA.C.CCR....TT.Y.RT..T...K...G......T
P9-B7.HVR1          C.....T.TAC.........G..C.YT.CT..GGS.G..T....AA.C.CCR....TT.Y.RT..T...K...G......T
P9-B8.HVR1          C.....T.TAC.........G..C.YT.CT..GGS.G..T....AA.C.CCR....TT.Y.RT..T...K...G......T
```

B.

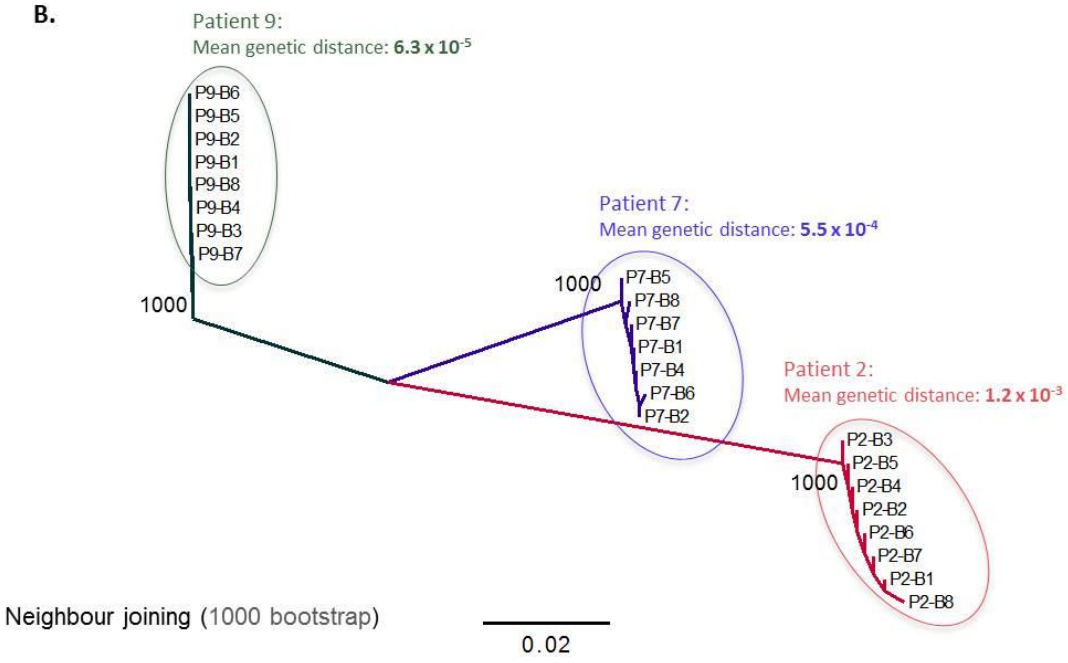


Figure 6.3 Analysis of genotype 1a hepatic E1, E2 consensus sequences.

A. Alignment of the liver-derived HVR1 consensus nucleotide sequences generated from each of the 8 segments (B1-8) from patients 2, 7 and 9 (P2, 7 or 9). The HCV H77 clone (genotype 1a) is used as a reference sequence. **B** Un-rooted Neighbour Joining phylogenetic tree generated with 1000 bootstraps. The branch lengths are a direct measure of the genetic distances. Mean HCV genetic distance within one liver is indicated next to the patient lineage. Only bootstrap-values above 750 are indicated.

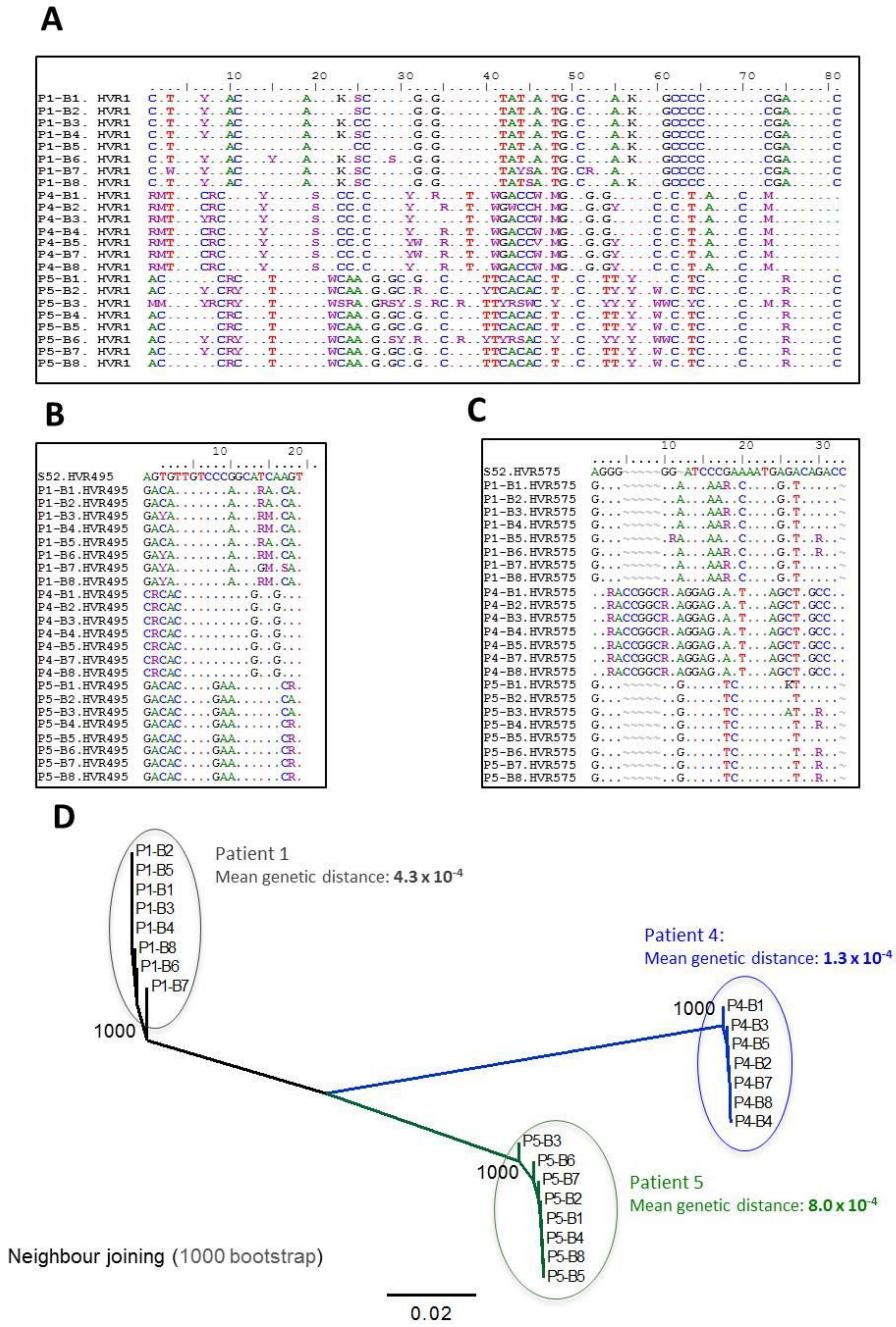


Figure 6.4 Analysis of genotype 3a hepatic E1, E2, p7 consensus sequences.

A. Alignment of the liver-derived HVR1 consensus nucleotide sequences detected in the 8 specimens from patients 1, 4 and 5 (P1, P4 and P5). **B. + C.** Nucleotide alignment of the hepatic HVR495 and HVR575 for the same patients. The HCV S52 clone (genotype 3a) is included as a reference. **D.** Un-rooted Neighbour Joining phylogenetic tree generated with 1000 bootstraps. The branch lengths are a direct measure of the genetic distances. Mean genetic distances are indicated next the patient lineage. Bootstrap-values above 750 are indicated in the phylogenetic tree.

6.3 Phylogenetic analysis of SGA-derived HCV sequences isolated from the liver

The low genetic variability in the HCV consensus sequences could potentially be the result of sequencing major HCV variants at the expense of the minor sequences (<10% abundance, see Figure 6.2A). To evaluate the efficiency of minor sequence amplifications, a spiking experiment was performed using RNA from the genotype 3a infected patient 1 (P1-B6.H3-1.6) and patient 5 (P5-B3.F3-B9) HCV clones (see Table 3.5). Decreasing amounts of P5-B3.F3-B9 RNA were spiked into a constant volume of P1-B6.H3-1.6 RNA together with a constant concentration of total hepatic RNA from an ALD liver.

Mixtures of the three RNA templates were subject to genotype 3a specific cDNA synthesis, PCR amplification for the 3.4kb fragment (Figure 6.5A) and finally Sanger sequencing of the consensus HVR1 region (Figure 6.5B). Detection of the P5-B3.F3-B9 sequences in the chromatogram was only possible at a HCV RNA mixing ratio of 1:1. Lower dilutions of P5-B3.F3-B9 RNA resulted in loss of the minor sequence (Figure 6.5B). These data indicate that HCV variants present at a concentration below 1 : 1 will either not be reversed transcribed into cDNA or not amplified during PCR.

Failure to detect HCV sequences present at low concentrations justifies the approach of single genome amplifications (SGAs) for the investigation of HCV

genetic compartmentalisation. In total 120 HCV clones were generated from liver and plasma specimens collected from patients 1, 2, 4, 5, 7 and 9 (Table 3.5). In average 20 clones were generated per patient. The E1 and E2 sequences isolated from patients infected with genotype 1 and the E1, E2 and p7 from patients infected with genotype 3a can be seen in Appendix 1.

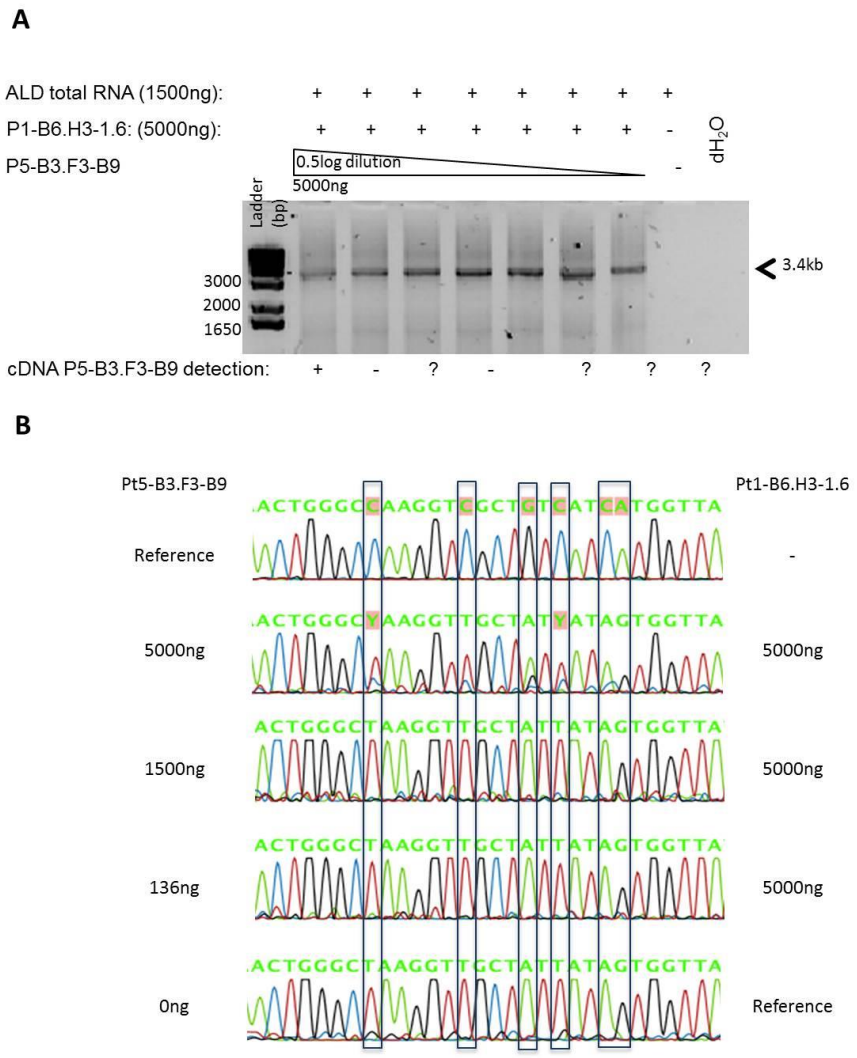


Figure 6.5 Amplification of major and minor sequences *in vitro.*

A. Synthetic RNA was transcribed from patient 5, biopsy 3 (P5-B3.F3-B9) and patient 1, biopsy 6 (P1-B6.H3-1.6) clones. A 0.5-log dilution series of P5-B3.F3-B9 RNA was spiked into a constant level of P1-B6.H3-1.6 RNA (5000ng/reaction) and total hepatic RNA from an alcoholic liver disease (ALD) patient (1500ng/reaction). HCV cDNA was generated with reverse primers against the NS2-NS3A junction, followed by a nested PCR with genotype 3a specific primers for the 3.4kb fragment. The PCR product where separated by gel electrophoresis. PCR products, which failed sequencing, are indicated with a question mark. **B.** Chromatograms from consensus Sanger sequencing of the PCR products in A. The consensus sequencing failed to detect the minor sequence at 1500ng (dilution 1:3.3 with P1-B6.H3-1.6) and 136ng (dilution 1 : 36.8 with P1-B6.H3-1.6) template RNA.

Comparative sequence analyses were performed for HVR1 in the hepatic SGA sequences. The Nearest Neighbour sequence analysis revealed that the HVR1 sequences formed patient-specific lineages and that sequences isolated from the same patient had shorter genetic distance than sequences isolated from different patients (Figure 6.6A). Interestingly, patient 5 showed up to 2-fold larger internal genetic distance than the other patients, which is due to a sub-clustering of P5-B1.D9-A12 and P5-B3.F3-B9.

Several identical HVR1 sequences were detected in the same liver, with the highest number found for P2-B1 (60%). In general, patients infected with genotype 3a, showed significantly higher HVR1 variation, with multiple unique sequences, than did patients infected with genotype 1a (Figure 6.6B).

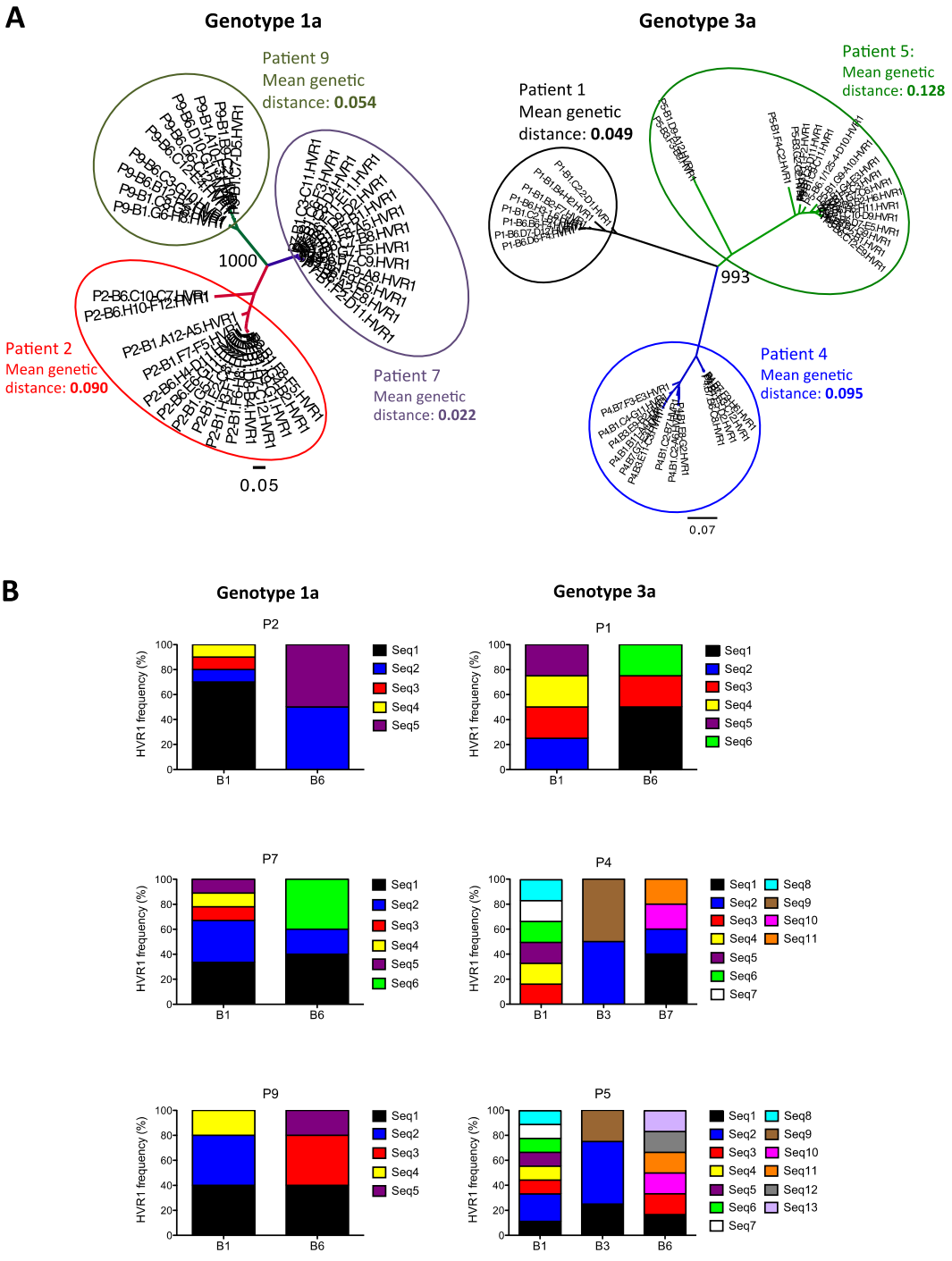


Figure 6.6 HVR1 sequence analysis.

A. The 81 nucleotide long HVR1 was isolated from the patient SGAs and used in Nearest Neighbour sequence analyses. The phylogenies are here shown as un-rooted Neighbour Joining trees for each of the genotypes (1000 bootstraps). The mean genetic distances can be seen for each patient lineage and a scale for the distances can be seen underneath the trees. The bootstrap value is indicated for the

central node. **B.** The HVR1 sequences (Seq) isolated from the same patient were aligned and the number of unique HVR1 noted. The fraction of SGAs with identical HVR1 sequences are presented as a percentage of the total number of sequences isolated from the particular site in each patient.

Neighbour joining sequencing analysis was also performed for the HCV genotype 3a specific HVRs (HVR495 and HVR575) (Humphreys et al., 2009). As for HVR1, the HVR495 and HVR575 were highly comparable within the same liver, indicated by low mean genetic distances (Figure 6.7A).

The frequency of identical HVR495 and HVR575 sequences within one liver was considerable higher than for the HVR1 sequences and in some instances the isolated sequences were all identical (Figure 6.7B). These observations led to the decision not to perform further phylogenetic analysis on the SGA-derived HVR1, HVR495 and HVR575.

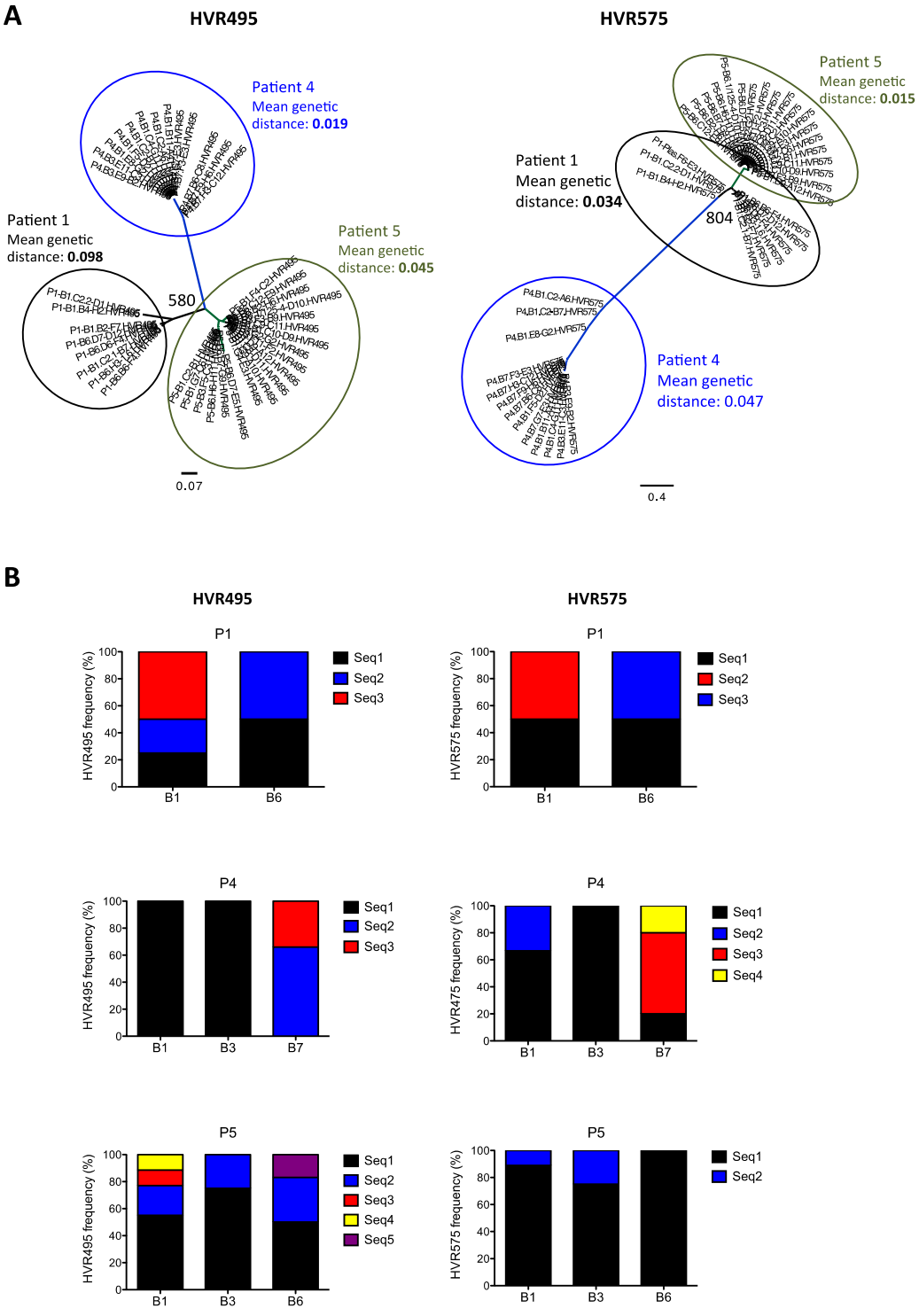


Figure 6.7 Genotype 3a HVR495 and HVR575 sequence analysis.

A. Nearest Neighbour analyses were applied to the SGA-derived HVR495 and HVR575 sequences from the genotype 3a patients. The phylogenies are here shown as un-rooted Neighbour Joining trees generated with 1000 bootstraps. The patient-specific HVRs are indicated in colour; P1: black, P4: blue and P5 green. The

boostrap-value is indicated for the central node. **B.** The intra-hepatic HVR495 and HVR575 sequences were aligned and the number of unique sequences found for each patient. The fraction of identical HVR495 and HVR575 sequences are indicated as a percentage of the total number of sequences from the particular site.

As mentioned above, all 6 patients showed low genetic variability in the detected HVR sequences. To date, there is no evidence that polymorphisms in HVR1 directly predict the number of variations outside this region and it was decided to analyse the genetic variability of the E1, E2 structural genes for genotype 1a infected patients and the E1, E2 together with p7 for genotype 3a. Initial alignment analysis revealed that the patient-specific SGAs had a mean of 113 ±39.5 polymorphic sites for genotype 1a sequences (1.7kb) and 120 ±67.6 sites for genotype 3a (in 1.8kb). Interestingly, the number of changes per nucleotide in each viral gene varied between patients. Sequences isolated from patients 2, 4 and 5 contained a higher number of differences in E2, whereas patients 7 and 9 contain more in E1. Patient 1 showed the highest number of changes per nucleotide in p7 (Figure 6.8).

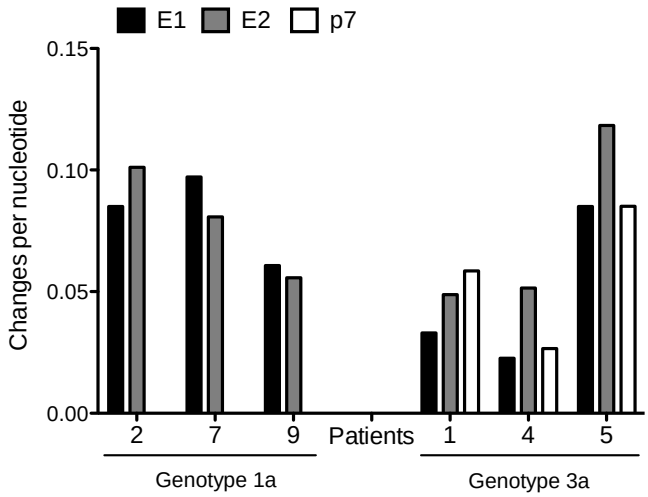


Figure 6.8 Distribution of polymorphisms for SGA-derived hepatic HCV sequences.

The number of changes per nucleotide was quantified for the patient-specific E1, E2 and p7 SGA sequences. The number of changes per nucleotide is normalised to the size of the particular gene (E1 = 576nt, E2 = 1089nt and p7 = 189nt).

Comparing the SGA-derived sequences with the consensus found for the same cDNA template revealed that most of the ambiguities detected in the consensus sequence chromatograms were represented among the SGA-derived clones. However, in some instances the polymorphisms among the consensus sequences were not represented in the clones, indicating that not all HCV variants had been isolated. In addition, several polymorphisms detected among the SGAs, were not represented by the consensus sequences (Appendix 1).

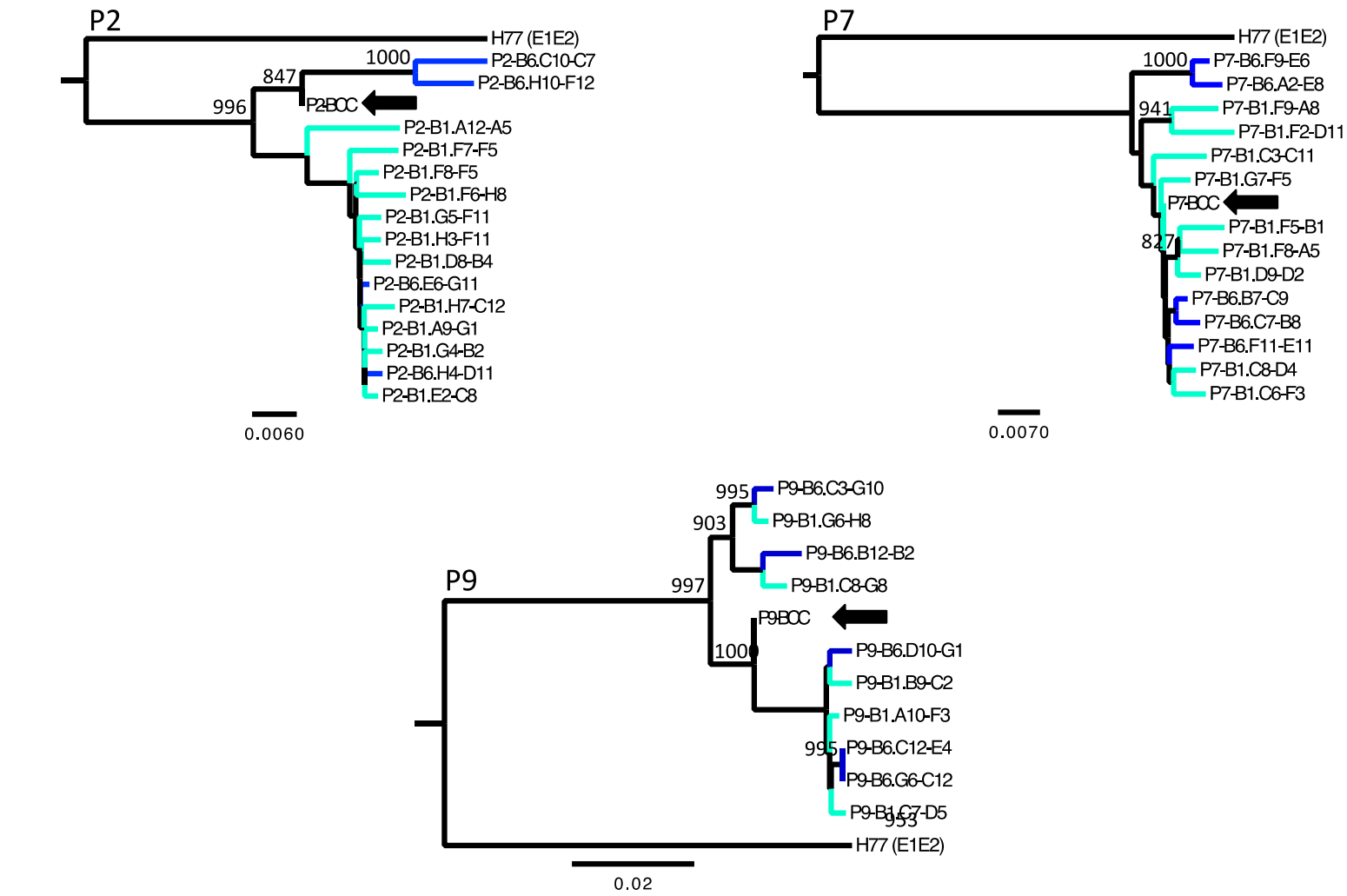
Phylogenetic analysis of the SGA-derived HCV sequences were performed by Nearest Neighbour with a bootstrapping of 1000. The data are presented as rooted Neighbour Joining trees instead of un-rooted, in order to better distinguish between highly similar sequences. The inferred biopsy consensus sequence of the aligned intra-hepatic consensus “bulk” sequences (BCC: Biopsy Consensus Consensus) was included in the analysis as a representative for the bulk sequences. The molecular clone H77 (genotype 1a) and S52 (genotype 3a) was included as reference sequences (Figure 6.9A + B). The phylogenetic tree places the BCC sequence in the middle of the trees with approximately equal genetic distance to the patient-specific SGAs. The analysis further revealed that patient 5, P5-B1-C2.B1 and P5-B6.B7-G9 and patient 9 P9-B6.C12-E4 and P9-B6.G6-C11 were identical.

Although the remaining 116 HCV SGAs differed from each other, they did not form site-specific lineages. For Patient 1, 2, 4 and 9 clusters were formed of

two to four SGA sequences originating from the same hepatic site, however in most cases these clusters had shorter branch length to SGAs from a different hepatic site, than to sequences derived from the same biopsy (Figure 6.9A +B). Evaluating the mean genetic distance between sequences identified at difference hepatic compartments revealed that these were either smaller or equal to the mean genetic distances between SGAs originating from the same site (Appendix 2). HCV sequences from patient 7 showed more defined formation of site related lineages, demonstrated by the formation of four distinct clusters of sequences isolated from biopsy 1 and 6. However, the positioning of the clusters revealed interruption of site-specific lineages (Figure 6.9A). The observation is supported by the equal genetic distance between HCV sequences detected within biopsy 1 and 6 and between the two sites (Appendix 2). In comparison, the 20 SGAs generated from patient 5 did not form any clusters. Instead the site-specific SGAs were scattered across the phylogenetic tree (Figure 6.9B) and the mean genetic distance were larger between sequences originating from the same biopsy than from different ones (Appendix 2).

Additional phylogenetic analysis of the hepatic HCV sequences was performed with the tree-based Hasegawa-Kishino-Yano (HKY) substitution model. The generated Maximum Likelihood phylogenetic tree was highly comparable with the Neighbour Joining model and it was therefore decided not to use the more complex HKY substitution model in further analyses (data not shown).

A


P2
H77 (E1E2)
1000
P2-B6.C10-C7
P2-B6.H10-F12
847
P2BCC
996
P2-B1.A12-A5
P2-B1.F7-F5
P2-B1.F8-F5
P2-B1.F6-H8
P2-B1.G5-F11
P2-B1.H3-F11
P2-B1.D8-B4
P2-B6.E6-G11
P2-B1.H7-C12
P2-B1.A9-G1
P2-B1.G4-B2
P2-B6.H4-D11
P2-B1.E2-C8
0.0060
P7
H77 (E1E2)
1000
P7-B6.F9-E6
P7-B6.A2-E8
941
P7-B1.F9-A8
P7-B1.F2-D11
P7-B1.C3-C11
P7-B1.G7-F5
P7BCC
827
P7-B1.F5-B1
P7-B1.F8-A5
P7-B1.D9-D2
P7-B6.B7-C9
P7-B6.C7-B8
P7-B6.F11-E11
P7-B1.C8-D4
P7-B1.C6-F3
0.0070
P9
995
P9-B6.C3-G10
903
P9-B1.G6-H8
P9-B6.B12-B2
997
P9-B1.C8-G8
P9BCC
1000
P9-B6.D10-G1
P9-B1.B9-C2
P9-B1.A10-F3
995
P9-B6.C12-E4
P9-B6.G6-C12
P9-B1.C7-D5
953
H77 (E1E2)
0.02

B

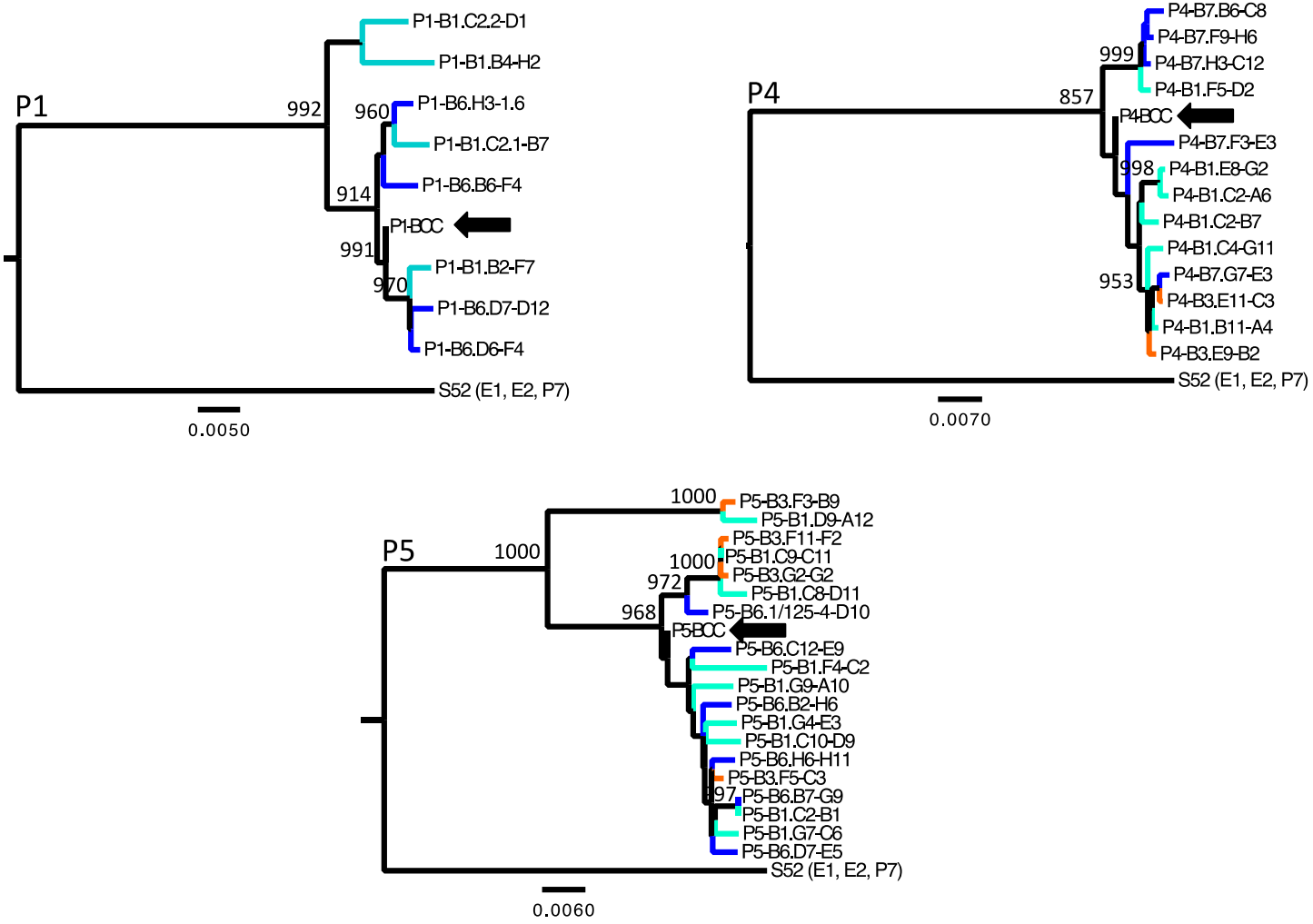

P1
992
960
914
991
970
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.H3-1.6
P1-B1.C2.1-B7
P1-B6.B6-F4
P1-BCC
P1-B1.B2-F7
P1-B6.D7-D12
P1-B6.D6-F4
S52 (E1, E2, P7)
0.0050
P4
857
999
998
953
P4-B7.B6-C8
P4-B7.F9-H6
P4-B7.H3-C12
P4-B1.F5-D2
P4BCC
P4-B7.F3-E3
P4-B1.E8-G2
P4-B1.C2-A6
P4-B1.C2-B7
P4-B1.C4-G11
P4-B7.G7-E3
P4-B3.E11-C3
P4-B1.B11-A4
P4-B3.E9-B2
S52 (E1, E2, P7)
0.0070
P5
1000
1000
1000
972
968
97
P5-B3.F3-B9
P5-B1.D9-A12
P5-B3.F11-F2
P5-B1.C9-C11
P5-B3.G2-G2
P5-B1.C8-D11
P5-B6.1/125-4-D10
P5BCC
P5-B6.C12-E9
P5-B1.F4-C2
P5-B1.G9-A10
P5-B6.B2-H6
P5-B1.G4-E3
P5-B1.C10-D9
P5-B6.H6-H11
P5-B3.F5-C3
P5-B6.B7-G9
P5-B1.C2-B1
P5-B1.G7-C6
P5-B6.D7-E5
S52 (E1, E2, P7)
0.0060

Figure 6.9 Phylogenetic analysis of hepatic HCV nucleotide sequences.

Nearest Neighbour analyses were applied to the SGA sequences isolated from each patient and represented as rooted Neighbour Joining trees with 1000 bootstraps. Branches supported with more than 750 bootstraps are indicated in the trees. The branch lengths are a direct measure of the genetic distances between the SGAs and a scale for the length can be seen underneath each tree. The branches from biopsy 1 (B1) are turquoise, biopsy 3 (B3) SGAs are orange and biopsy 6 (B6) and 7 (B7) blue. The branch for the consensus sequence (BCC: biopsy consensus consensus) generated from the 8 hepatic consensus sequences are coloured black and are indicated with an arrow. **A.** Genotype 1a patients 2, 7 an 9. The E1E2 nucleotide sequence for the genotype 1 H77 variant is used as a reference. **B.** Genotype 3a patients 1, 4 and 5. The E1E2, P7 nucleotide sequence for the genotype 3a S52 variant is used as a reference.

6.4 Phylogenetic analysis of SGA-derived HCV sequences isolated from plasma

In addition to the intra-hepatic SGA-derived HCV sequences, phylogenetic analyses were carried out on HCV SGAs detected in the plasma from patients 1, 2, 4, 7 and 9 (list of SGA clones can be seen in Table 3.5). For unknown reasons HCV sequences could not be amplified from the plasma of patient 5.

Alignments of the plasma-derived HCV sequences revealed that each patient had variables level of nucleotide changes for the investigated genes. Sequences from patient 2, 9 and 4 showed higher level of changes per nucleotide for the viral encoded E2, whereas patient 7 had higher levels in E1 and patient 1 in p7 (Figure 6.10). Notably, all patients except for patient 1 showed a lower number of polymorphic sites between the plasma-derived HCV sequences compared to the liver-derived. However the difference is believed to be the result of more liver- than plasma-derived sequences being included in the analyses. The analysis of patient 1 included more sequences from the plasma than the liver, which explains why this patient show more polymorphic sites in the HCV sequences isolated from the plasma.

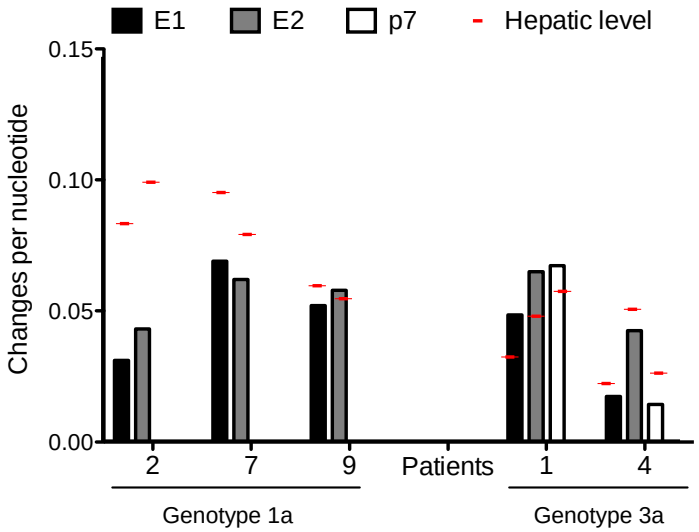


Figure 6.10 Distribution of polymorphisms for SGA-derived plasma HCV sequences.

Number of gene-specific differences between SGA sequences isolated from the same patient. The value is expressed as changes per nucleotide for the genes and are normalised to the size of the individual genes (E1 = 576nt, E2 = 1089nt, p7 = 189nt). The number of differences in the patient's hepatic SGA sequences is indicated with a red line.

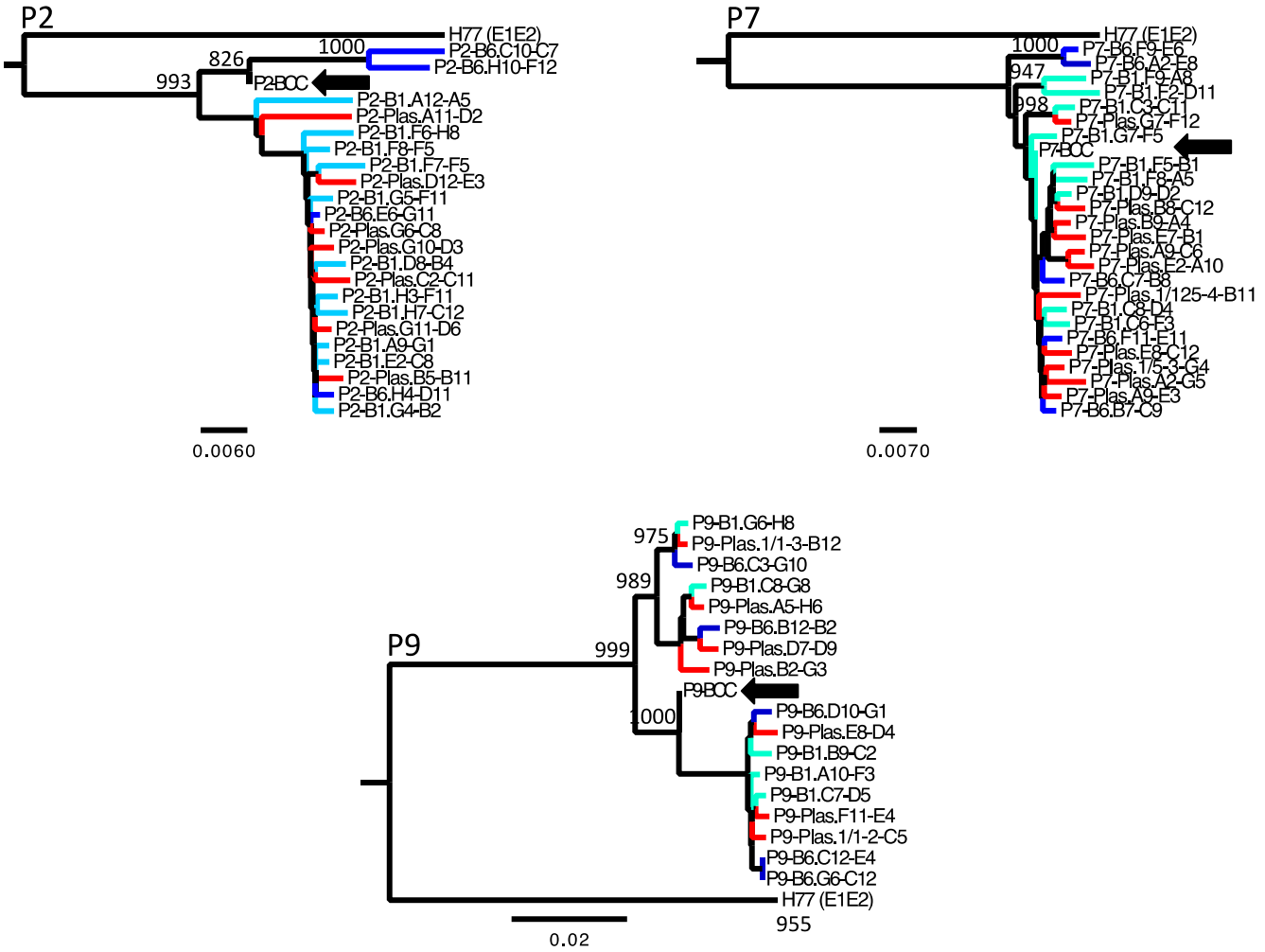
Neighbour Joining phylogenetic trees were generated for the plasma- and liver-derived sequences isolated from patient 1, 2, 4, 7 and 9 (Figure 6.11A+B). The analysis revealed that none of the plasma-derived SGA clones were identical within one patient. Furthermore none of them were identical to the liver-derived sequences.

Although the plasma- and the liver-derived HCV sequences represent different anatomical compartment, we did not observe distinct site-specific lineage. Instead most plasma-derived sequences revealed closer genetic relation with sequences from hepatic sites (Figure 6.11A+B). This pattern was especially seen for the sequences isolated from patients 2, 4 and 9, who also showed lower mean genetic distances between plasma- and liver specific sequences than between the individual plasma sequences (Appendix 2). Small clusters of two - four plasma sequences were seen for patient 1 and 7, however the clusters was scattered in between hepatic sequences and the mean genetic distance between the site-specific sequences and within the plasma-derived sequence alignments were highly comparable (Appendix 2).

Since the hepatocytes are the main site for HCV replication we decided to investigate any potential links between plasma- and site-specific liver-derived HCV sequences. Interestingly, plasma sequences isolated from patient 2 showed significantly closer relation to HCV sequences from biopsy 1 than 6 and for patient 9 the plasma-derived HCV sequences demonstrated shorter

genetic distance to sequences isolated from biopsy 6 than for biopsy 1 (Appendix 1). The remaining patients had comparable distances between the plasma-derived and the hepatic HCV sequences.

A


P2
H77 (E1E2)
1000
P2-B6.C10-C7
826
P2-B6.H10-F12
993
P2BCC
P2-B1.A12-A5
P2-Plas.A11-D2
P2-B1.F6-H8
P2-B1.F8-F5
P2-B1.F7-F5
P2-Plas.D12-E3
P2-B1.G5-F11
P2-B6.E6-G11
P2-Plas.G6-C8
P2-Plas.G10-D3
P2-B1.D8-B4
P2-Plas.C2-C11
P2-B1.H3-F11
P2-B1.H7-C12
P2-Plas.G11-D6
P2-B1.A9-G1
P2-B1.E2-C8
P2-Plas.B5-B11
P2-B6.H4-D11
P2-B1.G4-B2
0.0060
P7
H77 (E1E2)
1000
P7-B6.F9-E6
P7-B6.A2-F8
947
P7-B1.F9-A8
P7-B1.F2-D11
998
P7-B1.C3-C11
P7-Plas.G7-F12
P7-B1.G7-F5
P7BCC
P7-B1.E5-B1
P7-B1.F8-A5
P7-B1.D9-D2
P7-Plas.B8-C12
P7-Plas.B9-A4
P7-Plas.E7-B1
P7-Plas.A9-C6
P7-Plas.E2-A10
P7-B6.C7-B8
P7-Plas.1/125-4-B11
P7-B1.C8-D4
P7-B1.C6-E3
P7-B6.F11-E11
P7-Plas.E8-C12
P7-Plas.1/5-3-G4
P7-Plas.A2-G5
P7-Plas.A9-E3
P7-B6.B7-C9
0.0070
P9
975
P9-B1.G6-H8
P9-Plas.1/1-3-B12
P9-B6.C3-G10
989
P9-B1.C8-G8
P9-Plas.A5-H6
P9-B6.B12-B2
999
P9-Plas.D7-D9
P9-Plas.B2-G3
P9BCC
1000
P9-B6.D10-G1
P9-Plas.E8-D4
P9-B1.B9-C2
P9-B1.A10-F3
P9-B1.C7-D5
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
P9-B6.C12-E4
P9-B6.G6-C12
H77 (E1E2)
955
0.02

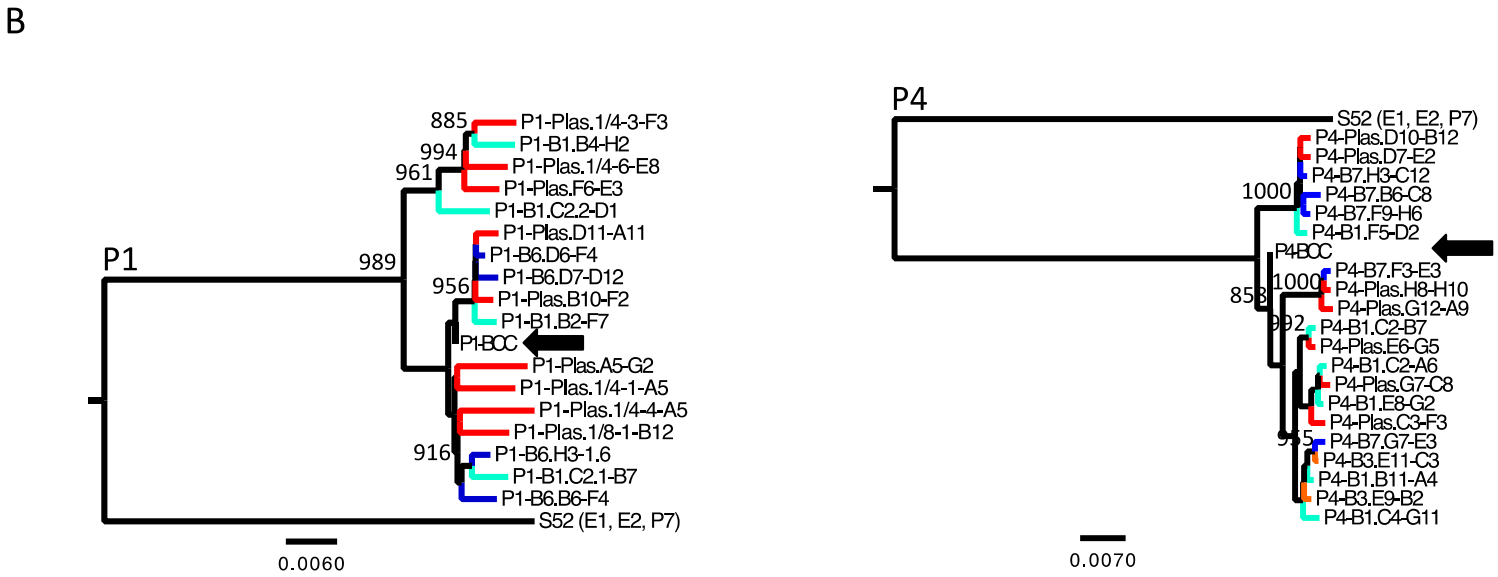


Figure 6.11 Phylogenetic analysis of hepatic- and plasma-derived HCV nucleotide sequences.

Nearest Neighbour analyses were applied to the SGA sequences isolated from each patient and the data is here presented as rooted Neighbour Joining trees with 1000 bootstraps. Branches supported with more than 750 bootstraps are indicated at the branchings. The branch lengths are a direct measure of the genetic distances between the SGAs and a scale for the length can be seen underneath each tree. The branches representing plasma (Plas) are red, biopsy 1 (B1) turquoise, biopsy 3 (B3) SGAs are orange and biopsy 6 (B6) and 7 (B7) blue. The branch for the consensus sequence (BCC: biopsy consensus consensus) generated from the 8 hepatic consensus sequences are black and are indicated with an arrow. **A.** Genotype 1a patients 2, 7 and 9. The E1E2 nucleotide sequence for the genotype 1a H77 variant is used as a reference. **B.** Genotype 3a patients 1 and 4 The E1E2, P7 nucleotide sequence for the genotype 3a S52 variant is used as a reference.

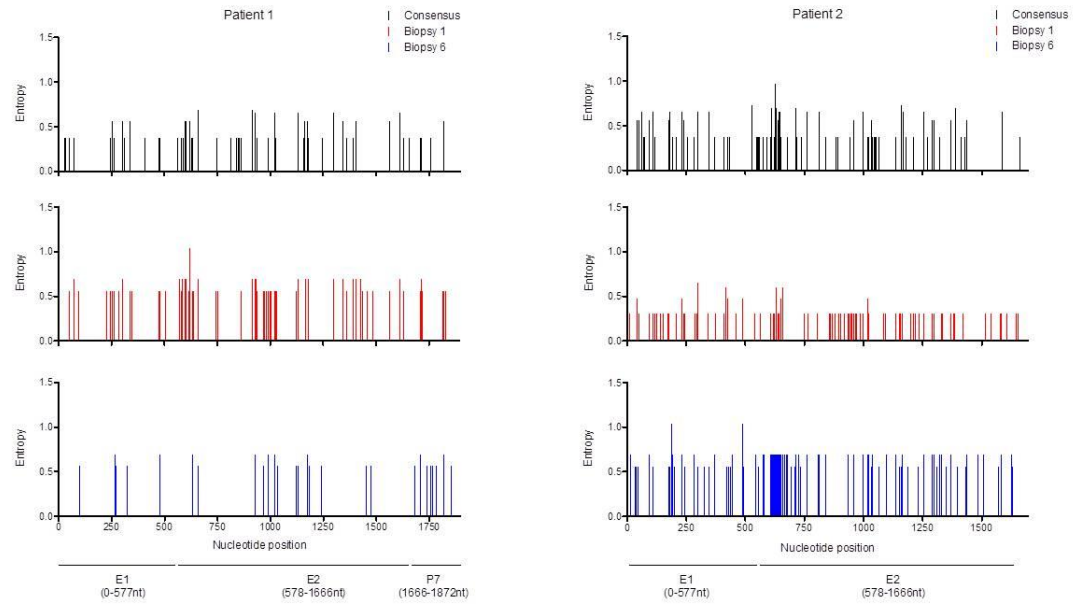
6.5 Analysis of HCV entropy and migration events

Shannon entropy analysis was performed for the SGA-derived sequences originating from the same hepatic site or from the plasma (One-way Shannon entropy). The analysis revealed that the HCV sequences shared a high number of identical nucleotides (entropy = 0) and that the positive entropies were comparable between site-specific sequence alignments. The highest density of entropies was found in the 5'-end of E2, consistent with the location of HVR1 (here position 573-654nt) (Figure 6.12A).

The hepatic sequences generated from patient 2, 5, 7 and 9 had overall more sites with polymorphisms (more number of sites with entropies above 0) and higher complexity at the polymorphic sites (high level of entropy at variable sites) (Figure 6.12A right, represented by patient 2) than seen for patient 1 and 4 (Figure 6.12A left, represented by patient 1), leading to higher mean entropies for the patient 2, 5, 7 and 9 (Appendix 3). These results are supported by the observation that hepatic SGA alignments for patient 2, 5, 7 and 9 show higher levels of changes per nucleotide than patient 1 and 4 (Figure 6.8). Plasma sequences isolated from patient 1, 2, , 7 and 9 had higher numbers of polymorphic sites than the patient 4 plasma sequences. However patient 4 had a higher level of complexity, with entropies up to 0.995, than seen for any of the other patients (data not shown). The mean entropy for the patient 4 plasma SGA alignment is therefore significantly higher than for the other patient-derived plasma sequences (Appendix 3).

Performing Shannon entropy between the biopsy- and plasma-derived HCV sequence alignments (two-way Shannon entropy) demonstrated significantly more variable sites than for the one-way Shannon entropy analyses (Figure 6.12B). Furthermore, patient 2 and 7 showed higher levels of complexity at the variable position between site-specific sequence alignments, indicating that the site-specific sequences do not share the same sequence diversity. In contrast, patient 9 showed decreased levels of two-way entropies compared to the one-way, suggesting that the HCV sequences isolated from the different compartments shares the same polymorphisms (data not shown). Patient 1, 4 and 5 showed similar complexity between the one- and two-way Shannon entropies. The mean two-way Shannon entropies can be found in Appendix 3.

A


Patient 1
Consensus
Biopsy 1
Biopsy 6
Entropy
Nucleotide position
E1
(0-577nt)
E2
(578-1666nt)
P7
(1666-1872nt)
Patient 2
Consensus
Biopsy 1
Biopsy 6
Entropy
Nucleotide position
E1
(0-577nt)
E2
(578-1666nt)

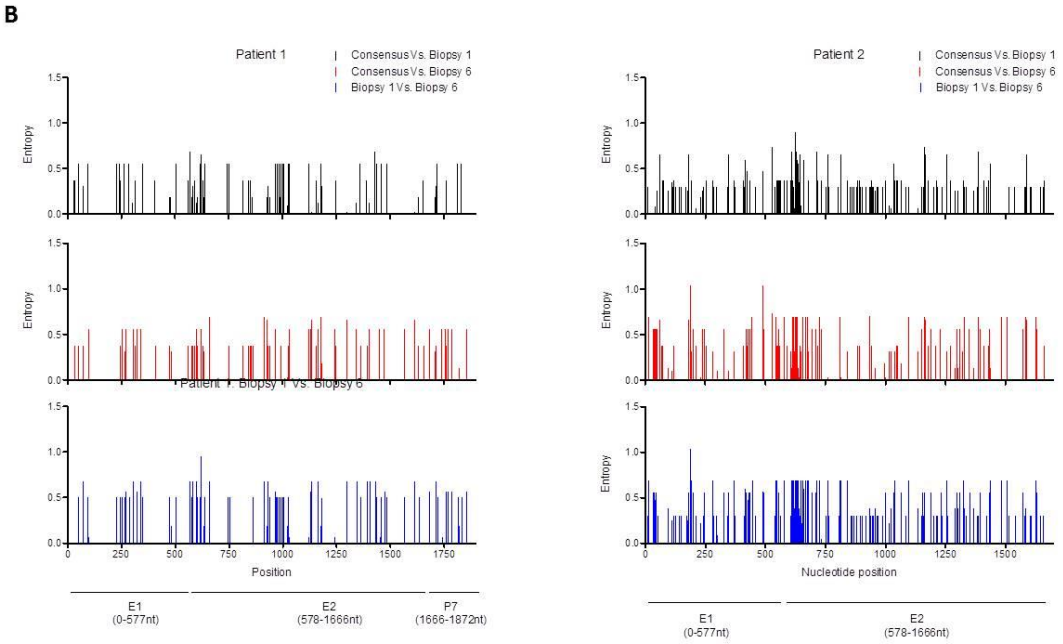


Figure 6.12 Shannon entropies.

Shannon entropies were calculated for each nucleotide position in the SGA sequence alignments. Entropies above zero indicate polymorphisms at the particular site in the alignment. Here patient 1 and 2 are shown as representatives for the two different entropy patterns seen for all patients. Patients 5, 7, 9 showed similar patterns as patient 2, whereas patient 4 is comparable with patient 1. **A.** One-way Shannon entropies for SGA sequences isolated from the same anatomical site **B.** Two-way Shannon entropies between sequences alignments generated from different sites.

Estimation of the minimum number of HCV migration required for the observed spatial distribution of sequences between sites was done by Slatkin-Maddison modelling. In this model a probability (*p*) below 0.05 is considered significant evidence for genetic compartmentalisation (Rodrigo et al., 2001). Performing Slatkin-Maddison modelling for the SGA-derived HCV sequences revealed that migration may have occurred between the hepatic sites within one liver. However for patients 2, 4, 5, 7 and 9, the migration from biopsy 1 only occurred in one direction (Table 6.1). Statistical significance for the number of migration events was only seen for the HCV sequences isolated from patient 7 ($p = 0.02$). Fourteen hepatic HCV sequences were analysis for patient 7 and of these only two were required to predict the distribution of sequences between the hepatic compartments. This observation confirms the formation of four site-specific clusters in the Nearest Neighbour phylogenetic tree (Figure 6.9) and it is concluded that the liver collected from patient 7 may demonstrate genetic compartmentalisation of the viral-encoded E1 and E2 genes.

Slatkin-Maddison modelling for the plasma-derived HCV sequences indicated that HCV migration events occurred both from plasma to the individual biopsies and in reverse to the plasma, but in none of the case was the number of migration events significant ($p > 0.05$). Comparing the plasma-derived sequences with the combined intra-hepatic sequences showed that migration is more likely to occur from the liver to the plasma (patient 2, 7, 9 and 4). However the data was not statistical significant (Table 6.1).

Table 6.1: Slatkin-Maddison modelling of HCV sequence migration between the liver and plasma

Genotype	Patient	Direction of Migration:	Number of migration events:	P-value
1a	2	Biopsy 1 → Biopsy 6	3	0.25
		Biopsy 1 → Plasma	7	1
		Plasma → Biopsy 6	3	0.45
		Liver → Plasma	7	1
	7	Biopsy 1 → Biopsy 6	2	0.02
		Plasma → Biopsy 1	1	0.11
		Biopsy 1 → Plasma	4	0.11
		Plasma → Biopsy 6	4	0.43
		Liver → Plasma	7	0.22
	9	Biopsy 1 → Biopsy 6	4	0.88
		Plasma → Biopsy 1	5	1
		Plasma → Biopsy 6	4	0.71
		Liver → Plasma	7	1
3a	1	Biopsy 1 → Biopsy 6	2	0.80
		Biopsy 6 → Biopsy 1	1	0.08
		Plasma → Biopsy 1	3	1
		Biopsy 1 → Plasma	1	1
		Plasma → Biopsy 6	3	1
		Biopsy 6 → Plasma	1	1
		Plasma → Liver	2	0.33
		Liver → Plasma	3	0.33
	4	Biopsy 1 → Biopsy 3	2	1
		Biopsy 1 → Biopsy 7	3	0.34
		Biopsy 3 → Biopsy 7	2	1
		Biopsy 1 → Plasma	5	0.86
		Plasma → Biopsy 3	1	0.11
		Plasma → Biopsy 7	4	0.68
		Liver → Plasma	7	1
	5	Biopsy 1 → Biopsy 3	4	1
		Biopsy 1 → Biopsy 6	6	1
		Biopsy 3 → Biopsy 6	1	0.51
		Biopsy 6 → Biopsy 3	2	0.51

Slatkin-Maddison was calculated by the HyPhy2.220130723 software.

6.6 Analysis of selective pressure on SGA-derived HCV sequences

The identification of polymorphisms among the HCV SGA nucleotide sequences suggests potential differences between the inferred SGA-derived HCV amino acid sequences. The presence of amino acid polymorphism depends on whether the detected nucleotide changes are synonymous (*ds*) or non-synonymous (*dn*). Prior to measuring *ds* and *dn* changes, clones with deletions (P2-B1.A9-G1*, P2-B1.F6-H8* and P7-B6.C3-C11*) were edited by inserting nucleotides identical to the consensus of the other SGAs in the alignment. At the same time the guanine insertion in P7-Plas.E2-A10* was deleted to avoid out of frame codons.

Evaluation of the mean *ds* and *dn* changes per codon for SGA-derived HCV sequences isolated from different compartments revealed low levels of changes (*ds*: 0.007 – 0.084 changes per nucleotide and *dn:* 0.003 - 0.013 changes per nucleotide) with most of the codons having no changes. Those codons that did differ between sequences were more likely to contain *ds* than *dn* changes (two-tailed paired t-test: $p < 0.0001$), suggesting higher levels of random mutation than non-random at the investigating sites (Table 6.2 and 6.3). Similar observations were made when comparing SGA-derived sequences originating from the same compartment (Figure 6.13A), leading to a mean *ds*/*dn*-ratio per codon of above 1 for the site-specific sequence alignments (Table 6.2 and 6.3).

The position of the *ds* and *dn* changes were further evaluated for the SGA-derived HCV sequences. (Figure 6.13A and B). As mentioned above, SGA-derived sequences from the same and from different anatomical sites show higher numbers of *ds* than *dn* changes across the investigated sequences. Interestingly, high numbers of *dn* changes were mainly observed between positions 155 and 210 in the 5'UTR-end of E2, consistent with the HVR1 (Figure 16.13A + B, here represented by patient 1. Sequences isolated from the remaining patients showed similar distribution of *ds* and *dn* changes). Furthermore, patient 1, 4 and 5 showed ratios below 1 at positions 390 to 435, which potentially coincide with the HVR575, which is located between codon positions 384 and 392 (Figure 16.13A + B).

Table 6.2: ds/dn ratio for HCV genotype 1a sequences

Patient 2 (±STDV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	0.000	0.028 (± 0.018)	0.023 (± 0.013)	0.029 (±0.018)
Biopsy 1	0.005 (± 0.003)	0.06 (± 0.05)	0.084 (± 0.056)	0.037 (±0.018)
Biopsy 6	0.004 (± 0.003)	0.012 (± 0.011)	0.13 (± 0.05)	0.084 (±0.060)
Plasma	0.004 (±0.003)	0.003 (±0.003)	0.012 (±0.060)	0.52 (±0.30)

Patient 7 (±STDV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	0.07 (± 0.12)	0.035 (± 0.019)	0.036 (± 0.026)	0.027 (±0.009)
Biopsy 1	0.001 (± 0.001)	0.06 (± 0.03)	0.083 (± 0.037)	0.064 (±0.022)
Biopsy 6	0.001 (± 0.001)	0.005 (± 0.004)	0.04 (± 0.02)	0.069 (±0.034)
Plasma	0.001 (±0.0009)	0.003 (±0.001)	0.003 (±0.001)	0.05 (± 0.03)

Patient 9 (±STDV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	0.000	0.014 (± 0.007)	0.013 (± 0.009)	0.018 (±0.009)
Biopsy 1	0.001 (± 0.000)	0.08 (± 0.02)	0.075 (± 0.051)	0.078 (±0.051)
Biopsy 6	0.001 (± 0.001)	0.006 (± 0.004)	0.06 (±0.03)	0.080 (± 0.052)
Plasma	0.001 (±0.001)	0.006 (±0.003)	0.006 (±0.004)	0.09 (±0.04)

Green boxes: mean number of synonymous (*ds*) changes; Blue boxes: mean number of non-synonymous (*dn*) changes per nucleotide; Diagonal white boxes: *dn/ds*-ratio.
STDV: Standard deviation

Table 6.3: ds/dn ratio for HCV genotype 3a sequences

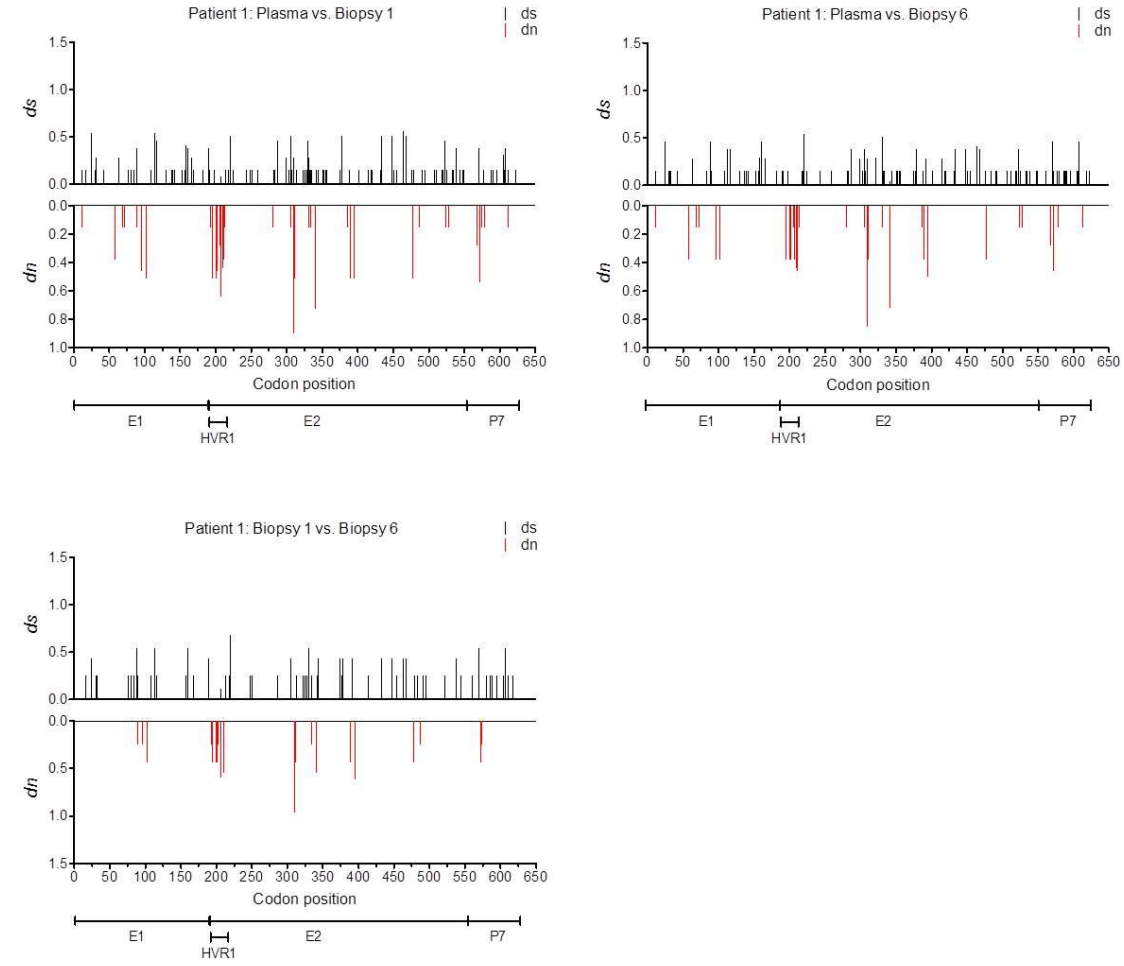
Patient 1 (±STDV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	0.000	0.023 (± 0.012)	0.012 (± 0.004)	0.024 (±0.010)
Biopsy 1	0.004 (± 0.003)	0.16 (± 0.06)	0.047 (± 0.019)	0.053 (±0.016)
Biopsy 6	0.001 (± 0.001)	0.007 (± 0.005)	0.05 (± 0.04)	0.045 (±0.015)
Plasma	0.003 (±0.003)	0.008 (±0.005)	0.006 (±0.005)	0.19 (± 0.07)

Patient 4 (±STDV)	Consensus	Biopsy 1	Biopsy 3	Biopsy 7	Plasma
Consensus	0.000	0.003 (± 0.002)	0.001 (± 0.001)	0.008 (± 0.004)	0.008 (± 0.005)
Biopsy 1	0.002 (± 0.001)	0.91 (± 0.47)	0.007 (± 0.006)	0.020 (± 0.008)	0.017 (± 0.08)
Biopsy 3	0.001 (± 0.001)	0.008 (± 0.005)	NP	0.018 (± 0.008)	0.014 (± 0.007)
Biopsy 7	0.001 (± 0.001)	0.012 (± 0.005)	0.013 (± 0.006)	0.42 (± 0.28)	0.023 (± 0.010)
Plasma	0.001 (± 0..001)	0.010 (± 0.005)	0.011 (± 0.004)	0.011 (± 0.006)	0.50 (± 0.15)

Patient 5 (±STDV)	Consensus	Biopsy 1	Biopsy 3	Biopsy 6
Consensus	0.001 (± 0.001)	0.025 (± 0.027)	0.025 (± 0.035)	0.014 (± 0.002)
Biopsy 1	0.003 (± 0.005)	0.14 (± 0.05)	0.069 (± 0.048)	0.052 (± 0.038)
Biopsy 3	0.005 (± 0.008)	0.013 (± 0.011)	0.34 (± 3.0)	0.066 (± 0.045)
Biopsy 6	0.001 (± 0.001)	0.007 (± 0.008)	0.013 (± 0.010)	0.09 (± 0.03)

Green boxes: mean number of synonymous (*ds*) changes; Blue boxes: mean number of non-synonymous (*dn*) changes per nucleotide; Diagonal white boxes: *dn/ds*-ratio.
STDV: Standard deviation

A


Patient 1: Plasma vs. Biopsy 1
ds
dn
ds
dn
Codon position
E1
HVR1
E2
P7
Patient 1: Plasma vs. Biopsy 6
ds
dn
ds
dn
Codon position
E1
HVR1
E2
P7
Patient 1: Biopsy 1 vs. Biopsy 6
ds
dn
ds
dn
Codon position
E1
HVR1
E2
P7

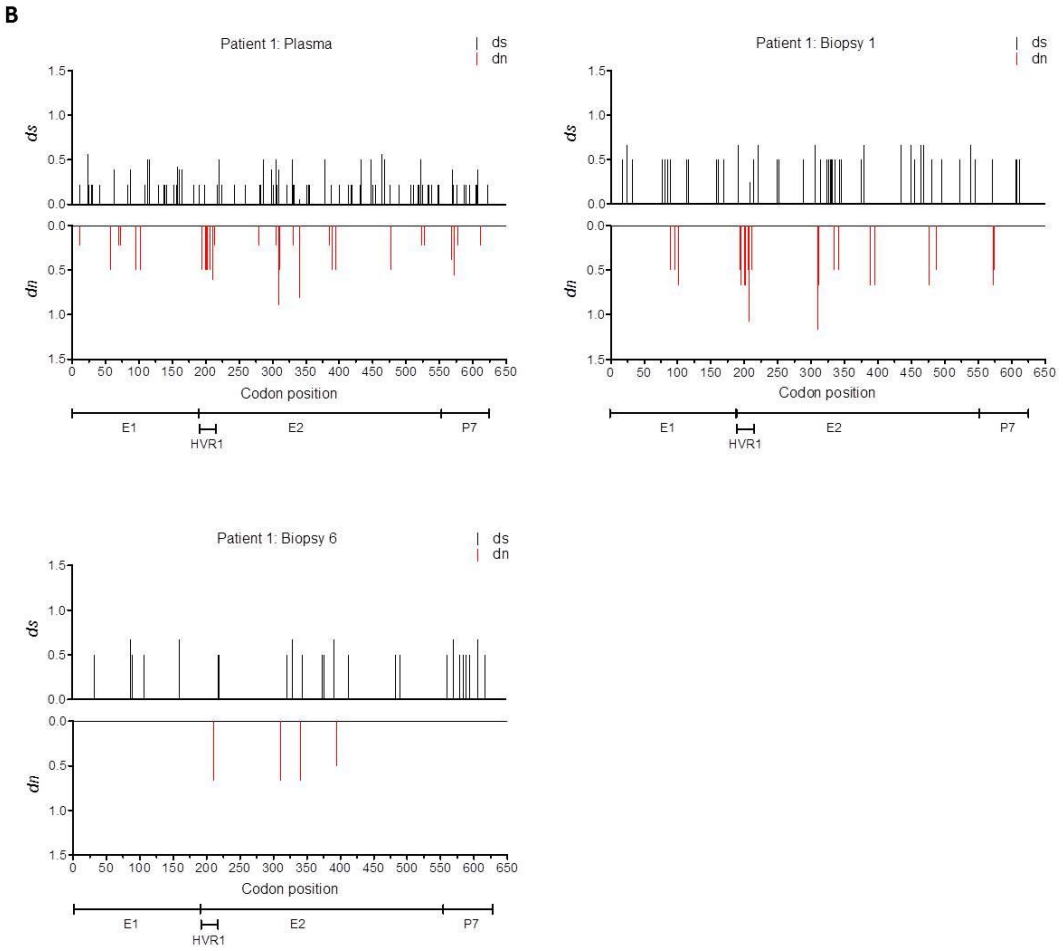


Figure 6.13 *ds* and *dn* distribution across E1 E2 and p7.

The *ds* and *dn* values were calculated for each codon by SNAP v1.1.1. The position of the codons can be seen underneath the graph. Patient 1 is here shown as representative for the typical distribution of *ds* and *dn* changes observed for all patients. **A**. *ds* and *dn* changes found for the site-specific SGA sequences. **B**. *ds* and *dn* changes found between SGA site-specific alignments.

6.7 Analysis of inferred HCV amino acid sequences isolated from plasma and liver

To investigate potential compartmentalisation of the inferred HCV amino acid and their theoretic phenotype the SGA-derived HCV sequences were translated to amino acid using the CLC main workbench. Initial analysis of the amino acid sequences revealed that P9-Plas.E8-D4, P4-B7.G7-F3, P4-B7.B6-C8 and P5-B3.F3-B9 contained *in-frame* stop codons in the E1 E2 region. Due to genotype variability in the length of the E1 and E2 sequences the stop codon in P9-Plas.A5-G2 (W425*) is the same as seen in P4-B7.B6-C8 (W433*). Same stop codons were also detected in P5.B3-F3-B9 (Y603*) and in P4-B7.G7-F3 (Y605*) (Table 6.4). Whether these terminations reflect *in vivo* defective particles or artefacts in cDNA synthesis cannot be established using current methods.

During post-translation modifications oligosaccharyltransferase may add oligosaccharides to the asparagine in the tri-peptide motif, Asn – x – Ser/Thr (NxS/T). This glycosylation process does not occur at all NxS/T motifs, but is essential for specific sites in the E1 and E2 proteins in order to maintaining stability and for correct folding (Goffard et al., 2005; Helle et al., 2010; Meunier et al., 1999). The SGA-derived amino acid sequences were manually screened for polymorphisms in the potential N-linked glycosylation (PNG) sites. None of the HCV sequences from patients 1, 5 and 9 had NxS/T polymorphisms, however polymorphic sites were identified in P2-B6.H10-F12 (N385) and P7-B1.F2.D11 (T387I). Surprisingly, 12 out of 20 SGA-derived

sequences from patient 4 showed polymorphisms in PNG sites independently of the anatomical origin of the sequences. All except for 1 clone had more than one effected PNG and the most frequently altered sites were N287T and N342K (Table 6.4).

Table 6.4: Sequence analysis of HCV amino acid sequences*

Genotype	Patient	Clones	Stop Codons	Potential N-glycosylation site
1a	2	P2-B6.H10-F12	-	N385D
	7	P7-B1.F2-D11	-	T387I
	9	P9-Plas.E8-D4	W425*	-
3a	1	-	-	-
	4	P4-Plas.G7-C3	-	N287T, N342K, T368M
		P4-Plas.C3-F3	-	N287T, N342K, T368M
		P4-Plas.E6-G5	-	N287T, T368M
		P4-B1.C2-A6	-	N287T, T368M
		P4-B1.E8-G2	-	N287T, N342K, T368M
		P4-B1.C4-G11	-	N287T, N342K, T368M
		P4-B1.B11-A4	-	N287T, T368M
		P4-B1.F5-D2	-	N287T, N366D, T368M
		P4-B3.E9-B2	-	N287T, N366D, T368M
		P4-B7.G7-F3	Y605*	N287T, T368M
		P4-B7.F3-E3	-	N287T, N366D, T368M
		P4-B7.B6-C8	W433*	S395L
	5	P5-B3.F3-B9	Y603*	-

*All noted amino acid positions refer to the SGA sequence:
Patient 2: E1: 1-193, E2: 194-555aa.
Patient 7 and 9: E1: 1-192, E2: 193-555aa.
Patient 1 and 5: E1: 1-192, E2: 193-561, p7: 562-624aa.
Patient 4: E1: 1-192, E2: 193-563, p7: 564 - 626aa.

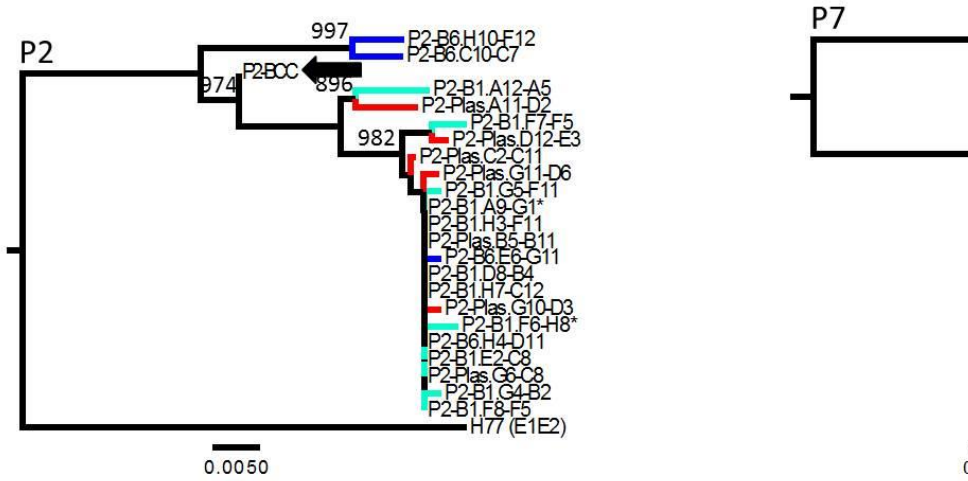
The identification of relatively few polymorphic sites, suggested that the inferred SGA-derived amino acid sequences would be similar. To evaluate this Nearest Neighbour analyses were performed on the translated HCV sequences. The phylogenetic analysis revealed an increase in the number of identical sequences isolated from patient 1, 2, 4, 5 and 7. However, the identical sequences had derived from different anatomical sites (Figure 6.14A+B). As a result the mean distance between amino acid sequences were comparable with the distances between the original nucleotide sequences (Figure 6.15) (unpaired two-tailed t test: $p > 0.85$). In addition, we did not observe any significant differences for the mean distance between amino acid sequences deriving from the same anatomical site and between sites (Figure 6.15). Interestingly, the analysis for patient 9 revealed no increase in the number for identical amino acid sequences compared to the analysis of the nucleotide sequences (data not shown). This observation indicates that the SGA-derived sequences detected in patient 9 mainly contain *ds* changes or that *dn* changes potentially were located at the same positions. As a consequence the distance between the translated SGA sequences remained unchanged from the genetic distance between nucleotide sequences isolated from patient 9.

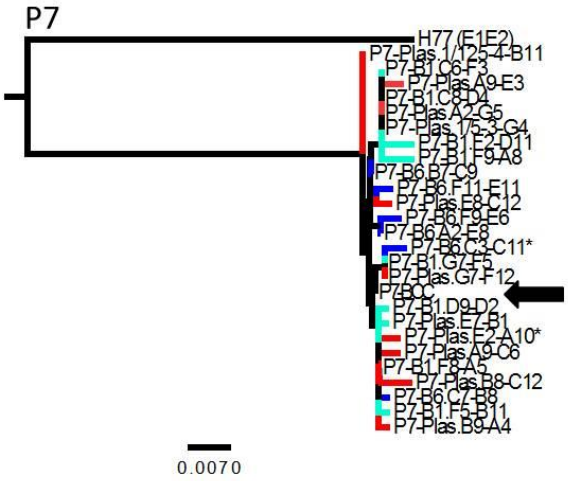
Nearest Neighbour analysis assumes equal frequency of nucleotide and amino acid changes and will weigh all substitutions the same. In nature amino acid substitutions do not occur randomly, some are more likely to happen than others. Identifying rare amino acid substitutions between sequences would

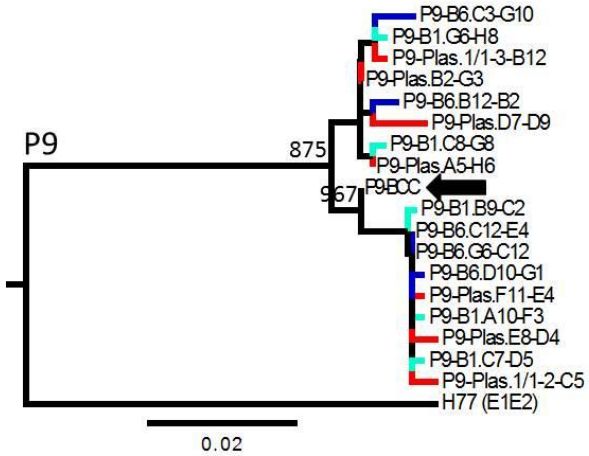
therefore suggest that they are related. We decided to evaluate the similarity between sequences by evaluating the weighted substitutions by BLOcks SUbstitution Matrix (BLOSUM) modelling. Unsurprisingly, the SGA derived sequences scored up to 99.4% similarity between SGA-derived HCV sequences originating from the same site and 99.3% similarity for sequences from different sites (Appendix 4). The minimum level of identity between sites were 97.4%, suggesting that the SGA-derived HCV sequences are highly comparable and that they contain similar levels of weighted substitutions.

Slatkin-Maddison modelling of the patient-derived nucleotide sequences previously showed significant evidence of genetic intra-hepatic compartmentalisation for patient 7. The analysis was repeated to see if the differences were still apparent for the amino acid sequences. This time the Slatkin-Maddison model estimated 4 migration events; two from biopsy 1 to 6 and two in the opposite direction, however the result was no longer significant ($p = 0.34$).

A


P2
997
974
896
982
P2BCC
P2-B6.H10-F12
P2-B6.C10-C7
P2-B1.A12-A5
P2-Plas.A11-D2
P2-B1.F7-F5
P2-Plas.D12-E3
P2-Plas.C2-C11
P2-Plas.G11-D6
P2-B1.G5-F11
P2-B1.A9-G1*
P2-B1.H3-F11
P2-Plas.B5-B11
P2-B6.E6-G11
P2-B1.D8-B4
P2-B1.H7-C12
P2-Plas.G10-D3
P2-B1.F6-H8*
P2-B6.H4-D11
P2-B1.E2-C8
P2-Plas.G6-C8
P2-B1.G4-B2
P2-B1.F8-F5
H77 (E1E2)
0.0050


P7
H77 (E1E2)
P7-Plas.1/125-4-B11
P7-B1.C6-F3
P7-Plas.A9-E3
P7-B1.C8-D4
P7-Plas.A2-G5
P7-Plas.1/5-3-G4
P7-B1.E2-D11
P7-B1.F9-A8
P7-B6.B7-C9
P7-B6.F11-E11
P7-Plas.E8-C12
P7-B6.F8-E6
P7-B6.A2-E8
P7-B6.C3-C11*
P7-B1.G7-F5
P7-Plas.G7-F12
P7BCC
P7-B1.D9-D2
P7-Plas.E7-B1
P7-Plas.E2-A10*
P7-Plas.A9-C6
P7-B1.F8-A5
P7-Plas.B8-C12
P7-B6.C7-B8
P7-B1.F5-B11
P7-Plas.B9-A4
0.0070


P9
875
967
P9BCC
P9-B6.C3-G10
P9-B1.G6-H8
P9-Plas.1/1-3-B12
P9-Plas.B2-G3
P9-B6.B12-B2
P9-Plas.D7-D9
P9-B1.C8-G8
P9-Plas.A5-H6
P9-B1.B9-C2
P9-B6.C12-E4
P9-B6.G6-C12
P9-B6.D10-G1
P9-Plas.F11-E4
P9-B1.A10-F3
P9-Plas.E8-D4
P9-B1.C7-D5
P9-Plas.1/1-2-C5
H77 (E1E2)
0.02

B

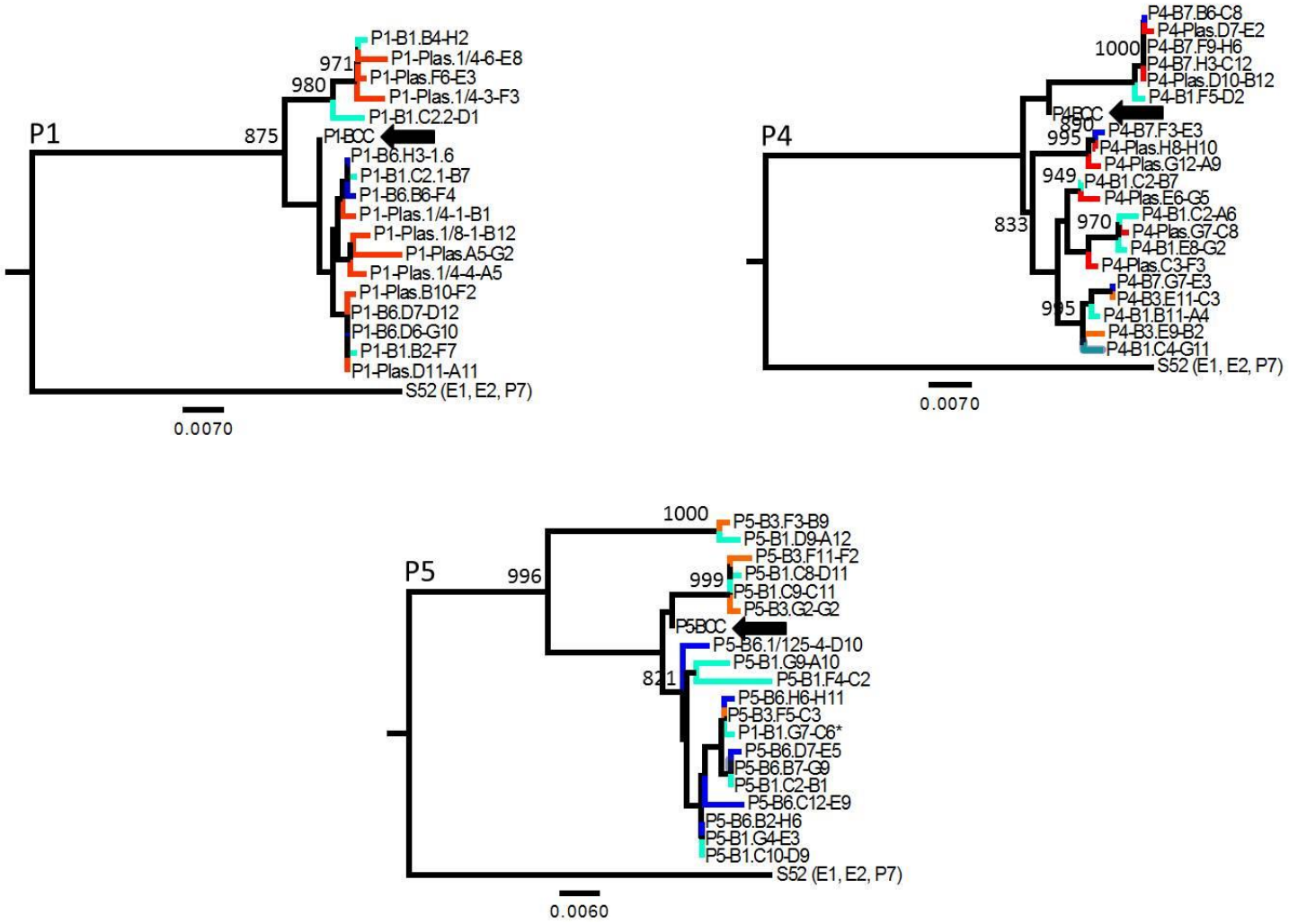

P1
875
P1-BCC
P1-B1.B4-H2
971
P1-Plas.1/4-6-E8
980
P1-Plas.F6-E3
P1-Plas.1/4-3-F3
P1-B1.C2.2-D1
P1-B6.H3-1.6
P1-B1.C2.1-B7
P1-B6.B6-F4
P1-Plas.1/4-1-B1
P1-Plas.1/8-1-B12
P1-Plas.A5-G2
P1-Plas.1/4-4-A5
P1-Plas.B10-F2
P1-B6.D7-D12
P1-B6.D6-G10
P1-B1.B2-F7
P1-Plas.D11-A11
S52 (E1, E2, P7)
0.0070
P4
P4-B7.B6-C8
P4-Plas.D7-E2
1000
P4-B7.F9-H6
P4-B7.H3-C12
P4-Plas.D10-B12
P4-B1.F5-D2
P4BCC
890
995
P4-B7.F3-E3
P4-Plas.H8-H10
P4-Plas.G12-A9
949
P4-B1.C2-B7
P4-Plas.E6-G5
833
970
P4-B1.C2-A6
P4-Plas.G7-C8
P4-B1.E8-G2
P4-Plas.C3-F3
P4-B7.G7-E3
P4-B3.E11-C3
995
P4-B1.B11-A4
P4-B3.E9-B2
P4-B1.C4-G11
S52 (E1, E2, P7)
0.0070
P5
1000
P5-B3.F3-B9
P5-B1.D9-A12
996
P5-B3.F11-F2
999
P5-B1.C8-D11
P5-B1.C9-C11
P5-B3.G2-G2
P5BCC
P5-B6.1/125-4-D10
P5-B1.G9-A10
821
P5-B1.F4-C2
P5-B6.H6-H11
P5-B3.F5-C3
P1-B1.G7-C6*
P5-B6.D7-E5
P5-B6.B7-G9
P5-B1.C2-B1
P5-B6.C12-E9
P5-B6.B2-H6
P5-B1.G4-E3
P5-B1.C10-D9
S52 (E1, E2, P7)
0.0060

Figure 6.14 Phylogenetic analysis of hepatic and plasma HCV amino acid sequences.

Nearest Neighbour analyses with 1000 bootstraps were applied to the SGA-derived sequences from patient 1, 2, 4, 5, 7 and 9 and represented as rooted Neighbour Joining trees. Branches supported with more than 750 bootstraps are indicated in the trees. The branch lengths are a direct measure of the genetic distances between the SGAs and a scale for the length can be seen underneath each tree. The branches from plasma (Plas) are red, biopsy 1 (B1) turquoise, biopsy 3 (B3) SGAs are orange and biopsy 6 (B6) and 7 (B7) blue. The branch for the consensus sequence (BCC: biopsy consensus consensus) generated from the 8 hepatic consensus sequences are coloured black and are indicated with an arrow. Sequences corrected for nucleotide deletions and insertions are indicated with *. **A.** Genotype 1a patients 2, 7 and 9. The E1E2 amino acid sequence for the genotype 3a S52 variant is used as a reference. **B.** Genotype 3a patients 1, 4 and 5. The E1E2-P7 amino acid sequence for the genotype 3a S52 variant is used as a reference.

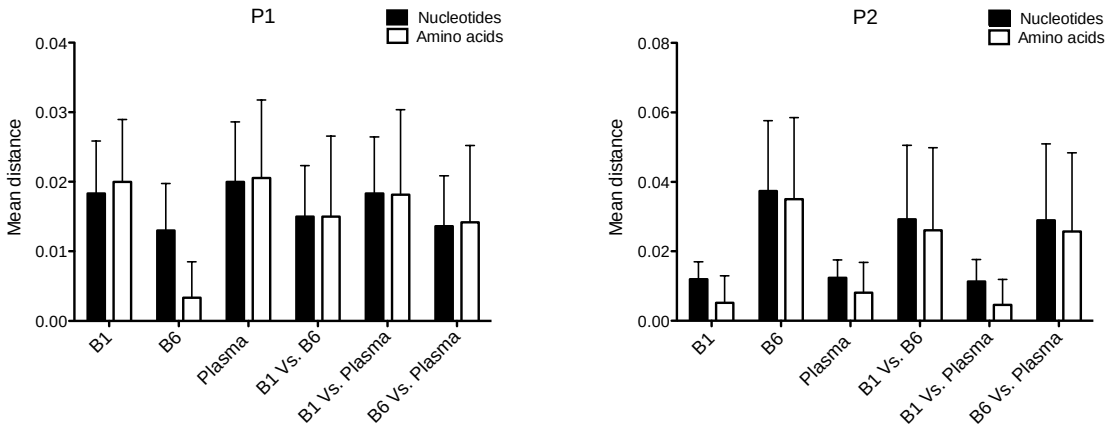


Figure 6.15 Distances between SGA-derived nucleotide and amino acid sequences.

The distances between SGAs were found for the nucleotide and amino acid sequences and the mean distance calculated between SGAs isolated from the same site and between sites. The data is here represented for patient 1 and 2 as a mean with standard deviation. Two-way unpaired t-test applied to the data set revealed no statistical differences in the mean distance between nucleotide and amino acid sequences (unpaired two-tailed t test: $p > 0.85$).

6.8 HCV sequences identified by next generation 454 pyro-sequencing

Analysing quansispecies distribution by the SGA sequencing approach requires extensive cloning of the PCR products to ensure coverage of major and minor sequences in the population. One of the concerns with the SGA sequencing approach is whether enough HCV variants are investigated to reach valid conclusions.

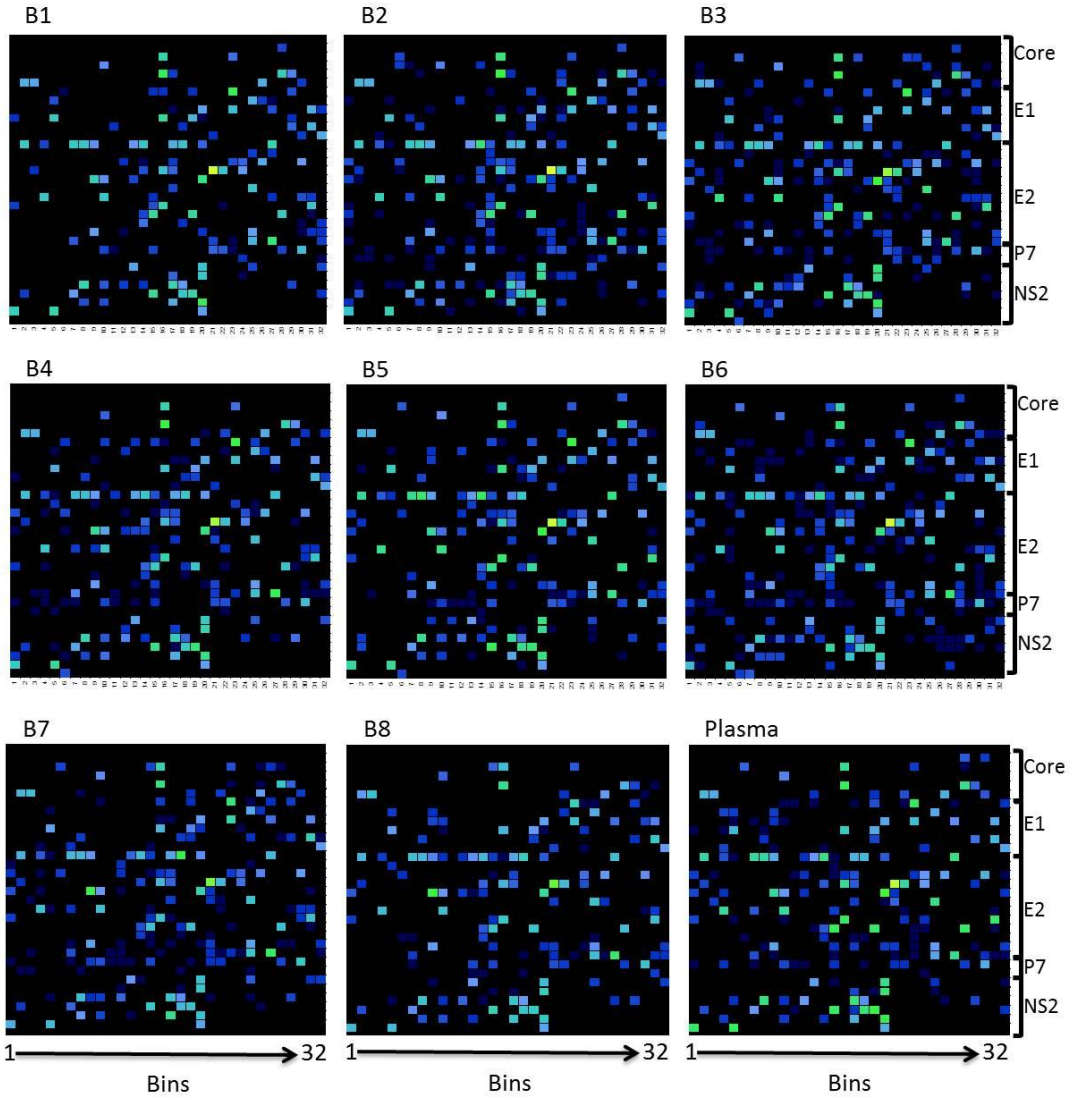
Within recent years next-generation sequencing platforms have been developed to investigate viral quasispecies in much greater detail. Its high throughput approach allows the generation of thousands of sequence reads in a single sample and facilitate in-depth sequencing of even low frequency variants (Beerenwinkel et al., 2012; Henn et al., 2012; Yang et al., 2012). To generate a more detailed investigation of the HCV quasispecies population at end-stage liver disease, total liver RNA and plasma from patients 1 and 2 were subjected to next generation 454 pyrosequencing (NGS). Our collaborators Dr. Damien Tully, Dr. Karen Power and Prof. Todd Allen at Harvard University, USA, performed the NGS and the following data processing.

Total liver RNA from all 8 biopsies and the plasma from patients 1 (genotype 3a) and 2 (genotype 1a) were sequenced across core, E1, E2, p7 and NS2 region (3.1kb). The efficiency of the emulsion PCR and the pyro-sequencing can be found in Table 3.6. The sequence reads were organised into 32 bins according to their size and the codon diversity were determined at 32

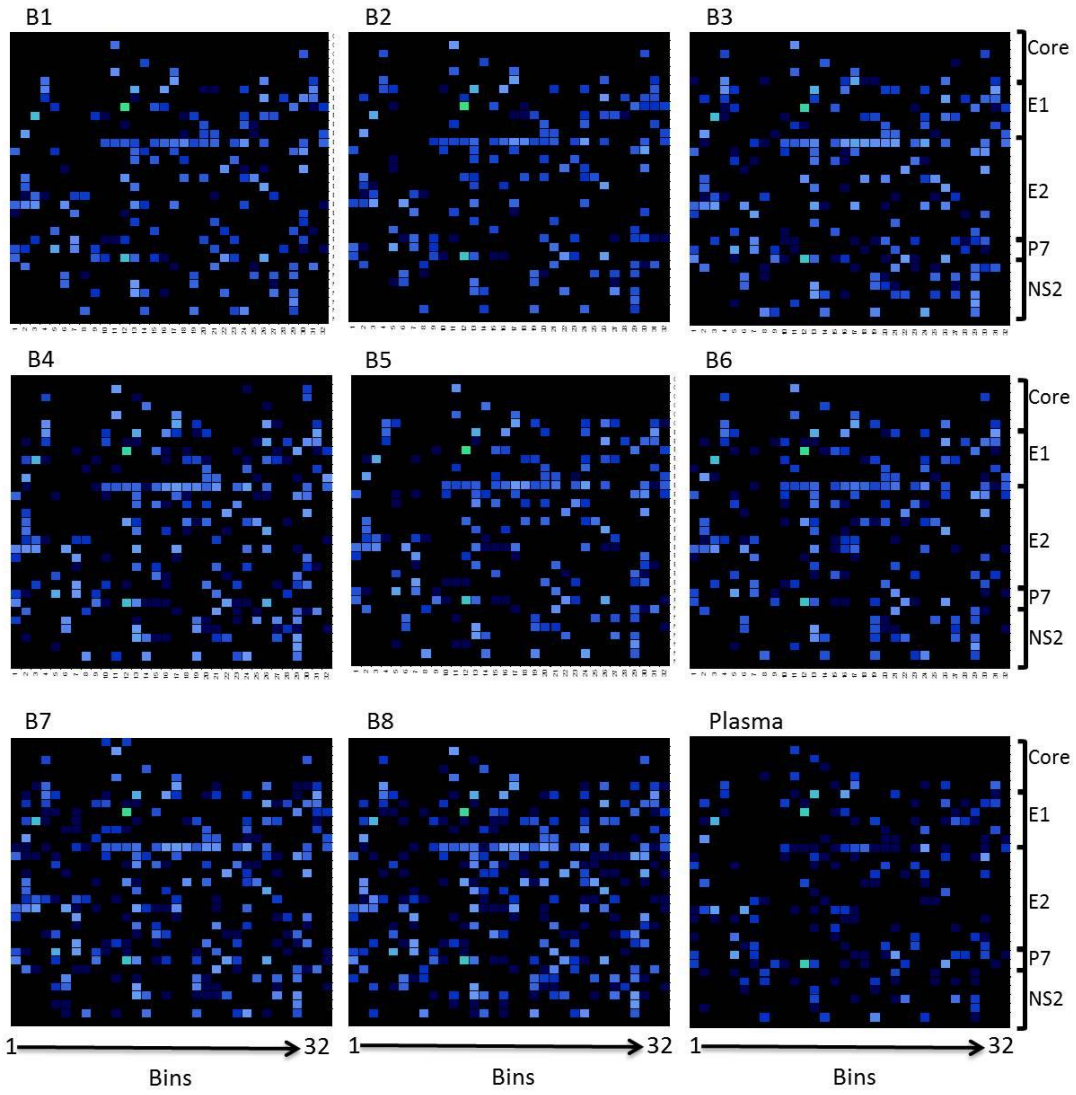
nucleotide long fragments (Figure 6.16A + B). The heat-maps of the hepatic sequence diversity revealed heterogeneity at each of the investigated sites. Not surprisingly, the highest level of diversity was found at the HVR1 site in the 5'UTR-end of E2 (Figure 6.16A + B). In addition, the fraction of polymorphic codons between sequences isolated from different hepatic sites only differed with approximately 1.4-folds (Figure 6.16C).

Comparing the sequence diversity between liver and plasma NGS data sets showed that the plasma-derived sequences from patient 1 shared similar levels of heterogeneity across the core-NS2 region as the liver (Figure 6.16A). However, for patient 2 the plasma sequences had lower numbers of polymorphic bins and an overall lower diversity across the investigated region (Figure 6.16B). This difference in sequence diversity for patient 2 plasma and liver became more appearent when comparing the number of changes per codon between the site-specific sequences. For patient 2 the number of polymorphism differs with approximately 1.5-fold between liver- and plasma-derived sequences, whereas patient 1 had a similar level of changes per codon (1.0-fold) (Figure 6.16C).

A Patient 1


B1
B2
B3
Core
E1
E2
P7
NS2
B4
B5
B6
Core
E1
E2
P7
NS2
B7
B8
Plasma
Core
E1
E2
P7
NS2
1
32
Bins
1
32
Bins
1
32
Bins
Color key
0
20
40
60
Diversity (%)

B Patient 2


B1
B2
B3
Core
E1
E2
P7
NS2
B4
B5
B6
Core
E1
E2
P7
NS2
B7
B8
Plasma
Core
E1
E2
P7
NS2
1
32
Bins
1
32
Bins
1
32
Bins
Color key
0
20
40
60
Diversity (%)

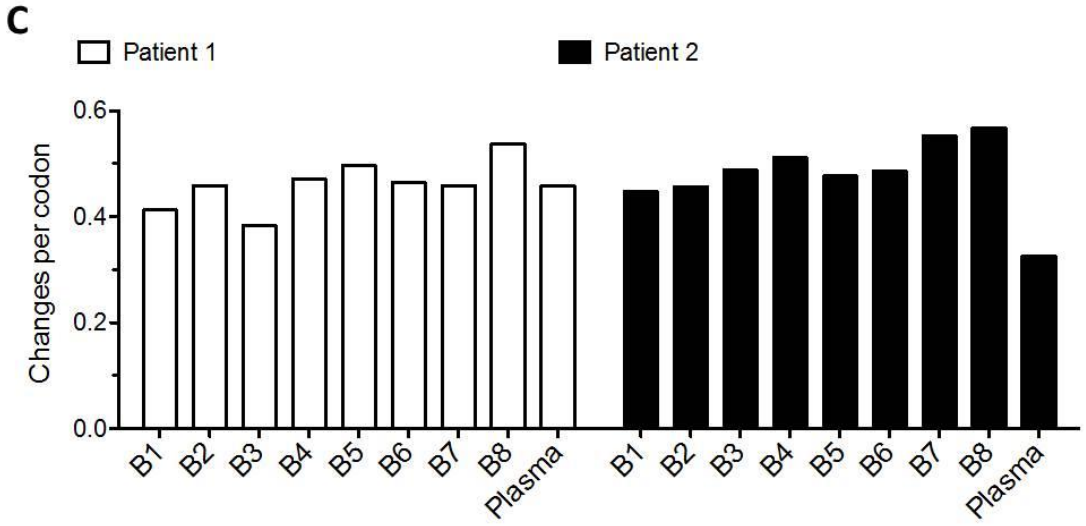


Figure 6.16 Heat-maps for HCV codon diversity detected by NGS.

Our collaborators Dr. Damien Tully, Dr. Karen Power and Prof. Todd Allen at Harvard University, USA, performed the NGS and data processing. **A** + **B.** The heat-maps show the codon diversities across the HCV core – NS2 region (Y-axes: bins of ~32 codons) for different sizes of read length (X-axes: bins of 6 bp). The diversity is indicated as a percentage **(A.** Patient 1, **B.** Patent 2). **C.** Quantifications of the total number of polymorphism per codon between the core-NS2 sequences identified in the plasma and liver specimens collected from patients 1 and 2.

The polymorphisms observed for the HCV sequences isolated from patient 1 and 2 were further divided into *ds* and *dn* changes. HCV diversity heat-maps for each of the collected samples revealed that the sequences contained both types of changes, however there was up to 4-fold more *ds* than *dn* changes (Figure 6.17A and B). As seen for the SGA-derived sequences (section 6.6), HVR1 was the only region with higher levels of *dn* than *ds* changes (Figure 6.17B).

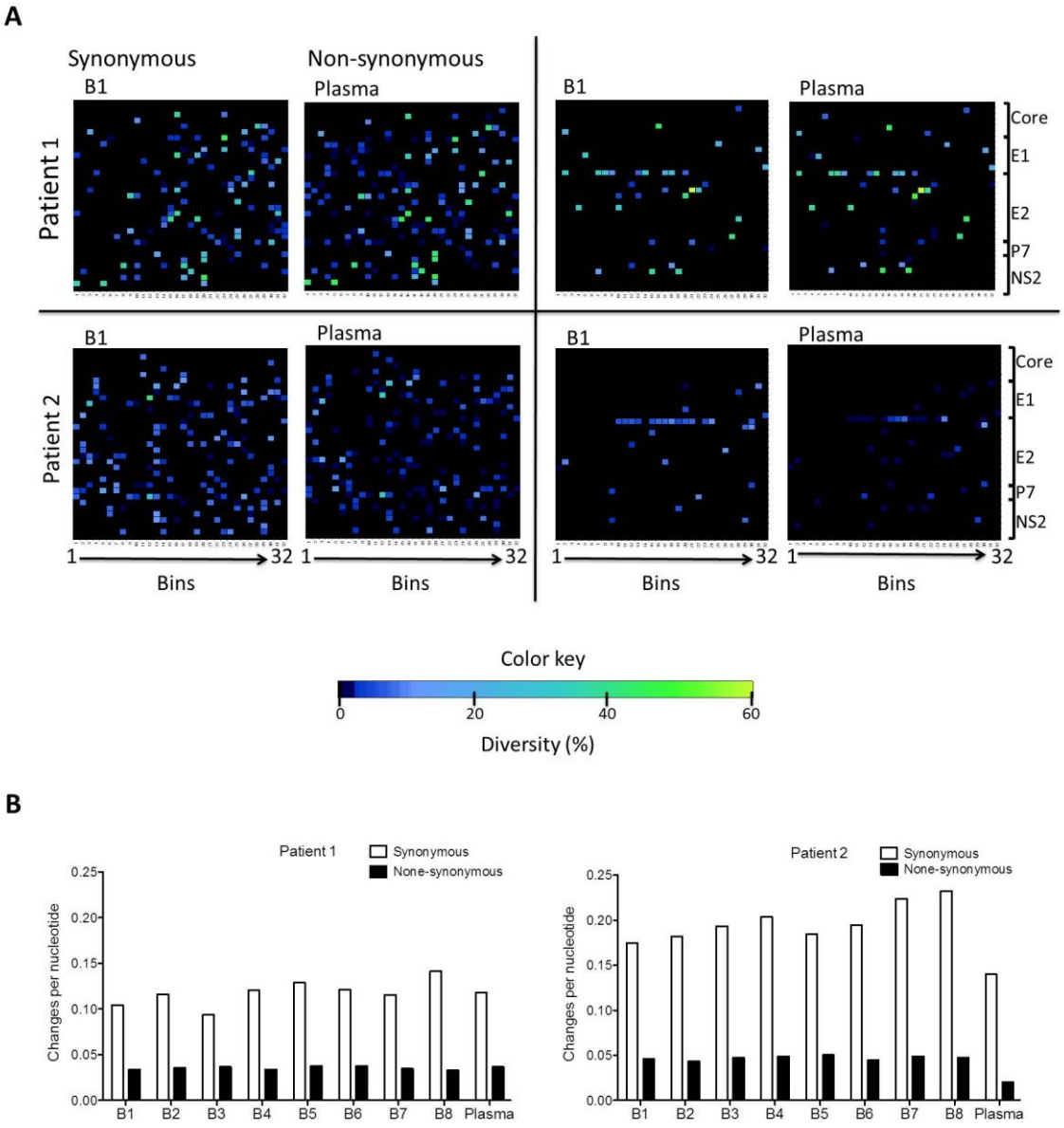


Figure 6.17 Heat-maps for HCV *ds* and *dn* diversity detected by NGS.

Our collaborators Dr. Damien Tully, Dr. Karen Power and Prof. Todd Allen at Harvard University, USA, performed the NGS and data processing. **A.** Heat-maps of *ds and dn* changes observed for each codon (Y-axes: bins of ~32 codons) in the different sized sequence reads (X-axes: bins of 6 bp). The diversity is indicated as a percentage. **B.** Quantifications of the *ds* and *dn* changes per codon were calculated for patient 1 and patient 2 specimens.

Statistical analyses of the viral population distribution require alignments of the isolated sequences. One of the disadvantages with NGS is that sequencing is performed on relatively short fragments (~350 base) and we can therefore not assume true linkage of the 3.1kb region. Instead of risking erroneous alignments of the full-length core-NS2 fragment, our collaborators decided to use HVR1 for the compartmentalisation analysis. Alignments of the HVR1 amino acid sequences showed that patient 1 and patient 2 had identical major sequence in the specimens collected from one liver, but whereas patient 1 had four dominating strains with prevalence above 10%, patient two had one (Figure 6.18). Instead, patient 2 had far more unique minor variants at the different sites.

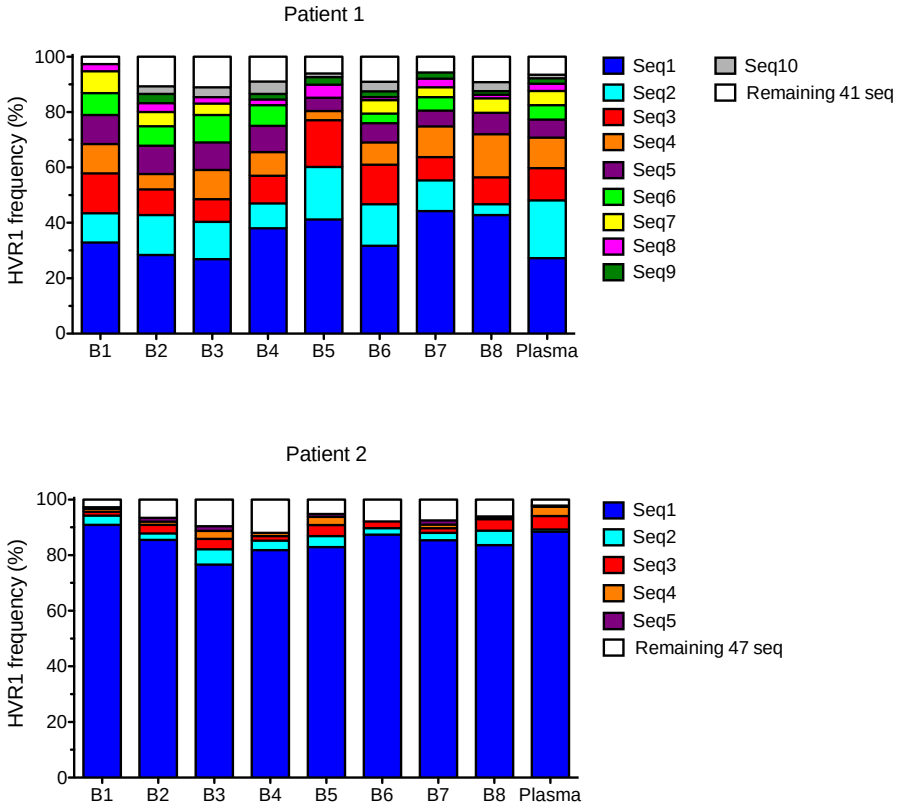


Figure 6.18 HVR1 amino acid diversity across liver and plasma specimens.
The frequencies of the different HVR1 sequences (Seq) were calculated as a percentage of the total number of sequences detected. All HVR1 sequences were specific to each patient. Sequences with frequencies below 5% are presented together in a white bar. The data has been provided by Dr. Damien Tully, Dr. Karen Power and Prof. Todd Allen at Harvard University, USA.

Performing four independent statistical analyses (Slatkin-Maddison, correlation coefficient of branch number, Wright's measure of population subdivision and Nearest neighbour statistics) reveal no significant evidence for HVR1 genetic compartmentalisation between biopsy and plasma for patient 1 and 2 (Table 6.5). Statistical compartmentalisation analysis between the intra-hepatic sequences has not yet been received, but recent conversations with our collaborators reveals that the HVR1 sequences do not form compartmentalisation between different sites in the liver.

Table 6.5: Statistical analysis of HVR1

	Slatkin-Maddison (p-value)	Correlation coeffient by branch numbers (p-value)	Wright's measure of population subdivision (p-value)	Nearest neighbor statistics (p-value)
Patient 1	0.65	0.88	0.92	0.98
Patient 2	0.23	0.79	0.78	0.85

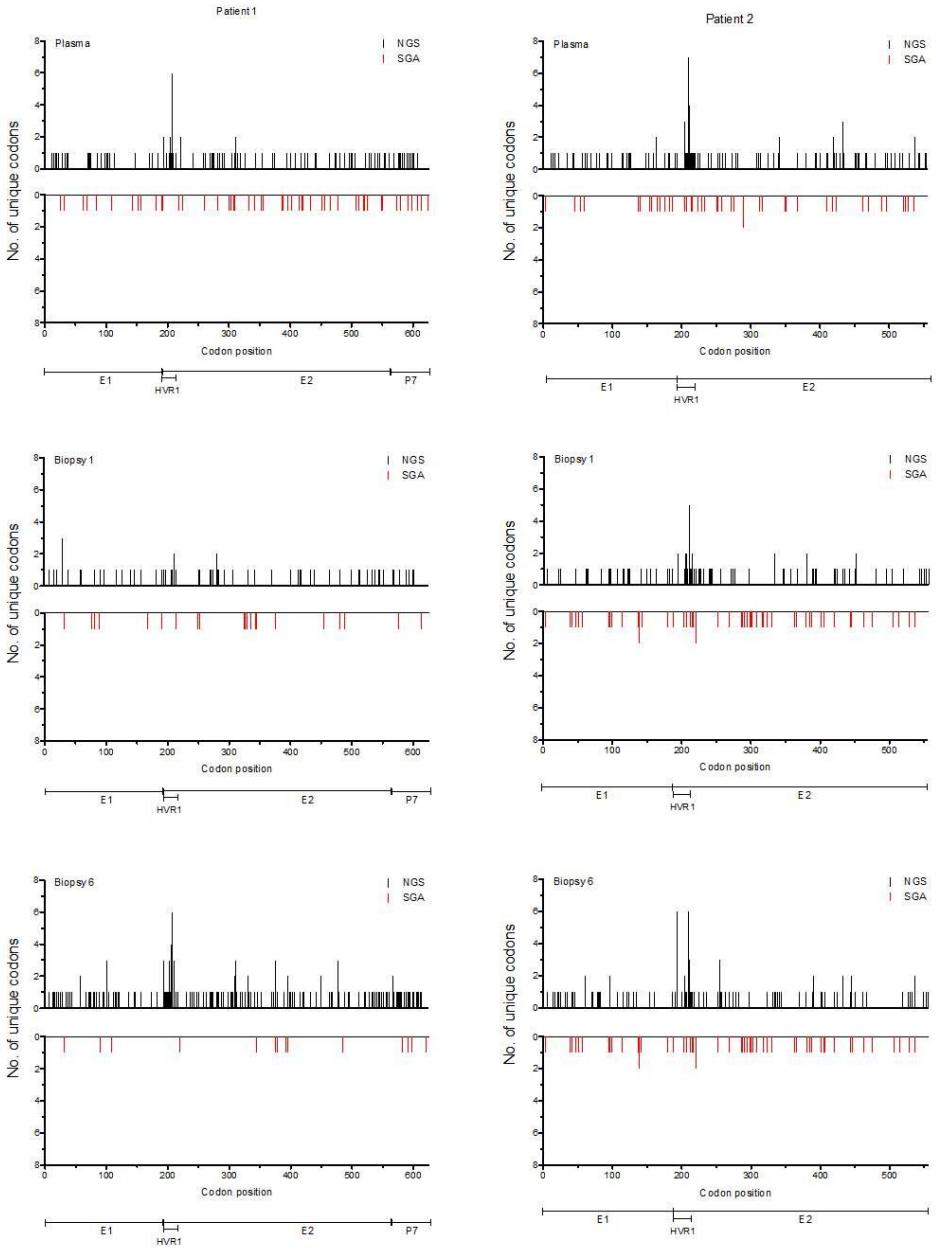
6.9 Comparing data from next generation sequencing to the single molecule amplification

For this study two sequence amplification methods and sequencing strategies were used for analyzing the HCV sequence diversity in samples collected from different anatomical sites in patients with end-stage liver disease. This allows us to evaluate and compare the efficiency between the two techniques.

The NGS- and SGA-derived HCV sequences collected from patient 1 and 2 were compared for the E1, E2 and p7 region or the E1 and E2, respectively. Polymorphisms detected by one of the techniques and not the other were noted and the number of unique codons displayed across the investigated region (Figure 6.19A). As expected, NGS identified more unique codons across the fragment than the SGA and Sanger sequencing had. Quantifying the fraction of unique codons detected by the two approaches, demonstrated that up to 89% of all polymorphisms detected by NGS were also found by SGA. In comparison, up to 97% of the SGA detected codons were also found by NGS (Table 6.6). We also detected multiple amino acid polymorphisms unique to each sequencing approaches, but their frequencies and the level of variety was far less than for the codons (Figure 6.19B). As a consequence NGS had coverage 99% of the SGA amino acid sequences and SGA had covered 98% of the amino acids detected by NGS (Table 6.6). The difference in NGS and SGA coverage for nucleotide and amino acid coverage is likely to be the result of minor HCV variants, which are more efficiently detected by NGS. However, it is important to note that NGS did not have 100% coverage of all SGA

sequences, indicating that although NGS is resourceful in detecting many different variants it might not detect extremely rare variants, which are below the detection threshold. These will only be detected by random chance in the SGA.

A


Patient 1
Patient 2
Plasma
Biopsy 1
Biopsy 6
NGS
SGA
No. of unique codons
Codon position
E1
HVR1
E2
P7

B

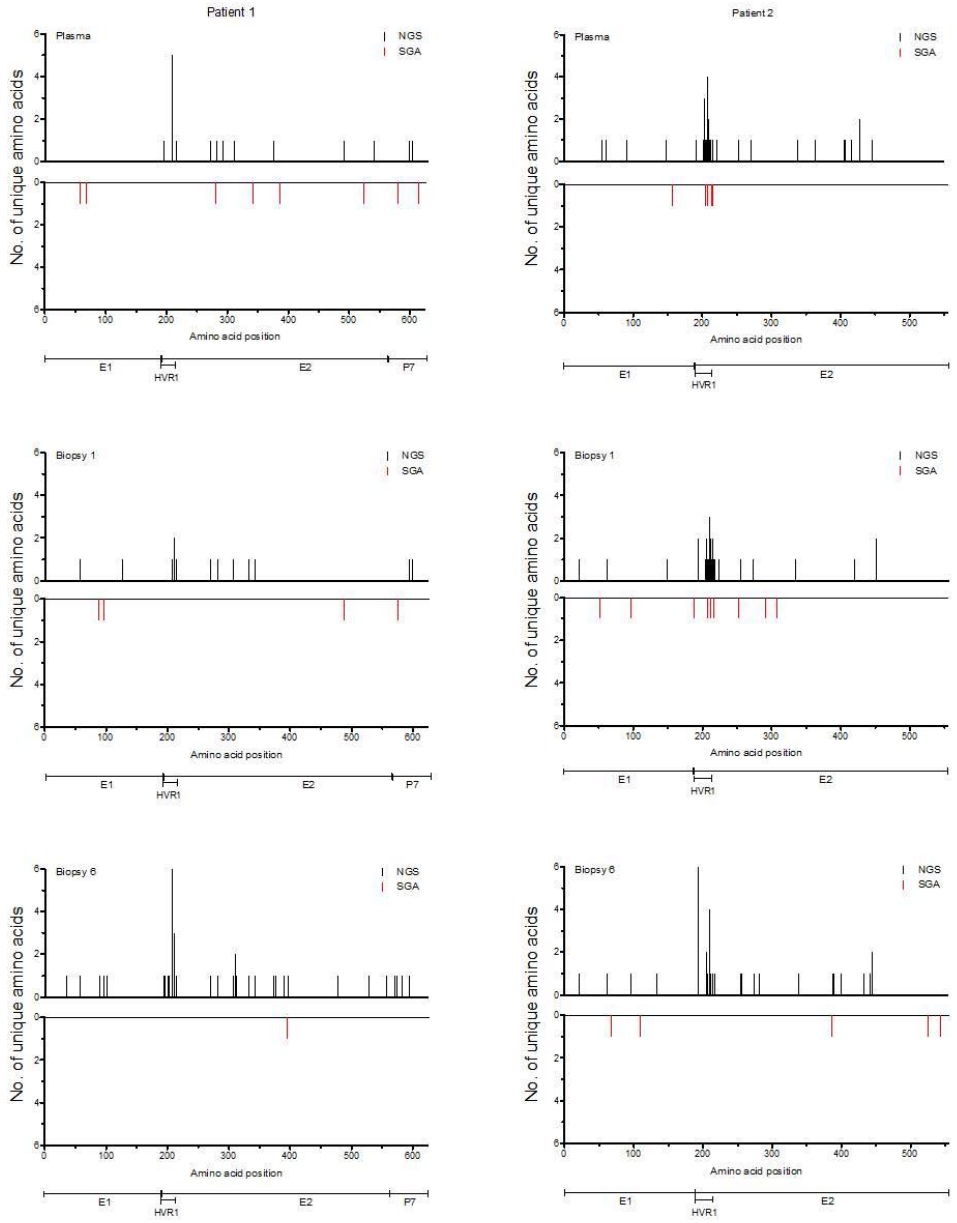


Figure 6.19 Detection of unique codons and amino acids by SGA and NGS.
Diversities were compared between the sequences detected by SGA and NGS and the number of unique sites, not detected by the competitive approach, was noted. The value zero indicates that the same codon(s) were detected by both techniques. The position of the sites can be seen underneath the graphs. **A.** Quantification of unique codons for plasma, biopsy 1 and biopsy 6 from patient 1 and 2. **B.** Quantification of the unique amino acids detected by SGA and NGS are found for the same samples as in A.

Table 6.6: Identification of unique codons and amino acids by SGA and NGS

Patient	Sample	Sites with unique codons (%)		Sequencing coverage (%)		Sites with unique aa (%)		Sequencing coverage (%)	
		NGS	SGA	NGS	SGA	NGS	SGA	NGS	SGA
1	Plasma	15	8	92	85	1.9	1.3	98.7	98.1
	Biopsy 1	10	3	97	90	1.9	0.6	99.4	98.1
	Biopsy 6	26	2	98	74	5.3	0.2	99.8	94.7
2	Plasma	17	7	93	83	4.8	0.9	99.1	95.2
	Biopsy 1	14	9	91	86	3.8	1.6	98.4	96.2
	Biopsy 6	15	9	91	84	4.3	0.9	99.1	95.7

6.10 Discussion

Limited information is available on the distribution of HCV quasispecies within the liver and between liver and plasma for patients with end-stage liver disease. Increasing our understanding of the spatial distribution of HCV genetic variants within infected tissue will help elucidate the factors controlling viral replication and dissemination routes. The goal of this study was to investigate the genetic compartmentalisation of HCV variants in the chronically infected liver and to compare these with the virus circulating in the periphery. Detection of the HCV quasispecies was performed with three different approaches: (*i*) bulk amplification followed by consensus Sanger sequencing; (*ii*) Single molecule amplification (SGA) followed by Sanger sequencing, and (*iii*) 454 next-generation pyro-sequencing (NGS).

HCV bulk amplifications were made from each of the 8 hepatic sites from 6 patients and the PCR products sequenced by consensus sequencing. The HCV consensus sequences contained relatively few polymorphisms across the E1, E2 (and p7) region and were comparable between sites of the same liver. This high level of similarity is considered a result of low sensitivity of the assay, since a spiking experiment with different concentrations of HCV RNA revealed that sequences present at low concentrations (below a ratio of 1:1) were not efficiently detected (Figure 6.5). This hypothesis would predict that consensus sequencing underestimates the HCV genetic variability for a sample and favours the detection of the dominating sequences. As a result any comparison of consensus sequences could create false inferences. Our

phylogenetic analyses of the SGA-derived HVR1 sequence also underestimated the HCV complexity of the samples, indicating that analyses of small regions should only be performed on large data sets, such as the data generated by NGS, which will have a greater level of sequence coverage.

The SGA-derived HCV sequences from selected hepatic sites revealed that only 2 pairs of patient-specific sequences were 100% identical and that the remaining sequences in each patient differed with up to 105 substitutions. The high number of non-identical sequences from the same patient is consistent with previous sequencing studies on serum-derived HCV ORFs isolated from acutely infected HCV patients and chimpanzees (Yanagi et al., 1997; Yanagi et al., 1998). As predicted, the SGA-derived sequences from different sites had significantly more substitutions than the HCV consensus sequences generated from the 8 hepatic sections, strengthening our earlier conclusion that consensus sequencing underestimated the HCV complexity in the hepatic samples.

The phylogenetic and Shannon entropy analyses revealed a high level of similarity between SGA-derived sequences independent of the hepatic site they had been isolated from and in most cases it was impossible to differentiate between the site-specific sequences in the patient-specific phylogenetic trees. Statistical analysis of the minimum migration events required to explain the distribution of HCV sequence within one liver (Slatkin-

Maddison modelling) revealed no evidence of statistical compartmentalisation for five of six patients. Furthermore, the NGS data sets for patient 1 and 2 showed limited evidence of compartmentalisation across all 8 investigated sites in the explanted liver. From these observations we conclude that intra-hepatic HCV genetic compartmentalisation is limited during end-stage liver disease.

The SGA-derived sequences isolated from the liver of patient 7, showed statistical evidence of compartmentalisation between biopsies 1 and 6. The non-random distribution of HCV variants between individual sites in the liver may be the result of high RIG-I and TLR activation (see Chapter 5), leading to higher turnover of infected hepatocytes than seen for the other patients. This would result in new naïve hepatocytes being generated, which in principle could be infected by genetically distinct HCV founder viruses and eventually the formation of genetic compartmentalisation (Heim, 2013b). An alternative explanations may be that the HCV variants isolated from the patient 7 liver have a higher fraction of cell-cell transmission compared to cell-free, leading to the formation of genetically distinct foci in the infected patient. This hypothesis is supported by previous reports of HCV antigen and RNA positive foci in the chronically infected liver (Kandathil et al., 2013; Liang et al., 2009; Wieland et al., 2013).

In this study, all SGA-derived HCV sequences were subjected to Sanger sequencing, however it is possible that we only sampled HCV sequences that were significantly different between the sites in patient 7 liver or that only highly similar sequences were sampled from the livers of patient 1, 2, 4, 5 and 9. This phenomenon is known as “re-sampling bias” and may result in false interpretation of the results. One of the major concerns when sequencing individual HCV clones is to ensure that enough variants are included in the analyses to support valid conclusions. Previous studies suggested that 99 HVR1 clones were needed to identify 95% of all variants present at a frequency of at least 3% (McCaughan et al., 2003) and for HIV compartmentalisation studies at least 20 HIV envelope clones should be studied from each site (Zarate et al., 2007). These high numbers are not achieved for the SGA analyses of any of the studied patients and one may argue that more SGAs are required in our analyses. The NGS data for the 8 hepatic samples collected from patient 1 and 2 confirms the limited detection of HCV compartmentalisation between hepatic sites and suggests that analysing at least 4 SGA-derived HCV envelope sequences from each hepatic specimen are enough to avoid “re-sampling bias” of samples collected during end-stage liver disease. The observation is further confirmed by SGA-derived sequences covering up to 89% of the polymorphisms detected by NGS. It should be noted that more sequences might be acquired from earlier HCV-associated disease stages. Performing NGS for the samples collected from patient 7 would be required to confirm valid sequence representation among the SGA-derived sequences in this patient.

It should be mentioned that the statistical evidence for HCV genetic compartmentalisation found for hepatic SGA sequences isolated from patient 7 may be a result of overestimation of the degree of segregation between branches by the statistical model. The Slatkin-Maddison model estimates the tree topology without bootstrapping and the compartmentalisation score may lead to an overestimation because short interior branches may have a higher (false) weight, which is achieved through a low (not accounted for) bootstrapping value (Zarate et al., 2007). Performing additional compartmentalisation analysis for the SGA-derived sequences, such as Mantel's or Wright's test would potentially lead to non-statistically significant results for the hepatic sequences isolated from patient 7.

Investigating the internal genetic between SGA-derived plasma sequences, collected at the time of transplantation, revealed that the mean genetic distance did not statistically differ from the mean distance found between the liver- and plasma-derived sequences detected in the same patient. In addition, we did not find any statistical evidence of HCV genetic compartmentalisation between the two compartments by Slatkin-Maddison analysis. The observation was confirmed by NGS of the HVR1 amino acid sequences isolated from patient 1 and 2 and it is concluded that HCV genetic compartmentalisation between plasma and liver is limited during end-stage liver disease.

NGS provided in depth information about the HCV diversity and complexity within the samples collected from patient 1 and 2. Not surprisingly, the liver-derived HVR1 sequences had higher diversity than the sequences isolated from the plasma and contained several unique minor sequences, which were not detected in the plasma specimens. Similar observation has previously been reported for patients with early chronic infection (Fan et al., 1999). This finding suggests that not all HCV sequences are released into the periphery. It is likely that these minor sequences either were not packed into virus particles or that the viral progeny were not released. Interestingly, the NGS data revealed that 2% of the HVR1 sequences detected in the plasma of patient 1 were not found in the hepatic samples. Their low frequency implies that the sequences do not form a distinct viral population within the plasma, which is also confirmed by the lack of statistical evidence for HVR1 compartmentalisation between plasma and liver. The observation of unique plasma-derived HVR1 sequences suggests that the 8 hepatic specimens do not cover all the HCV variants found in the periphery. It is possible that these plasma-derived minor sequences would be identified in the liver if NGS was applied to additional hepatic specimens, however the sequences would probably remain rare since the 8 investigated sections already showed identical dominant sequences. Alternatively, the unique plasma-derived HVR1 sequences may have originated from HCV replication in extra-hepatic sites, such as the CNS (Maggi et al., 1999), PBMCs (Natarajan et al., 2010; Willems et al., 1994) or lymph nodes (Pal et al., 2006). All of these studies have used negative-strand qPCR to determine HCV replication; a technique that is considered to have low specificity and the prediction of active replication at

these sites is debateable. At the time of writing, there is no reliable evidence for or against extra-hepatic HCV replication *in vivo* and it is not possible to reject the theory that these NGS detected minor sequences in the plasma may have derived from HCV replication outside the liver. To determine whether the identified plasma-derived sequences were present in the liver, additional hepatic samples would be required and the NGS continued until all HVR1 sequences in the plasma had been accounted for in the liver.

In this study we found limited statistical evidence of genetic compartmentalisation between HCV sequences deriving from different sites in the liver and between liver and plasma, when using the SGA and NGS approaches. These observations contrast with Li *et al.* who previously reported that 78% of patients with chronic HCV infection had statistical evidence for HVR1 genetic compartmentalisation between liver and plasma (Li et al., 2010) and with Sobesky *et al.* who reported that the HCV core sequences formed distinct compartments between micro-dissected hepatic sections (Sobesky et al., 2007). The statistical analysis were in both studies performed by the Mantel's test, which compares matrixes of genetic and phenotypic distances to determine whether sequences in one compartment are more similar with each other or with sequences in another compartments (Beerli et al., 2001). Difference in the statistical approach may explain our contradicting results. It is however considered more likely that either the different technical approach taken by Li *et al*. and the dissimilar disease stage of the collected samples in the Sobesky *et al.* study are the source of the discrepancy between our

observations. In the Li *et al.* study HVR1 bulk amplifications from each compartment were cloned and the HVR1 positive colonies were afterwards screened by heteroduplex gel mobility to identify unique HVR1 sequences. Only unique HVR1 sequences were included in the statistical analysis and the complexity of HVR1 minor and major sequences are not taken into consideration (Li et al., 2010). It should be mentioned that the same researchers found close to identical HVR1 sequence complexity, when performing the heteroduplex gel mobility assay for 10 hepatic specimens sampled from two explanted livers (Li et al., 2010). Although a conclusion on intra-hepatic HCV compartmentalisation cannot be made from that analysis, it is notable that Li *et al.* reported high HVR1 variability between hepatic and plasma sequences and low HVR1 variability between hepatic specimens from patient's with end-stage liver disease. In contrast, our cohort showed comparable HCV diversity and complexity between all investigated compartments. The differences observed between ours and the Sobesky *et al.* intra-hepatic compartmentalisation study may be the result of our samples being collected from patients with end-stage liver disease whereas the patient's in the Sobesky's study had earlier stages of liver disease, which required hepatoctomy and not liver transplantation (Sobesky et al., 2007). To the best of our knowledge no other studies have reported statistical evidence for or against HCV genetic compartmentalisation. Instead several studies have reported that identical dominant and sub-dominant HCV variants were found between plasma- and liver-derived HCV clones (Cabot et al., 2000; Ramirez et al., 2009; Sakai et al., 1999) and between multiple hepatic sites (Cabot et al., 1997). In the Cabot *et al.* study researchers reported that HCV does not form

large-scale intra-hepatic compartmentalisation based on the detection of identical dominant and sub-dominant 5'UTR and E2-NS2 junction sequences, despite each sample containing multiple unique minor sequences (Cabot et al., 1997). Since these researchers did not account for the minor sequences their conclusion disputes the fundamental concept behind quasispecies modelling; "*A quasispecies population acts as a whole unit rather than an individual sequence acting in isolation*" (Holmes, 2010a). According to this concept, conclusions on population distribution should be based on the whole unit of HCV sequences and not on individual clones.

The SGA and NGS data sets revealed that most substitutions were found in the E2 region, followed by E1. The identification of a high substitution rate in E2 is not surprising, since this region is believed to contain multiple nAb epitopes, leading to a positive selection (Fournillier et al., 2001; Helle et al., 2011; Law et al., 2008; Logvinoff et al., 2004; Tarr et al., 2011). However, when investigating the translated sequences we found higher fractions of synonymous (*ds*) than non-synonymous changes (*dn*) for most of the investigated regions, except within the HVR1 region (*ds/dn* < 1). The high number of *ds* changes outside the HVR1 supports a model of high turnover of viral variants with deleterious *dn* changes, leading to an excess of *ds* mutations that do not have an effect on the phenotype (Wang et al., 2007b). It is concluded that the overall selective pressure across the envelope genes are limited to the HVR1 during end-stage stage liver disease. The identification of *ds*/*dn* ratios below 1 in HVR1 may be driven by selective pressure from nAbs

targeting this region (Booth et al., 1998; Kato et al., 1994; von Hahn et al., 2007) and/or from variability in the SR-BI affinity (Bankwitz et al., 2010). The exact source of the selective pressure on the HVR1 evolutionary path in this cohort is unknown.

The main focus of this study was to investigate the HCV genetic distribution during end-stage liver disease. A potential next step would be to biologically characterise the clones generated by SGA. Although analysis of the clone's phenotypes has not yet been done, the examination of essential N-linked glycosylation sites can reveal important information about the potential phenotype (Bungyoku et al., 2009; Helle et al., 2007; Helle et al., 2010; Wei et al., 2003). Investigation of the SGA sequences revealed that HCV variants isolated from patient 4 had polymorphic changes in E2N6 and E2N8, resulting in loss of the N-linked glycosylation. These variants are likely to have less stable E1E2 heterodimerization, leading to defective viral assembly (Helle et al., 2010) and a higher sensitivity to neutralisation by antibodies targeting the CD81 binding site, but potentially also a stronger interaction with CD81 since glycans is no longer shielding the binding site (Bungyoku et al., 2009). The SGA-derived HCV sequences isolated from patient 4 also contained polymorphisms in E2N5 and the P7-B1.F2-D11 had a substitution in E2N9, however, these sites do not appear essential for a specific phenotype (Helle et al., 2007; Helle et al., 2010). These postulated biological characterisations are at this point only speculations, empirical testing is required to confirm the phenotypes. All SGA sequences are cloned into the pcDNA3.1D expression

vector for producing infectious virus particles in the HCVpp system, but due to limited time these experiments are still to be performed. In the future we also hope to have SGAs isolated from the same patients following transplantation in order to characterise potential bottleneck effects during the liver-graft re-infection, as previously reported by Fafi-Kramer *et al.* (Fafi-Kremer et al., 2010) and identification of potential patient-derived ITX5061 escape variants (Zhu et al., 2012).

In conclusion, we found that HCV sequences isolated from different sites in the liver and plasma are highly comparable and found limited evidence for genetic compartmentalisation between different sampling sites collected from patients with end-stage liver disease. Identification of nearly identical HCV variants across the liver and between plasma and liver may have an impact on the choice of DAA treatment for patient with end-stage liver disease.

7 GENERAL DISCUSSION

The main site of HCV replication is hepatocytes in the liver. From here thousands of virus particles are produced and released. During the replication process the error-prone viral encoded RNA-dependent polymerase (NS5B) make mistakes, especially in the E1 and E2 envelope encoding genes, and the end-result is a 2nd generation of virus variants with these mutational changes expressed in the envelope proteins. Several studies have shown that patient-derived HCV ORFs encoding variable E1 and E2 amino acid sequences escapes the immune response (Keck et al., 2009; von Hahn et al., 2007) and anti-viral treatments (Catanese et al., 2013; Wu et al., 2013). They may also show different dependencies on the four viral entry factors (Grove et al., 2008; Prentoe et al., 2013), explaining why some HCV variants are more successful in infect new hosts (Brown et al., 2012; Bull et al., 2011; Wang et al., 2010) and re-infect the new liver graft (Fafi-Kremer et al., 2010). The observations of variable phenotypes among the patient-derived HCV particles could hypothetically be expanded to virus variants having diverse abilities to spread via a cell-free or cell-cell route, resulting in some variants being restricted to neighbouring cells and others released to the periphery. Recent *in vivo* data have found HCV antigen and/or RNA positive foci, surrounded by naïve hepatocytes in needle biopsies from HCV chronically infected patients (Kandathil et al., 2013; Liang et al., 2009; Stiffler et al., 2009; Wieland et al., 2013), supporting a model of cell-cell spread. These observations suggest that the HCV genetic pool at these foci may vary due to independent evolution at particular sites.

In this study we investigate the HCV quasispecies distribution within different sites of the liver and between plasma and liver for patients with end-stage liver disease and at the same time investigated host factors that may determine the HCV population dynamics. Comparing the viral load and the HCV quasispeceis sequences between different sites in the liver indicated an equal distribution of HCV RNA copies and a low genetic distance between the HCV sequences for all patients. This observation may be the result of having sampled 8 sites with comparable viral burdens and genetic complexities (re-sampling bias) and so sampling other sites potentially would have shown higher variability. This hypothesis is considered unlikely due to the high number of study subjects and the high number of hepatic sites sampled for each liver. In addition, the heat-denaturation of two biopsies from patients 1 and 5 implied no differences in the negative : positive HCV RNA strand ratios, which makes it unlikely that the virus would have replication hotspots with high content of detectable positive strands and independent viral evolution. The observation of highly comparable viral burden and HCV complexity between patient-specific specimens could, furthermore, be the result of having detected plasma- and not hepatocyte-derived sequences from the liver specimens. However, the finding of limited correlation between liver and plasma HCV RNA and the identification of unique plasma- and liver-derived sequences by NGS makes it unlikely that this would be the case. Instead, the result obtained in this study suggests that the observed HCV distribution is real and that macroscopic HCV compartmentalisation is limited during end-stage liver disease.

Similar viral burden and HCV sequence complexity across the liver suggests that the liver is exposed to comparable degrees of immune stimulation, leading to a uniform selective pressure on viral replication and HCV sequence evolution. Measuring the expression level of the 4-gene classifiers across the livers of 23 patients demonstrated that most patients had comparable ISG expressions. This even distribution of ISG may associate with the observed hepatic viral burden and sequence dynamics. In this model the uniform distribution of HCV RNA would lead to a concurrent stimulation of the innate and adaptive immune response, resulting in comparable ISG expression levels across the liver. The uniform immune activation may form comparable selective pressure on the HCV variants leading to genetically similar HCV sequences across the liver. Although we did not observe any correlation between hepatic viral load and ISG expression, the even expression of ISG may be stimulated by the mere presence of the virus and not its quantity. This posed model offers an attractive explanation to some of the *ex vivo* results obtain in this PhD study, however due to the complexity of the scenario it would be impossible to test the hypothesis *in vitro*. An alternative option would be to investigate the association between HCV viral burden, innate immune response and HCV sequence complexity in the humanized urokinase-type plasminogen activator severe combined immunedeficienct (uPA-SCID) mouse with transplanted human hepatocytes, however this model would only provide information about the type I and III ISGs since the mouse model does not have an adaptive immune system (Billerbeck et al., 2013; Bukh, 2012).

It is important to note that the data presented in this thesis were generated from livers with end-stage liver disease and many of these subjects may have been infected for up to 40 years. The HCV quasispecies dynamics in the liver and plasma may differ considerably during earlier stages of infection as a result of a more vigorous immune response or a higher hepatocyte turnover. This may explain the different observations made in our study and the studies reported by others (Kandathil et al., 2013; Li et al., 2010; Liang et al., 2009; Sobesky et al., 2007; Wieland et al., 2013).

8 CONCLUDING REMARKS

This study is the first to investigate the distribution of HCV at multiple sites during end-stage liver disease. The genetic and viral load data indicate that infectious HCV variants do not target specific sites in the liver, but that the liver is uniformly infected during end stage-liver disease.

The viral load between the plasma and liver show limited correlation, suggesting that the plasma is a poor predictor of the hepatic viral burden. Detection of higher HCV sequence complexity in the liver compared to the plasma, suggests that not all liver residential HCV sequences are released to the periphery. This observation may be clinically important when monitoring viral infections or assessing viral respond to treatments.

Furthermore, we conclude that the expression pattern of the 4-gene classifiers is comparable across the liver, however independent of the viral load in both the liver and the plasma. It is likely that the expression levels of ISG are not initiated in a dose-responsive manner to viral replication or the frequency of infected cells. In addition we found that the expression level of IFI27, but not ISG15, RSAD2 or HTATIP2 are unique to adults with HCV-related end-stage liver disease.

Investigating the dynamics of the HCV quasispecies are important in order to understand how the virus can persist despite the presence of a vigorous anti-HCV immune response and therapeutic therapies.

9 REFERNCES

Abdel-Hady, M., Bunn, S.K., Sira, J., Brown, R.M., Brundler, M.A., Davies, P., Kelly, D.A., 2011. Chronic hepatitis C in children--review of natural history at a National Centre. Journal of viral hepatitis 18, e535-540.

Acton, S., Rigotti, A., Landschulz, K.T., Xu, S., Hobbs, H.H., Krieger, M., 1996. Identification of scavenger receptor SR-BI as a high density lipoprotein receptor. Science 271, 518-520.

Afdhal, N.H., 2004. The natural history of hepatitis C. Semin Liver Dis 24 Suppl 2, 3-8.

Aggarwal, A., Iemma, T.L., Shih, I., Newsome, T.P., McAllery, S., Cunningham, A.L., Turville, S.G., 2012. Mobilization of HIV spread by diaphanous 2 dependent filopodia in infected dendritic cells. PLoS Pathog 8, e1002762.

Akira, S., Takeda, K., 2004. Toll-like receptor signalling. Nat Rev Immunol 4, 499-511.

Alter, H.J., Holland, P.V., Morrow, A.G., Purcell, R.H., Feinstone, S.M., Moritsugu, Y., 1975a. Clinical and serological analysis of transfusion-associated hepatitis. Lancet 2, 838-841.

Alter, H.J., Holland, P.V., Purcell, R.H., 1975b. The emerging pattern of post-transfusion hepatitis. Am J Med Sci 270, 329-334.

Anderson, J.A., Ping, L.H., Dibben, O., Jabara, C.B., Arney, L., Kincer, L., Tang, Y., Hobbs, M., Hoffman, I., Kazembe, P., Jones, C.D., Borrow, P., Fiscus, S., Cohen, M.S., Swanstrom, R., 2010. HIV-1 Populations in Semen Arise through Multiple Mechanisms. PLoS Pathog 6, e1001053.

Andre, P., Komurian-Pradel, F., Deforges, S., Perret, M., Berland, J.L., Sodoyer, M., Pol, S., Brechot, C., Paranhos-Baccala, G., Lotteau, V., 2002. Characterization of low- and very-low-density hepatitis C virus RNA-containing particles. Journal of virology 76, 6919-6928.

Angus, A.G., Dalrymple, D., Boulant, S., McGivern, D.R., Clayton, R.F., Scott, M.J., Adair, R., Graham, S., Owsianka, A.M., Targett-Adams, P., Li, K., Wakita, T., McLauchlan, J., Lemon, S.M., Patel, A.H., 2010. Requirement of cellular DDX3 for hepatitis C virus replication is unrelated to its interaction with the viral core protein. The Journal of general virology 91, 122-132.

Ariumi, Y., Kuroki, M., Abe, K., Dansako, H., Ikeda, M., Wakita, T., Kato, N., 2007. DDX3 DEAD-box RNA helicase is required for hepatitis C virus RNA replication. Journal of virology 81, 13922-13926.

Arnaud, N., Dabo, S., Akazawa, D., Fukasawa, M., Shinkai-Ouchi, F., Hugon, J., Wakita, T., Meurs, E.F., 2011. Hepatitis C virus reveals a novel early control in acute immune response. PLoS Pathog 7, e1002289.

Arshad, M., El-Kamary, S.S., Jhaveri, R., 2011. Hepatitis C virus infection during pregnancy and the newborn period--are they opportunities for treatment? Journal of viral hepatitis 18, 229-236.

Ashfaq, U.A., Javed, T., Rehman, S., Nawaz, Z., Riazuddin, S., 2011. An overview of HCV molecular biology, replication and immune responses. Virol J 8, 161.

Asselah, T., Marcellin, P., 2011. New direct-acting antivirals' combination for the treatment of chronic hepatitis C. Liver Int 31 Suppl 1, 68-77.

Baker, R.D., Dee, D., Baker, S.S., 2007. Response to pegylated interferon alpha-2b and ribavirin in children with chronic hepatitis C. J Clin Gastroenterol 41, 111-114.

Balfe, P., Simmonds, P., Ludlam, C.A., Bishop, J.O., Brown, A.J., 1990. Concurrent evolution of human immunodeficiency virus type 1 in patients infected from the same source: rate of sequence change and low frequency of inactivating mutations. Journal of virology 64, 6221-6233.

Ball, J.K., Tarr, A.W., McKeating, J.A., 2014. The past, present and future of neutralizing antibodies for hepatitis C virus. Antiviral research 105C, 100-111.

Ballardini, G., Manzin, A., Giostra, F., Francesconi, R., Groff, P., Grassi, A., Solforosi, L., Ghetti, S., Zauli, D., Clementi, M., Bianchi, F.B., 1997. Quantitative liver parameters of HCV infection: relation to HCV genotypes, viremia and response to interferon treatment. Journal of hepatology 26, 779-786.

Bancroft, J.D., Floyd, A.D., Suvarna, S.K., 2013. Bancroft's theory and practice of histological techniques, 7th ed. Churchill Livingstone Elsevier, [Oxford], p. 1 online resource.

Bankwitz, D., Steinmann, E., Bitzegeio, J., Ciesek, S., Friesland, M., Herrmann, E., Zeisel, M.B., Baumert, T.F., Keck, Z.Y., Foung, S.K., Pecheur, E.I., Pietschmann, T., 2010. Hepatitis C virus hypervariable region 1 modulates receptor interactions, conceals the CD81 binding site, and protects conserved neutralizing epitopes. Journal of virology 84, 5751-5763.

Barretto, N., Sainz, B., Jr., Hussain, S., Uprichard, S.L., 2014. Determining the involvement and therapeutic implications of host cellular factors in hepatitis C virus cell-to-cell spread. Journal of virology 88, 5050-5061.

Bartenschlager, R., Penin, F., Lohmann, V., Andre, P., 2011. Assembly of infectious hepatitis C virus particles. Trends Microbiol 19, 95-103.

Barth, H., Schafer, C., Adah, M.I., Zhang, F., Linhardt, R.J., Toyoda, H., Kinoshita-Toyoda, A., Toida, T., Van Kuppevelt, T.H., Depla, E., Von Weizsacker, F., Blum, H.E., Baumert, T.F., 2003. Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. The Journal of biological chemistry 278, 41003-41012.

Bartosch, B., Dubuisson, J., Cosset, F.L., 2003. Infectious hepatitis C virus pseudo-particles containing functional E1-E2 envelope protein complexes. J Exp Med 197, 633-642.

Becquart, P., Chomont, N., Roques, P., Ayouba, A., Kazatchkine, M.D., Belec, L., Hocini, H., 2002. Compartmentalization of HIV-1 between breast milk and blood of HIV-infected mothers. Virology 300, 109-117.

Beerenwinkel, N., Gunthard, H.F., Roth, V., Metzner, K.J., 2012. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. Front Microbiol 3, 329.

Beerli, P., Grassly, N.C., Kuhner, M.K., Nickle, D., Pybus, O., Rain, M., Rambaut, A., Rodrigo, A.G., Wang, Y., 2001. Population Genetics of HIV: Parameter estimation using genealogy-based methods. Kluwer Academia Publishers.

Behrens, S.E., Tomei, L., De Francesco, R., 1996. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. Embo J 15, 12-22.
Bellecave, P., Sarasin-Filipowicz, M., Donze, O., Kennel, A., Gouttenoire, J., Meylan, E., Terracciano, L., Tschopp, J., Sarrazin, C., Berg, T., Moradpour, D., Heim, M.H., 2010. Cleavage of mitochondrial antiviral signaling protein in the liver of patients with chronic hepatitis C correlates with a reduced activation of the endogenous interferon system. Hepatology 51, 1127-1136.
Benga, W.J., Krieger, S.E., Dimitrova, M., Zeisel, M.B., Parnot, M., Lupberger, J., Hildt, E., Luo, G., McLauchlan, J., Baumert, T.F., Schuster, C., 2010. Apolipoprotein E interacts with hepatitis C virus nonstructural protein 5A and determines assembly of infectious particles. Hepatology 51, 43-53.
Berg, T., Sarrazin, C., Herrmann, E., Hinrichsen, H., Gerlach, T., Zachoval, R., Wiedenmann, B., Hopf, U., Zeuzem, S., 2003. Prediction of treatment outcome in patients with chronic hepatitis C: significance of baseline parameters and viral dynamics during therapy. Hepatology 37, 600-609.
Bigger, C.B., Brasky, K.M., Lanford, R.E., 2001. DNA microarray analysis of chimpanzee liver during acute resolving hepatitis C virus infection. Journal of virology 75, 7059-7066.
Bigger, C.B., Guerra, B., Brasky, K.M., Hubbard, G., Beard, M.R., Luxon, B.A., Lemon, S.M., Lanford, R.E., 2004. Intrahepatic gene expression during chronic hepatitis C virus infection in chimpanzees. Journal of virology 78, 13779-13792.
Billerbeck, E., de Jong, Y., Dorner, M., de la Fuente, C., Ploss, A., 2013. Animal models for hepatitis C. Curr Top Microbiol Immunol 369, 49-86.
Blackard, J.T., Hiasa, Y., Smeaton, L., Jamieson, D.J., Rodriguez, I., Mayer, K.H., Chung, R.T., 2007. Compartmentalization of hepatitis C virus (HCV) during HCV/HIV coinfection. The Journal of infectious diseases 195, 1765-1773.
Blackard, J.T., Kemmer, N., Sherman, K.E., 2006. Extrahepatic replication of HCV: insights into clinical manifestations and biological consequences. Hepatology 44, 15-22.
Blackard, J.T., Ma, G., Welge, J.A., Martin, C.M., Sherman, K.E., Taylor, L.E., Mayer, K.H., Jamieson, D.J., 2012. Analysis of a non-structural gene reveals evidence of possible hepatitis C virus (HCV) compartmentalization. Journal of medical virology 84, 242-252.
Blanchard, E., Belouzard, S., Goueslain, L., Wakita, T., Dubuisson, J., Wychowski, C., Rouille, Y., 2006. Hepatitis C virus entry depends on clathrin-mediated endocytosis. Journal of virology 80, 6964-6972.
Blumberg, B.S., Alter, H.J., Visnich, S., 1965. A "New" Antigen in Leukemia Sera. Jama 191, 541-546.
Booth, J.C., Kumar, U., Webster, D., Monjardino, J., Thomas, H.C., 1998. Comparison of the rate of sequence variation in the hypervariable region of E2/NS1 region of hepatitis C virus in normal and hypogammaglobulinemic patients. Hepatology 27, 223-227.
Bornman, P.C., Beckingham, I.J., 2001. ABC of diseases of liver, pancreas, and biliary system. Pancreatic tumours. Bmj 322, 721-723.
Bortolotti, F., Verucchi, G., Camma, C., Cabibbo, G., Zancan, L., Indolfi, G., Giacchino, R., Marcellini, M., Marazzi, M.G., Barbera, C., Maggiore, G.,

Vajro, P., Bartolacci, S., Balli, F., Maccabruni, A., Guido, M., 2008. Long-term course of chronic hepatitis C in children: from viral clearance to end-stage liver disease. Gastroenterology 134, 1900-1907.
Bostan, N., Mahmood, T., 2010. An overview about hepatitis C: a devastating virus. Crit Rev Microbiol 36, 91-133.
Bowen, D.G., Walker, C.M., 2005a. Adaptive immune responses in acute and chronic hepatitis C virus infection. Nature 436, 946-952.
Bowen, D.G., Walker, C.M., 2005b. Mutational escape from CD8+ T cell immunity: HCV evolution, from chimpanzees to man. J Exp Med 201, 1709-1714.
Boxall, E., Baumann, K., Price, N., Sira, J., Brown, M., Kelly, D., 2007. Discordant outcome of perinatal transmission of hepatitis C in twin pregnancies. J Clin Virol 38, 91-95.
Bradley, D., McCaustland, K., Krawczynski, K., Spelbring, J., Humphrey, C., Cook, E.H., 1991. Hepatitis C virus: buoyant density of the factor VIII-derived isolate in sucrose. Journal of medical virology 34, 206-208.
Bradley, D.W., Maynard, J.E., 1986. Etiology and natural history of post-transfusion and enterically-transmitted non-A, non-B hepatitis. Semin Liver Dis 6, 56-66.
Brimacombe, C.L., Grove, J., Meredith, L.W., Hu, K., Syder, A.J., Flores, M.V., Timpe, J.M., Krieger, S.E., Baumert, T.F., Tellinghuisen, T.L., Wong-Staal, F., Balfe, P., McKeating, J.A., 2011. Neutralizing antibody-resistant hepatitis C virus cell-to-cell transmission. J Virol 85, 596-605.
Broering, R., Zhang, X., Kottilil, S., Trippler, M., Jiang, M., Lu, M., Gerken, G., Schlaak, J.F., 2010. The interferon stimulated gene 15 functions as a proviral factor for the hepatitis C virus and as a regulator of the IFN response. Gut 59, 1111-1119.
Brown, R.J., Hudson, N., Wilson, G., Rehman, S.U., Jabbari, S., Hu, K., Tarr, A.W., Borrow, P., Joyce, M., Lewis, J., Zhu, L.F., Law, M., Kneteman, N., Tyrrell, D.L., McKeating, J.A., Ball, J.K., 2012. Hepatitis C virus envelope glycoprotein fitness defines virus population composition following transmission to a new host. Journal of virology 86, 11956-11966.
Brown, R.J., Juttla, V.S., Tarr, A.W., Finnis, R., Irving, W.L., Hemsley, S., Flower, D.R., Borrow, P., Ball, J.K., 2005. Evolutionary dynamics of hepatitis C virus envelope genes during chronic infection. The Journal of general virology 86, 1931-1942.
Brown, R.S., 2005. Hepatitis C and liver transplantation. Nature 436, 973-978.
Bruggmann, 2014. Historical epidemiology of hepatitis C virus (HCV) in selective countries. J Viral Hepat.
Bukh, J., 2012. Animal models for the study of hepatitis C virus infection and related liver disease. Gastroenterology 142, 1279-1287 e1273.
Bull, R.A., Luciani, F., McElroy, K., Gaudieri, S., Pham, S.T., Chopra, A., Cameron, B., Maher, L., Dore, G.J., White, P.A., Lloyd, A.R., 2011. Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. PLoS Pathog 7, e1002243.
Bungyoku, Y., Shoji, I., Makine, T., Adachi, T., Hayashida, K., Nagano-Fujii, M., Ide, Y.H., Deng, L., Hotta, H., 2009. Efficient production of infectious hepatitis C virus with adaptive mutations in cultured hepatoma cells. The Journal of general virology 90, 1681-1691.

Burbelo, P.D., Dubovi, E.J., Simmonds, P., Medina, J.L., Henriquez, J.A., Mishra, N., Wagner, J., Tokarz, R., Cullen, J.M., Iadarola, M.J., Rice, C.M., Lipkin, W.I., Kapoor, A., 2012. Serology-enabled discovery of genetically diverse hepaciviruses in a new host. Journal of virology 86, 6171-6178.

Busch, M.P., 2001. Insights into the epidemiology, natural history and pathogenesis of hepatitis C virus infection from studies of infected donors and blood product recipients. Transfus Clin Biol 8, 200-206.

Cabanski, C.R., Cavin, K., Bizon, C., Wilkerson, M.D., Parker, J.S., Wilhelmsen, K.C., Perou, C.M., Marron, J.S., Hayes, D.N., 2012. ReQON: a Bioconductor package for recalibrating quality scores from next-generation sequencing data. BMC Bioinformatics 13, 221.

Cabot, B., Esteban, J.I., Martell, M., Genesca, J., Vargas, V., Esteban, R., Guardia, J., Gomez, J., 1997. Structure of replicating hepatitis C virus (HCV) quasispecies in the liver may not be reflected by analysis of circulating HCV virions. Journal of virology 71, 1732-1734.

Cabot, B., Martell, M., Esteban, J.I., Sauleda, S., Otero, T., Esteban, R., Guardia, J., Gomez, J., 2000. Nucleotide and amino acid complexity of hepatitis C virus quasispecies in serum and liver. Journal of virology 74, 805-811.

Canobio, S., Guilbert, C.M., Troesch, M., Samson, J., Lemay, M., Pelletier, V.A., Bernard-Bonnin, A.C., Kozielski, R., Lapointe, N., Martin, S.R., Soudeyns, H., 2004. Differing patterns of liver disease progression and hepatitis C virus (HCV) quasispecies evolution in children vertically coinfected with HCV and human immunodeficiency virus type 1. J Clin Microbiol 42, 4365-4369.

Caraballo Cortes, K., Laskus, T., Bukowska-Osko, I., Pawelczyk, A., Berak, H., Horban, A., Fic, M., Radkowski, M., 2012. Variability of hepatitis C virus hypervariable region 1 (HVR-1) during the early phase of pegylated interferon and ribavirin therapy. Adv Med Sci 57, 370-374.

Caraballo Cortes, K., Zagordi, O., Laskus, T., Ploski, R., Bukowska-Osko, I., Pawelczyk, A., Berak, H., Radkowski, M., 2013. Ultradeep pyrosequencing of hepatitis C virus hypervariable region 1 in quasispecies analysis. Biomed Res Int 2013, 626083.

Casino, C., McAllister, J., Davidson, F., Power, J., Lawlor, E., Yap, P.L., Simmonds, P., Smith, D.B., 1999. Variation of hepatitis C virus following serial transmission: multiple mechanisms of diversification of the hypervariable region and evidence for convergent genome evolution. The Journal of general virology 80 (Pt 3), 717-725.

Catanese, M.T., Loureiro, J., Jones, C.T., Dorner, M., von Hahn, T., Rice, C.M., 2013. Different requirements for SR-BI in hepatitis C virus cell-free versus cell-to-cell transmission. Journal of virology.

Chang, M., Williams, O., Mittler, J., Quintanilla, A., Carithers, R.L., Jr., Perkins, J., Corey, L., Gretch, D.R., 2003. Dynamics of hepatitis C virus replication in human liver. The American journal of pathology 163, 433-444.

Charlton, M., Ruppert, K., Belle, S.H., Bass, N., Schafer, D., Wiesner, R.H., Detre, K., Wei, Y., Everhart, J., 2004. Long-term results and modeling to predict outcomes in recipients with HCV infection: results of the NIDDK liver transplantation database. Liver Transpl 10, 1120-1130.

Chen, L., Borozan, I., Feld, J., Sun, J., Tannis, L.L., Coltescu, C., Heathcote, J., Edwards, A.M., McGilvray, I.D., 2005. Hepatic gene expression discriminates responders and nonresponders in treatment of chronic hepatitis C viral infection. Gastroenterology 128, 1437-1444.

Chen, L., Borozan, I., Sun, J., Guindi, M., Fischer, S., Feld, J., Anand, N., Heathcote, J., Edwards, A.M., McGilvray, I.D., 2010. Cell-type specific gene expression signature in liver underlies response to interferon therapy in chronic hepatitis C infection. Gastroenterology 138, 1123-1133 e1121-1123.

Cheynier, R., Langlade-Demoyen, P., Marescot, M.R., Blanche, S., Blondin, G., Wain-Hobson, S., Griscelli, C., Vilmer, E., Plata, F., 1992. Cytotoxic T lymphocyte responses in the peripheral blood of children born to human immunodeficiency virus-1-infected mothers. Eur J Immunol 22, 2211-2217.

Chisari, F.V., 2005. Unscrambling hepatitis C virus-host interactions. Nature 436, 930-932.

Chmielewska, A.M., Naddeo, M., Capone, S., Ammendola, V., Hu, K., Meredith, L., Verhoye, L., Rychlowska, M., Rappuoli, R., Ulmer, J.B., Colloca, S., Nicosia, A., Cortese, R., Leroux-Roels, G., Balfe, P., Bienkowska-Szewczyk, K., Meuleman, P., McKeating, J.A., Folgori, A., 2014. Combined adenovirus vector and hepatitis C virus envelope protein prime-boost regimen elicits T cell and neutralizing antibody immune responses. Journal of virology 88, 5502-5510.

Choo, Q.L., Kuo, G., Weiner, A.J., Overby, L.R., Bradley, D.W., Houghton, M., 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 244, 359-362.

Choo, Q.L., Richman, K.H., Han, J.H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, R., Barr, P.J., et al., 1991. Genetic organization and diversity of the hepatitis C virus. Proceedings of the National Academy of Sciences of the United States of America 88, 2451-2455.

Christie, J.M., Healey, C.J., Watson, J., Wong, V.S., Duddridge, M., Snowden, N., Rosenberg, W.M., Fleming, K.A., Chapel, H., Chapman, R.W., 1997. Clinical outcome of hypogammaglobulinaemic patients following outbreak of acute hepatitis C: 2 year follow up. Clinical and experimental immunology 110, 4-8.

Chua, P.K., McCown, M.F., Rajyaguru, S., Kular, S., Varma, R., Symons, J., Chiu, S.S., Cammack, N., Najera, I., 2009. Modulation of alpha interferon anti-hepatitis C virus activity by ISG15. The Journal of general virology 90, 2929-2939.

Ciesek, S., Westhaus, S., Wicht, M., Wappler, I., Henschen, S., Sarrazin, C., Hamdi, N., Abdelaziz, A.I., Strassburg, C.P., Wedemeyer, H., Manns, M.P., Pietschmann, T., von Hahn, T., 2011. Impact of intra- and interspecies variation of occludin on its function as coreceptor for authentic hepatitis C virus particles. Journal of virology 85, 7613-7621.

Cocquerel, L., Wychowski, C., Minner, F., Penin, F., Dubuisson, J., 2000. Charged residues in the transmembrane domains of hepatitis C virus glycoproteins play a major role in the processing, subcellular localization, and assembly of these envelope proteins. Journal of virology 74, 3623-3633.

Coller, K.E., Heaton, N.S., Berger, K.L., Cooper, J.D., Saunders, J.L., Randall, G., 2012. Molecular determinants and dynamics of hepatitis C virus secretion. PLoS Pathog 8, e1002466.
Cosset, F.L., Dreux, M., 2014. HCV transmission by hepatic exosomes establishes a productive infection. Journal of hepatology 60, 674-675.
Cox, A.L., Mosbruger, T., Lauer, G.M., Pardoll, D., Thomas, D.L., Ray, S.C., 2005. Comprehensive analyses of CD8+ T cell responses during longitudinal study of acute human hepatitis C. Hepatology 42, 104-112.
Cukierman, L., Meertens, L., Bertaux, C., Kajumo, F., Dragic, T., 2009. Residues in a highly conserved claudin-1 motif are required for hepatitis C virus entry and mediate the formation of cell-cell contacts. Journal of virology 83, 5477-5484.
Curran, R., Jameson, C.L., Craggs, J.K., Grabowska, A.M., Thomson, B.J., Robins, A., Irving, W.L., Ball, J.K., 2002. Evolutionary trends of the first hypervariable region of the hepatitis C virus E2 protein in individuals with differing liver disease severity. The Journal of general virology 83, 11-23.
Dane, D.S., Cameron, C.H., Briggs, M., 1970. Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. Lancet 1, 695-698.
Davis, G.L., Wong, J.B., McHutchison, J.G., Manns, M.P., Harvey, J., Albrecht, J., 2003. Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in patients with chronic hepatitis C. Hepatology 38, 645-652.
Day, C.L., Lauer, G.M., Robbins, G.K., McGovern, B., Wurcel, A.G., Gandhi, R.T., Chung, R.T., Walker, B.D., 2002. Broad specificity of virus-specific CD4+ T-helper-cell responses in resolved hepatitis C virus infection. Journal of virology 76, 12584-12595.
Delassus, S., Cheynier, R., Wain-Hobson, S., 1992. Nonhomogeneous distribution of human immunodeficiency virus type 1 proviruses in the spleen. Journal of virology 66, 5642-5645.
Descamps, V., Op de Beeck, A., Plassart, C., Brochot, E., Francois, C., Helle, F., Adler, M., Bourgeois, N., Degre, D., Duverlie, G., Castelain, S., 2012. Strong correlation between liver and serum levels of hepatitis C virus core antigen and RNA in chronically infected patients. J Clin Microbiol 50, 465-468.
Di Bisceglie, A.M., Hoofnagle, J.H., 2002. Optimal therapy of hepatitis C. Hepatology 36, S121-127.
Dill, M.T., Duong, F.H., Vogt, J.E., Bibert, S., Bochud, P.Y., Terracciano, L., Papassotiropoulos, A., Roth, V., Heim, M.H., 2011. Interferon-induced gene expression is a stronger predictor of treatment response than IL28B genotype in patients with hepatitis C. Gastroenterology 140, 1021-1031.
Dill, M.T., Makowska, Z., Duong, F.H., Merkofer, F., Filipowicz, M., Baumert, T.F., Tornillo, L., Terracciano, L., Heim, M.H., 2012. Interferon-gamma-stimulated genes, but not USP18, are expressed in livers of patients with acute hepatitis C. Gastroenterology 143, 777-786 e771-776.
Domagalski, K., Pawlowska, M., Tretyn, A., Halota, W., Pilarczyk, M., Smukalska, E., Linkowska, K., Grzybowski, T., 2013. Impact of IL-28B polymorphisms on pegylated interferon plus ribavirin treatment

response in children and adolescents infected with HCV genotypes 1 and 4. Eur J Clin Microbiol Infect Dis 32, 745-754.
Dooley, J.S., Lok, A., Burroughs, A., Heathcote, J., 2011. Sherlock's Diseases of the Liver and Biliary System, 12th ed. Wiley, Hoboken, p. 1 online resource (2907 p.).
Dowd, K.A., Netski, D.M., Wang, X.H., Cox, A.L., Ray, S.C., 2009. Selection pressure from neutralizing antibodies drives sequence evolution during acute infection with hepatitis C virus. Gastroenterology 136, 2377-2386.
Dreux, M., Garaigorta, U., Boyd, B., Decembre, E., Chung, J., Whitten-Bauer, C., Wieland, S., Chisari, F.V., 2012. Short-range exosomal transfer of viral RNA from infected cells to plasmacytoid dendritic cells triggers innate immunity. Cell Host Microbe 12, 558-570.
Drexler, J.F., Corman, V.M., Muller, M.A., Lukashev, A.N., Gmyl, A., Coutard, B., Adam, A., Ritz, D., Leijten, L.M., van Riel, D., Kallies, R., Klose, S.M., Gloza-Rausch, F., Binger, T., Annan, A., Adu-Sarkodie, Y., Oppong, S., Bourgarel, M., Rupp, D., Hoffmann, B., Schlegel, M., Kummerer, B.M., Kruger, D.H., Schmidt-Chanasit, J., Setien, A.A., Cottontail, V.M., Hemachudha, T., Wacharapluesadee, S., Osterrieder, K., Bartenschlager, R., Matthee, S., Beer, M., Kuiken, T., Reusken, C., Leroy, E.M., Ulrich, R.G., Drosten, C., 2013. Evidence for novel hepaciviruses in rodents. PLoS Pathog 9, e1003438.
Drummer, H.E., Boo, I., Maerz, A.L., Poumbourios, P., 2006. A conserved Gly436-Trp-Leu-Ala-Gly-Leu-Phe-Tyr motif in hepatitis C virus glycoprotein E2 is a determinant of CD81 binding and viral entry. Journal of virology 80, 7844-7853.
Drummer, H.E., Maerz, A., Poumbourios, P., 2003. Cell surface expression of functional hepatitis C virus E1 and E2 glycoproteins. FEBS Lett 546, 385-390.
Dubuisson, J., 2007. Hepatitis C virus proteins. World J Gastroenterol 13, 2406-2415.
Duffy, M., Salemi, M., Sheehy, N., Vandamme, A.M., Hegarty, J., Curry, M., Nolan, N., Kelleher, D., McKiernan, S., Hall, W.W., 2002. Comparative rates of nucleotide sequence variation in the hypervariable region of E1/E2 and the NS5b region of hepatitis C virus in patients with a spectrum of liver disease resulting from a common source of infection. Virology 301, 354-364.
Eigen, M., 1971. Molecular self-organization and the early stages of evolution. Q Rev Biophys 4, 149-212.
El-Shamy, A., Shindo, M., Shoji, I., Deng, L., Okuno, T., Hotta, H., 2013. Polymorphisms of the core, NS3, and NS5A proteins of hepatitis C virus genotype 1b associate with development of hepatocellular carcinoma. Hepatology 58, 555-563.
Elazar, M., Liu, P., Rice, C.M., Glenn, J.S., 2004. An N-terminal amphipathic helix in hepatitis C virus (HCV) NS4B mediates membrane association, correct localization of replication complex proteins, and HCV RNA replication. Journal of virology 78, 11393-11400.
Evans, M.J., von Hahn, T., Tscherne, D.M., Syder, A.J., Panis, M., Wolk, B., Hatziioannou, T., McKeating, J.A., Bieniasz, P.D., Rice, C.M., 2007. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. Nature 446, 801-805.

Fafi-Kremer, S., Fofana, I., Soulier, E., Carolla, P., Meuleman, P., Leroux-Roels, G., Patel, A.H., Cosset, F.L., Pessaux, P., Doffoel, M., Wolf, P., Stoll-Keller, F., Baumert, T.F., 2010. Viral entry and escape from antibody-mediated neutralization influence hepatitis C virus reinfection in liver transplantation. J Exp Med 207, 2019-2031.

Fan, X., Solomon, H., Poulos, J.E., Neuschwander-Tetri, B.A., Di Bisceglie, A.M., 1999. Comparison of genetic heterogeneity of hepatitis C viral RNA in liver tissue and serum. Am J Gastroenterol 94, 1347-1354.

Fanning, L., Loane, J., Kenny-Walsh, E., Sheehan, M., Whelton, M., Kirwan, W., Collins, J.K., Shanahan, F., 2001. Tissue viral load variability in chronic hepatitis C. Am J Gastroenterol 96, 3384-3389.

Farci, P., 2011. New insights into the HCV quasispecies and compartmentalization. Semin Liver Dis 31, 356-374.

Farci, P., Alter, H.J., Wong, D., Miller, R.H., Shih, J.W., Jett, B., Purcell, R.H., 1991. A long-term study of hepatitis C virus replication in non-A, non-B hepatitis. N Engl J Med 325, 98-104.

Farci, P., Shimoda, A., Coiana, A., Diaz, G., Peddis, G., Melpolder, J.C., Strazzera, A., Chien, D.Y., Munoz, S.J., Balestrieri, A., Purcell, R.H., Alter, H.J., 2000. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. Science 288, 339-344.

Farci, P., Shimoda, A., Wong, D., Cabezon, T., De Gioannis, D., Strazzera, A., Shimizu, Y., Shapiro, M., Alter, H.J., Purcell, R.H., 1996. Prevention of hepatitis C virus infection in chimpanzees by hyperimmune serum against the hypervariable region 1 of the envelope 2 protein. Proceedings of the National Academy of Sciences of the United States of America 93, 15394-15399.

Farquhar, M.J., Hu, K., Harris, H.J., Davis, C., Brimacombe, C.L., Fletcher, S.J., Baumert, T.F., Rappoport, J.Z., Balfe, P., McKeating, J.A., 2012. Hepatitis C virus induces CD81 and claudin-1 endocytosis. Journal of virology 86, 4305-4316.

Fauvelle, C., Felmlee, D.J., Baumert, T.F., 2014. Unraveling hepatitis C virus structure. Cell Res 24, 385-386.

Feinstone, S.M., Kapikian, A.Z., Purcell, R.H., Alter, H.J., Holland, P.V., 1975. Transfusion-associated hepatitis not due to viral hepatitis type A or B. N Engl J Med 292, 767-770.

Fishman, S.L., Branch, A.D., 2009. The quasispecies nature and biological implications of the hepatitis C virus. Infect Genet Evol 9, 1158-1167.

Fitzgerald, K.A., 2011. The interferon inducible gene: Viperin. J Interferon Cytokine Res 31, 131-135.

Flint, M., Logvinoff, C., Rice, C.M., McKeating, J.A., 2004. Characterization of infectious retroviral pseudotype particles bearing hepatitis C virus glycoproteins. J Virol 78, 6875-6882.

Flint, M., Maidens, C., Loomis-Price, L.D., Shotton, C., Dubuisson, J., Monk, P., Higginbottom, A., Levy, S., McKeating, J.A., 1999. Characterization of hepatitis C virus E2 glycoprotein interaction with a putative cellular receptor, CD81. J Virol 73, 6235-6244.

Fonseca-Coronado, S., Escobar-Gutierrez, A., Ruiz-Tovar, K., Cruz-Rivera, M.Y., Rivera-Osorio, P., Vazquez-Pichardo, M., Carpio-Pedroza, J.C., Ruiz-Pacheco, J.A., Cazares, F., Vaughan, G., 2012. Specific detection of naturally occurring hepatitis C virus mutants with resistance to

telaprevir and boceprevir (protease inhibitors) among treatment-naive infected individuals. J Clin Microbiol 50, 281-287.

Forns, X., Bukh, J., Purcell, R.H., Emerson, S.U., 1997. How Escherichia coli can bias the results of molecular cloning: preferential selection of defective genomes of hepatitis C virus during the cloning procedure. Proceedings of the National Academy of Sciences of the United States of America 94, 13909-13914.

Forns, X., Purcell, R.H., Bukh, J., 1999. Quasispecies in viral persistence and pathogenesis of hepatitis C virus. Trends Microbiol 7, 402-410.

Forton, D.M., Karayiannis, P., Mahmud, N., Taylor-Robinson, S.D., Thomas, H.C., 2004. Identification of unique hepatitis C virus quasispecies in the central nervous system and comparative analysis of internal translational efficiency of brain, liver, and serum variants. Journal of virology 78, 5170-5183.

Foster, T.L., Gallay, P., Stonehouse, N.J., Harris, M., 2011. Cyclophilin A interacts with domain II of hepatitis C virus NS5A and stimulates RNA binding in an isomerase-dependent manner. Journal of virology 85, 7460-7464.

Fournillier, A., Wychowski, C., Boucreux, D., Baumert, T.F., Meunier, J.C., Jacobs, D., Muguet, S., Depla, E., Inchauspe, G., 2001. Induction of hepatitis C virus E1 envelope protein-specific immune response can be enhanced by mutation of N-glycosylation sites. Journal of virology 75, 12088-12097.

Foy, E., Li, K., Wang, C., Sumpter, R., Jr., Ikeda, M., Lemon, S.M., Gale, M., Jr., 2003. Regulation of interferon regulatory factor-3 by the hepatitis C virus serine protease. Science 300, 1145-1148.

Friebe, P., Bartenschlager, R., 2002. Genetic analysis of sequences in the 3' nontranslated region of hepatitis C virus that are important for RNA replication. Journal of virology 76, 5326-5338.

Friebe, P., Boudet, J., Simorre, J.P., Bartenschlager, R., 2005. Kissing-loop interaction in the 3' end of the hepatitis C virus genome essential for RNA replication. Journal of virology 79, 380-392.

Fried, M.W., Buti, M., Dore, G.J., Flisiak, R., Ferenci, P., Jacobson, I., Marcellin, P., Manns, M., Nikitin, I., Poordad, F., Sherman, M., Zeuzem, S., Scott, J., Gilles, L., Lenz, O., Peeters, M., Sekar, V., De Smedt, G., Beumont-Mauviel, M., 2013. Once-daily simeprevir (TMC435) with pegylated interferon and ribavirin in treatment-naive genotype 1 hepatitis C: the randomized PILLAR study. Hepatology 58, 1918-1929.

Fried, M.W., Shiffman, M.L., Reddy, K.R., Smith, C., Marinos, G., Goncales, F.L., Jr., Haussinger, D., Diago, M., Carosi, G., Dhumeaux, D., Craxi, A., Lin, A., Hoffman, J., Yu, J., 2002. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 347, 975-982.

Frost, S.D., Dumaurier, M.J., Wain-Hobson, S., Brown, A.J., 2001. Genetic drift and within-host metapopulation dynamics of HIV-1 infection. Proceedings of the National Academy of Sciences of the United States of America 98, 6975-6980.

Gastaminza, P., Cheng, G., Wieland, S., Zhong, J., Liao, W., Chisari, F.V., 2008. Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. Journal of virology 82, 2120-2129.

Ge, D., Fellay, J., Thompson, A.J., Simon, J.S., Shianna, K.V., Urban, T.J., Heinzen, E.L., Qiu, P., Bertelsen, A.H., Muir, A.J., Sulkowski, M., McHutchison, J.G., Goldstein, D.B., 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 461, 399-401.

Ghany, M.G., Strader, D.B., Thomas, D.L., Seeff, L.B., 2009. Diagnosis, management, and treatment of hepatitis C: an update. Hepatology 49, 1335-1374.

Gibb, D.M., Goodall, R.L., Dunn, D.T., Healy, M., Neave, P., Cafferkey, M., Butler, K., 2000. Mother-to-child transmission of hepatitis C virus: evidence for preventable peripartum transmission. Lancet 356, 904-907.

Glynn, S.A., Wright, D.J., Kleinman, S.H., Hirschkorn, D., Tu, Y., Heldebrant, C., Smith, R., Giachetti, C., Gallarda, J., Busch, M.P., 2005. Dynamics of viremia in early hepatitis C virus infection. Transfusion 45, 994-1002.

Goffard, A., Callens, N., Bartosch, B., Wychowski, C., Cosset, F.L., Montpellier, C., Dubuisson, J., 2005. Role of N-linked glycans in the functions of hepatitis C virus envelope glycoproteins. Journal of virology 79, 8400-8409.

Gomatos, P.J., Tamm, I., 1963a. Animal and Plant Viruses with Double-Helical Rna. Proceedings of the National Academy of Sciences of the United States of America 50, 878-885.

Gomatos, P.J., Tamm, I., 1963b. The Secondary Structure of Reovirus Rna. Proceedings of the National Academy of Sciences of the United States of America 49, 707-714.

Gomez, J., Martell, M., Quer, J., Cabot, B., Esteban, J.I., 1999. Hepatitis C viral quasispecies. Journal of viral hepatitis 6, 3-16.

Gottwein, J.M., Scheel, T.K., Hoegh, A.M., Lademann, J.B., Eugen-Olsen, J., Lisby, G., Bukh, J., 2007. Robust hepatitis C genotype 3a cell culture releasing adapted intergenotypic 3a/2a (S52/JFH1) viruses. Gastroenterology 133, 1614-1626.

Grakoui, A., Shoukry, N.H., Woollard, D.J., Han, J.H., Hanson, H.L., Ghrayeb, J., Murthy, K.K., Rice, C.M., Walker, C.M., 2003. HCV persistence and immune evasion in the absence of memory T cell help. Science 302, 659-662.

Gratton, S., Cheynier, R., Dumaurier, M.J., Oksenhendler, E., Wain-Hobson, S., 2000. Highly restricted spread of HIV-1 and multiply infected cells within splenic germinal centers. Proceedings of the National Academy of Sciences of the United States of America 97, 14566-14571.

Gray, R.R., Parker, J., Lemey, P., Salemi, M., Katzourakis, A., Pybus, O.G., 2011. The mode and tempo of hepatitis C virus evolution within and among hosts. BMC Evol Biol 11, 131.

Gregori, J., Salicru, M., Domingo, E., Sanchez, A., Rodriguez-Frias, F., Quer, J., 2014. Inference with viral quasispecies diversity indices: Clonal and NGS approaches. Bioinformatics.

Grove, J., Huby, T., Stamataki, Z., Vanwolleghem, T., Meuleman, P., Farquhar, M., Schwarz, A., Moreau, M., Owen, J.S., Leroux-Roels, G., Balfe, P., McKeating, J.A., 2007. Scavenger receptor BI and BII expression levels modulate hepatitis C virus infectivity. J Virol 81, 3162-3169.

Grove, J., Nielsen, S., Zhong, J., Bassendine, M.F., Drummer, H.E., Balfe, P., McKeating, J.A., 2008. Identification of a residue in hepatitis C virus E2

glycoprotein that determines scavenger receptor BI and CD81 receptor dependency and sensitivity to neutralizing antibodies. J Virol 82, 12020-12029.
Gruner, N.H., Gerlach, T.J., Jung, M.C., Diepolder, H.M., Schirren, C.A., Schraut, W.W., Hoffmann, R., Zachoval, R., Santantonio, T., Cucchiarini, M., Cerny, A., Pape, G.R., 2000. Association of hepatitis C virus-specific CD8+ T cells with viral clearance in acute hepatitis C. The Journal of infectious diseases 181, 1528-1536.
Gu, M., Rice, C.M., 2013. Structures of hepatitis C virus nonstructural proteins required for replicase assembly and function. Curr Opin Virol 3, 129-136.
Guedj, J., Dahari, H., Rong, L., Sansone, N.D., Nettles, R.E., Cotler, S.J., Layden, T.J., Uprichard, S.L., Perelson, A.S., 2013. Modeling shows that the NS5A inhibitor daclatasvir has two modes of action and yields a shorter estimate of the hepatitis C virus half-life. Proceedings of the National Academy of Sciences of the United States of America 110, 3991-3996.
Gupta, A.C., Trehanpati, N., Sukriti, S., Hissar, S., Midha, V., Sood, A., Sarin, S.K., 2014. Interleukin-28b CC genotype predicts early treatment response and CT/TT genotypes predicts non-response in patients infected with HCV genotype 3. Journal of medical virology.
Harris, H.J., Davis, C., Mullins, J.G., Hu, K., Goodall, M., Farquhar, M.J., Mee, C.J., McCaffrey, K., Young, S., Drummer, H., Balfe, P., McKeating, J.A., 2010. Claudin association with CD81 defines hepatitis C virus entry. J Biol Chem 285, 21092-21102.
Harris, H.J., Farquhar, M.J., Mee, C.J., Davis, C., Reynolds, G.M., Jennings, A., Hu, K., Yuan, F., Deng, H., Hubscher, S.G., Han, J.H., Balfe, P., McKeating, J.A., 2008. CD81 and claudin 1 coreceptor association: role in hepatitis C virus entry. J Virol 82, 5007-5020.
Hayashi, J., Kishihara, Y., Yamaji, K., Furusyo, N., Yamamoto, T., Pae, Y., Etoh, Y., Ikematsu, H., Kashiwagi, S., 1997. Hepatitis C viral quasispecies and liver damage in patients with chronic hepatitis C virus infection. Hepatology 25, 697-701.
Heim, M.H., 2013a. 25 years of interferon-based treatment of chronic hepatitis C: an epoch coming to an end. Nat Rev Immunol 13, 535-542.
Heim, M.H., 2013b. Innate immunity and HCV. Journal of hepatology 58, 564-574.
Helbig, K.J., Eyre, N.S., Yip, E., Narayana, S., Li, K., Fiches, G., McCartney, E.M., Jangra, R.K., Lemon, S.M., Beard, M.R., 2011. The antiviral protein viperin inhibits hepatitis C virus replication via interaction with nonstructural protein 5A. Hepatology 54, 1506-1517.
Helle, F., Duverlie, G., Dubuisson, J., 2011. The hepatitis C virus glycan shield and evasion of the humoral immune response. Viruses 3, 1909-1932.
Helle, F., Goffard, A., Morel, V., Duverlie, G., McKeating, J., Keck, Z.Y., Foung, S., Penin, F., Dubuisson, J., Voisset, C., 2007. The neutralizing activity of anti-hepatitis C virus antibodies is modulated by specific glycans on the E2 envelope protein. J Virol 81, 8101-8111.
Helle, F., Vieyres, G., Elkrief, L., Popescu, C.I., Wychowski, C., Descamps, V., Castelain, S., Roingeard, P., Duverlie, G., Dubuisson, J., 2010. Role of N-linked glycans in the functions of hepatitis C virus envelope proteins incorporated into infectious virions. Journal of virology 84, 11905-11915.

Hemler, M.E., 2005. Tetraspanin functions and associated microdomains. Nat Rev Mol Cell Biol 6, 801-811.

Henn, M.R., Boutwell, C.L., Charlebois, P., Lennon, N.J., Power, K.A., Macalalad, A.R., Berlin, A.M., Malboeuf, C.M., Ryan, E.M., Gnerre, S., Zody, M.C., Erlich, R.L., Green, L.M., Berical, A., Wang, Y., Casali, M., Streeck, H., Bloom, A.K., Dudek, T., Tully, D., Newman, R., Axten, K.L., Gladden, A.D., Battis, L., Kemper, M., Zeng, Q., Shea, T.P., Gujja, S., Zedlack, C., Gasser, O., Brander, C., Hess, C., Gunthard, H.F., Brumme, Z.L., Brumme, C.J., Bazner, S., Rychert, J., Tinsley, J.P., Mayer, K.H., Rosenberg, E., Pereyra, F., Levin, J.Z., Young, S.K., Jessen, H., Altfeld, M., Birren, B.W., Walker, B.D., Allen, T.M., 2012. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. PLoS Pathog 8, e1002529.

Hijikata, M., Shimizu, Y.K., Kato, H., Iwamoto, A., Shih, J.W., Alter, H.J., Purcell, R.H., Yoshikura, H., 1993. Equilibrium centrifugation studies of hepatitis C virus: evidence for circulating immune complexes. Journal of virology 67, 1953-1958.

Hinson, E.R., Cresswell, P., 2009. The antiviral protein, viperin, localizes to lipid droplets via its N-terminal amphipathic alpha-helix. Proceedings of the National Academy of Sciences of the United States of America 106, 20452-20457.

Hishiki, T., Shimizu, Y., Tobita, R., Sugiyama, K., Ogawa, K., Funami, K., Ohsaki, Y., Fujimoto, T., Takaku, H., Wakita, T., Baumert, T.F., Miyanari, Y., Shimotohno, K., 2010. Infectivity of hepatitis C virus is influenced by association with apolipoprotein E isoforms. Journal of virology 84, 12048-12057.

Hoffmann, C., Minkah, N., Leipzig, J., Wang, G., Arens, M.Q., Tebas, P., Bushman, F.D., 2007. DNA bar coding and pyrosequencing to identify rare HIV drug resistance mutations. Nucleic Acids Res 35, e91.

Holmes, E.C., 2010a. Does hepatitis C virus really form quasispecies? Infect Genet Evol 10, 431-432.

Holmes, E.C., 2010b. The RNA virus quasispecies: fact or fiction? J Mol Biol 400, 271-273.

Holmes, E.C., Grenfell, B.T., 2009. Discovering the phylodynamics of RNA viruses. PLoS Comput Biol 5, e1000505.

Honda, M., Kaneko, S., Sakai, A., Unoura, M., Murakami, S., Kobayashi, K., 1994. Degree of diversity of hepatitis C virus quasispecies and progression of liver disease. Hepatology 20, 1144-1151.

Honda, M., Sakai, A., Yamashita, T., Nakamoto, Y., Mizukoshi, E., Sakai, Y., Nakamura, M., Shirasaki, T., Horimoto, K., Tanaka, Y., Tokunaga, K., Mizokami, M., Kaneko, S., 2010. Hepatic ISG expression is associated with genetic variation in interleukin 28B and the outcome of IFN therapy for chronic hepatitis C. Gastroenterology 139, 499-509.

Hoofnagle, J.H., 2002. Course and outcome of hepatitis C. Hepatology 36, S21-29.

Hsu, M., Zhang, J., Flint, M., Logvinoff, C., Cheng-Mayer, C., Rice, C.M., McKeating, J.A., 2003a. Hepatitis C virus glycoproteins mediate pH-dependent cell entry of pseudotyped retroviral particles. Proceedings of

the National Academy of Sciences of the United States of America 100, 7271-7276.
Hsu, M., Zhang, J., Flint, M., Logvinoff, C., Cheng-Mayer, C., Rice, C.M., McKeating, J.A., 2003b. Hepatitis C virus glycoproteins mediate pH-dependent cell entry of pseudotyped retroviral particles. Proc Natl Acad Sci U S A 100, 7271-7276.
Huang, L., Hwang, J., Sharma, S.D., Hargittai, M.R., Chen, Y., Arnold, J.J., Raney, K.D., Cameron, C.E., 2005. Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. The Journal of biological chemistry 280, 36417-36428.
Hughes, E.S., Bell, J.E., Simmonds, P., 1997. Investigation of the dynamics of the spread of human immunodeficiency virus to brain and other tissues by evolutionary analysis of sequences from the p17gag and env genes. Journal of virology 71, 1272-1280.
Humphreys, I., Fleming, V., Fabris, P., Parker, J., Schulenberg, B., Brown, A., Demetriou, C., Gaudieri, S., Pfafferott, K., Lucas, M., Collier, J., Huang, K.H., Pybus, O.G., Klenerman, P., Barnes, E., 2009. Full-length characterization of hepatitis C virus subtype 3a reveals novel hypervariable regions under positive selection during acute infection. Journal of virology 83, 11456-11466.
Hussy, P., Faust, H., Wagner, J.C., Schmid, G., Mous, J., Jacobsen, H., 1997. Evaluation of hepatitis C virus envelope proteins expressed in E. coli and insect cells for use as tools for antibody screening. Journal of hepatology 26, 1179-1186.
Idrovo, V., Dailey, P.J., Jeffers, L.J., Coelho-Little, E., Bernstein, D., Bartholomew, M., Alvarez, L., Urdea, M.S., Collins, M.L., Schiff, E.R., 1996. Hepatitis C virus RNA quantification in right and left lobes of the liver in patients with chronic hepatitis C. Journal of viral hepatitis 3, 239-246.
Imhof, I., Simmonds, P., 2011. Genotype differences in susceptibility and resistance development of hepatitis C virus to protease inhibitors telaprevir (VX-950) and danoprevir (ITMN-191). Hepatology 53, 1090-1099.
Indolfi, G., Resti, M., 2009. Perinatal transmission of hepatitis C virus infection. Journal of medical virology 81, 836-843.
Inui, A., Fujisawa, T., Sogo, T., Komatsu, H., Isozaki, A., Sekine, I., 2002. Different outcomes of vertical transmission of hepatitis C virus in a twin pregnancy. J Gastroenterol Hepatol 17, 617-619.
Iorio, R., Giannattasio, A., Sepe, A., Terracciano, L.M., Vecchione, R., Vegnente, A., 2005. Chronic hepatitis C in childhood: an 18-year experience. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 41, 1431-1437.
Irshad, M., Ansari, M.A., Singh, A., Nag, P., Raghvendra, L., Singh, S., Badhal, S.S., 2010. HCV-genotypes: a review on their origin, global status, assay system, pathogenecity and response to treatment. Hepatogastroenterology 57, 1529-1538.
Itsui, Y., Sakamoto, N., Kurosaki, M., Kanazawa, N., Tanabe, Y., Koyama, T., Takeda, Y., Nakagawa, M., Kakinuma, S., Sekine, Y., Maekawa, S., Enomoto, N., Watanabe, M., 2006. Expressional screening of

interferon-stimulated genes for antiviral activity against hepatitis C virus replication. Journal of viral hepatitis 13, 690-700.
Jacka, B., Lamoury, F., Simmonds, P., Dore, G.J., Grebely, J., Applegate, T., 2013. Sequencing of the Hepatitis C Virus: A Systematic Review. PLoS One 8, e67073.
Jacobson, I.M., McHutchison, J.G., Dusheiko, G., Di Bisceglie, A.M., Reddy, K.R., Bzowej, N.H., Marcellin, P., Muir, A.J., Ferenci, P., Flisiak, R., George, J., Rizzetto, M., Shouval, D., Sola, R., Terg, R.A., Yoshida, E.M., Adda, N., Bengtsson, L., Sankoh, A.J., Kieffer, T.L., George, S., Kauffman, R.S., Zeuzem, S., 2011. Telaprevir for previously untreated chronic hepatitis C virus infection. N Engl J Med 364, 2405-2416.
Jara, P., Hierro, L., de la Vega, A., Diaz, C., Camarena, C., Frauca, E., Minos-Bartolo, G., Diez-Dorado, R., de Guevara, C.L., Larrauri, J., Rueda, M., 2008. Efficacy and safety of peginterferon-alpha2b and ribavirin combination therapy in children with chronic hepatitis C infection. Pediatr Infect Dis J 27, 142-148.
Jenkins, G.M., Worobey, M., Woelk, C.H., Holmes, E.C., 2001. Evidence for the non-quasispecies evolution of RNA viruses [corrected]. Mol Biol Evol 18, 987-994.
Jensen, T.B., Gottwein, J.M., Scheel, T.K., Hoegh, A.M., Eugen-Olsen, J., Bukh, J., 2008. Highly efficient JFH1-based cell-culture system for hepatitis C virus genotype 5a: failure of homologous neutralizing-antibody treatment to control infection. The Journal of infectious diseases 198, 1756-1765.
Jonas, M.M., 2002. Children with hepatitis C. Hepatology 36, S173-178.
Josefsson, L., King, M.S., Makitalo, B., Brannstrom, J., Shao, W., Maldarelli, F., Kearney, M.F., Hu, W.S., Chen, J., Gaines, H., Mellors, J.W., Albert, J., Coffin, J.M., Palmer, S.E., 2011. Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. Proceedings of the National Academy of Sciences of the United States of America 108, 11199-11204.
Joyce, M.A., Tyrrell, D.L., 2010. The cell biology of hepatitis C virus. Microbes Infect 12, 263-271.
Kage, M., Fujisawa, T., Shiraki, K., Tanaka, T., Kimura, A., Shimamatsu, K., Nakashima, E., Kojiro, M., Koike, M., Tazawa, Y., Abukawa, D., Okaniwa, M., Takita, H., Matsui, A., Hayashi, T., Etou, T., Terasawa, S., Sugiyama, K., Tajiri, H., Yoden, A., Kajiwara, Y., Sata, M., Uchimura, Y., 1997. Pathology of chronic hepatitis C in children. Child Liver Study Group of Japan. Hepatology 26, 771-775.
Kandathil, A.J., Graw, F., Quinn, J., Hwang, H.S., Torbenson, M., Perelson, A.S., Ray, S.C., Thomas, D.L., Ribeiro, R.M., Balagopal, A., 2013. Use of laser capture microdissection to map hepatitis C virus-positive hepatocytes in human liver. Gastroenterology 145, 1404-1413 e1410.
Kapoor, A., Simmonds, P., Gerold, G., Qaisar, N., Jain, K., Henriquez, J.A., Firth, C., Hirschberg, D.L., Rice, C.M., Shields, S., Lipkin, W.I., 2011. Characterization of a canine homolog of hepatitis C virus. Proceedings of the National Academy of Sciences of the United States of America 108, 11608-11613.
Kapoor, A., Simmonds, P., Scheel, T.K., Hjelle, B., Cullen, J.M., Burbelo, P.D., Chauhan, L.V., Duraisamy, R., Sanchez Leon, M., Jain, K., Vandegrift,

K.J., Calisher, C.H., Rice, C.M., Lipkin, W.I., 2013. Identification of rodent homologs of hepatitis C virus and pegiviruses. MBio 4, e00216-00213.
Kato, N., Ootsuyama, Y., Sekiya, H., Ohkoshi, S., Nakazawa, T., Hijikata, M., Shimotohno, K., 1994. Genetic drift in hypervariable region 1 of the viral genome in persistent hepatitis C virus infection. Journal of virology 68, 4776-4784.
Kato, T., Wakita, T., 2005. [Production of infectious hepatitis C virus in cell culture]. Uirusu 55, 287-295.
Ke, P.Y., Chen, S.S., 2012. Hepatitis C virus and cellular stress response: implications to molecular pathogenesis of liver diseases. Viruses 4, 2251-2290.
Keck, Z.Y., Li, S.H., Xia, J., von Hahn, T., Balfe, P., McKeating, J.A., Witteveldt, J., Patel, A.H., Alter, H., Rice, C.M., Foung, S.K., 2009. Mutations in hepatitis C virus E2 located outside the CD81 binding sites lead to escape from broadly neutralizing antibodies but compromise virus infectivity. Journal of virology 83, 6149-6160.
Keck, Z.Y., Li, T.K., Xia, J., Bartosch, B., Cosset, F.L., Dubuisson, J., Foung, S.K., 2005. Analysis of a highly flexible conformational immunogenic domain a in hepatitis C virus E2. Journal of virology 79, 13199-13208.
Khaliq, S., Jahan, S., Hassan, S., 2011. Hepatitis C virus p7: molecular function and importance in hepatitis C virus life cycle and potential antiviral target. Liver Int 31, 606-617.
Khan, A.G., Whidby, J., Miller, M.T., Scarborough, H., Zatorski, A.V., Cygan, A., Price, A.A., Yost, S.A., Bohannon, C.D., Jacob, J., Grakoui, A., Marcotrigiano, J., 2014. Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2. Nature.
Kim, C.W., Chang, K.M., 2013. Hepatitis C virus: virology and life cycle. Clin Mol Hepatol 19, 17-25.
Kim, K.I., Zhang, D.E., 2003. ISG15, not just another ubiquitin-like protein. Biochem Biophys Res Commun 307, 431-434.
Koizumi, K., Enomoto, N., Kurosaki, M., Murakami, T., Izumi, N., Marumo, F., Sato, C., 1995. Diversity of quasispecies in various disease stages of chronic hepatitis C virus infection and its significance in interferon treatment. Hepatology 22, 30-35.
Komatsu, H., Inui, A., Tsunoda, T., Sogo, T., Fujisawa, T., 2013. Association between an IL-28B genetic polymorphism and the efficacy of the response-guided pegylated interferon therapy in children with chronic hepatic C infection. Hepatol Res 43, 327-338.
Kong, L., Giang, E., Nieusma, T., Kadam, R.U., Cogburn, K.E., Hua, Y., Dai, X., Stanfield, R.L., Burton, D.R., Ward, A.B., Wilson, I.A., Law, M., 2013. Hepatitis C virus E2 envelope glycoprotein core structure. Science 342, 1090-1094.
Krause, G., Winkler, L., Mueller, S.L., Haseloff, R.F., Piontek, J., Blasig, I.E., 2008. Structure and function of claudins. Biochim Biophys Acta 1778, 631-645.
Krey, T., d'Alayer, J., Kikuti, C.M., Saulnier, A., Damier-Piolle, L., Petitpas, I., Johansson, D.X., Tawar, R.G., Baron, B., Robert, B., England, P., Persson, M.A., Martin, A., Rey, F.A., 2010. The disulfide bonds in

glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule. PLoS Pathog 6, e1000762.
Krugman, S., 1976. Viral hepatitis: overview and historical perspectives. Yale J Biol Med 49, 199-203.
Lamers, S.L., Gray, R.R., Salemi, M., Huysentruyt, L.C., McGrath, M.S., 2011. HIV-1 phylogenetic analysis shows HIV-1 transits through the meninges to brain and peripheral tissues. Infect Genet Evol 11, 31-37.
Lavillette, D., Pecheur, E.I., Donot, P., Fresquet, J., Molle, J., Corbau, R., Dreux, M., Penin, F., Cosset, F.L., 2007. Characterization of fusion determinants points to the involvement of three discrete regions of both E1 and E2 glycoproteins in the membrane fusion process of hepatitis C virus. Journal of virology 81, 8752-8765.
Law, M., Maruyama, T., Lewis, J., Giang, E., Tarr, A.W., Stamataki, Z., Gastaminza, P., Chisari, F.V., Jones, I.M., Fox, R.I., Ball, J.K., McKeating, J.A., Kneteman, N.M., Burton, D.R., 2008. Broadly neutralizing antibodies protect against hepatitis C virus quasispecies challenge. Nat Med 14, 25-27.
Lawitz, E., Gane, E.J., 2013. Sofosbuvir for previously untreated chronic hepatitis C infection. N Engl J Med 369, 678-679.
Lawitz, E., Lalezari, J.P., Hassanein, T., Kowdley, K.V., Poordad, F.F., Sheikh, A.M., Afdhal, N.H., Bernstein, D.E., Dejesus, E., Freilich, B., Nelson, D.R., Dieterich, D.T., Jacobson, I.M., Jensen, D., Abrams, G.A., Darling, J.M., Rodriguez-Torres, M., Reddy, K.R., Sulkowski, M.S., Bzowej, N.H., Hyland, R.H., Mo, H., Lin, M., Mader, M., Hindes, R., Albanis, E., Symonds, W.T., Berrey, M.M., Muir, A., 2013. Sofosbuvir in combination with peginterferon alfa-2a and ribavirin for non-cirrhotic, treatment-naive patients with genotypes 1, 2, and 3 hepatitis C infection: a randomised, double-blind, phase 2 trial. Lancet Infect Dis 13, 401-408.
Layden, T.J., Lam, N.P., Wiley, T.E., 1999. Hepatitis C viral dynamics. Clinics in liver disease 3, 793-810.
Le Campion, A., Larouche, A., Fauteux-Daniel, S., Soudeyns, H., 2012. Pathogenesis of hepatitis C during pregnancy and childhood. Viruses 4, 3531-3550.
Lechner, F., Wong, D.K., Dunbar, P.R., Chapman, R., Chung, R.T., Dohrenwend, P., Robbins, G., Phillips, R., Klenerman, P., Walker, B.D., 2000. Analysis of successful immune responses in persons infected with hepatitis C virus. J Exp Med 191, 1499-1512.
Lenschow, D.J., Lai, C., Frias-Staheli, N., Giannakopoulos, N.V., Lutz, A., Wolff, T., Osiak, A., Levine, B., Schmidt, R.E., Garcia-Sastre, A., Leib, D.A., Pekosz, A., Knobeloch, K.P., Horak, I., Virgin, H.W.t., 2007. IFN-stimulated gene 15 functions as a critical antiviral molecule against influenza, herpes, and Sindbis viruses. Proceedings of the National Academy of Sciences of the United States of America 104, 1371-1376.
Lenz, O., Vijgen, L., Berke, J.M., Cummings, M.D., Fevery, B., Peeters, M., De Smedt, G., Moreno, C., Picchio, G., 2013. Virologic response and characterisation of HCV genotype 2-6 in patients receiving TMC435 monotherapy (study TMC435-C202). Journal of hepatology 58, 445-451.
Li, H., Hughes, A.L., Bano, N., McArdle, S., Livingston, S., Deubner, H., McMahon, B.J., Townshend-Bulson, L., McMahan, R., Rosen, H.R.,

Gretch, D.R., 2011. Genetic diversity of near genome-wide hepatitis C virus sequences during chronic infection: evidence for protein structural conservation over time. PLoS One 6, e19562.
Li, H., Sullivan, D.G., Feuerborn, N., McArdle, S., Bekele, K., Pal, S., Yeh, M., Carithers, R.L., Perkins, J.D., Gretch, D.R., 2010. Genetic diversity of hepatitis C virus predicts recurrent disease after liver transplantation. Virology 402, 248-255.
Li, H., Thomassen, L.V., Majid, A., McMahon, B.J., Bruden, D., McArdle, S., Bano, N., Chung, M., Carithers, R.L., Perkins, J.D., Sullivan, D.G., Gretch, D.R., 2008. Investigation of putative multisubtype hepatitis C virus infections in vivo by heteroduplex mobility analysis of core/envelope subgenomes. Journal of virology 82, 7524-7532.
Li, K., Foy, E., Ferreon, J.C., Nakamura, M., Ferreon, A.C., Ikeda, M., Ray, S.C., Gale, M., Jr., Lemon, S.M., 2005. Immune evasion by hepatitis C virus NS3/4A protease-mediated cleavage of the Toll-like receptor 3 adaptor protein TRIF. Proceedings of the National Academy of Sciences of the United States of America 102, 2992-2997.
Liang, Y., Shilagard, T., Xiao, S.Y., Snyder, N., Lau, D., Cicalese, L., Weiss, H., Vargas, G., Lemon, S.M., 2009. Visualizing hepatitis C virus infections in human liver by two-photon microscopy. Gastroenterology 137, 1448-1458.
Liberal, R., Grant, C.R., Mieli-Vergani, G., Vergani, D., 2013. Autoimmune hepatitis: a comprehensive review. J Autoimmun 41, 126-139.
Lim, Y.S., Hwang, S.B., 2011. Hepatitis C virus NS5A protein interacts with phosphatidylinositol 4-kinase type IIIalpha and regulates viral propagation. The Journal of biological chemistry 286, 11290-11298.
Lin, C., Lindenbach, B.D., Pragai, B.M., McCourt, D.W., Rice, C.M., 1994a. Processing in the hepatitis C virus E2-NS2 region: identification of p7 and two distinct E2-specific products with different C termini. Journal of virology 68, 5063-5073.
Lin, H.H., Kao, J.H., Hsu, H.Y., Ni, Y.H., Yeh, S.H., Hwang, L.H., Chang, M.H., Hwang, S.C., Chen, P.J., Chen, D.S., 1994b. Possible role of high-titer maternal viremia in perinatal transmission of hepatitis C virus. The Journal of infectious diseases 169, 638-641.
Lin, L., Libbrecht, L., Verbeeck, J., Verslype, C., Roskams, T., van Pelt, J., Van Ranst, M., Fevery, J., 2009. Quantitation of replication of the HCV genome in human livers with end-stage cirrhosis by strand-specific real-time RT-PCR assays: methods and clinical relevance. Journal of medical virology 81, 1569-1575.
Lindenbach, B.D., 2013. Virion assembly and release. Curr Top Microbiol Immunol 369, 199-218.
Lindenbach, B.D., Evans, M.J., Syder, A.J., Wolk, B., Tellinghuisen, T.L., Liu, C.C., Maruyama, T., Hynes, R.O., Burton, D.R., McKeating, J.A., Rice, C.M., 2005. Complete replication of hepatitis C virus in cell culture. Science 309, 623-626.
Lindenbach, B.D., Rice, C.M., 2005. Unravelling hepatitis C virus replication from genome to function. Nature 436, 933-938.
Lindenbach, B.D., Rice, C.M., 2013. The ins and outs of hepatitis C virus entry and assembly. Nat Rev Microbiol 11, 688-700.

Linton, P.J., Dorshkind, K., 2004. Age-related changes in lymphocyte development and function. Nat Immunol 5, 133-139.
Liu, L., Fisher, B.E., Dowd, K.A., Astemborski, J., Cox, A.L., Ray, S.C., 2010. Acceleration of hepatitis C virus envelope evolution in humans is consistent with progressive humoral immune selection during the transition from acute to chronic infection. Journal of virology 84, 5067-5077.
Liu, S., Thaler, D.S., Libchaber, A., 2002. Signal and noise in bridging PCR. BMC Biotechnol 2, 13.
Logvinoff, C., Major, M.E., Oldach, D., Heyward, S., Talal, A., Balfe, P., Feinstone, S.M., Alter, H., Rice, C.M., McKeating, J.A., 2004. Neutralizing antibody response during acute and chronic hepatitis C virus infection. Proc Natl Acad Sci U S A 101, 10149-10154.
Lohmann, V., Bartenschlager, R., 2014. On the history of hepatitis C virus cell culture systems. J Med Chem 57, 1627-1642.
Lohmann, V., Korner, F., Koch, J., Herian, U., Theilmann, L., Bartenschlager, R., 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science 285, 110-113.
London, W.T., Sutnick, A.I., Blumberg, B.S., 1969. Australia antigen and acute viral hepatitis. Ann Intern Med 70, 55-59.
Lorent, E., Bierau, H., Engelborghs, Y., Verheyden, G., Bosman, F., 2008. Structural characterisation of the hepatitis C envelope glycoprotein E1 ectodomain derived from a mammalian and a yeast expression system. Vaccine 26, 399-410.
Lorenzo, L.J., Garcia, O., Acosta-Rivero, N., Duenas-Carrera, S., Martinez, G., Alvarez-Obregon, J., Pichardo, D., Ramos, A., Guerra, I., Morales, J., 2000. Expression and immunological evaluation of the Escherichia coli-derived hepatitis C virus envelope E1 protein. Biotechnol Appl Biochem 32 (Pt 2), 137-143.
Lupberger, J., Zeisel, M.B., Xiao, F., Thumann, C., Fofana, I., Zona, L., Davis, C., Mee, C.J., Turek, M., Gorke, S., Royer, C., Fischer, B., Zahid, M.N., Lavillette, D., Fresquet, J., Cosset, F.L., Rothenberg, S.M., Pietschmann, T., Patel, A.H., Pessaux, P., Doffoel, M., Raffelsberger, W., Poch, O., McKeating, J.A., Brino, L., Baumert, T.F., 2011. EGFR and EphA2 are host factors for hepatitis C virus entry and possible targets for antiviral therapy. Nature medicine 17, 589-595.
Lyons, S., Kapoor, A., Sharp, C., Schneider, B.S., Wolfe, N.D., Culshaw, G., Corcoran, B., McGorum, B.C., Simmonds, P., 2012. Nonprimate hepaciviruses in domestic horses, United kingdom. Emerg Infect Dis 18, 1976-1982.
Macalalad, A.R., Zody, M.C., Charlebois, P., Lennon, N.J., Newman, R.M., Malboeuf, C.M., Ryan, E.M., Boutwell, C.L., Power, K.A., Brackney, D.E., Pesko, K.N., Levin, J.Z., Ebel, G.D., Allen, T.M., Birren, B.W., Henn, M.R., 2012. Highly sensitive and specific detection of rare variants in mixed viral populations from massively parallel sequence data. PLoS Comput Biol 8, e1002417.
MacCallum, F.O., 1972. 1971 International Symposium on Viral Hepatitis. Historical perspectives. Can Med Assoc J 106, Suppl:423-426.
MacCallum, F.O., Bradley, W.H., 1944. Transmission of infective hepatitis to human volunteers. Lancet 2.

Mack, C.L., Gonzalez-Peralta, R.P., Gupta, N., Leung, D., Narkewicz, M.R., Roberts, E.A., Rosenthal, P., Schwarz, K.B., 2012. NASPGHAN practice guidelines: Diagnosis and management of hepatitis C infection in infants, children, and adolescents. J Pediatr Gastroenterol Nutr 54, 838-855.

Maggi, F., Giorgi, M., Fornai, C., Morrica, A., Vatteroni, M.L., Pistello, M., Siciliano, G., Nuccorini, A., Bendinelli, M., 1999. Detection and quasispecies analysis of hepatitis C virus in the cerebrospinal fluid of infected patients. J Neurovirol 5, 319-323.

Major, M.E., Dahari, H., Mihalik, K., Puig, M., Rice, C.M., Neumann, A.U., Feinstone, S.M., 2004. Hepatitis C virus kinetics and host responses associated with disease and outcome of infection in chimpanzees. Hepatology 39, 1709-1720.

Mamedov, T.G., Pienaar, E., Whitney, S.E., TerMaat, J.R., Carvill, G., Goliath, R., Subramanian, A., Viljoen, H.J., 2008. A fundamental study of the PCR amplification of GC-rich DNA templates. Comput Biol Chem 32, 452-457.

Mangia, A., Mottola, L., Santoro, R., 2013. Interleukin 28B polymorphisms as predictor of response in hepatitis C virus genotype 2 and 3 infected patients. World J Gastroenterol 19, 8924-8928.

Manzin, A., Solforosi, L., Petrelli, E., Macarri, G., Tosone, G., Piazza, M., Clementi, M., 1998. Evolution of hypervariable region 1 of hepatitis C virus in primary infection. Journal of virology 72, 6271-6276.

Mardis, E.R., 2008. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 9, 387-402.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376-380.

Marine-Barjoan, E., Berrebi, A., Giordanengo, V., Favre, S.F., Haas, H., Moreigne, M., Izopet, J., Tricoire, J., Tran, A., Pradier, C., Bongain, A., 2007. HCV/HIV co-infection, HCV viral load and mode of delivery: risk factors for mother-to-child transmission of hepatitis C virus? AIDS 21, 1811-1815.

Markarian, S.A., Asatryan, A.M., Grigoryan, K.R., Sargsyan, H.R., 2006. Effect of diethylsulfoxide on the thermal denaturation of DNA. Biopolymers 82, 1-5.

Martin, D.N., Uprichard, S.L., 2013. Identification of transferrin receptor 1 as a hepatitis C virus entry factor. Proceedings of the National Academy of Sciences of the United States of America 110, 10777-10782.

Mas, A., Lopez-Galindez, C., Cacho, I., Gomez, J., Martinez, M.A., 2010. Unfinished stories on viral quasispecies and Darwinian views of evolution. J Mol Biol 397, 865-877.

Masaki, T., Suzuki, R., Murakami, K., Aizaki, H., Ishii, K., Murayama, A., Date, T., Matsuura, Y., Miyamura, T., Wakita, T., Suzuki, T., 2008. Interaction of hepatitis C virus nonstructural protein 5A with core protein is critical for the production of infectious virus particles. Journal of virology 82, 7964-7976.

Masoud, S.A., Johnson, L.B., White, F.F., 1992. The sequence within two primers influences the optimum concentration of dimethyl sulfoxide in the PCR. PCR Methods Appl 2, 89-90.

McCaffrey, K., Gouklani, H., Boo, I., Poumbourios, P., Drummer, H.E., 2011. The variable regions of hepatitis C virus glycoprotein E2 have an essential structural role in glycoprotein assembly and virion infectivity. The Journal of general virology 92, 112-121.

McCaughan, G.W., Laskus, T., Vargas, H.E., 2003. Hepatitis C virus quasispecies: misunderstood and mistreated? Liver Transpl 9, 1048-1052.

McGilvray, I., Feld, J.J., Chen, L., Pattullo, V., Guindi, M., Fischer, S., Borozan, I., Xie, G., Selzner, N., Heathcote, E.J., Siminovitch, K., 2012. Hepatic cell-type specific gene expression better predicts HCV treatment outcome than IL28B genotype. Gastroenterology 142, 1122-1131 e1121.

McHutchison, J.G., Lawitz, E.J., Shiffman, M.L., Muir, A.J., Galler, G.W., McCone, J., Nyberg, L.M., Lee, W.M., Ghalib, R.H., Schiff, E.R., Galati, J.S., Bacon, B.R., Davis, M.N., Mukhopadhyay, P., Koury, K., Noviello, S., Pedicone, L.D., Brass, C.A., Albrecht, J.K., Sulkowski, M.S., 2009. Peginterferon alfa-2b or alfa-2a with ribavirin for treatment of hepatitis C infection. N Engl J Med 361, 580-593.

McKeating, J.A., Zhang, L.Q., Logvinoff, C., Flint, M., Zhang, J., Yu, J., Butera, D., Ho, D.D., Dustin, L.B., Rice, C.M., Balfe, P., 2004. Diverse hepatitis C virus glycoproteins mediate viral infection in a CD81-dependent manner. Journal of virology 78, 8496-8505.

McLauchlan, J., 2000. Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. Journal of viral hepatitis 7, 2-14.

McLauchlan, J., Lemberg, M.K., Hope, G., Martoglio, B., 2002. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. Embo J 21, 3980-3988.

McMahan, R.H., Golden-Mason, L., Nishimura, M.I., McMahon, B.J., Kemper, M., Allen, T.M., Gretch, D.R., Rosen, H.R., 2010. Tim-3 expression on PD-1+ HCV-specific human CTLs is associated with viral persistence, and its blockade restores hepatocyte-directed in vitro cytotoxicity. J Clin Invest 120, 4546-4557.

McOmish, F., Yap, P.L., Dow, B.C., Follett, E.A., Seed, C., Keller, A.J., Cobain, T.J., Krusius, T., Kolho, E., Naukkarinen, R., et al., 1994. Geographical distribution of hepatitis C virus genotypes in blood donors: an international collaborative survey. J Clin Microbiol 32, 884-892.

McWilliam Leitch, E.C., McLauchlan, J., 2013. Determining the cellular diversity of hepatitis C virus quasispecies by single-cell viral sequencing. Journal of virology 87, 12648-12655.
Meertens, L., Bertaux, C., Dragic, T., 2006. Hepatitis C virus entry requires a critical postinternalization step and delivery to early endosomes via clathrin-coated vesicles. Journal of virology 80, 11571-11578.
Meredith, L.W., Harris, H.J., Wilson, G.K., Fletcher, N.F., Balfe, P., McKeating, J.A., 2013a. Early infection events highlight the limited transmissibility of hepatitis C virus in vitro. Journal of hepatology 58, 1074-1080.
Meredith, L.W., Wilson, G.K., Fletcher, N.F., McKeating, J.A., 2012. Hepatitis C virus entry: beyond receptors. Reviews in medical virology 22, 182-193.
Meredith, L.W., Zitzmann, N., McKeating, J.A., 2013b. Differential effect of p7 inhibitors on hepatitis C virus cell-to-cell transmission. Antiviral research 100, 636-639.
Meunier, J.C., Fournillier, A., Choukhi, A., Cahour, A., Cocquerel, L., Dubuisson, J., Wychowski, C., 1999. Analysis of the glycosylation sites of hepatitis C virus (HCV) glycoprotein E1 and the influence of E1 glycans on the formation of the HCV glycoprotein complex. The Journal of general virology 80 (Pt 4), 887-896.
Meylan, E., Curran, J., Hofmann, K., Moradpour, D., Binder, M., Bartenschlager, R., Tschopp, J., 2005. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature 437, 1167-1172.
Miyanari, Y., Atsuzawa, K., Usuda, N., Watashi, K., Hishiki, T., Zayas, M., Bartenschlager, R., Wakita, T., Hijikata, M., Shimotohno, K., 2007. The lipid droplet is an important organelle for hepatitis C virus production. Nat Cell Biol 9, 1089-1097.
Mohd Hanafiah, K., Groeger, J., Flaxman, A.D., Wiersma, S.T., 2013. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. Hepatology 57, 1333-1342.
Mok, J., Pembrey, L., Tovo, P.A., Newell, M.L., 2005. When does mother to child transmission of hepatitis C virus occur? Arch Dis Child Fetal Neonatal Ed 90, F156-160.
Moradpour, D., Penin, F., Rice, C.M., 2007. Replication of hepatitis C virus. Nat Rev Microbiol 5, 453-463.
Moskophidis, D., Lechner, F., Pircher, H., Zinkernagel, R.M., 1993. Virus persistence in acutely infected immunocompetent mice by exhaustion of antiviral cytotoxic effector T cells. Nature 362, 758-761.
Murray, K.F., Finn, L.S., Taylor, S.L., Seidel, K.D., Larson, A.M., 2005. Liver histology and alanine aminotransferase levels in children and adults with chronic hepatitis C infection. J Pediatr Gastroenterol Nutr 41, 634-638.
Naito, M., Hayashi, N., Moribe, T., Hagiwara, H., Mita, E., Kanazawa, Y., Kasahara, A., Fusamoto, H., Kamada, T., 1995. Hepatitis C viral quasispecies in hepatitis C virus carriers with normal liver enzymes and patients with type C chronic liver disease. Hepatology 22, 407-412.
Natarajan, V., Kottilil, S., Hazen, A., Adelsberger, J., Murphy, A.A., Polis, M.A., Kovacs, J.A., 2010. HCV in peripheral blood mononuclear cells are

predominantly carried on the surface of cells in HIV/HCV co-infected individuals. Journal of medical virology 82, 2032-2037.
Navas, S., Martin, J., Quiroga, J.A., Castillo, I., Carreno, V., 1998. Genetic diversity and tissue compartmentalization of the hepatitis C virus genome in blood mononuclear cells, liver, and serum from chronic hepatitis C patients. Journal of virology 72, 1640-1646.
Neefe, J.R., Stokes, J.J., Reinhold, J.G., Lukens, D.D.W., 1944. Hepatitis due to the infection of homologous blood products in human volunteers. J. Clin. Invest. 23, 20.
Neumann, A.U., Lam, N.P., Dahari, H., Gretch, D.R., Wiley, T.E., Layden, T.J., Perelson, A.S., 1998. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. Science 282, 103-107.
Neumann-Haefelin, C., Timm, J., Spangenberg, H.C., Wischniowski, N., Nazarova, N., Kersting, N., Roggendorf, M., Allen, T.M., Blum, H.E., Thimme, R., 2008. Virological and immunological determinants of intrahepatic virus-specific CD8+ T-cell failure in chronic hepatitis C virus infection. Hepatology 47, 1824-1836.
Ogawa, K., Hishiki, T., Shimizu, Y., Funami, K., Sugiyama, K., Miyanari, Y., Shimotohno, K., 2009. Hepatitis C virus utilizes lipid droplet for production of infectious virus. Proc Jpn Acad Ser B Phys Biol Sci 85, 217-228.
Okumura, A., Lu, G., Pitha-Rowe, I., Pitha, P.M., 2006. Innate antiviral response targets HIV-1 release by the induction of ubiquitin-like protein ISG15. Proceedings of the National Academy of Sciences of the United States of America 103, 1440-1445.
Osburn, W.O., Snider, A.E., Wells, B.L., Latanich, R., Bailey, J.R., Thomas, D.L., Cox, A.L., Ray, S.C., 2014. Clearance of hepatitis C infection is associated with the early appearance of broad neutralizing antibody responses. Hepatology.
Owen, D.M., Huang, H., Ye, J., Gale, M., Jr., 2009. Apolipoprotein E on hepatitis C virion facilitates infection through interaction with low-density lipoprotein receptor. Virology 394, 99-108.
Owsianka, A.M., Timms, J.M., Tarr, A.W., Brown, R.J., Hickling, T.P., Szwejk, A., Bienkowska-Szewczyk, K., Thomson, B.J., Patel, A.H., Ball, J.K., 2006. Identification of conserved residues in the E2 envelope glycoprotein of the hepatitis C virus that are critical for CD81 binding. Journal of virology 80, 8695-8704.
Pal, S., Sullivan, D.G., Kim, S., Lai, K.K., Kae, J., Cotler, S.J., Carithers, R.L., Jr., Wood, B.L., Perkins, J.D., Gretch, D.R., 2006. Productive replication of hepatitis C virus in perihepatic lymph nodes in vivo: implications of HCV lymphotropism. Gastroenterology 130, 1107-1116.
Pene, V., Hernandez, C., Vauloup-Fellous, C., Garaud-Aunis, J., Rosenberg, A.R., 2009. Sequential processing of hepatitis C virus core protein by host cell signal peptidase and signal peptide peptidase: a reassessment. Journal of viral hepatitis 16, 705-715.
Penin, F., Dubuisson, J., Rey, F.A., Moradpour, D., Pawlotsky, J.M., 2004. Structural biology of hepatitis C virus. Hepatology 39, 5-19.
Penton, P.K., Blackard, J.T., 2013. Analysis of HIV Quasispecies Suggests Compartmentalization in the Liver. AIDS research and human retroviruses.

Pestka, J.M., Zeisel, M.B., Blaser, E., Schurmann, P., Bartosch, B., Cosset, F.L., Patel, A.H., Meisel, H., Baumert, J., Viazov, S., Rispeter, K., Blum, H.E., Roggendorf, M., Baumert, T.F., 2007. Rapid induction of virus-neutralizing antibodies and viral clearance in a single-source outbreak of hepatitis C. Proceedings of the National Academy of Sciences of the United States of America 104, 6025-6030.

Peters, T., Schlayer, H.J., Hiller, B., Rosler, B., Blum, H., Rasenack, J., 1997. Quasispecies analysis in hepatitis C virus infection by fluorescent single strand conformation polymorphism. Journal of virological methods 64, 95-102.

Philpott, S., Burger, H., Tsoukas, C., Foley, B., Anastos, K., Kitchen, C., Weiser, B., 2005. Human immunodeficiency virus type 1 genomic RNA sequences in the female genital tract and blood: compartmentalization and intrapatient recombination. Journal of virology 79, 353-363.

Pietschmann, T., Kaul, A., Koutsoudakis, G., Shavinskaya, A., Kallis, S., Steinmann, E., Abid, K., Negro, F., Dreux, M., Cosset, F.L., Bartenschlager, R., 2006. Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. Proceedings of the National Academy of Sciences of the United States of America 103, 7408-7413.

Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., Weiner, A.J., Houghton, M., Rosa, D., Grandi, G., Abrignani, S., 1998. Binding of hepatitis C virus to CD81. Science 282, 938-941.

Ploss, A., Evans, M.J., Gaysinskaya, V.A., Panis, M., You, H., de Jong, Y.P., Rice, C.M., 2009. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. Nature 457, 882-886.

Ploss, A., Khetani, S.R., Jones, C.T., Syder, A.J., Trehan, K., Gaysinskaya, V.A., Mu, K., Ritola, K., Rice, C.M., Bhatia, S.N., 2010. Persistent hepatitis C virus infection in microscale primary human hepatocyte cultures. Proceedings of the National Academy of Sciences of the United States of America 107, 3141-3145.

Polyak, S.J., Khabar, K.S., Paschal, D.M., Ezelle, H.J., Duverlie, G., Barber, G.N., Levy, D.E., Mukaida, N., Gretch, D.R., 2001. Hepatitis C virus nonstructural 5A protein induces interleukin-8, leading to partial inhibition of the interferon-induced antiviral response. Journal of virology 75, 6095-6106.

Poordad, F., McCone, J., Jr., Bacon, B.R., Bruno, S., Manns, M.P., Sulkowski, M.S., Jacobson, I.M., Reddy, K.R., Goodman, Z.D., Boparai, N., DiNubile, M.J., Sniukiene, V., Brass, C.A., Albrecht, J.K., Bronowicki, J.P., 2011. Boceprevir for untreated chronic HCV genotype 1 infection. N Engl J Med 364, 1195-1206.

Popescu, C.I., Dubuisson, J., 2010. Role of lipid metabolism in hepatitis C virus assembly and entry. Biol Cell 102, 63-74.

Powers, D.C., 1993. Influenza A virus-specific cytotoxic T lymphocyte activity declines with advancing age. J Am Geriatr Soc 41, 1-5.

Prentoe, J., Serre, S.B., Ramirez, S., Nicosia, A., Gottwein, J.M., Bukh, J., 2013. Hypervariable Region 1 Deletion and Required Adaptive Envelope Mutations Confer Decreased Dependency on Scavenger Receptor Class B Type I and Low Density Lipoprotein Receptor for Hepatitis C Virus. Journal of virology.

Prentoe, J., Serre, S.B., Ramirez, S., Nicosia, A., Gottwein, J.M., Bukh, J., 2014. Hypervariable region 1 deletion and required adaptive envelope mutations confer decreased dependency on scavenger receptor class B type I and low-density lipoprotein receptor for hepatitis C virus. Journal of virology 88, 1725-1739.

Prince, A.M., 1968. An antigen detected in the blood during the incubation period of serum hepatitis. Proceedings of the National Academy of Sciences of the United States of America 60, 814-821.

Prince, A.M., Brotman, B., Lee, D.H., Ren, L., Moore, B.S., Scheffel, J.W., 1999. Significance of the anti-E2 response in self-limited and chronic hepatitis C virus infections in chimpanzees and in humans. The Journal of infectious diseases 180, 987-991.

Qashqari, H., Al-Mars, A., Chaudhary, A., Abuzenadah, A., Damanhouri, G., Alqahtani, M., Mahmoud, M., El Sayed Zaki, M., Fatima, K., Qadri, I., 2013. Understanding the molecular mechanism(s) of hepatitis C virus (HCV) induced interferon resistance. Infect Genet Evol 19C, 113-119.

Qin, Y., Zhang, J., Garcia, T., Ito, K., Gutelius, D., Li, J., Wands, J., Tong, S., 2011. Improved method for rapid and efficient determination of genome replication and protein expression of clinical hepatitis B virus isolates. J Clin Microbiol 49, 1226-1233.

Ramakrishnaiah, V., Thumann, C., Fofana, I., Habersetzer, F., Pan, Q., de Ruiter, P.E., Willemsen, R., Demmers, J.A., Stalin Raj, V., Jenster, G., Kwekkeboom, J., Tilanus, H.W., Haagmans, B.L., Baumert, T.F., van der Laan, L.J., 2013. Exosome-mediated transmission of hepatitis C virus between human hepatoma Huh7.5 cells. Proceedings of the National Academy of Sciences of the United States of America 110, 13109-13113.

Ramirez, C., Gregori, J., Buti, M., Tabernero, D., Camos, S., Casillas, R., Quer, J., Esteban, R., Homs, M., Rodriguez-Frias, F., 2013. A comparative study of ultra-deep pyrosequencing and cloning to quantitatively analyze the viral quasispecies using hepatitis B virus infection as a model. Antiviral research 98, 273-283.

Ramirez, S., Perez-Del-Pulgar, S., Carrion, J.A., Costa, J., Gonzalez, P., Massaguer, A., Fondevila, C., Garcia-Valdecasas, J.C., Navasa, M., Forns, X., 2009. Hepatitis C virus compartmentalization and infection recurrence after liver transplantation. Am J Transplant 9, 1591-1601.

Rauch, A., Kutalik, Z., Descombes, P., Cai, T., Di Iulio, J., Mueller, T., Bochud, M., Battegay, M., Bernasconi, E., Borovicka, J., Colombo, S., Cerny, A., Dufour, J.F., Furrer, H., Gunthard, H.F., Heim, M., Hirschel, B., Malinverni, R., Moradpour, D., Mullhaupt, B., Witteck, A., Beckmann, J.S., Berg, T., Bergmann, S., Negro, F., Telenti, A., Bochud, P.Y., 2010. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. Gastroenterology 138, 1338-1345, 1345 e1331-1337.

Ray, S.C., Fanning, L., Wang, X.H., Netski, D.M., Kenny-Walsh, E., Thomas, D.L., 2005. Divergent and convergent evolution after a common-source outbreak of hepatitis C virus. J Exp Med 201, 1753-1759.

Ray, S.C., Wang, Y.M., Laeyendecker, O., Ticehurst, J.R., Villano, S.A., Thomas, D.L., 1999. Acute hepatitis C virus structural gene sequences

as predictors of persistent viremia: hypervariable region 1 as a decoy. Journal of virology 73, 2938-2946.
Resti, M., Azzari, C., Mannelli, F., Moriondo, M., Novembre, E., de Martino, M., Vierucci, A., 1998. Mother to child transmission of hepatitis C virus: prospective study of risk factors and timing of infection in children born to women seronegative for HIV-1. Tuscany Study Group on Hepatitis C Virus Infection. Bmj 317, 437-441.
Reuben, A., 2002. The thin red line. Hepatology 36, 770-773.
Ribeiro, R.M., Li, H., Wang, S., Stoddard, M.B., Learn, G.H., Korber, B.T., Bhattacharya, T., Guedj, J., Parrish, E.H., Hahn, B.H., Shaw, G.M., Perelson, A.S., 2012. Quantifying the diversification of hepatitis C virus (HCV) during primary infection: estimates of the in vivo mutation rate. PLoS Pathog 8, e1002881.
Roccasecca, R., Ansuini, H., Vitelli, A., Meola, A., Scarselli, E., Acali, S., Pezzanera, M., Ercole, B.B., McKeating, J., Yagnik, A., Lahm, A., Tramontano, A., Cortese, R., Nicosia, A., 2003. Binding of the hepatitis C virus E2 glycoprotein to CD81 is strain specific and is modulated by a complex interplay between hypervariable regions 1 and 2. J Virol 77, 1856-1867.
Rodrigo, A., Hanley, W., Goracke, P., Learn, G.H., 2001. Sampling and processing HIV molecular sequences: A computational evolutionary biologist's perspective. Kluwer Academia Publishers.
Rong, X., Lu, L., Wang, J., Xiong, H., Huang, J., Chen, J., Huang, K., Xu, R., Wang, M., Zhang, X., Guo, T., Liu, Y., Gao, G., Fu, Y., Nelson, K.E., 2012. Correlation of viral loads with HCV genotypes: higher levels of virus were revealed among blood donors infected with 6a strains. PLoS One 7, e52467.
Roque-Afonso, A.M., Ducoulombier, D., Di Liberto, G., Kara, R., Gigou, M., Dussaix, E., Samuel, D., Feray, C., 2005. Compartmentalization of hepatitis C virus genotypes between plasma and peripheral blood mononuclear cells. Journal of virology 79, 6349-6357.
Rosen, H.R., 2013. Emerging concepts in immunity to hepatitis C virus infection. J Clin Invest 123, 4121-4130.
Rouille, Y., Helle, F., Delgrange, D., Roingeard, P., Voisset, C., Blanchard, E., Belouzard, S., McKeating, J., Patel, A.H., Maertens, G., Wakita, T., Wychowski, C., Dubuisson, J., 2006. Subcellular localization of hepatitis C virus structural proteins in a cell culture system that efficiently replicates the virus. Journal of virology 80, 2832-2841.
Sainz, B., Jr., Barretto, N., Martin, D.N., Hiraga, N., Imamura, M., Hussain, S., Marsh, K.A., Yu, X., Chayama, K., Alrefai, W.A., Uprichard, S.L., 2012. Identification of the Niemann-Pick C1-like 1 cholesterol absorption receptor as a new hepatitis C virus entry factor. Nature medicine 18, 281-285.
Sakai, A., Kaneko, S., Honda, M., Matsushita, E., Kobayashi, K., 1999. Quasispecies of hepatitis C virus in serum and in three different parts of the liver of patients with chronic hepatitis. Hepatology 30, 556-561.
Salmeron, J., Casado, J., Rueda, P.M., Lafuente, V., Diago, M., Romero-Gomez, M., Palacios, A., Leon, J., Gila, A., Quiles, R., Rodriguez, L., Ruiz-Extremera, A., 2008. Quasispecies as predictive factor of rapid,

early and sustained virological responses in chronic hepatitis C, genotype 1, treated with peginterferon-ribavirin. J Clin Virol 41, 264-269.
Santolini, E., Migliaccio, G., La Monica, N., 1994. Biosynthesis and biochemical properties of the hepatitis C virus core protein. Journal of virology 68, 3631-3641.
Sarasin-Filipowicz, M., Oakeley, E.J., Duong, F.H., Christen, V., Terracciano, L., Filipowicz, W., Heim, M.H., 2008. Interferon signaling and treatment outcome in chronic hepatitis C. Proceedings of the National Academy of Sciences of the United States of America 105, 7034-7039.
Sardanyes, J., Elena, S.F., Sole, R.V., 2008. Simple quasispecies models for the survival-of-the-flattest effect: The role of space. J Theor Biol 250, 560-568.
Scarselli, E., Ansuini, H., Cerino, R., Roccasecca, R.M., Acali, S., Filocamo, G., Traboni, C., Nicosia, A., Cortese, R., Vitelli, A., 2002. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. Embo J 21, 5017-5025.
Scheel, T.K., Gottwein, J.M., Jensen, T.B., Prentoe, J.C., Hoegh, A.M., Alter, H.J., Eugen-Olsen, J., Bukh, J., 2008. Development of JFH1-based cell culture systems for hepatitis C virus genotype 4a and evidence for cross-genotype neutralization. Proceedings of the National Academy of Sciences of the United States of America 105, 997-1002.
Scheel, T.K., Rice, C.M., 2013. Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. Nature medicine 19, 837-849.
Schoggins, J.W., Rice, C.M., 2011. Interferon-stimulated genes and their antiviral effector functions. Curr Opin Virol 1, 519-525.
Schroder, M., Baran, M., Bowie, A.G., 2008. Viral targeting of DEAD box protein 3 reveals its role in TBK1/IKKepsilon-mediated IRF activation. Embo J 27, 2147-2157.
Schulze zur Wiesch, J., Lauer, G.M., Day, C.L., Kim, A.Y., Ouchi, K., Duncan, J.E., Wurcel, A.G., Timm, J., Jones, A.M., Mothe, B., Allen, T.M., McGovern, B., Lewis-Ximenez, L., Sidney, J., Sette, A., Chung, R.T., Walker, B.D., 2005. Broad repertoire of the CD4+ Th cell response in spontaneously controlled hepatitis C virus infection includes dominant and highly promiscuous epitopes. J Immunol 175, 3603-3613.
Schuster, S.C., 2008. Next-generation sequencing transforms today's biology. Nat Methods 5, 16-18.
Schwarz, K.B., Gonzalez-Peralta, R.P., Murray, K.F., Molleston, J.P., Haber, B.A., Jonas, M.M., Rosenthal, P., Mohan, P., Balistreri, W.F., Narkewicz, M.R., Smith, L., Lobritto, S.J., Rossi, S., Valsamakis, A., Goodman, Z., Robuck, P.R., Barton, B.A., 2011. The combination of ribavirin and peginterferon is superior to peginterferon and placebo for children and adolescents with chronic hepatitis C. Gastroenterology 140, 450-458 e451.
Schwimmer, J.B., Balistreri, W.F., 2000. Transmission, natural history, and treatment of hepatitis C virus infection in the pediatric population. Semin Liver Dis 20, 37-46.
Schwinefus, J.J., Menssen, R.J., Kohler, J.M., Schmidt, E.C., Thomas, A.L., 2013. Quantifying the temperature dependence of glycine-betaine RNA duplex destabilization. Biochemistry 52, 9339-9346.

Sharma, S.A., Feld, J.J., 2014. Acute hepatitis C: management in the rapidly evolving world of HCV. Curr Gastroenterol Rep 16, 371.
Sherwin, W.B., 2010. Entrophy and informational approaches to genetic diversity and its expression:genomic geofraphy. Entropy 12, 34.
Shulla, A., Randall, G., 2012. Hepatitis C virus-host interactions, replication, and viral assembly. Curr Opin Virol 2, 725-732.
Simmonds, P., 1995. Variability of hepatitis C virus. Hepatology 21, 570-583.
Simmonds, P., 2004. Genetic diversity and evolution of hepatitis C virus--15 years on. The Journal of general virology 85, 3173-3188.
Simmonds, P., Balfe, P., Peutherer, J.F., Ludlam, C.A., Bishop, J.O., Brown, A.J., 1990. Human immunodeficiency virus-infected individuals contain provirus in small numbers of peripheral mononuclear cells and at low copy numbers. Journal of virology 64, 864-872.
Sobesky, R., Feray, C., Rimlinger, F., Derian, N., Dos Santos, A., Roque-Afonso, A.M., Samuel, D., Brechot, C., Thiers, V., 2007. Distinct hepatitis C virus core and F protein quasispecies in tumoral and nontumoral hepatocytes isolated via microdissection. Hepatology 46, 1704-1712.
Sokal, E.M., Bourgois, A., Stephenne, X., Silveira, T., Porta, G., Gardovska, D., Fischler, B., Kelly, D., 2010. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection in children and adolescents. Journal of hepatology 52, 827-831.
Soriano, V., Mocroft, A., Rockstroh, J., Ledergerber, B., Knysz, B., Chaplinskas, S., Peters, L., Karlsson, A., Katlama, C., Toro, C., Kupfer, B., Vogel, M., Lundgren, J., 2008. Spontaneous viral clearance, viral load, and genotype distribution of hepatitis C virus (HCV) in HIV-infected patients with anti-HCV antibodies in Europe. The Journal of infectious diseases 198, 1337-1344.
Spengler, U., Nattermann, J., 2007. Immunopathogenesis in hepatitis C virus cirrhosis. Clin Sci (Lond) 112, 141-155.
Spiess, A.N., Mueller, N., Ivell, R., 2004. Trehalose is a potent PCR enhancer: lowering of DNA melting temperature and thermal stabilization of taq polymerase by the disaccharide trehalose. Clin Chem 50, 1256-1259.
Stapleford, K.A., Lindenbach, B.D., 2011. Hepatitis C virus NS2 coordinates virus particle assembly through physical interactions with the E1-E2 glycoprotein and NS3-NS4A enzyme complexes. Journal of virology 85, 1706-1717.
Steininger, C., Kundi, M., Jatzko, G., Kiss, H., Lischka, A., Holzmann, H., 2003. Increased risk of mother-to-infant transmission of hepatitis C virus by intrapartum infantile exposure to maternal blood. The Journal of infectious diseases 187, 345-351.
Stiffler, J.D., Nguyen, M., Sohn, J.A., Liu, C., Kaplan, D., Seeger, C., 2009. Focal distribution of hepatitis C virus RNA in infected livers. PLoS One 4, e6661.
Su, A.I., Pezacki, J.P., Wodicka, L., Brideau, A.D., Supekova, L., Thimme, R., Wieland, S., Bukh, J., Purcell, R.H., Schultz, P.G., Chisari, F.V., 2002. Genomic analysis of the host response to hepatitis C virus infection. Proceedings of the National Academy of Sciences of the United States of America 99, 15669-15674.

Sugano, M., Hayashi, Y., Yoon, S., Kinoshita, M., Ninomiya, T., Ohta, K., Itoh, H., Kasuga, M., 1995. Quantitation of hepatitis C viral RNA in liver and serum samples using competitive polymerase chain reaction. Journal of clinical pathology 48, 820-825.

Sulkowski, M.S., Kang, M., Matining, R., Wyles, D., Johnson, V.A., Morse, G.D., Amorosa, V., Bhattacharya, D., Coughlin, K., Wong-Staal, F., Glesby, M.J., 2013. Safety and Antiviral Activity of the HCV Entry Inhibitor ITX5061 in Treatment-Naive HCV-Infected Adults: A Randomized, Double-Blind, Phase 1b Study. The Journal of infectious diseases.

Sullivan, D.G., Bruden, D., Deubner, H., McArdle, S., Chung, M., Christensen, C., Hennessy, T., Homan, C., Williams, J., McMahon, B.J., Gretch, D.R., 2007. Hepatitis C virus dynamics during natural infection are associated with long-term histological outcome of chronic hepatitis C disease. The Journal of infectious diseases 196, 239-248.

Sumida, K., Shimoda, S., Iwasaka, S., Hisamoto, S., Kawanaka, H., Akahoshi, T., Ikegami, T., Shirabe, K., Shimono, N., Maehara, Y., Selmi, C., Gershwin, M.E., Akashi, K., 2013. Characteristics of splenic CD8+ T cell exhaustion in patients with hepatitis C. Clinical and experimental immunology 174, 172-178.

Suzuki, T., Ishii, K., Aizaki, H., Wakita, T., 2007. Hepatitis C viral life cycle. Adv Drug Deliv Rev 59, 1200-1212.

Syder, A.J., Lee, H., Zeisel, M.B., Grove, J., Soulier, E., Macdonald, J., Chow, S., Chang, J., Baumert, T.F., McKeating, J.A., McKelvy, J., Wong-Staal, F., 2011. Small molecule scavenger receptor BI antagonists are potent HCV entry inhibitors. Journal of hepatology 54, 48-55.

Tan, X., Lu, Z.J., Gao, G., Xu, Q., Hu, L., Fellmann, C., Li, M.Z., Qu, H., Lowe, S.W., Hannon, G.J., Elledge, S.J., 2012. Tiling genomes of pathogenic viruses identifies potent antiviral shRNAs and reveals a role for secondary structure in shRNA efficacy. Proceedings of the National Academy of Sciences of the United States of America 109, 869-874.

Tanaka, Y., Nishida, N., Sugiyama, M., Kurosaki, M., Matsuura, K., Sakamoto, N., Nakagawa, M., Korenaga, M., Hino, K., Hige, S., Ito, Y., Mita, E., Tanaka, E., Mochida, S., Murawaki, Y., Honda, M., Sakai, A., Hiasa, Y., Nishiguchi, S., Koike, A., Sakaida, I., Imamura, M., Ito, K., Yano, K., Masaki, N., Sugauchi, F., Izumi, N., Tokunaga, K., Mizokami, M., 2009. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 41, 1105-1109.

Tarr, A.W., Lafaye, P., Meredith, L., Damier-Piolle, L., Urbanowicz, R.A., Meola, A., Jestin, J.L., Brown, R.J., McKeating, J.A., Rey, F.A., Ball, J.K., Krey, T., 2013. An alpaca nanobody inhibits hepatitis C virus entry and cell-to-cell transmission. Hepatology 58, 932-939.

Tarr, A.W., Urbanowicz, R.A., Hamed, M.R., Albecka, A., McClure, C.P., Brown, R.J., Irving, W.L., Dubuisson, J., Ball, J.K., 2011. Hepatitis C patient-derived glycoproteins exhibit marked differences in susceptibility to serum neutralizing antibodies: genetic subtype defines antigenic but not neutralization serotype. Journal of virology 85, 4246-4257.

Terrault, N.A., Dailey, P.J., Ferrell, L., Collins, M.L., Wilber, J.C., Urdea, M.S., Bhandari, B.N., Wright, T.L., 1997. Hepatitis C virus: quantitation and distribution in liver. Journal of medical virology 51, 217-224.

Thimme, R., Binder, M., Bartenschlager, R., 2012. Failure of innate and adaptive immune responses in controlling hepatitis C virus infection. FEMS Microbiol Rev 36, 663-683.

Thimme, R., Oldach, D., Chang, K.M., Steiger, C., Ray, S.C., Chisari, F.V., 2001. Determinants of viral clearance and persistence during acute hepatitis C virus infection. J Exp Med 194, 1395-1406.

Thomas, S.L., Newell, M.L., Peckham, C.S., Ades, A.E., Hall, A.J., 1998. A review of hepatitis C virus (HCV) vertical transmission: risks of transmission to infants born to mothers with and without HCV viraemia or human immunodeficiency virus infection. Int J Epidemiol 27, 108-117.

Thorley, J.A., McKeating, J.A., Rappoport, J.Z., 2010. Mechanisms of viral entry: sneaking in the front door. Protoplasma 244, 15-24.

Tilg, H., Moschen, A.R., 2010. Evolution of inflammation in nonalcoholic fatty liver disease: the multiple parallel hits hypothesis. Hepatology 52, 1836-1846.

Timpe, J.M., McKeating, J.A., 2008. Hepatitis C virus entry: possible targets for therapy. Gut 57, 1728-1737.

Timpe, J.M., Stamataki, Z., Jennings, A., Hu, K., Farquhar, M.J., Harris, H.J., Schwarz, A., Desombere, I., Roels, G.L., Balfe, P., McKeating, J.A., 2008. Hepatitis C virus cell-cell transmission in hepatoma cells in the presence of neutralizing antibodies. Hepatology 47, 17-24.

Tong, X., Le Pogam, S., Li, L., Haines, K., Piso, K., Baronas, V., Yan, J.M., So, S.S., Klumpp, K., Najera, I., 2014. In Vivo Emergence of a Novel Mutant L159F/L320F in the NS5B Polymerase Confers Low-Level Resistance to the HCV Polymerase Inhibitors Mericitabine and Sofosbuvir. The Journal of infectious diseases 209, 668-675.

Tscherne, D.M., Jones, C.T., Evans, M.J., Lindenbach, B.D., McKeating, J.A., Rice, C.M., 2006. Time- and temperature-dependent activation of hepatitis C virus for low-pH-triggered entry. Journal of virology 80, 1734-1741.

Tuplin, A., Struthers, M., Simmonds, P., Evans, D.J., 2012. A twist in the tail: SHAPE mapping of long-range interactions and structural rearrangements of RNA elements involved in HCV replication. Nucleic Acids Res.

Urban, T.J., Thompson, A.J., Bradrick, S.S., Fellay, J., Schuppan, D., Cronin, K.D., Hong, L., McKenzie, A., Patel, K., Shianna, K.V., McHutchison, J.G., Goldstein, D.B., Afdhal, N., 2010. IL28B genotype is associated with differential expression of intrahepatic interferon-stimulated genes in patients with chronic hepatitis C. Hepatology 52, 1888-1896.

Vera-Otarola, J., Barria, M.I., Leon, U., Marsac, D., Carvallo, P., Soza, A., Lopez-Lastra, M., 2009. Hepatitis C virus quasispecies in plasma and peripheral blood mononuclear cells of treatment naive chronically infected patients. Journal of viral hepatitis 16, 633-643.

Vieyres, G., Thomas, X., Descamps, V., Duverlie, G., Patel, A.H., Dubuisson, J., 2010. Characterization of the envelope glycoproteins associated with infectious hepatitis C virus. Journal of virology 84, 10159-10168.

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., Andino, R., 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439, 344-348.

von Hahn, T., Yoon, J.C., Alter, H., Rice, C.M., Rehermann, B., Balfe, P., McKeating, J.A., 2007. Hepatitis C virus continuously escapes from neutralizing antibody and T-cell responses during chronic infection in vivo. Gastroenterology 132, 667-678.

Vona, G., Tuveri, R., Delpuech, O., Vallet, A., Canioni, D., Ballardini, G., Baptiste Trabut, J., Le Bail, B., Nalpas, B., Carnot, F., Pol, S., Brechot, C., Thiers, V., 2004. Intrahepatic hepatitis C virus RNA quantification in microdissected hepatocytes. Journal of hepatology 40, 682-688.

Wang, C., Mitsuya, Y., Gharizadeh, B., Ronaghi, M., Shafer, R.W., 2007a. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. Genome Res 17, 1195-1201.

Wang, G.P., Sherrill-Mix, S.A., Chang, K.M., Quince, C., Bushman, F.D., 2010. Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. Journal of virology 84, 6218-6228.

Wang, H.J., Gao, B., Zakhari, S., Nagy, L.E., 2012. Inflammation in alcoholic liver disease. Annu Rev Nutr 32, 343-368.

Wang, X.H., Netski, D.M., Astemborski, J., Mehta, S.H., Torbenson, M.S., Thomas, D.L., Ray, S.C., 2007b. Progression of fibrosis during chronic hepatitis C is associated with rapid virus evolution. Journal of virology 81, 6513-6522.

Wang, Y., Li, J., Wang, X., Sang, M., Ho, W., 2013. Hepatic stellate cells, liver innate immunity, and hepatitis C virus. J Gastroenterol Hepatol 28 Suppl 1, 112-115.

Waxman, S., Wurmbach, E., 2007. De-regulation of common housekeeping genes in hepatocellular carcinoma. BMC Genomics 8, 243.

Wei, X., Decker, J.M., Wang, S., Hui, H., Kappes, J.C., Wu, X., Salazar-Gonzalez, J.F., Salazar, M.G., Kilby, J.M., Saag, M.S., Komarova, N.L., Nowak, M.A., Hahn, B.H., Kwong, P.D., Shaw, G.M., 2003. Antibody neutralization and escape by HIV-1. Nature 422, 307-312.

Wherry, E.J., Blattman, J.N., Murali-Krishna, K., van der Most, R., Ahmed, R., 2003. Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. Journal of virology 77, 4911-4927.

White, P.A., Zhai, X., Carter, I., Zhao, Y., Rawlinson, W.D., 2000. Simplified hepatitis C virus genotyping by heteroduplex mobility analysis. J Clin Microbiol 38, 477-482.

Wieland, S., Makowska, Z., Campana, B., Calabrese, D., Dill, M.T., Chung, J., Chisari, F.V., Heim, M.H., 2013. Simultaneous detection of hepatitis C virus and interferon stimulated gene expression in infected human liver. Hepatology.

Wilke, C.O., 2005. Quasispecies theory in the context of population genetics. BMC Evol Biol 5, 44.

Wilke, C.O., Wang, J.L., Ofria, C., Lenski, R.E., Adami, C., 2001. Evolution of digital organisms at high mutation rates leads to survival of the flattest. Nature 412, 331-333.

Willems, M., Sheng, L., Roskams, T., Ramdani, B., Doutrelepont, J.M., Nevens, F., Durez, P., Treille, S., Adler, M., Desmet, V., et al., 1994.

Hepatitis C virus and its genotypes in patients suffering from chronic hepatitis C with or without a cryoglobulinemia-related syndrome. Journal of medical virology 44, 266-271.
Wilson, J.J., Polyak, S.J., Day, T.D., Gretch, D.R., 1995. Characterization of simple and complex hepatitis C virus quasispecies by heteroduplex gel shift analysis: correlation with nucleotide sequencing. The Journal of general virology 76 (Pt 7), 1763-1771.
Wirth, S., 2012. Current treatment options and response rates in children with chronic hepatitis C. World J Gastroenterol 18, 99-104.
Wirth, S., Pieper-Boustani, H., Lang, T., Ballauff, A., Kullmer, U., Gerner, P., Wintermeyer, P., Jenke, A., 2005. Peginterferon alfa-2b plus ribavirin treatment in children and adolescents with chronic hepatitis C. Hepatology 41, 1013-1018.
Wirth, S., Ribes-Koninckx, C., Calzado, M.A., Bortolotti, F., Zancan, L., Jara, P., Shelton, M., Kerkar, N., Galoppo, M., Pedreira, A., Rodriguez-Baez, N., Ciocca, M., Lachaux, A., Lacaille, F., Lang, T., Kullmer, U., Huber, W.D., Gonzalez, T., Pollack, H., Alonso, E., Broue, P., Ramakrishna, J., Neigut, D., Valle-Segarra, A.D., Hunter, B., Goodman, Z., Xu, C.R., Zheng, H., Noviello, S., Sniukiene, V., Brass, C., Albrecht, J.K., 2010. High sustained virologic response rates in children with chronic hepatitis C receiving peginterferon alfa-2b plus ribavirin. Journal of hepatology 52, 501-507.
Witteveldt, J., Evans, M.J., Bitzegeio, J., Koutsoudakis, G., Owsianka, A.M., Angus, A.G., Keck, Z.Y., Foung, S.K., Pietschmann, T., Rice, C.M., Patel, A.H., 2009. CD81 is dispensable for hepatitis C virus cell-to-cell transmission in hepatoma cells. The Journal of general virology 90, 48-58.
Wong, J.J., Pung, Y.F., Sze, N.S., Chin, K.C., 2006. HERC5 is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets. Proceedings of the National Academy of Sciences of the United States of America 103, 10735-10740.
Wozniak, A.L., Griffin, S., Rowlands, D., Harris, M., Yi, M., Lemon, S.M., Weinman, S.A., 2010. Intracellular proton conductance of the hepatitis C virus p7 protein and its contribution to infectious virus production. PLoS Pathog 6, e1001087.
Wu, S., Kanda, T., Nakamoto, S., Imazeki, F., Yokosuka, O., 2013. Hepatitis C virus protease inhibitor-resistance mutations: Our experience and review. World J Gastroenterol 19, 8940-8948.
Xiao, F., Fofana, I., Heydmann, L., Barth, H., Soulier, E., Habersetzer, F., Doffoel, M., Bukh, J., Patel, A.H., Zeisel, M.B., Baumert, T.F., 2014. Hepatitis C virus cell-cell transmission and resistance to direct-acting antiviral agents. PLoS Pathog 10, e1004128.
Yanagi, M., Purcell, R.H., Emerson, S.U., Bukh, J., 1997. Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. Proceedings of the National Academy of Sciences of the United States of America 94, 8738-8743.
Yanagi, M., St Claire, M., Shapiro, M., Emerson, S.U., Purcell, R.H., Bukh, J., 1998. Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo. Virology 244, 161-172.

Yang, X., Charlebois, P., Gnerre, S., Coole, M.G., Lennon, N.J., Levin, J.Z., Qu, J., Ryan, E.M., Zody, M.C., Henn, M.R., 2012. De novo assembly of highly diverse viral populations. BMC Genomics 13, 475.

Yeung, L.T., Roberts, E.A., 2010. Current issues in the management of paediatric viral hepatitis. Liver Int 30, 5-18.

Yi, M., Ma, Y., Yates, J., Lemon, S.M., 2007. Compensatory mutations in E1, p7, NS2, and NS3 enhance yields of cell culture-infectious intergenotypic chimeric hepatitis C virus. Journal of virology 81, 629-638.

Yuan, W., Krug, R.M., 2001. Influenza B virus NS1 protein inhibits conjugation of the interferon (IFN)-induced ubiquitin-like ISG15 protein. Embo J 20, 362-371.

Yuki, N., Matsumoto, S., Tadokoro, K., Mochizuki, K., Kato, M., Yamaguchi, T., 2006. Significance of liver negative-strand HCV RNA quantitation in chronic hepatitis C. Journal of hepatology 44, 302-309.

Zanetti, A.R., Tanzi, E., Paccagnini, S., Principi, N., Pizzocolo, G., Caccamo, M.L., D'Amico, E., Cambie, G., Vecchi, L., 1995. Mother-to-infant transmission of hepatitis C virus. Lombardy Study Group on Vertical HCV Transmission. Lancet 345, 289-291.

Zarate, S., Pond, S.L., Shapshak, P., Frost, S.D., 2007. Comparative study of methods for detecting sequence compartmentalization in human immunodeficiency virus type 1. Journal of virology 81, 6643-6651.

Zeisel, M.B., Koutsoudakis, G., Schnober, E.K., Haberstroh, A., Blum, H.E., Cosset, F.L., Wakita, T., Jaeck, D., Doffoel, M., Royer, C., Soulier, E., Schvoerer, E., Schuster, C., Stoll-Keller, F., Bartenschlager, R., Pietschmann, T., Barth, H., Baumert, T.F., 2007. Scavenger receptor class B type I is a key host factor for hepatitis C virus infection required for an entry step closely linked to CD81. Hepatology 46, 1722-1731.

Zhang, J., Randall, G., Higginbottom, A., Monk, P., Rice, C.M., McKeating, J.A., 2004. CD81 is required for hepatitis C virus glycoprotein-mediated viral infection. Journal of virology 78, 1448-1455.

Zhao, C., Denison, C., Huibregtse, J.M., Gygi, S., Krug, R.M., 2005. Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. Proceedings of the National Academy of Sciences of the United States of America 102, 10200-10205.

Zhong, J., Gastaminza, P., Cheng, G., Kapadia, S., Kato, T., Burton, D.R., Wieland, S.F., Uprichard, S.L., Wakita, T., Chisari, F.V., 2005. Robust hepatitis C virus infection in vitro. Proceedings of the National Academy of Sciences of the United States of America 102, 9294-9299.

Zhong, P., Agosto, L.M., Munro, J.B., Mothes, W., 2013. Cell-to-cell transmission of viruses. Curr Opin Virol 3, 44-50.

Zhu, H., Wong-Staal, F., Lee, H., Syder, A., McKelvy, J., Schooley, R.T., Wyles, D.L., 2012. Evaluation of ITX 5061, a scavenger receptor B1 antagonist: resistance selection and activity in combination with other hepatitis C virus antivirals. The Journal of infectious diseases 205, 656-662.

Zuccotti, G.V., Ribero, M.L., Giovannini, M., Fasola, M., Riva, E., Portera, G., Biasucci, G., Decarlis, S., Profeta, M.L., Tagger, A., 1995. Effect of

hepatitis C genotype on mother-to-infant transmission of virus. J Pediatr 127, 278-280.

10 APPENDIX

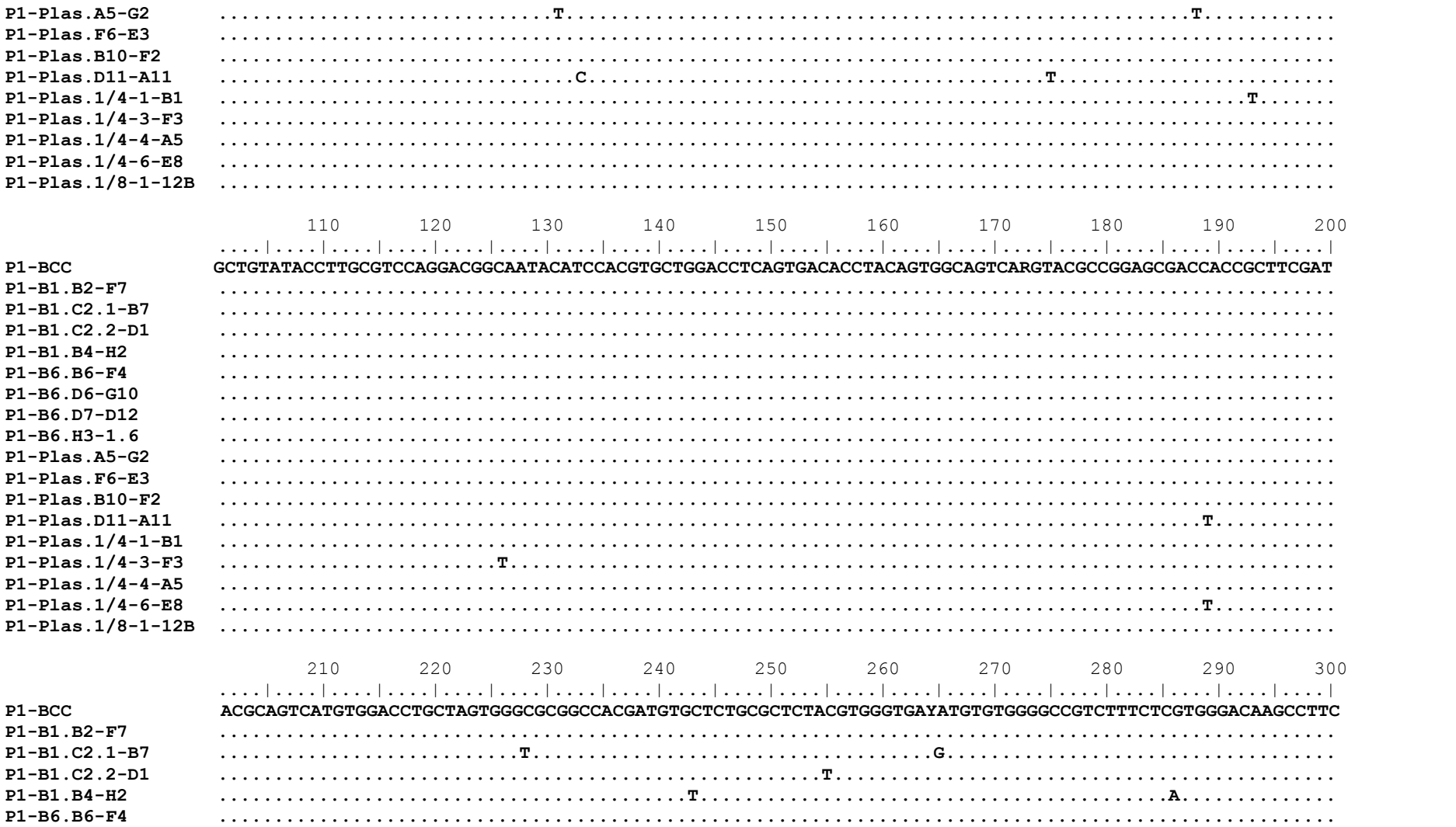
10.1 Appendix 1

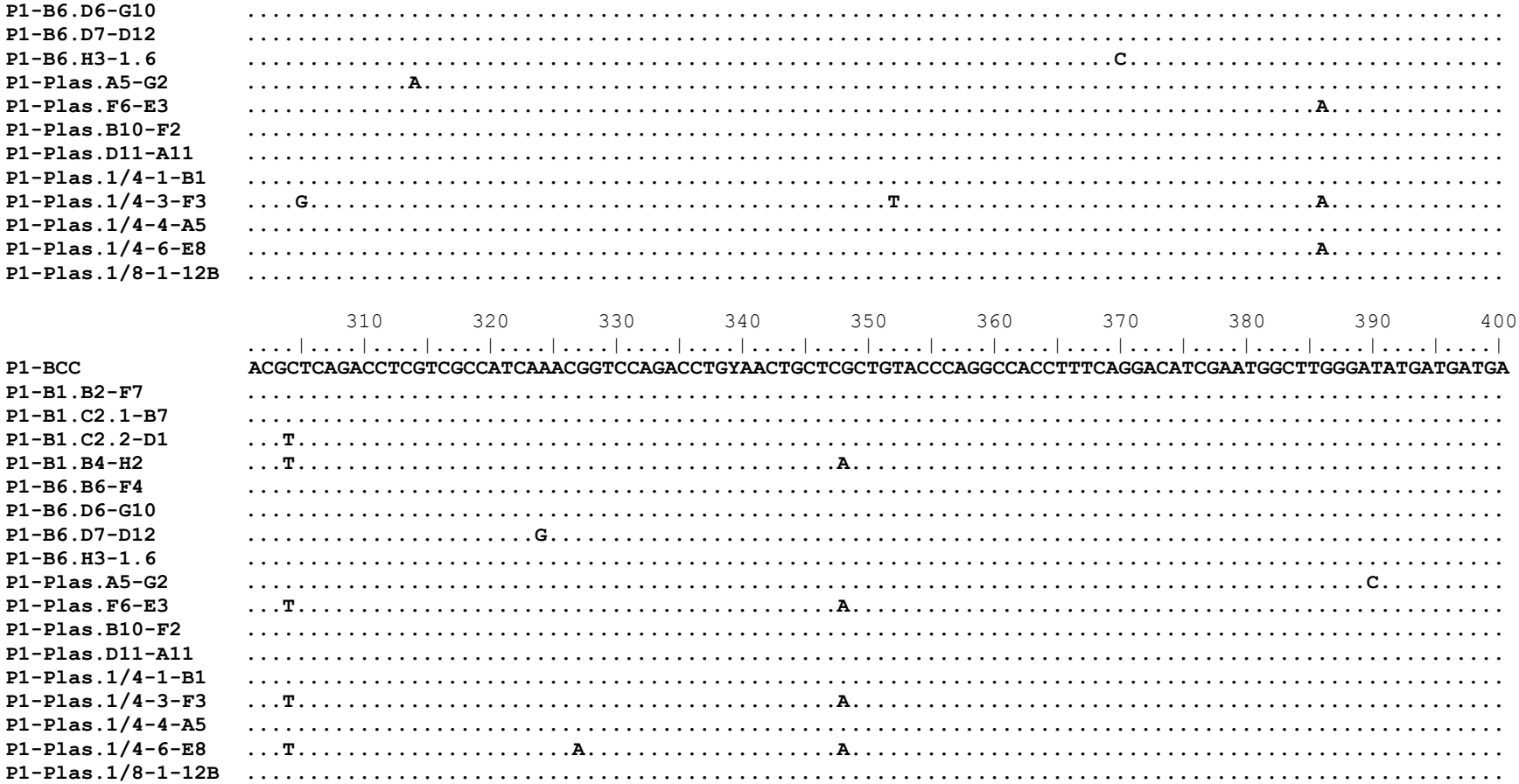
SGA-derived sequence alignments for patient 1, 2, 4, 5, 7 and 9. The position of E1, E2 and p7 differs between patients and is stated above each sequence alignment.

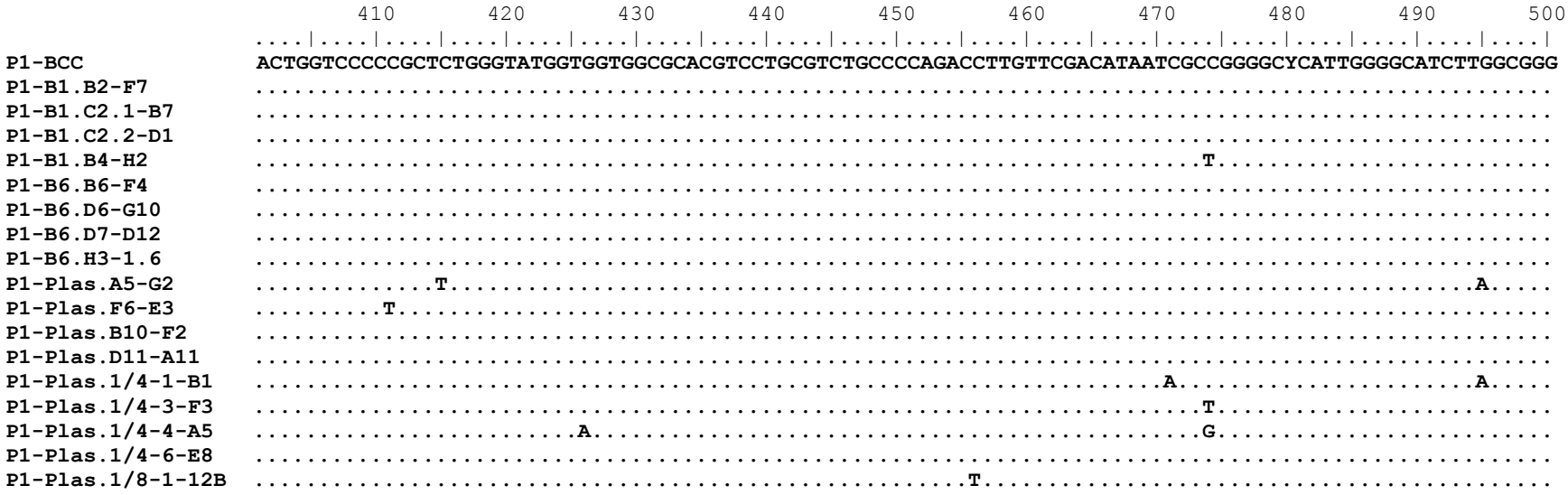
10.1.1 Patient 1

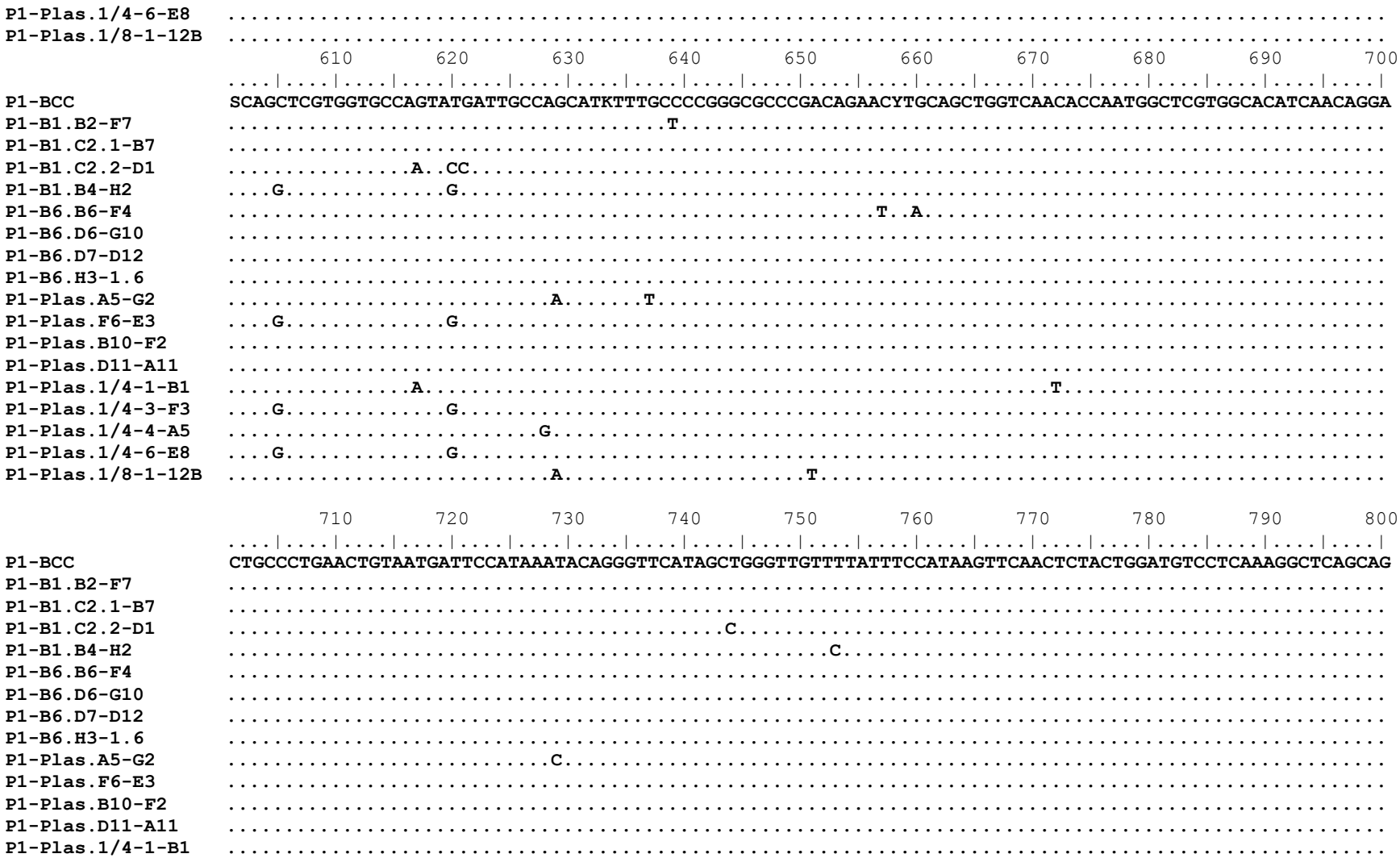
Gene position: E1: 1 - 576nt, E2: 577 - 1683nt, p7: 1683-1872nt , HVR1: 577 – 657nt, HVR494: 913 - 933nt, HVR595: 1153 - 1179nt.

```
                          10        20        30        40        50        60        70        80        90       100
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P1-BCC           TTAGAGTGGCGGAATACGTCCGGCCTCTACCTTCTTACCAACGACTGTTCTAATAGCAGTATTGTGTATGARGCCGATGACGTTATCCTGCACACACCTG
P1-B1.B2-F7      ....................................................................................................
P1-B1.C2.1-B7    ..................................................C.................................................
P1-B1.C2.2-D1    ....................................................................................................
P1-B1.B4-H2      ...................................................................................................T....
P1-B6.B6-F4      ....................................................................................................C.
P1-B6.D6-G10     ....................................................................................................
P1-B6.D7-D12     ....................................................................................................
P1-B6.H3-1.6     ....................................................................................................
```


P1-Plas.A5-G2 ..........T..........T.....
P1-Plas.F6-E3 ....
P1-Plas.B10-F2 ....
P1-Plas.D11-A11 ..........C..........T.....
P1-Plas.1/4-1-B1 ....T.....
P1-Plas.1/4-3-F3 ....
P1-Plas.1/4-4-A5 ....
P1-Plas.1/4-6-E8 ....
P1-Plas.1/8-1-12B ....
110 120 130 140 150 160 170 180 190 200
P1-BCC GCTGTATACCTTGCGTCCAGGACGGCAATACATCCACGTGCTGGACCTCAGTGACACCTACAGTGGCAGTCARGTACGCCGGAGCGACCACCGCTTCGAT
P1-B1.B2-F7 ....
P1-B1.C2.1-B7 ....
P1-B1.C2.2-D1 ....
P1-B1.B4-H2 ....
P1-B6.B6-F4 ....
P1-B6.D6-G10 ....
P1-B6.D7-D12 ....
P1-B6.H3-1.6 ....
P1-Plas.A5-G2 ....
P1-Plas.F6-E3 ....
P1-Plas.B10-F2 ....
P1-Plas.D11-A11 ....T.....
P1-Plas.1/4-1-B1 ....
P1-Plas.1/4-3-F3 ....T.....
P1-Plas.1/4-4-A5 ....
P1-Plas.1/4-6-E8 ....T.....
P1-Plas.1/8-1-12B ....
210 220 230 240 250 260 270 280 290 300
P1-BCC ACGCAGTCATGTGGACCTGCTAGTGGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGTGAYATGTGTGGGGCCGTCTTTCTCGTGGGACAAGCCTTC
P1-B1.B2-F7 ....
P1-B1.C2.1-B7 ..........T..........G.....
P1-B1.C2.2-D1 ....T.....
P1-B1.B4-H2 ..........T..........A.....
P1-B6.B6-F4 ....


P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5
P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B
310 320 330 340 350 360 370 380 390 400
P1-BCC
ACGCTCAGACCTCGTCGCCATCAAACGGTCCAGACCTGYAACTGCTCGCTGTACCCAGGCCACCTTTCAGGACATCGAATGGCTTGGGATATGATGATGA
P1-B1.B2-F7
P1-B1.C2.1-B7
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.B6-F4
P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5
P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B

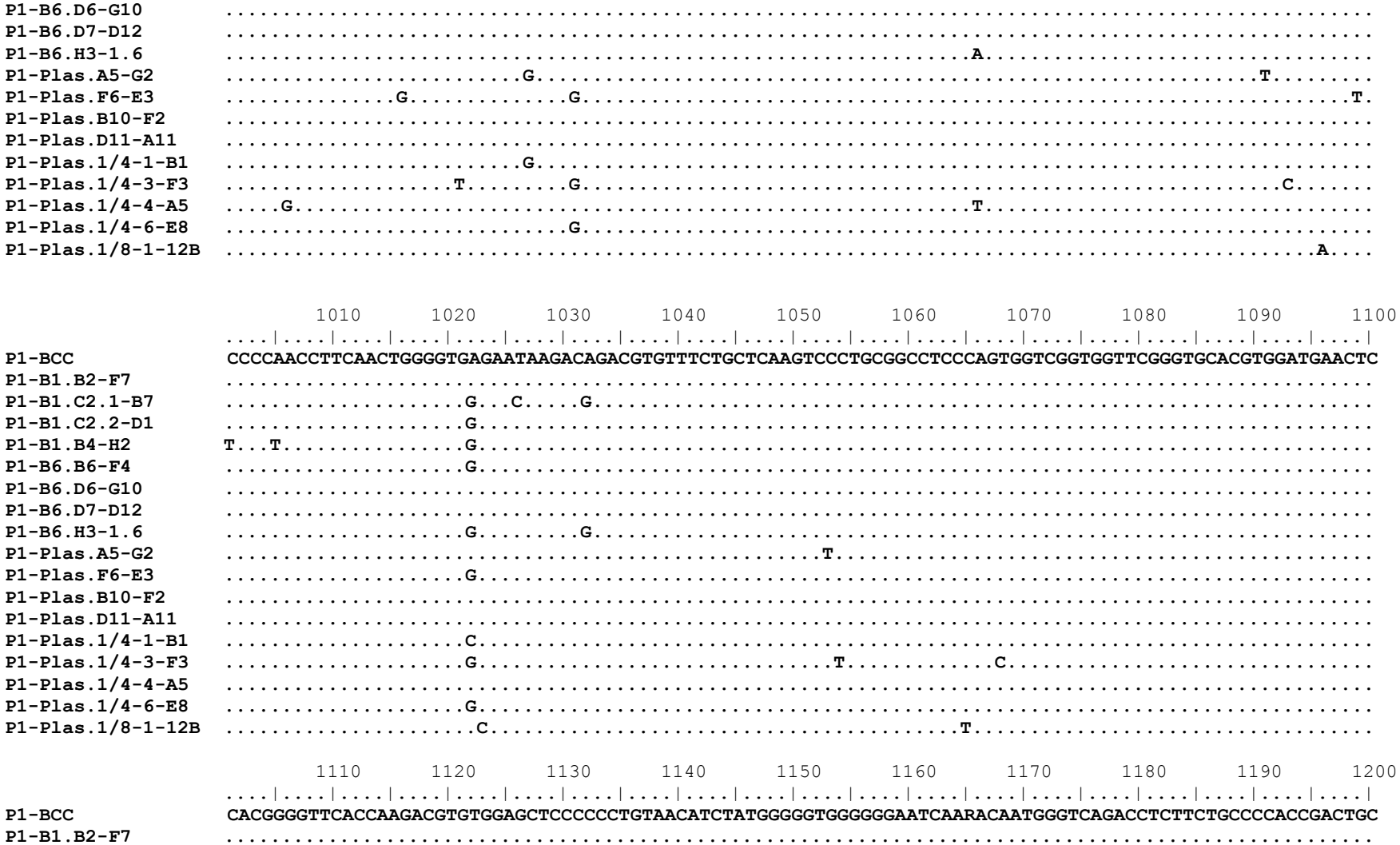

410 420 430 440 450 460 470 480 490 500
P1-BCC ACTGGTCCCCCGCTCTGGGTATGGTGGTGGCGCACGTCCTGCGTCTGCCCCAGACCTTGTTCGACATAATCGCCGGGGCYCATTGGGGCATCTTGGCGGG
P1-B1.B2-F7
P1-B1.C2.1-B7
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.B6-F4
P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5
P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B
510 520 530 540 550 560 570 580 590 600
P1-BCC CCTAGCCTATTACTCCATGCAGGGCAACTGGGCTAAGGTTGCTATTATAGTGGTTATGTTTTCAGGGGTCGATGCCCATACAYATACCACCGGTAGCAKT
P1-B1.B2-F7
P1-B1.C2.1-B7
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.B6-F4
P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5


P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B
610 620 630 640 650 660 670 680 690 700
P1-BCC SCAGCTCGTGGTGCCAGTATGATTGCCAGCATKTTTGCCCCGGGCGCCCGACAGAACYTGCAGCTGGTCAACACCAATGGCTCGTGGCACATCAACAGGA
P1-B1.B2-F7
P1-B1.C2.1-B7
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.B6-F4
P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5
P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B
710 720 730 740 750 760 770 780 790 800
P1-BCC CTGCCCTGAACTGTAATGATTCCATAAATACAGGGTTCATAGCTGGGTTGTTTTATTTCCATAAGTTCAACTCTACTGGATGTCCTCAAAGGCTCAGCAG
P1-B1.B2-F7
P1-B1.C2.1-B7
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.B6-F4
P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1

```
P1-Plas.1/4-3-F3   ................................................................................
P1-Plas.1/4-4-A5   ..............................................................C.................
P1-Plas.1/4-6-E8   ................................................................................
P1-Plas.1/8-1-12B  ................................................................................
```

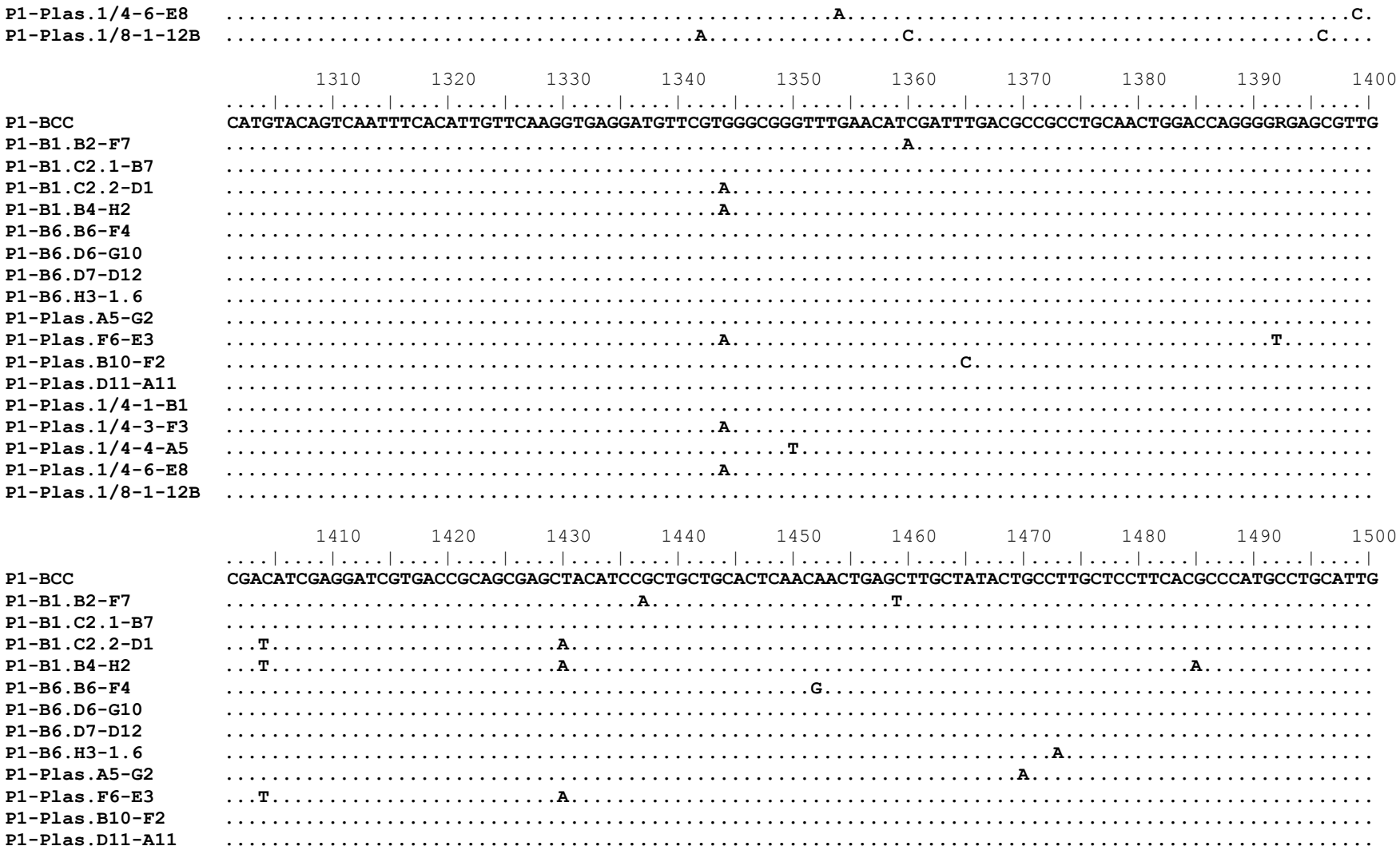
```
                          810       820       830       840       850       860       870       880       890       900
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P1-BCC            CTGCAAGCCCATCACTGCCTTTAGGCAGGGGTGGGGCTCCTTGACAGATGCTAACATCAGCGGTTCCTCTGAAGACAAACCATACTGCTGGCACTACGCA
P1-B1.B2-F7       ................................................................................................
P1-B1.C2.1-B7     ................................................................................................
P1-B1.C2.2-D1     ................................................................................................
P1-B1.B4-H2       ..........................................................T.....................................
P1-B6.B6-F4       ................................................................................................
P1-B6.D6-G10      ................................................................................................
P1-B6.D7-D12      ................................................................................................
P1-B6.H3-1.6      ................................................................................................
P1-Plas.A5-G2     ..................................................C.............................................
P1-Plas.F6-E3     ..........................................C.....................T.....................................
P1-Plas.B10-F2    ................................................................................................
P1-Plas.D11-A11   ................................................................................................
P1-Plas.1/4-1-B1  ................................................................................................
P1-Plas.1/4-3-F3  ..........................................G.......................T............R.....................
P1-Plas.1/4-4-A5  ..............................................................................................T...
P1-Plas.1/4-6-E8  ..........................................................T.....................................T...
P1-Plas.1/8-1-12B ................................................................................................
```

```
                          910       920       930       940       950       960       970       980       990       1000
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P1-BCC            CCTAGACCTTGTGAYATTGTCCCAGCARMACATGTCTGCGGCCCTGTGTACTGCTTCACACCATCGCCAGTGGTTGTAGGCACTACTGAYACTAAGGGCG
P1-B1.B2-F7       ................................................................................................
P1-B1.C2.1-B7     ..............................................................................C.................
P1-B1.C2.2-D1     ............................G.......T.............................T.....C.......................
P1-B1.B4-H2       ............................G.....................................................C.......
P1-B6.B6-F4       ................................................................................................
```

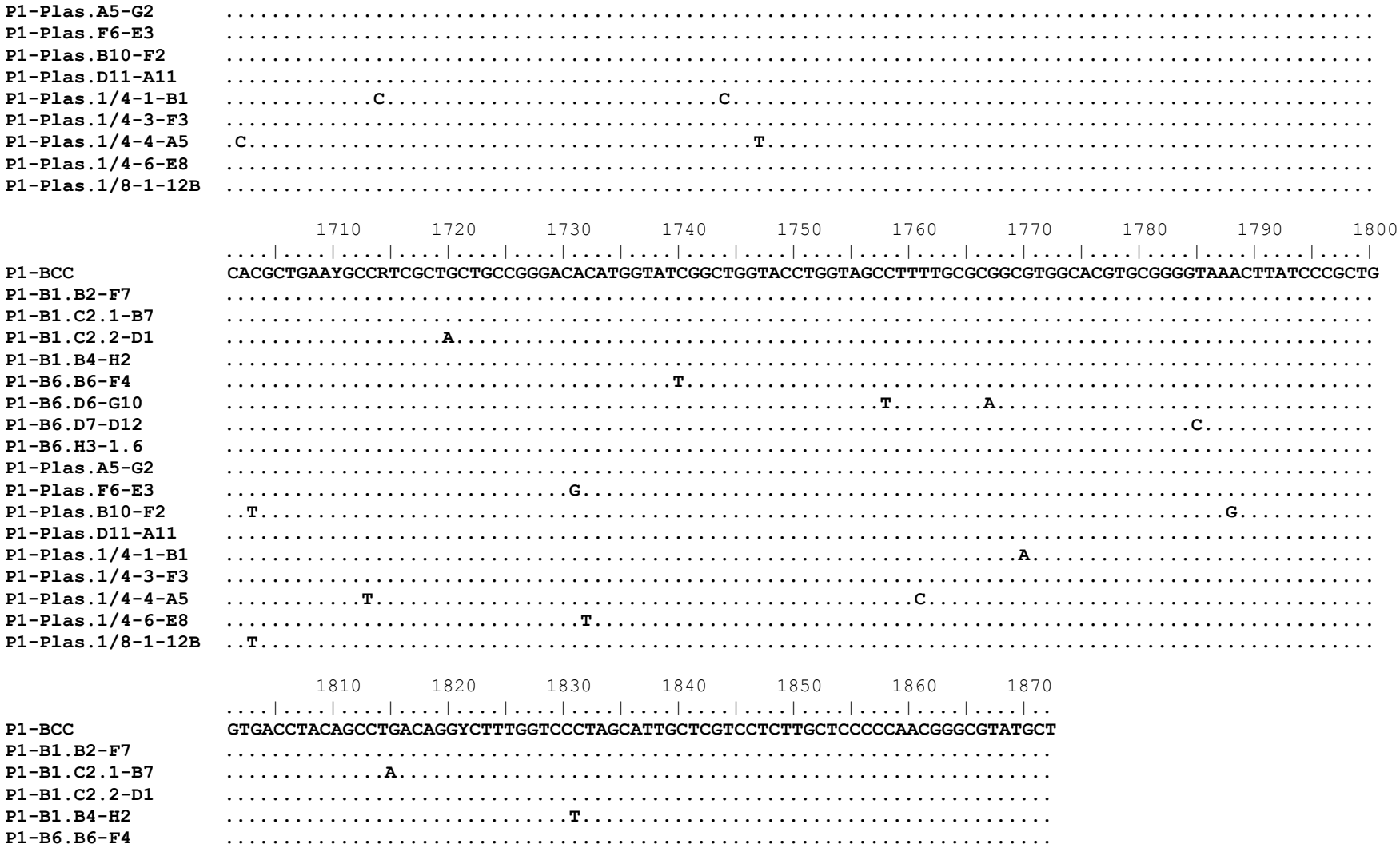

P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5
P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B
1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
P1-BCC CCCCAACCTTCAACTGGGGTGAGAATAAGACAGACGTGTTTCTGCTCAAGTCCCTGCGGCCTCCCAGTGGTCGGTGGTTCGGGTGCACGTGGATGAACTC
P1-B1.B2-F7
P1-B1.C2.1-B7
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.B6-F4
P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5
P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B
1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
P1-BCC CACGGGGTTCACCAAGACGTGTGGAGCTCCCCCCTGTAACATCTATGGGGGTGGGGGGAATCAARACAATGGGTCAGACCTCTTCTGCCCCACCGACTGC
P1-B1.B2-F7

```
P1-B1.C2.1-B7      .....................C.....................................................................
P1-B1.C2.2-D1      .....................................T.....................................A.................
P1-B1.B4-H2        .....................................T.....................................A.................
P1-B6.B6-F4        ..........................................................................G......C.................
P1-B6.D6-G10       ...........................................................................................
P1-B6.D7-D12       ................................T.....................................G.....................
P1-B6.H3-1.6       .....................C.....................................................................
P1-Plas.A5-G2      ...........................................................................................
P1-Plas.F6-E3      .....................................T.....................................A.................
P1-Plas.B10-F2     ...........................................................................................
P1-Plas.D11-A11    ...........................................................................................
P1-Plas.1/4-1-B1   ...........................................................................................
P1-Plas.1/4-3-F3   .....................................T.....................................A.................
P1-Plas.1/4-4-A5   ..........................................................G....................................
P1-Plas.1/4-6-E8   .....................................T.....................A.....................A.................
P1-Plas.1/8-1-12B  ...........................................................................................
```

```
                            1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P1-BCC             TTCAGGAAACATCCTGGGGCCACATACAGCCGGTGTGGTGCGGGGCCCTGGTTGACACCTCGATGCATGGTCGACTACCCATACCGGCTTTGGCATTATC
P1-B1.B2-F7        ...........................................................................................
P1-B1.C2.1-B7      ...........................................................................................
P1-B1.C2.2-D1      ...........................................................................................C.
P1-B1.B4-H2        ...........................................................................................C.
P1-B6.B6-F4        ..........................................A................................................
P1-B6.D6-G10       ...........................................................................................
P1-B6.D7-D12       ...........................................................................................
P1-B6.H3-1.6       ...........................................................................................
P1-Plas.A5-G2      ...........................................................................................
P1-Plas.F6-E3      ...........................................................................................C.
P1-Plas.B10-F2     ...........................................................................................
P1-Plas.D11-A11    ..T........................................................................................
P1-Plas.1/4-1-B1   ...........................................................................................
P1-Plas.1/4-3-F3   ...........................................................................................C.
P1-Plas.1/4-4-A5   ...........................................................................................
```


P1-Plas.1/4-6-E8 .....A.....C.
P1-Plas.1/8-1-12B .....A.....C.....C....
1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
P1-BCC CATGTACAGTCAATTTCACATTGTTCAAGGTGAGGATGTTCGTGGGCGGGTTTGAACATCGATTTGACGCCGCCTGCAACTGGACCAGGGGRGAGCGTTG
P1-B1.B2-F7 .....A.....
P1-B1.C2.1-B7 .....
P1-B1.C2.2-D1 .....A.....
P1-B1.B4-H2 .....A.....
P1-B6.B6-F4 .....
P1-B6.D6-G10 .....
P1-B6.D7-D12 .....
P1-B6.H3-1.6 .....
P1-Plas.A5-G2 .....
P1-Plas.F6-E3 .....A.....T.....
P1-Plas.B10-F2 .....C.....
P1-Plas.D11-A11 .....
P1-Plas.1/4-1-B1 .....
P1-Plas.1/4-3-F3 .....A.....
P1-Plas.1/4-4-A5 .....T.....
P1-Plas.1/4-6-E8 .....A.....
P1-Plas.1/8-1-12B .....
1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
P1-BCC CGACATCGAGGATCGTGACCGCAGCGAGCTACATCCGCTGCTGCACTCAACAACTGAGCTTGCTATACTGCCTTGCTCCTTCACGCCCATGCCTGCATTG
P1-B1.B2-F7 .....A.....T.....
P1-B1.C2.1-B7 .....
P1-B1.C2.2-D1 ...T.....A.....
P1-B1.B4-H2 ...T.....A.....A.....
P1-B6.B6-F4 .....G.....
P1-B6.D6-G10 .....
P1-B6.D7-D12 .....
P1-B6.H3-1.6 .....A.....
P1-Plas.A5-G2 .....A.....
P1-Plas.F6-E3 ...T.....A.....
P1-Plas.B10-F2 .....
P1-Plas.D11-A11 .....

```
P1-Plas.1/4-1-B1   ..........................A.....................................................................
P1-Plas.1/4-3-F3   ...T.......................A.....................................................................
P1-Plas.1/4-4-A5   .....................................................................................
P1-Plas.1/4-6-E8   ...T.......................A.....................................................................
P1-Plas.1/8-1-12B  .....................................................................................

                          1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P1-BCC             TCGACAGGTCTAATACACCTCCACCAAAACATCGTGGATGTCCAATACCTTTATGGTGTTGGATCYGGCATGGTGGGATGGGTGTTGAAATGGGAGTTTG
P1-B1.B2-F7        .....................................................................................
P1-B1.C2.1-B7      .....................................................................................
P1-B1.C2.2-D1      .....................................................................................
P1-B1.B4-H2        .....................................................................................
P1-B6.B6-F4        .....................................................................................
P1-B6.D6-G10       .....................................................................................
P1-B6.D7-D12       .....................................................................................
P1-B6.H3-1.6       .....................................................................................
P1-Plas.A5-G2      .....................................................................................
P1-Plas.F6-E3      ..........................................................C.....C.....................
P1-Plas.B10-F2     .....................................................................................
P1-Plas.D11-A11    .....................................................................................
P1-Plas.1/4-1-B1   ...................................................................................C.
P1-Plas.1/4-3-F3   ..............................................................................C......
P1-Plas.1/4-4-A5   ...........................................................................A.........
P1-Plas.1/4-6-E8   ....................A.....................................................G.............
P1-Plas.1/8-1-12B  ................................A....................................................

                          1610      1620      1630      1640      1650      1660      1670      1680      1690      1700
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P1-BCC             TTATCCTCGTTTTTCTCCTCCTAGCAGACGCACGCGTGTGCGTTGCCCTTTGGCTAATGCTGATGATATCACAAGCAGAAGCAGCCTTGGAGAACCTTGT
P1-B1.B2-F7        ............C........................................................................
P1-B1.C2.1-B7      ................................G....................................................
P1-B1.C2.2-D1      ............C........................................................................
P1-B1.B4-H2        .....................................................................................
P1-B6.B6-F4        .....................................................................................
P1-B6.D6-G10       .....................................................................................
P1-B6.D7-D12       ..............................................................................G......
P1-B6.H3-1.6       .....................................................................................
```

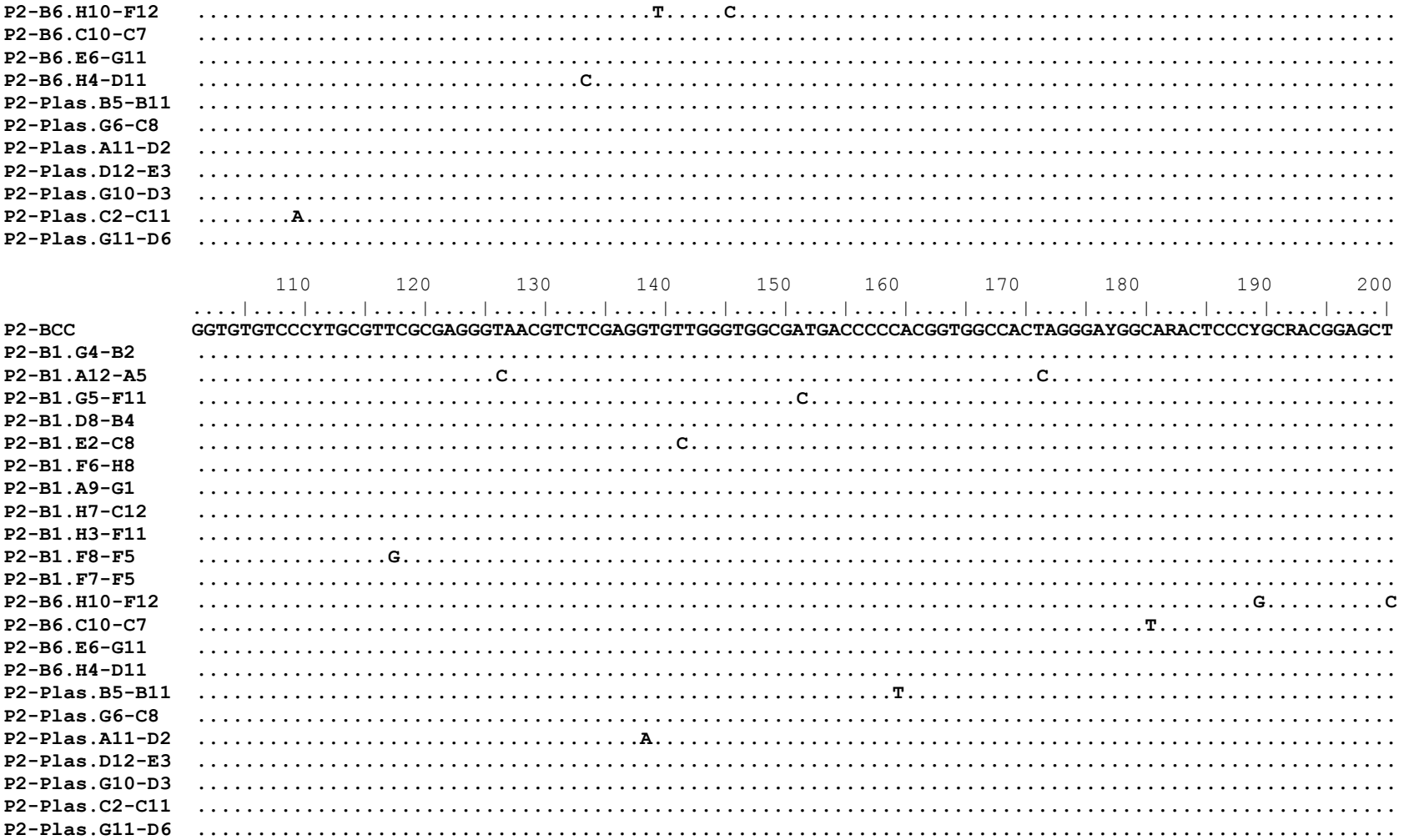

P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5
P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B
1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
P1-BCC CACGCTGAAYGCCRTCGCTGCTGCCGGGACACATGGTATCGGCTGGTACCTGGTAGCCTTTTGCGCGGCGTGGCACGTGCGGGGTAAACTTATCCCGCTG
P1-B1.B2-F7
P1-B1.C2.1-B7
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.B6-F4
P1-B6.D6-G10
P1-B6.D7-D12
P1-B6.H3-1.6
P1-Plas.A5-G2
P1-Plas.F6-E3
P1-Plas.B10-F2
P1-Plas.D11-A11
P1-Plas.1/4-1-B1
P1-Plas.1/4-3-F3
P1-Plas.1/4-4-A5
P1-Plas.1/4-6-E8
P1-Plas.1/8-1-12B
1810 1820 1830 1840 1850 1860 1870
P1-BCC GTGACCTACAGCCTGACAGGYCTTTGGTCCCTAGCATTGCTCGTCCTCTTGCTCCCCCAACGGGCGTATGCT
P1-B1.B2-F7
P1-B1.C2.1-B7
P1-B1.C2.2-D1
P1-B1.B4-H2
P1-B6.B6-F4

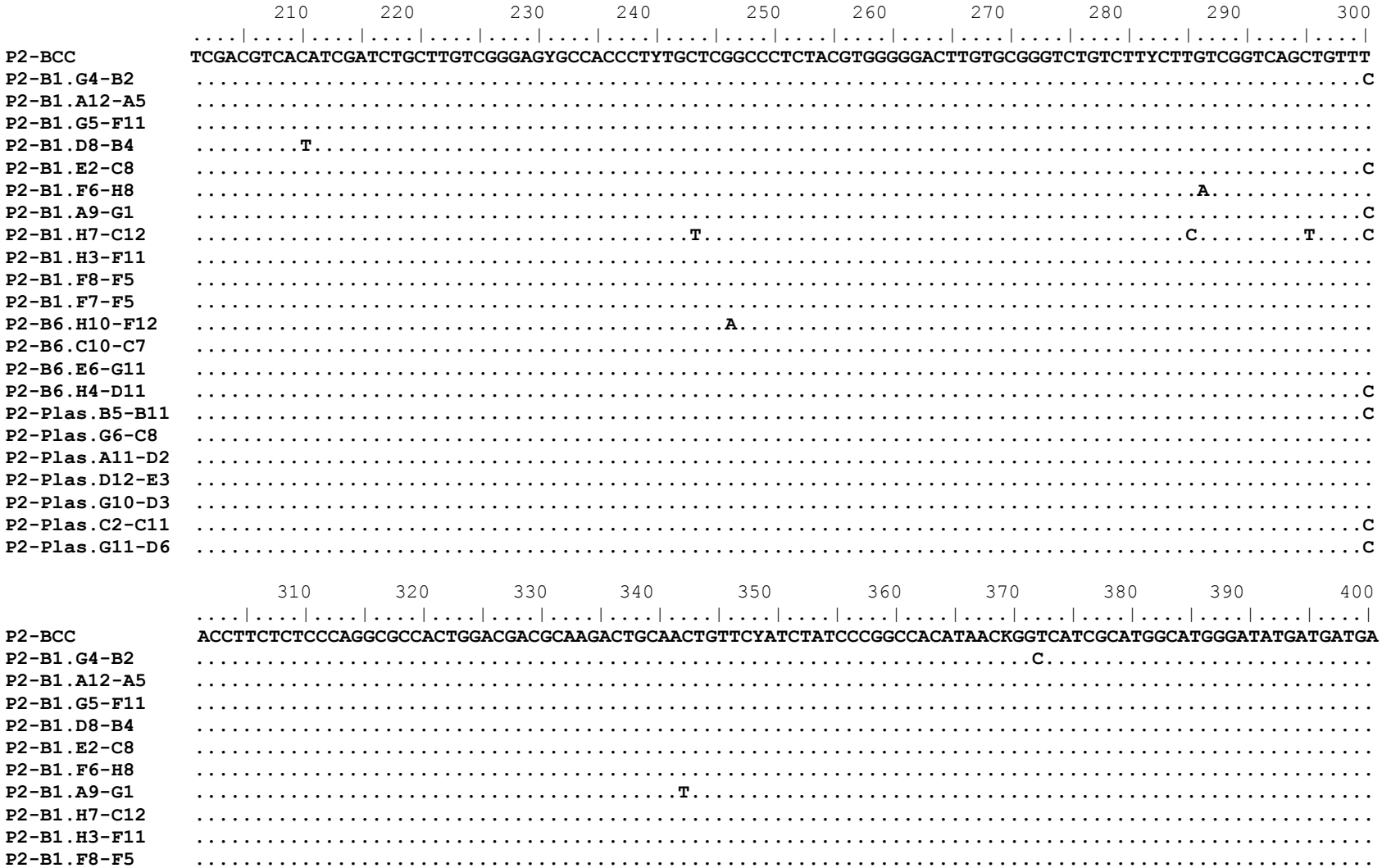
```
P1-B6.D6-G10        ............................................................
P1-B6.D7-D12        ............................................................
P1-B6.H3-1.6        ...................................................A........
P1-Plas.A5-G2       ............T................................................
P1-Plas.F6-E3       ............................................................
P1-Plas.B10-F2      ............................................................
P1-Plas.D11-A11     ............................................................
P1-Plas.1/4-1-B1    ............................................................
P1-Plas.1/4-3-F3    ............................................................
P1-Plas.1/4-4-A5    ................G.............T.............................
P1-Plas.1/4-6-E8    ...........................................................A......
P1-Plas.1/8-1-12B   ............................................................
```

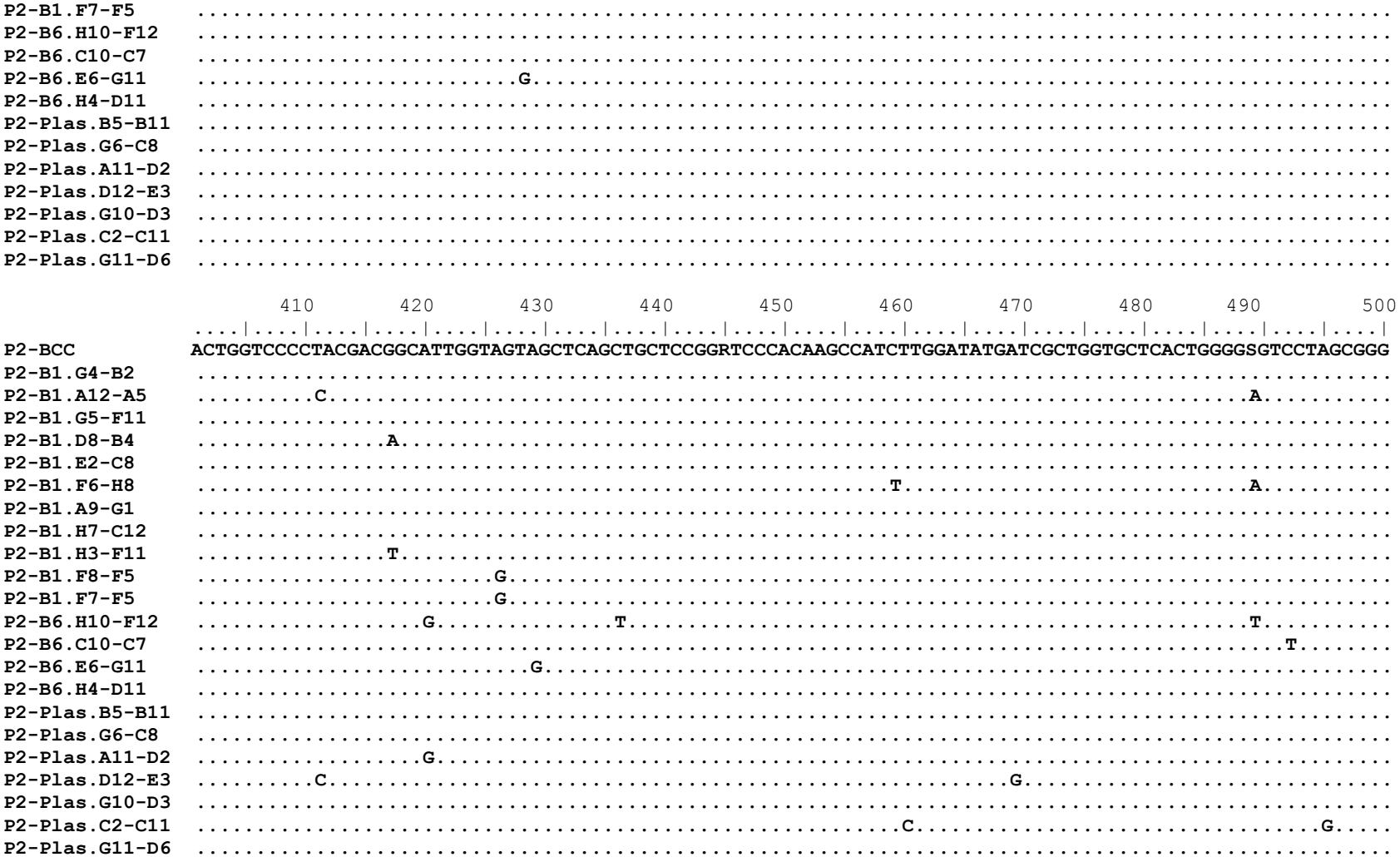
10.1.2 Patient 2

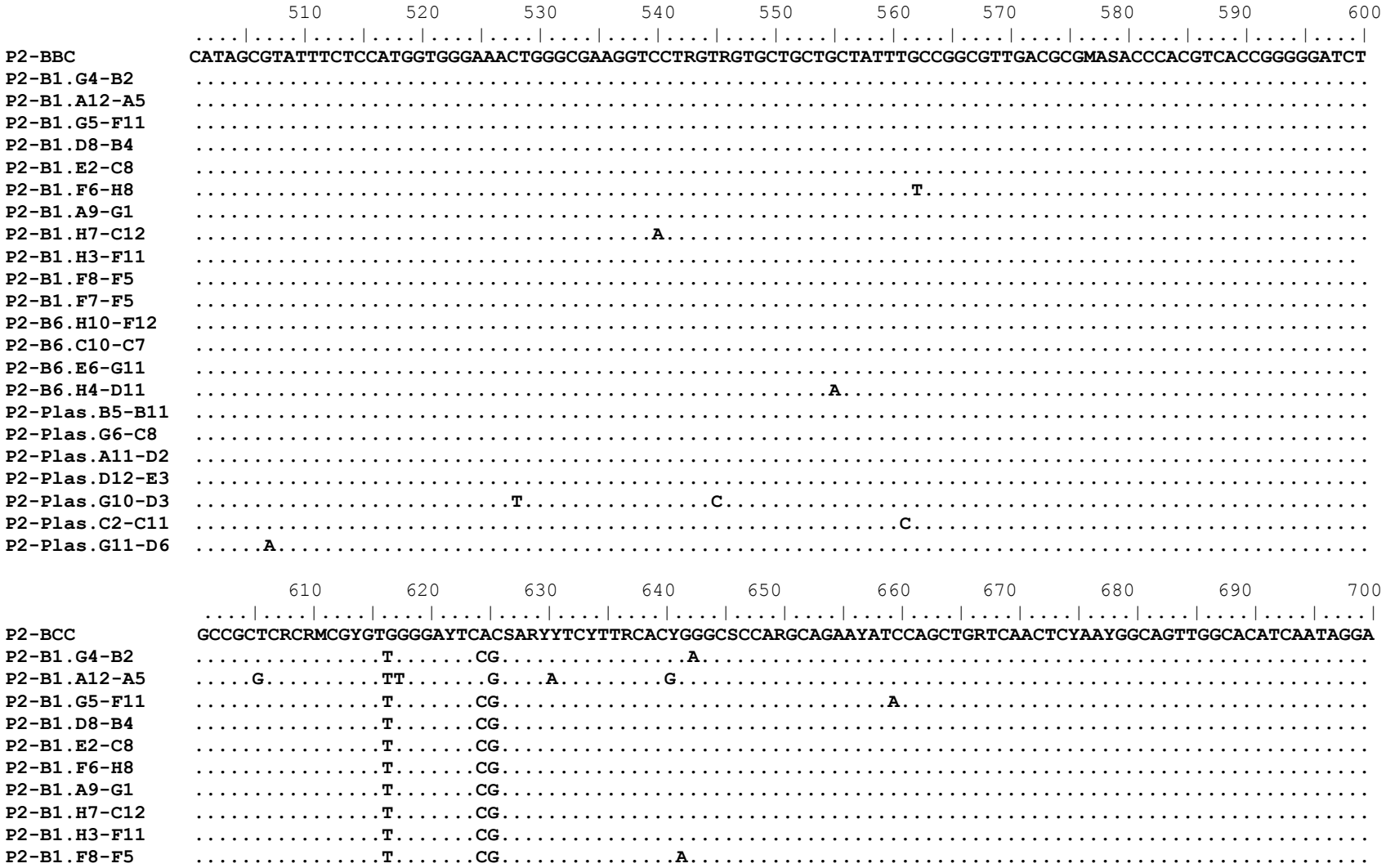
Gene position: E1: 1 - 576nt, E2: 577 - 1665nt, HVR1: 577 – 657nt.

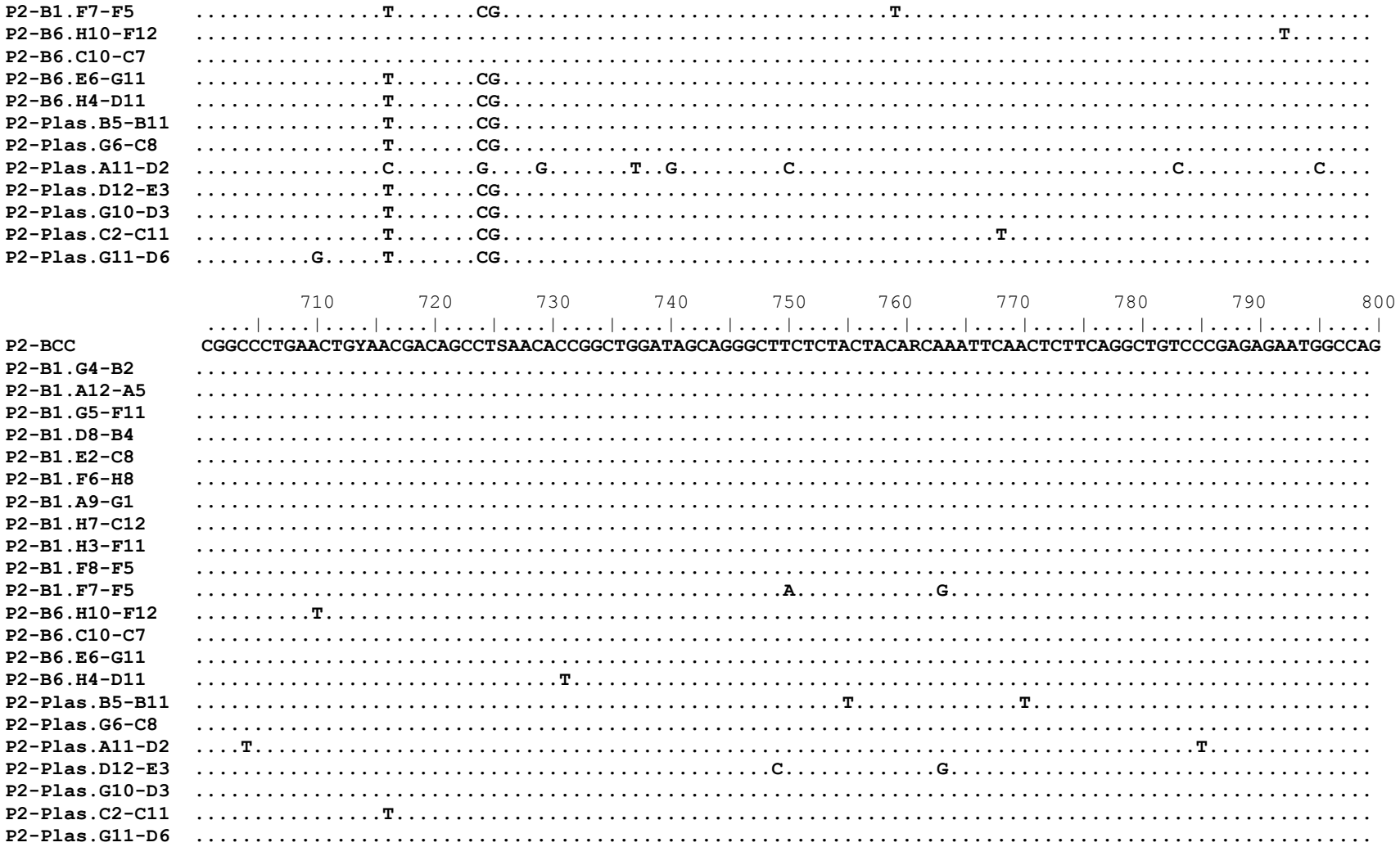
```
                          10        20        30        40        50        60        70        80        90        100
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P2-BCC              TACCAAGTGCGCAAYTCCACGGGGCTTTACCATGTCACCAAYGATTGCCCTAACTCGAGTATTGTGTACGAGACAGCTGATGCCATCCTACACRCTCCGG
P2-B1.G4-B2         ..........................................................................................
P2-B1.A12-A5        ..........................................................................................
P2-B1.G5-F11        ..........................................................................................
P2-B1.D8-B4         ........A.................................................................................
P2-B1.E2-C8         ..........................................................................................
P2-B1.F6-H8         ...................................................C......................................
P2-B1.A9-G1         ..........................................................................................
P2-B1.H7-C12        ..........................................................................................
P2-B1.H3-F11        ..........................................................................................
P2-B1.F8-F5         ..........................................................................................
P2-B1.F7-F5         ..........................................................................................
```

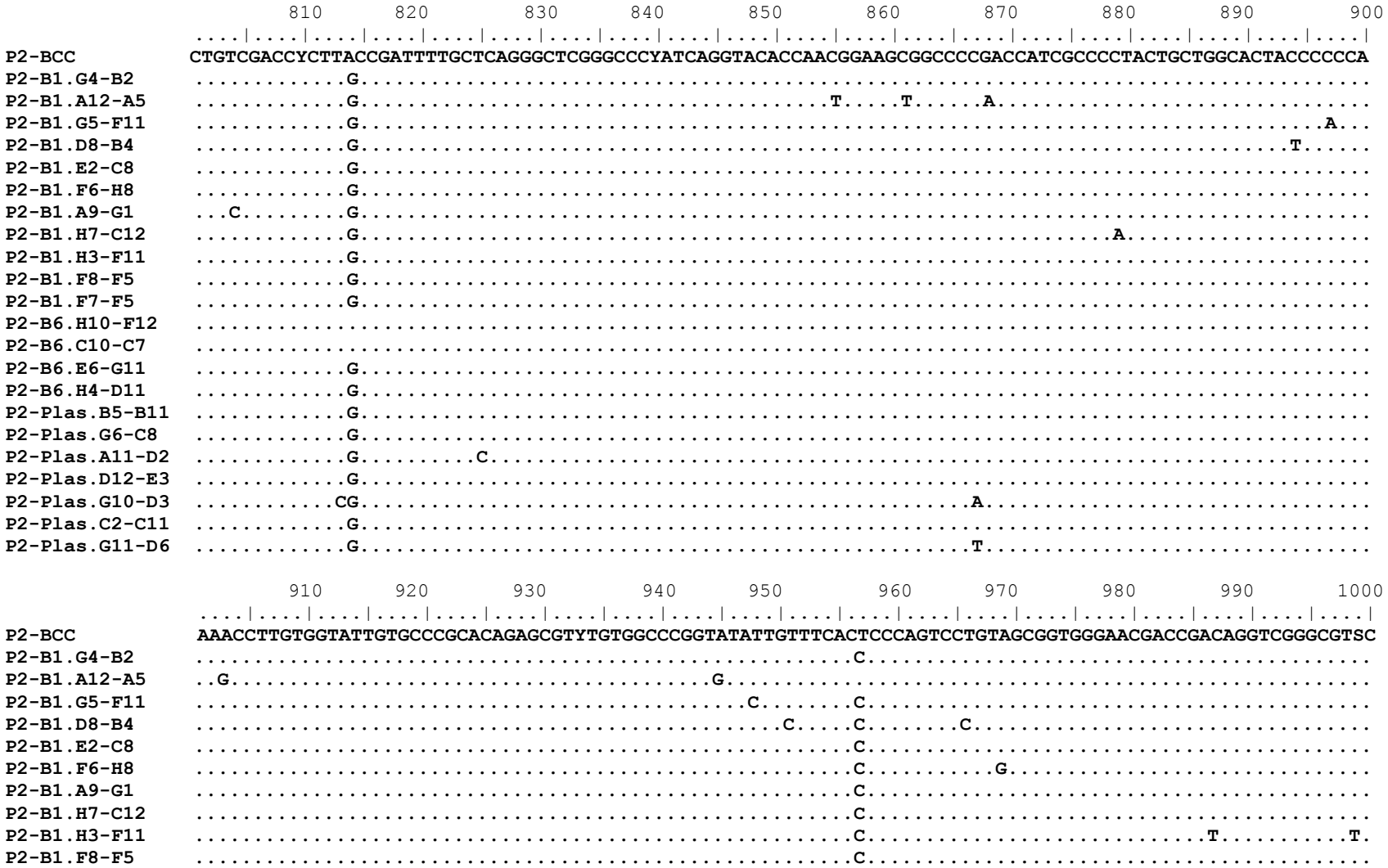

P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6
110
120
130
140
150
160
170
180
190
200
P2-BCC
GGTGTGTCCCYTGCGTTCGCGAGGGTAACGTCTCGAGGTGTTGGGTGGCGATGACCCCCACGGTGGCCACTAGGGAYGGCARACTCCCYGCRACGGAGCT
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5
P2-B1.F7-F5
P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6

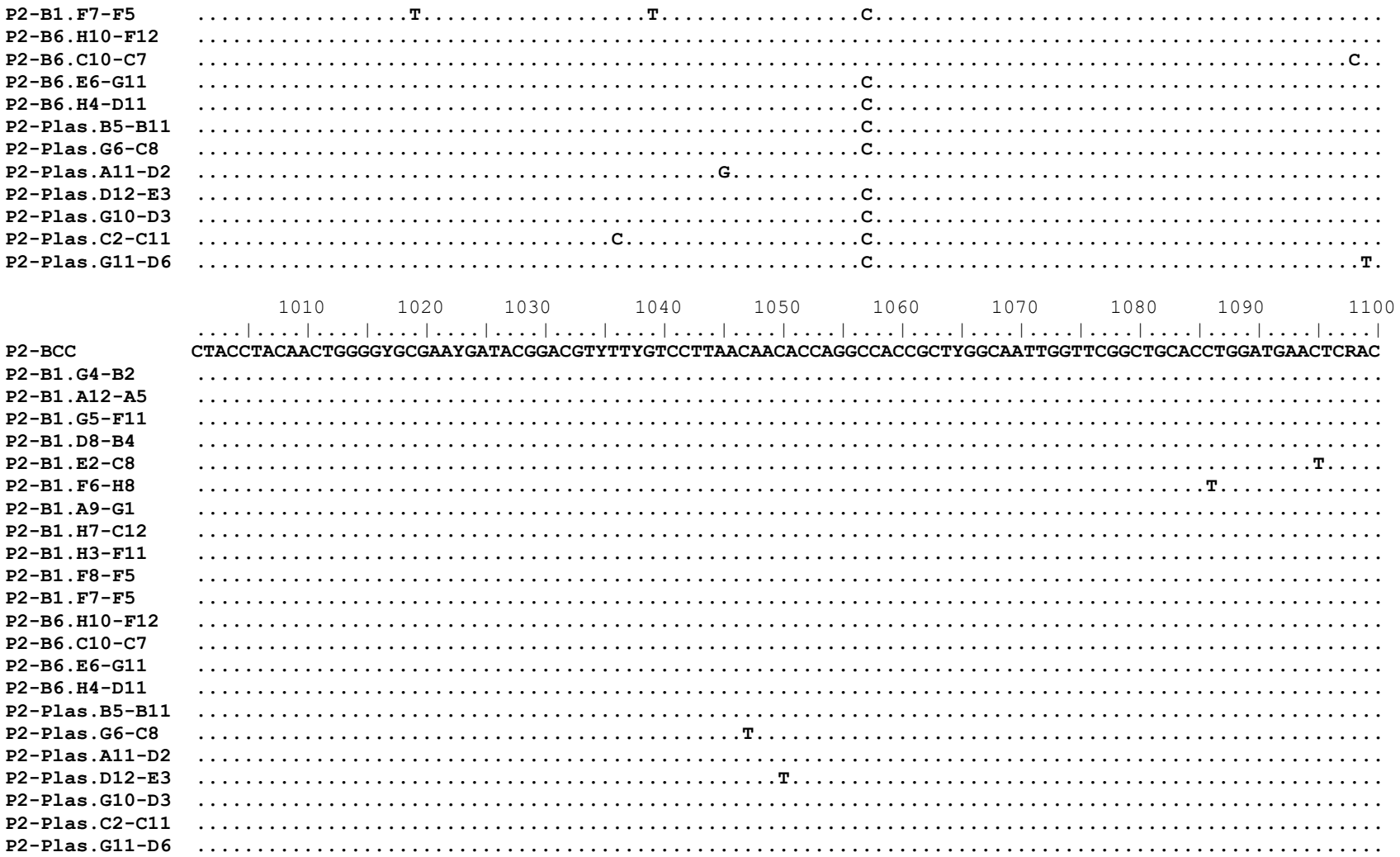

210 220 230 240 250 260 270 280 290 300
P2-BCC TCGACGTCACATCGATCTGCTTGTCGGGAGYGCCACCCTYTGCTCGGCCCTCTACGTGGGGGACTTGTGCGGGTCTGTCTTYCTTGTCGGTCAGCTGTTT
P2-B1.G4-B2 C
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4 T
P2-B1.E2-C8 C
P2-B1.F6-H8 A
P2-B1.A9-G1 C
P2-B1.H7-C12 T C T C
P2-B1.H3-F11
P2-B1.F8-F5
P2-B1.F7-F5
P2-B6.H10-F12 A
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11 C
P2-Plas.B5-B11 C
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11 C
P2-Plas.G11-D6 C
310 320 330 340 350 360 370 380 390 400
P2-BCC ACCTTCTCTCCCAGGCGCCACTGGACGACGCAAGACTGCAACTGTTCYATCTATCCCGGCCACATAACKGGTCATCGCATGGCATGGGATATGATGATGA
P2-B1.G4-B2 C
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1 T
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5

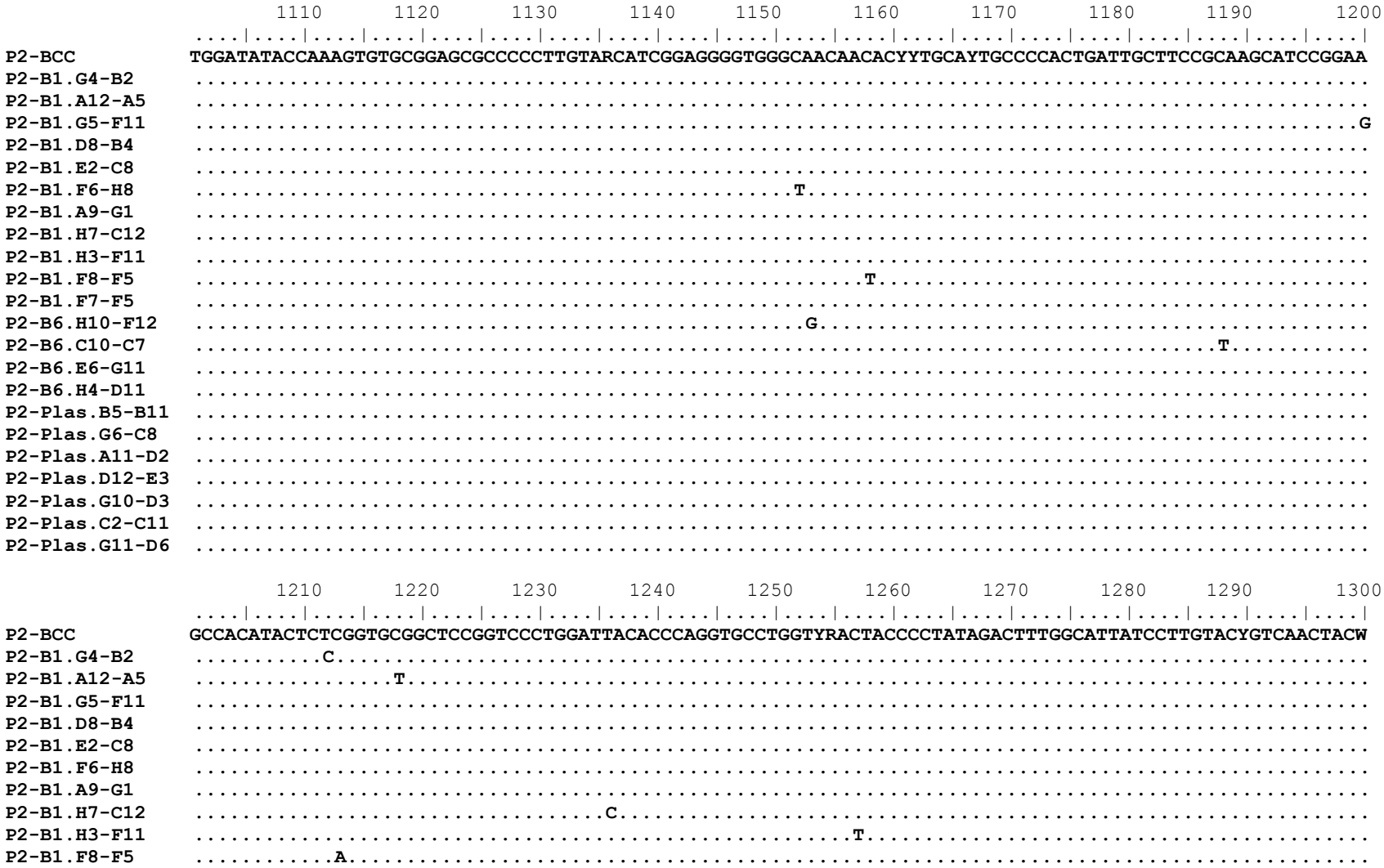

P2-B1.F7-F5
P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6
410
420
430
440
450
460
470
480
490
500
P2-BCC
ACTGGTCCCCTACGACGGCATTGGTAGTAGCTCAGCTGCTCCGGRTCCCACAAGCCATCTTGGATATGATCGCTGGTGCTCACTGGGGSGTCCTAGCGGG
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5
P2-B1.F7-F5
P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6

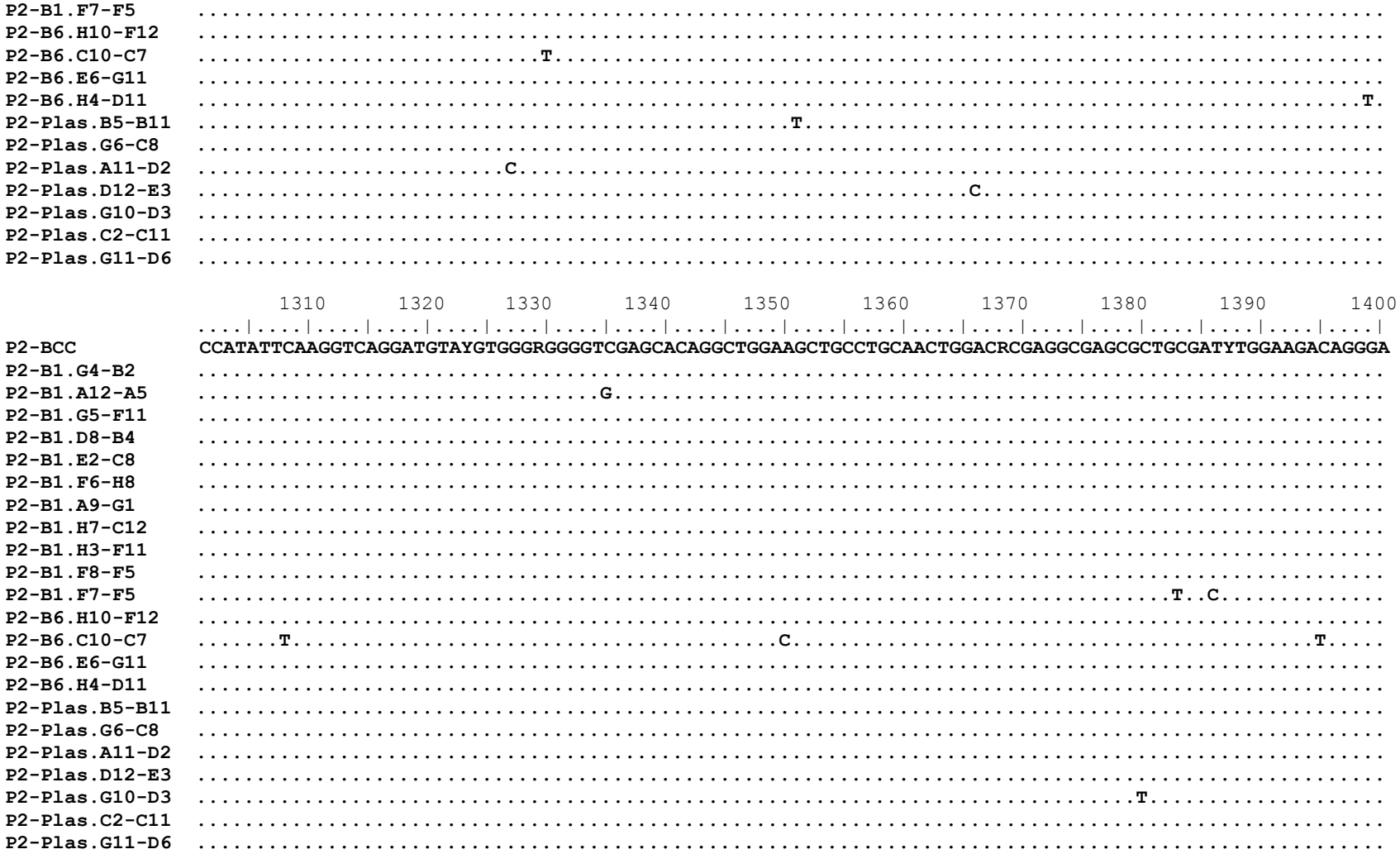

510 520 530 540 550 560 570 580 590 600
P2-BBC CATAGCGTATTTCTCCATGGTGGGAAACTGGGCGAAGGTCCTRGTRGTGCTGCTGCTATTTGCCGGCGTTGACGCGMASACCCACGTCACCGGGGGATCT
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5
P2-B1.F7-F5
P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6
610 620 630 640 650 660 670 680 690 700
P2-BCC GCCGCTCRCRMCGYGTGGGGAYTCACSARYYTCYTTRCACYGGGCSCCARGCAGAAYATCCAGCTGRTCAACTCYAAYGGCAGTTGGCACATCAATAGGA
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5

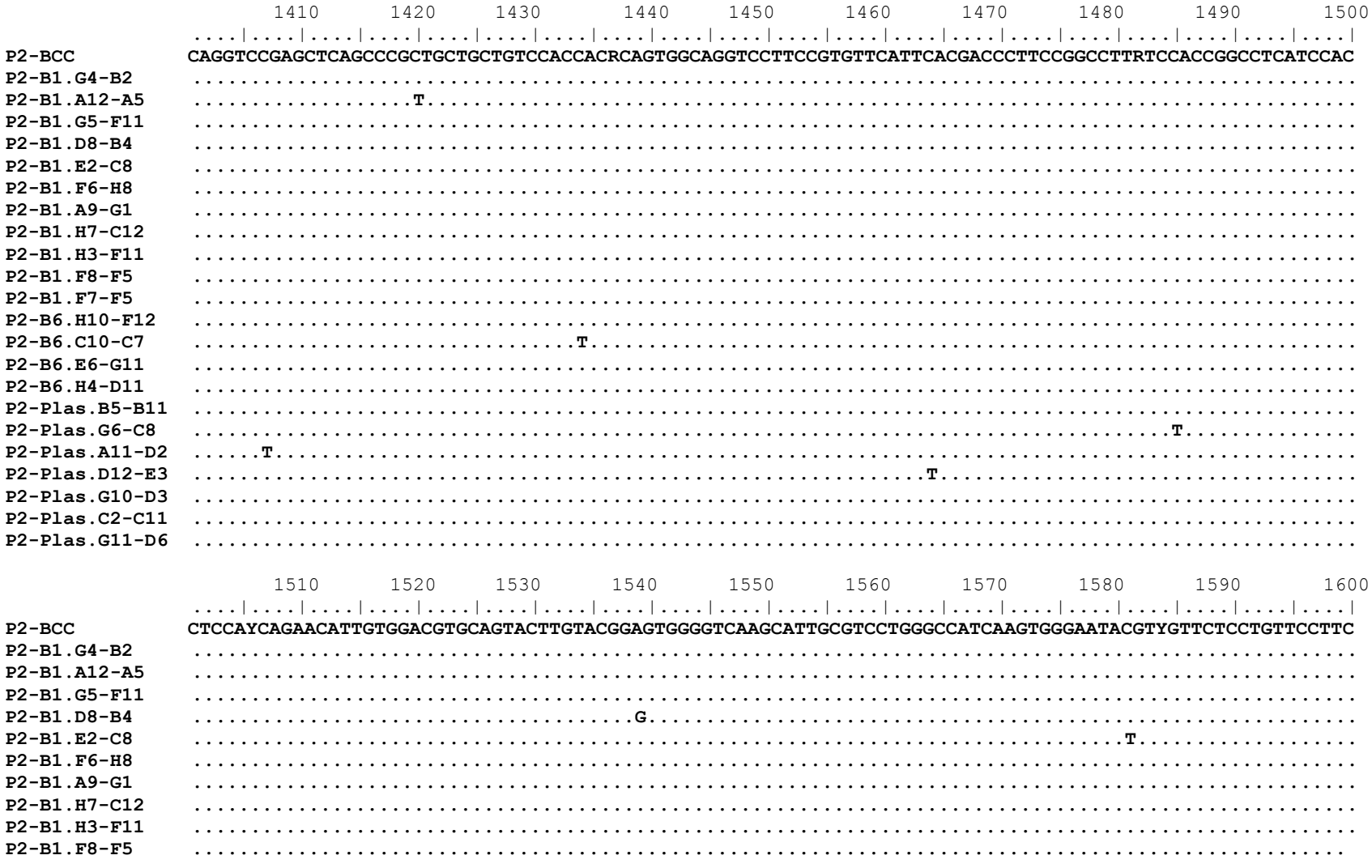

P2-B1.F7-F5 .....T.....CG.....T.....
P2-B6.H10-F12 .....T.....
P2-B6.C10-C7 .....
P2-B6.E6-G11 .....T.....CG.....
P2-B6.H4-D11 .....T.....CG.....
P2-Plas.B5-B11 .....T.....CG.....
P2-Plas.G6-C8 .....T.....CG.....
P2-Plas.A11-D2 .....C.....G....G.....T..G.....C.....C.....C....
P2-Plas.D12-E3 .....T.....CG.....
P2-Plas.G10-D3 .....T.....CG.....
P2-Plas.C2-C11 .....T.....CG.....T.....
P2-Plas.G11-D6 .....G.....T.....CG.....
710 720 730 740 750 760 770 780 790 800
P2-BCC CGGCCCTGAACTGYAACGACAGCCTSAACACCGGCTGGATAGCAGGGCTTCTCTACTACARCAAATTCAACTCTTCAGGCTGTCCCGAGAGAATGGCCAG
P2-B1.G4-B2 .....
P2-B1.A12-A5 .....
P2-B1.G5-F11 .....
P2-B1.D8-B4 .....
P2-B1.E2-C8 .....
P2-B1.F6-H8 .....
P2-B1.A9-G1 .....
P2-B1.H7-C12 .....
P2-B1.H3-F11 .....
P2-B1.F8-F5 .....
P2-B1.F7-F5 .....A.....G.....
P2-B6.H10-F12 .....T.....
P2-B6.C10-C7 .....
P2-B6.E6-G11 .....
P2-B6.H4-D11 .....T.....
P2-Plas.B5-B11 .....T.....T.....
P2-Plas.G6-C8 .....
P2-Plas.A11-D2 ....T.....T.....
P2-Plas.D12-E3 .....C.....G.....
P2-Plas.G10-D3 .....
P2-Plas.C2-C11 .....T.....
P2-Plas.G11-D6 .....

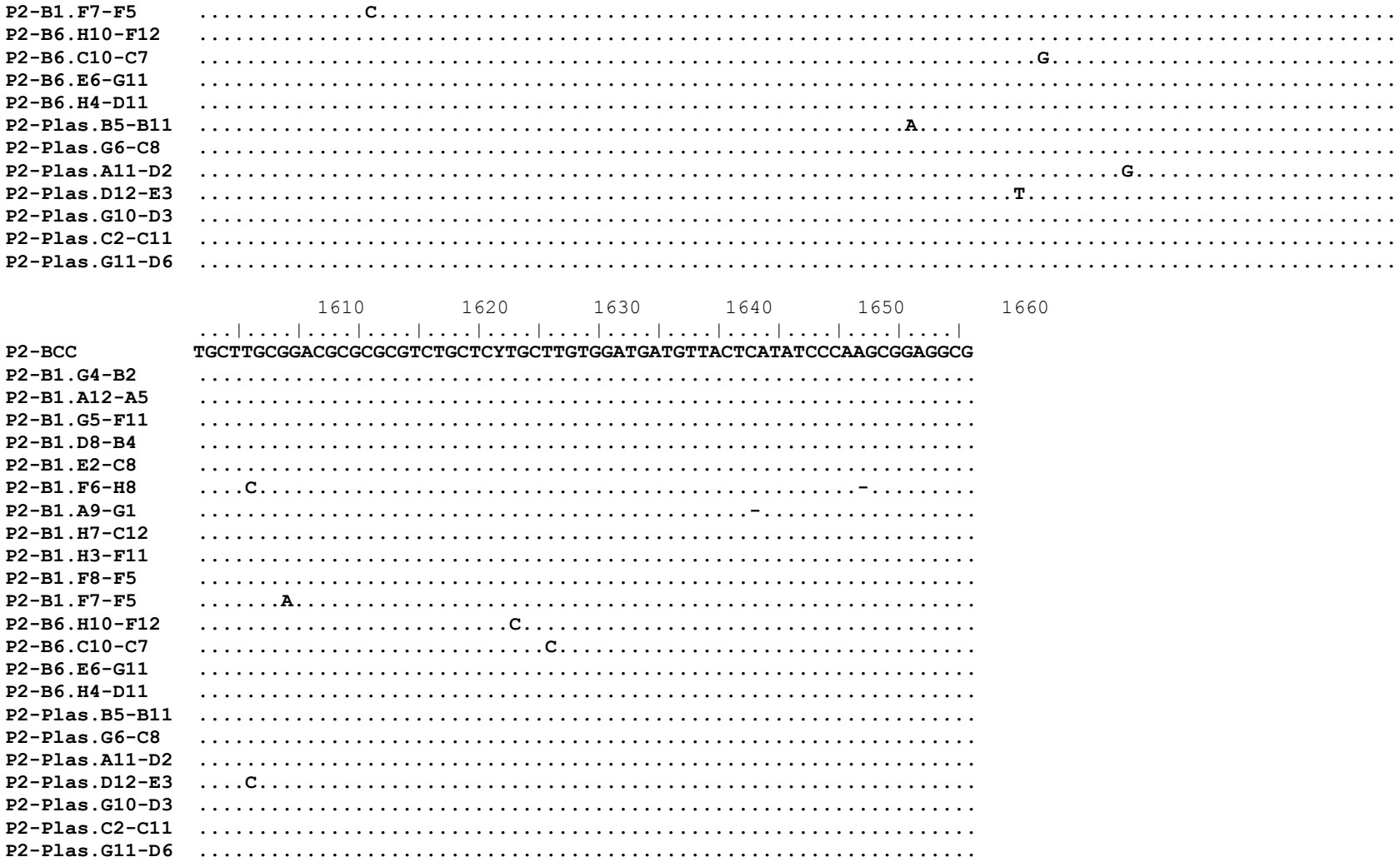

810 820 830 840 850 860 870 880 890 900
P2-BCC CTGTCGACCYCTTACCGATTTTGCTCAGGGCTCGGGCCCYATCAGGTACACCAACGGAAGCGGCCCCGACCATCGCCCCTACTGCTGGCACTACCCCCCA
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5
P2-B1.F7-F5
P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6
910 920 930 940 950 960 970 980 990 1000
P2-BCC AAACCTTGTGGTATTGTGCCCGCACAGAGCGTYTGTGGCCCGGTATATTGTTTCACTCCCAGTCCTGTAGCGGTGGGAACGACCGACAGGTCGGGCGTSC
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5


P2-B1.F7-F5 ..................T..................T.................C.....
P2-B6.H10-F12 .....
P2-B6.C10-C7 .....C..
P2-B6.E6-G11 .....C.....
P2-B6.H4-D11 .....C.....
P2-Plas.B5-B11 .....C.....
P2-Plas.G6-C8 .....C.....
P2-Plas.A11-D2 ....G.....
P2-Plas.D12-E3 .....C.....
P2-Plas.G10-D3 .....C.....
P2-Plas.C2-C11 ..........C.....C.....
P2-Plas.G11-D6 .....C.....T.
1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
P2-BCC CTACCTACAACTGGGGYGCGAAYGATACGGACGTYTTYGTCCTTAACAACACCAGGCCACCGCTYGGCAATTGGTTCGGCTGCACCTGGATGAACTCRAC
P2-B1.G4-B2 .....
P2-B1.A12-A5 .....
P2-B1.G5-F11 .....
P2-B1.D8-B4 .....
P2-B1.E2-C8 .....T.....
P2-B1.F6-H8 .....T.....
P2-B1.A9-G1 .....
P2-B1.H7-C12 .....
P2-B1.H3-F11 .....
P2-B1.F8-F5 .....
P2-B1.F7-F5 .....
P2-B6.H10-F12 .....
P2-B6.C10-C7 .....
P2-B6.E6-G11 .....
P2-B6.H4-D11 .....
P2-Plas.B5-B11 .....
P2-Plas.G6-C8 ....T.....
P2-Plas.A11-D2 .....
P2-Plas.D12-E3 .....T.....
P2-Plas.G10-D3 .....
P2-Plas.C2-C11 .....
P2-Plas.G11-D6 .....


1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
P2-BCC TGGATATACCAAAGTGTGCGGAGCGCCCCCTTGTARCATCGGAGGGGTGGGCAACAACACYYTGCAYTGCCCCACTGATTGCTTCCGCAAGCATCCGGAA
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11 G
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8 T
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5 T
P2-B1.F7-F5
P2-B6.H10-F12 G
P2-B6.C10-C7 T
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6
1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
P2-BCC GCCACATACTCTCGGTGCGGCTCCGGTCCCTGGATTACACCCAGGTGCCTGGTYRACTACCCCTATAGACTTTGGCATTATCCTTGTACYGTCAACTACW
P2-B1.G4-B2 C
P2-B1.A12-A5 T
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12 C
P2-B1.H3-F11 T
P2-B1.F8-F5 A


P2-B1.F7-F5
P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6
1310
1320
1330
1340
1350
1360
1370
1380
1390
1400
P2-BCC
CCATATTCAAGGTCAGGATGTAYGTGGGRGGGGTCGAGCACAGGCTGGAAGCTGCCTGCAACTGGACRCGAGGCGAGCGCTGCGATYTGGAAGACAGGGA
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5
P2-B1.F7-F5
P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6

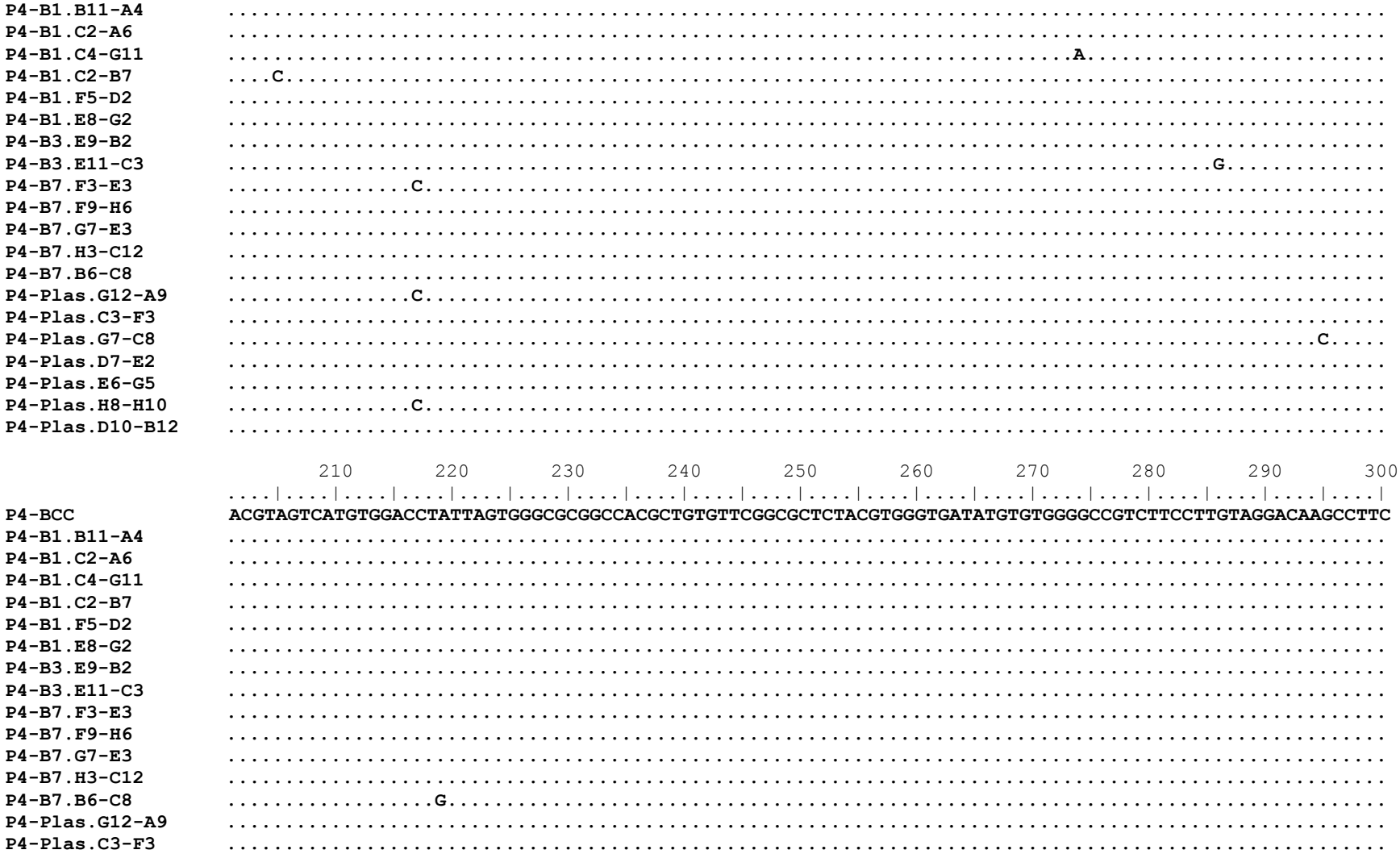

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
P2-BCC CAGGTCCGAGCTCAGCCCGCTGCTGCTGTCCACCACRCAGTGGCAGGTCCTTCCGTGTTCATTCACGACCCTTCCGGCCTTRTCCACCGGCCTCATCCAC
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5
P2-B1.F7-F5
P2-B6.H10-F12
P2-B6.C10-C7
P2-B6.E6-G11
P2-B6.H4-D11
P2-Plas.B5-B11
P2-Plas.G6-C8
P2-Plas.A11-D2
P2-Plas.D12-E3
P2-Plas.G10-D3
P2-Plas.C2-C11
P2-Plas.G11-D6
1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
P2-BCC CTCCAYCAGAACATTGTGGACGTGCAGTACTTGTACGGAGTGGGGTCAAGCATTGCGTCCTGGGCCATCAAGTGGGAATACGTYGTTCTCCTGTTCCTTC
P2-B1.G4-B2
P2-B1.A12-A5
P2-B1.G5-F11
P2-B1.D8-B4
P2-B1.E2-C8
P2-B1.F6-H8
P2-B1.A9-G1
P2-B1.H7-C12
P2-B1.H3-F11
P2-B1.F8-F5

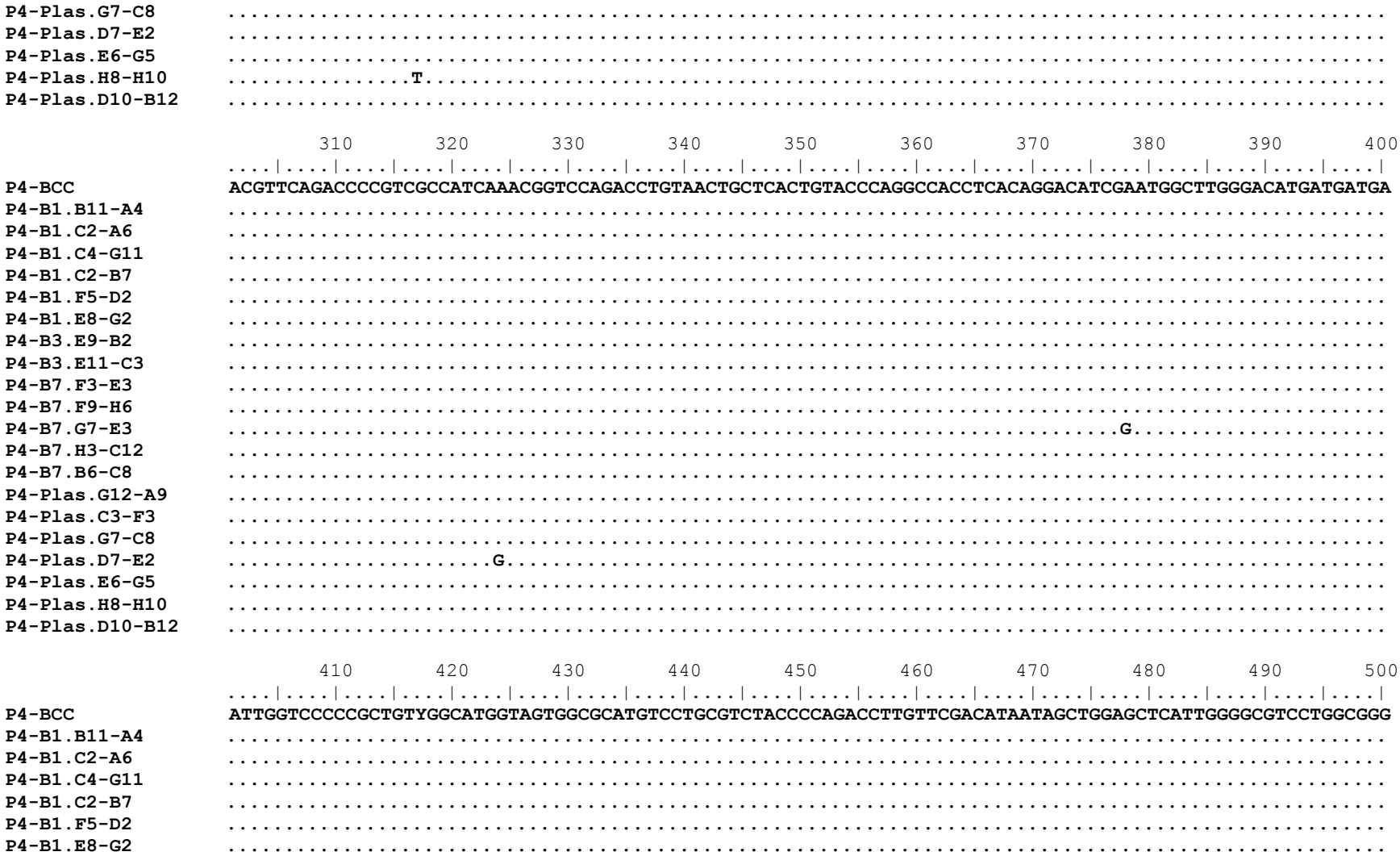

P2-B1.F7-F5 .....C.....
P2-B6.H10-F12 .....
P2-B6.C10-C7 .....G.....
P2-B6.E6-G11 .....
P2-B6.H4-D11 .....
P2-Plas.B5-B11 .....A.....
P2-Plas.G6-C8 .....
P2-Plas.A11-D2 .....G.....
P2-Plas.D12-E3 .....T.....
P2-Plas.G10-D3 .....
P2-Plas.C2-C11 .....
P2-Plas.G11-D6 .....
1610 1620 1630 1640 1650 1660
P2-BCC TGCTTGCGGACGCGCGCGTCTGCTCYTGCTTGTGGATGATGTTACTCATATCCCAAGCGGAGGCG
P2-B1.G4-B2 .....
P2-B1.A12-A5 .....
P2-B1.G5-F11 .....
P2-B1.D8-B4 .....
P2-B1.E2-C8 .....
P2-B1.F6-H8 ....C.....-.....
P2-B1.A9-G1 .....-.....
P2-B1.H7-C12 .....
P2-B1.H3-F11 .....
P2-B1.F8-F5 .....
P2-B1.F7-F5 .....A.....
P2-B6.H10-F12 .....C.....
P2-B6.C10-C7 .....C.....
P2-B6.E6-G11 .....
P2-B6.H4-D11 .....
P2-Plas.B5-B11 .....
P2-Plas.G6-C8 .....
P2-Plas.A11-D2 .....
P2-Plas.D12-E3 ....C.....
P2-Plas.G10-D3 .....
P2-Plas.C2-C11 .....
P2-Plas.G11-D6 .....

10.1.3 Patient 4

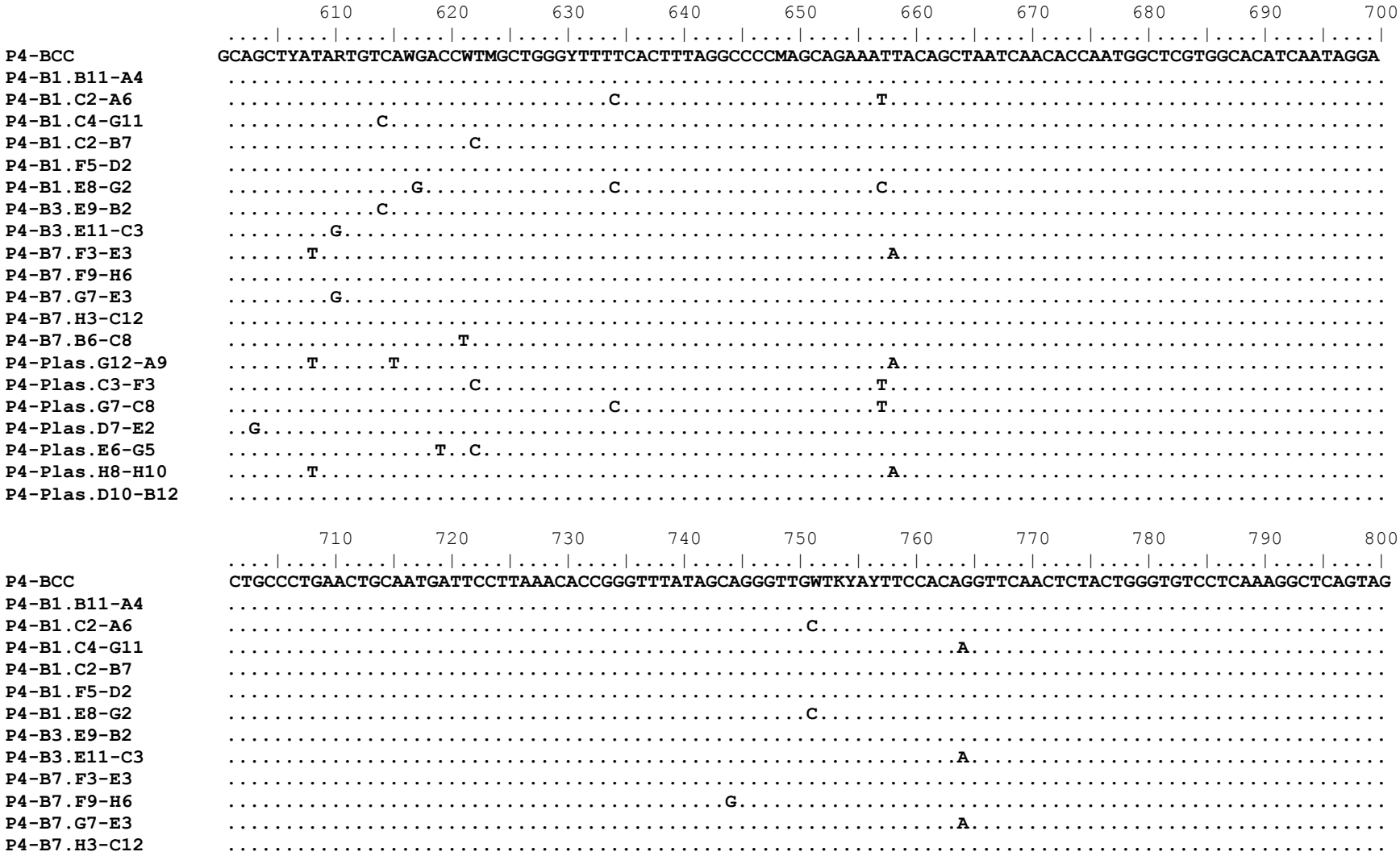
Gene position: E1: 1 - 576nt, E2: 577 - 1692nt, p7: 1689-1878nt, HVR1: 577 – 657nt, HVR494: 913 - 933nt, HVR595: 1156 - 1182nt.

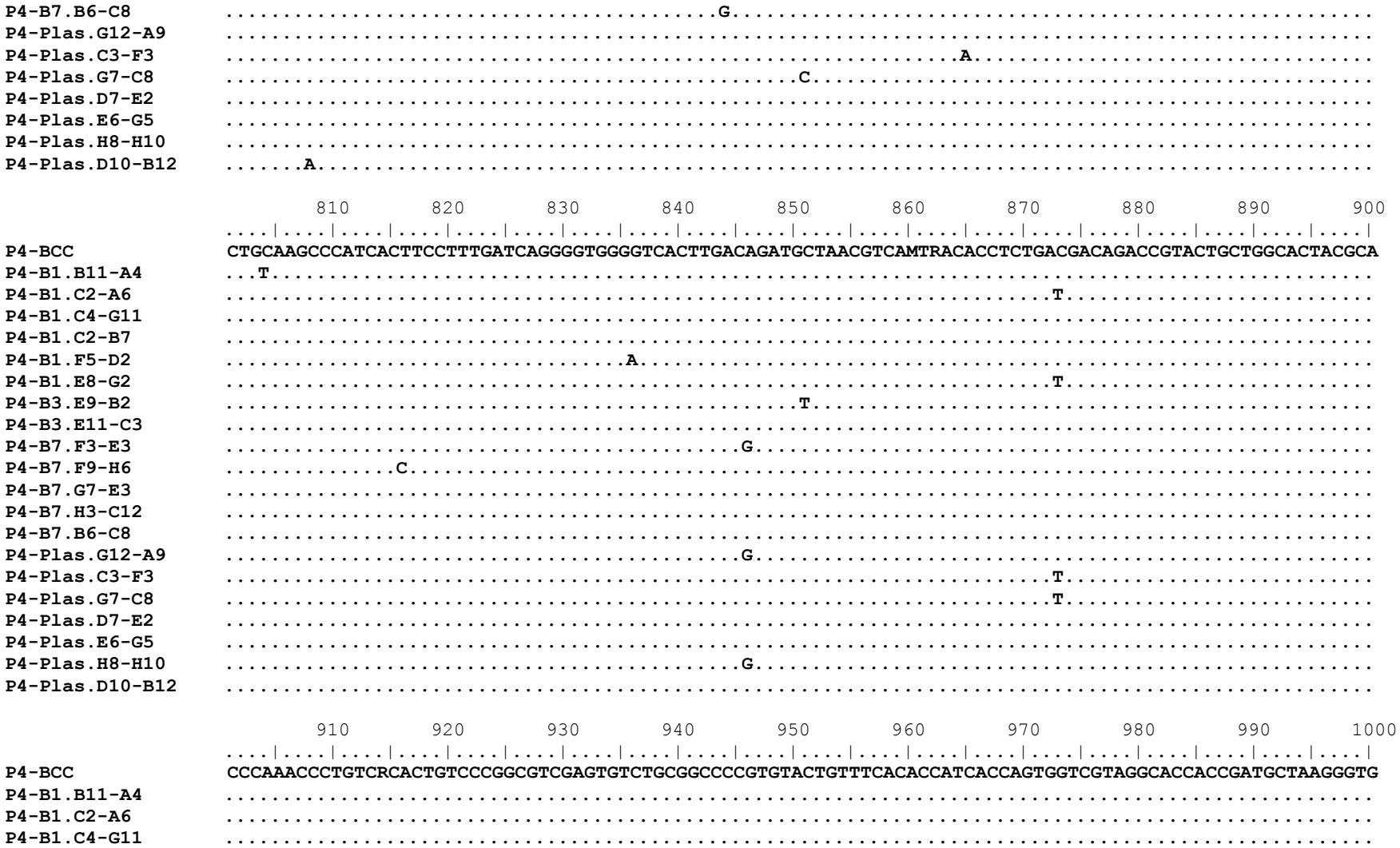
```
                        10        20        30        40        50        60        70        80        90       100
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P4-BCC              CTAGAGTGGCGGAATGTGTCTGGTCTTTATGTCCTTACTAACGACTGTTCCAATAGCAGTATTGTGTATGAGGCCGATGACGTYATTCTGCACTCACCTG
P4-B1.B11-A4        ....................................................................................................
P4-B1.C2-A6         ....................................................................................................
P4-B1.C4-G11        ....................................................................................................
P4-B1.C2-B7         ....................................................................................................
P4-B1.F5-D2         ....................................................................................................
P4-B1.E8-G2         ....................................................................................................
P4-B3.E9-B2         ....................................................................................................
P4-B3.E11-C3        ....................................................................................................
P4-B7.F3-E3         ..................................................C.................................................
P4-B7.F9-H6         ....................................................................................................
P4-B7.G7-E3         ...................................................................................................C.
P4-B7.H3-C12        ....................................................................................................
P4-B7.B6-C8         ....................................................................................................
P4-Plas.G12-A9      ....................................................................................................
P4-Plas.C3-F3       ....................................................................................................
P4-Plas.G7-C8       ....................................................................................................
P4-Plas.D7-E2       ....................................................................................................
P4-Plas.E6-G5       ....................................................................................................
P4-Plas.H8-H10      ....................................................................................................
P4-Plas.D10-B12     ..............................................................C.....................

                       110       120       130       140       150       160       170       180       190       200
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P4-BCC              GTTGTGTACCTTGTGTTCAGGACGGCAATACATCCACGTGCTGGACCTCAGTGACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACCACYGCTTCGAT
```

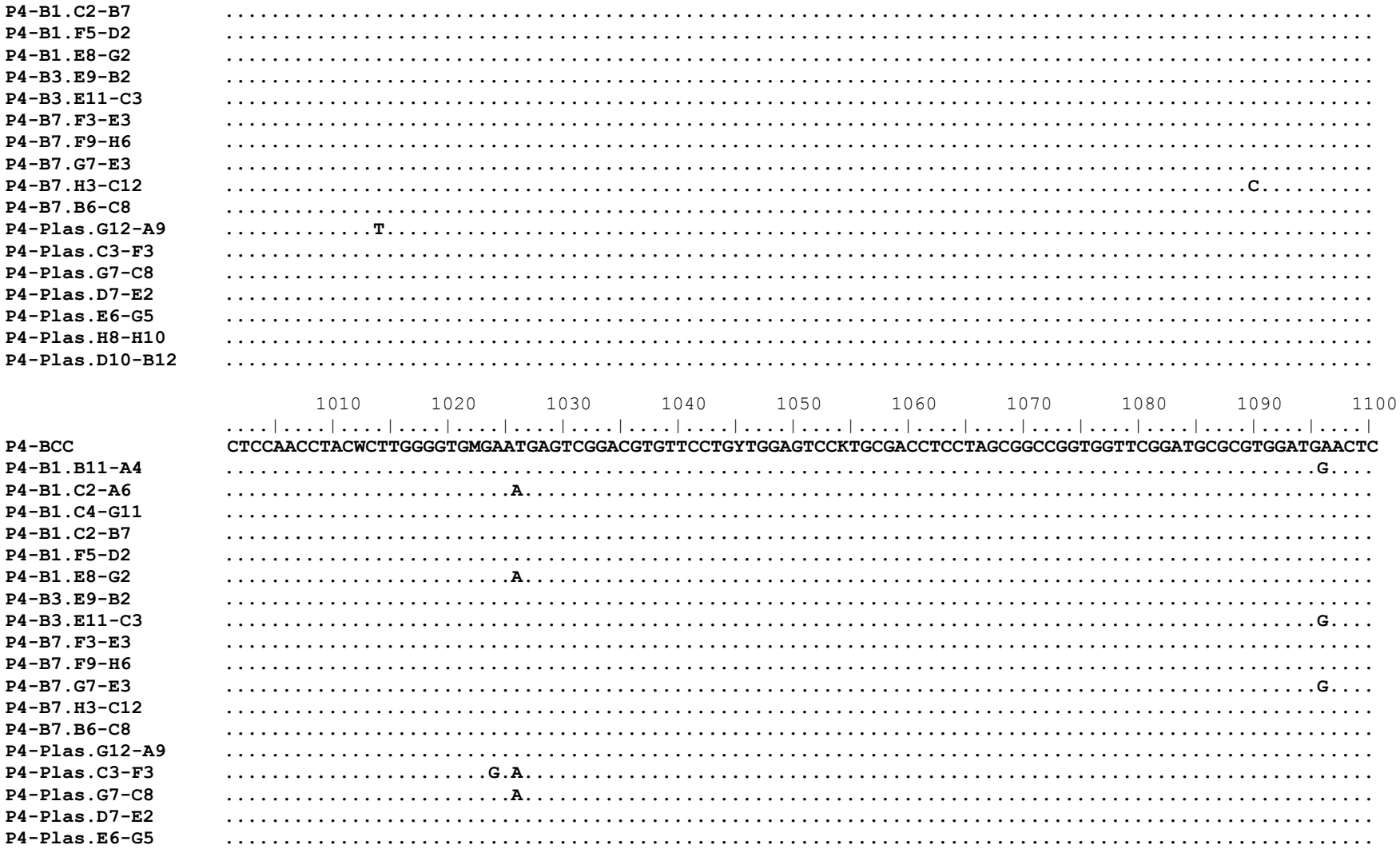

P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
210
220
230
240
250
260
270
280
290
300
P4-BCC
ACGTAGTCATGTGGACCTATTAGTGGGCGCGGCCACGCTGTGTTCGGCGCTCTACGTGGGTGATATGTGTGGGGCCGTCTTCCTTGTAGGACAAGCCTTC

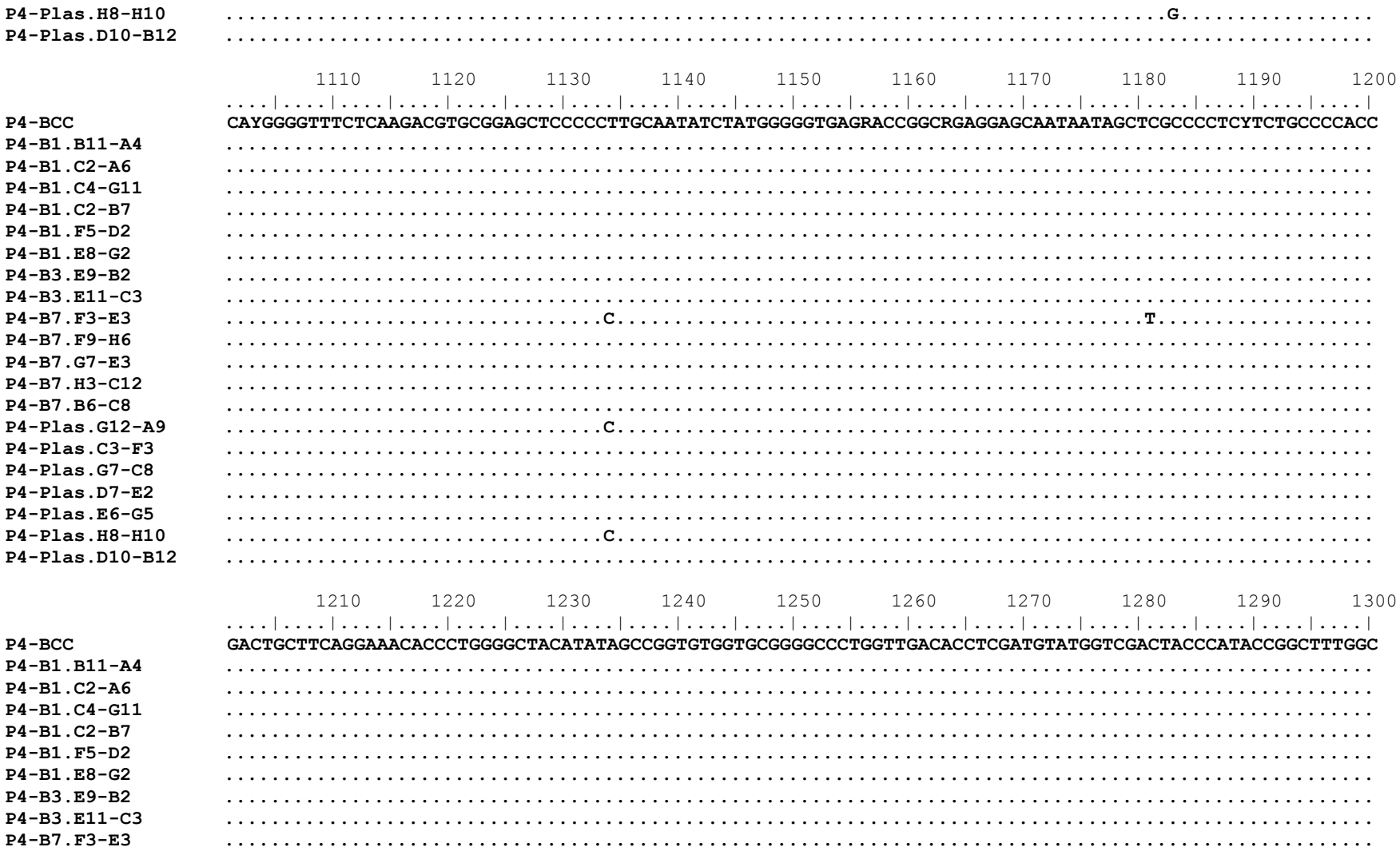

P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
310 320 330 340 350 360 370 380 390 400
P4-BCC ACGTTCAGACCCCGTCGCCATCAAACGGTCCAGACCTGTAACTGCTCACTGTACCCAGGCCACCTCACAGGACATCGAATGGCTTGGGACATGATGATGA
P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
410 420 430 440 450 460 470 480 490 500
P4-BCC ATTGGTCCCCCGCTGTYGGCATGGTAGTGGCGCATGTCCTGCGTCTACCCCAGACCTTGTTCGACATAATAGCTGGAGCTCATTGGGGCGTCCTGGCGGG
P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2

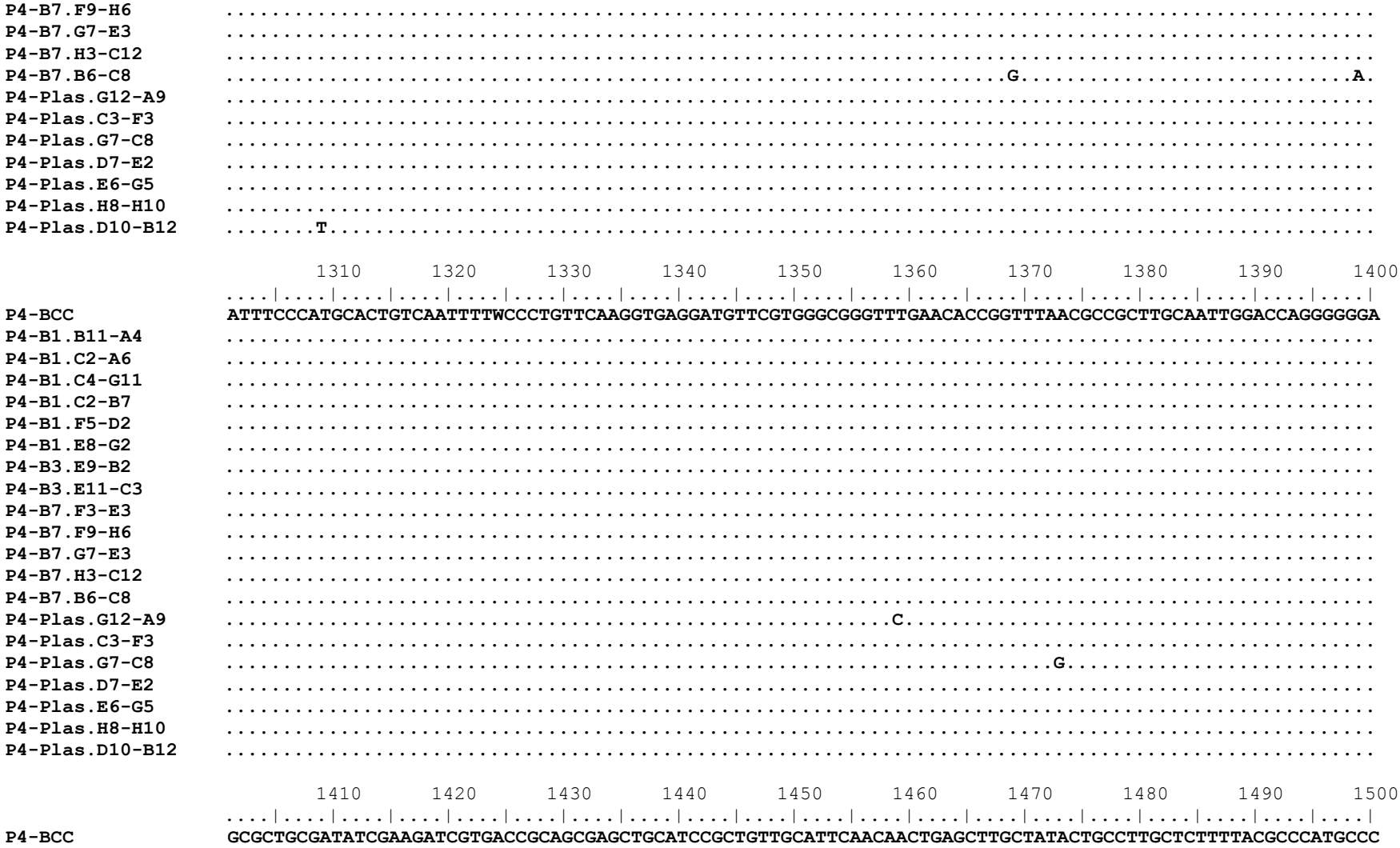
```
P4-B3.E9-B2       ..........................................................................................................
P4-B3.E11-C3      ..........................................................................................................
P4-B7.F3-E3       ..........................................................................................................
P4-B7.F9-H6       ..........................................................................................................
P4-B7.G7-E3       ..........................................................................................................
P4-B7.H3-C12      ..........................................................................................................
P4-B7.B6-C8       ..........................................................................................................
P4-Plas.G12-A9    ..........................................................................................................
P4-Plas.C3-F3     ..........................................................................................................
P4-Plas.G7-C8     ..........................................................................................................
P4-Plas.D7-E2     ..........................................................................................................
P4-Plas.E6-G5     ..........................................................................................................
P4-Plas.H8-H10    ..........................................................................................................
P4-Plas.D10-B12   ..........................................................................................................

                           510       520       530       540       550       560       570       580       590       600
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P4-BCC            CCTAGCCTATTATTCCATGCAGGGTAACTGGGCCAAGGTTGCTATTGTCATGGTCATGTTTTCAGGAGTCGATGCCRMTACACRCGTCAYCGGTGSCACC
P4-B1.B11-A4      ..........................................................................................................
P4-B1.C2-A6       ..................................................T.......................................................
P4-B1.C4-G11      ..........................................................................................................
P4-B1.C2-B7       ..............................................................................T...........................
P4-B1.F5-D2       ..........................................................................................C...............
P4-B1.E8-G2       ..........................................................................................................
P4-B3.E9-B2       ..........................................................................................................
P4-B3.E11-C3      ..........................................................................................................
P4-B7.F3-E3       ..........................................................................................................
P4-B7.F9-H6       ..........................................................................................................
P4-B7.G7-E3       ..........................................................................................................
P4-B7.H3-C12      ..........................................................................................................
P4-B7.B6-C8       ..........................................................................................................
P4-Plas.G12-A9    ..........................................................................................................
P4-Plas.C3-F3     ..........................................................................................................
P4-Plas.G7-C8     ..........................................................................................................
P4-Plas.D7-E2     ..........................................................................................................
P4-Plas.E6-G5     ..............................................................................T...........................
P4-Plas.H8-H10    ..........................................................................................................
P4-Plas.D10-B12   ..........................................................................................................
```


610 620 630 640 650 660 670 680 690 700
P4-BCC GCAGCTYATARTGTCAWGACCWTMGCTGGGYTTTTCACTTTAGGCCCCMAGCAGAAATTACAGCTAATCAACACCAATGGCTCGTGGCACATCAATAGGA
P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
710 720 730 740 750 760 770 780 790 800
P4-BCC CTGCCCTGAACTGCAATGATTCCTTAAACACCGGGTTTATAGCAGGGTTGWTKYAYTTCCACAGGTTCAACTCTACTGGGTGTCCTCAAAGGCTCAGTAG
P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12

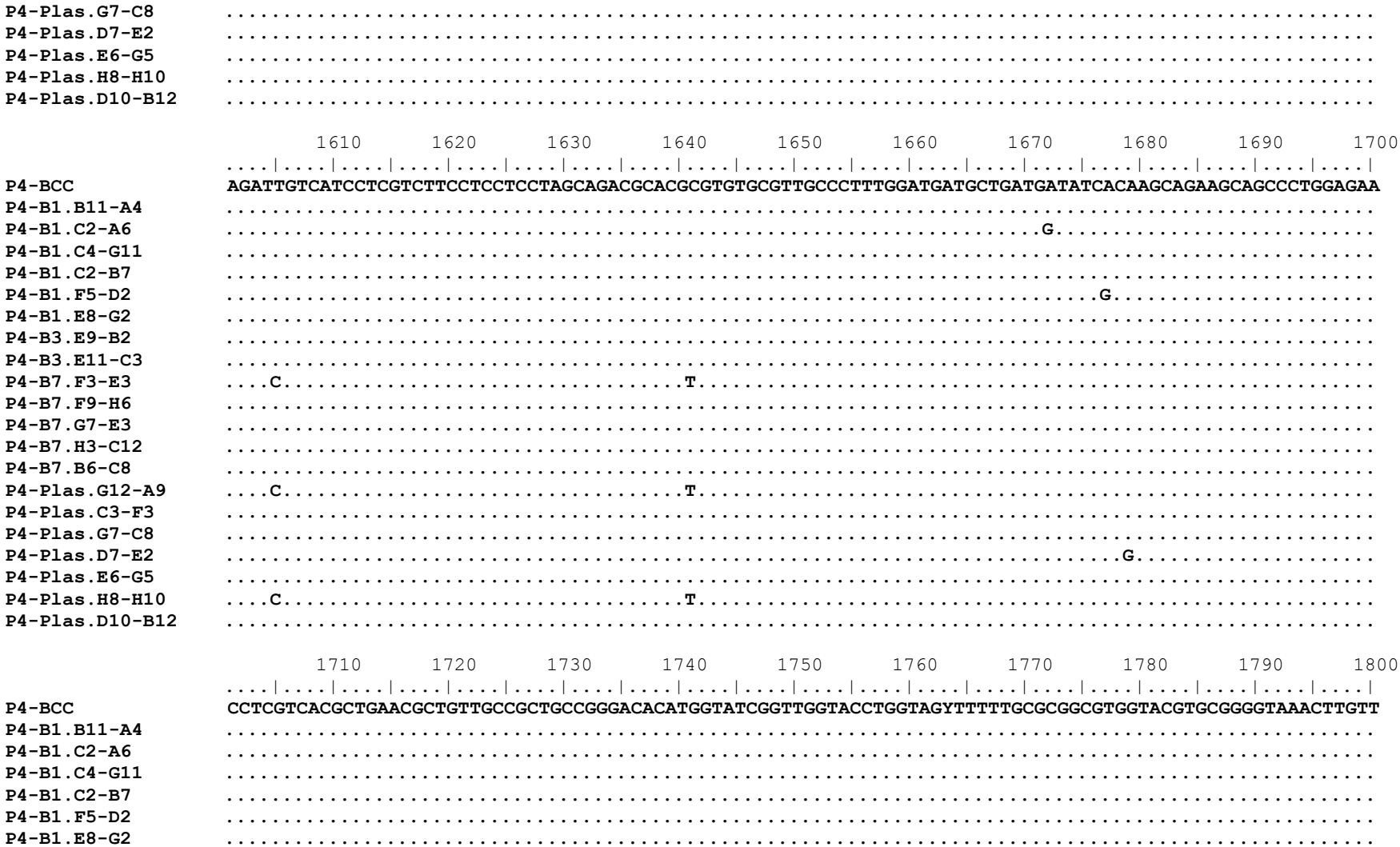

P4-B7.B6-C8 .....G.....
P4-Plas.G12-A9 .....
P4-Plas.C3-F3 .....A.....
P4-Plas.G7-C8 .....C.....
P4-Plas.D7-E2 .....
P4-Plas.E6-G5 .....
P4-Plas.H8-H10 .....
P4-Plas.D10-B12 .....A.....
810 820 830 840 850 860 870 880 890 900
P4-BCC CTGCAAGCCCATCACTTCCTTTGATCAGGGGTGGGGTCACTTGACAGATGCTAACGTCAMTRACACCTCTGACGACAGACCGTACTGCTGGCACTACGCA
P4-B1.B11-A4 ...T.....
P4-B1.C2-A6 .....T.....
P4-B1.C4-G11 .....
P4-B1.C2-B7 .....
P4-B1.F5-D2 .....A.....
P4-B1.E8-G2 .....T.....
P4-B3.E9-B2 .....T.....
P4-B3.E11-C3 .....
P4-B7.F3-E3 .....G.....
P4-B7.F9-H6 .....C.....
P4-B7.G7-E3 .....
P4-B7.H3-C12 .....
P4-B7.B6-C8 .....
P4-Plas.G12-A9 .....G.....
P4-Plas.C3-F3 .....T.....
P4-Plas.G7-C8 .....T.....
P4-Plas.D7-E2 .....
P4-Plas.E6-G5 .....
P4-Plas.H8-H10 .....G.....
P4-Plas.D10-B12 .....
910 920 930 940 950 960 970 980 990 1000
P4-BCC CCCAAACCCTGTCRCACTGTCCCGGCGTCGAGTGTCTGCGGCCCCGTGTACTGTTTCACACCATCACCAGTGGTCGTAGGCACCACCGATGCTAAGGGTG
P4-B1.B11-A4 .....
P4-B1.C2-A6 .....
P4-B1.C4-G11 .....

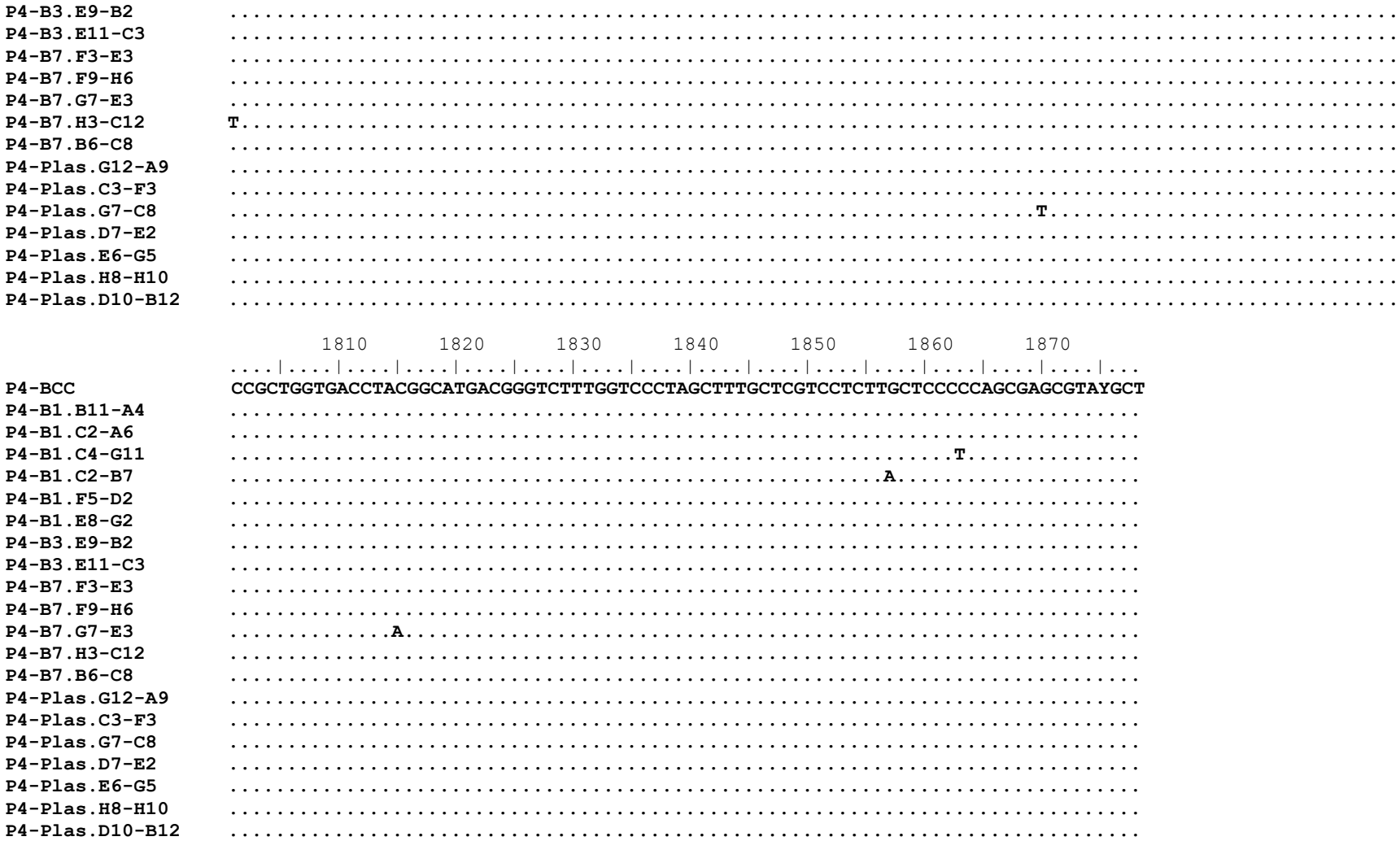

P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
1010
1020
1030
1040
1050
1060
1070
1080
1090
1100
P4-BCC
CTCCAACCTACWCTTGGGGTGMGAATGAGTCGGACGTGTTCCTGYTGGAGTCCKTGCGACCTCCTAGCGGCCGGTGGTTCGGATGCGCGTGGATGAACTC
P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5


P4-Plas.H8-H10 .....G.....
P4-Plas.D10-B12 .....
1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
P4-BCC CAYGGGGTTTCTCAAGACGTGCGGAGCTCCCCCTTGCAATATCTATGGGGGTGAGRACCGGCRGAGGAGCAATAATAGCTCGCCCCTCYTCTGCCCCACC
P4-B1.B11-A4 .....
P4-B1.C2-A6 .....
P4-B1.C4-G11 .....
P4-B1.C2-B7 .....
P4-B1.F5-D2 .....
P4-B1.E8-G2 .....
P4-B3.E9-B2 .....
P4-B3.E11-C3 .....
P4-B7.F3-E3 .....C.....T.....
P4-B7.F9-H6 .....
P4-B7.G7-E3 .....
P4-B7.H3-C12 .....
P4-B7.B6-C8 .....
P4-Plas.G12-A9 .....C.....
P4-Plas.C3-F3 .....
P4-Plas.G7-C8 .....
P4-Plas.D7-E2 .....
P4-Plas.E6-G5 .....
P4-Plas.H8-H10 .....C.....
P4-Plas.D10-B12 .....
1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
P4-BCC GACTGCTTCAGGAAACACCCTGGGGCTACATATAGCCGGTGTGGTGCGGGGCCCTGGTTGACACCTCGATGTATGGTCGACTACCCATACCGGCTTTGGC
P4-B1.B11-A4 .....
P4-B1.C2-A6 .....
P4-B1.C4-G11 .....
P4-B1.C2-B7 .....
P4-B1.F5-D2 .....
P4-B1.E8-G2 .....
P4-B3.E9-B2 .....
P4-B3.E11-C3 .....
P4-B7.F3-E3 .....


P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
P4-BCC ATTTCCCATGCACTGTCAATTTTWCCCTGTTCAAGGTGAGGATGTTCGTGGGCGGGTTTGAACACCGGTTTAACGCCGCTTGCAATTGGACCAGGGGGGA
P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
P4-BCC GCGCTGCGATATCGAAGATCGTGACCGCAGCGAGCTGCATCCGCTGTTGCATTCAACAACTGAGCTTGCTATACTGCCTTGCTCTTTTACGCCCATGCCC

```
P4-B1.B11-A4      ..........................................................................................................
P4-B1.C2-A6       ..........................................................................................................
P4-B1.C4-G11      ..........................................................................................................
P4-B1.C2-B7       ..........................................................................................................
P4-B1.F5-D2       ..........................................................................................................
P4-B1.E8-G2       ..........................................................................................................
P4-B3.E9-B2       ..........................................................................................................
P4-B3.E11-C3      ..........................................................................................................
P4-B7.F3-E3       ..........................................................................................................
P4-B7.F9-H6       ..........................................................................................................
P4-B7.G7-E3       ..........................................................................................................
P4-B7.H3-C12      ..........................................................................................................
P4-B7.B6-C8       .........C................................................................................................
P4-Plas.G12-A9    ..........................................................................................................
P4-Plas.C3-F3     ..................C.......................................................................................
P4-Plas.G7-C8     ..........................................................................................................
P4-Plas.D7-E2     ..........................................................................................................
P4-Plas.E6-G5     ..........................................................................................................
P4-Plas.H8-H10    ..........................................................................................................
P4-Plas.D10-B12   ..........................................................................................................

                           1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P4-BCC            GCATTGTCAACAGGTCTAATACACCTCCACCAAAACATCGTGGATGTCCAGTACCTTTATGGTGTTGGAGCTGGCATGGTGGGGTGGGCGCTTAAGTGGG
P4-B1.B11-A4      ..........................................................................................................
P4-B1.C2-A6       ..........................................................................................................
P4-B1.C4-G11      ..........................................................................................................
P4-B1.C2-B7       ..........................................................................................................
P4-B1.F5-D2       ..........................................................................................................
P4-B1.E8-G2       ..........................................................................................................
P4-B3.E9-B2       ..........................................................................................................
P4-B3.E11-C3      ..........................................................................................................
P4-B7.F3-E3       ..........................................................................................................
P4-B7.F9-H6       ..........................................................................................C.......
P4-B7.G7-E3       ..........................................................................................................
P4-B7.H3-C12      ..........................................................................................................
P4-B7.B6-C8       ..........................................................................................................
P4-Plas.G12-A9    ..........................................................................................................
P4-Plas.C3-F3     ..........................................................................................................
```

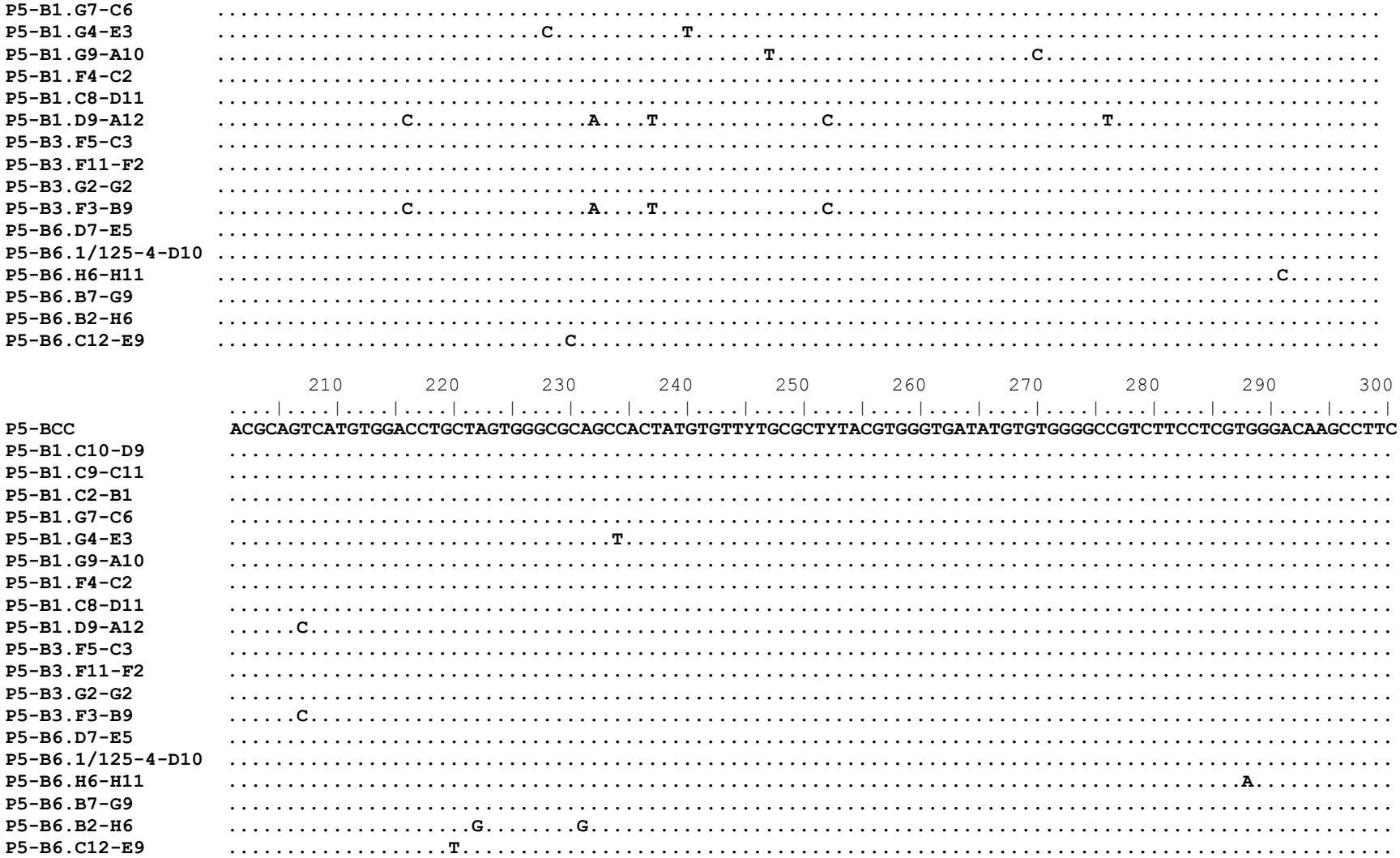

P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
P4-BCC AGATTGTCATCCTCGTCTTCCTCCTCCTAGCAGACGCACGCGTGTGCGTTGCCCTTTGGATGATGCTGATGATATCACAAGCAGAAGCAGCCCTGGAGAA
P4-B1.B11-A4
P4-B1.C2-A6 G
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2 G
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3 C T
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9 C T
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2 G
P4-Plas.E6-G5
P4-Plas.H8-H10 C T
P4-Plas.D10-B12
1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
P4-BCC CCTCGTCACGCTGAACGCTGTTGCCGCTGCCGGGACACATGGTATCGGTTGGTACCTGGTAGYTTTTTGCGCGGCGTGGTACGTGCGGGGTAAACTTGTT
P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2

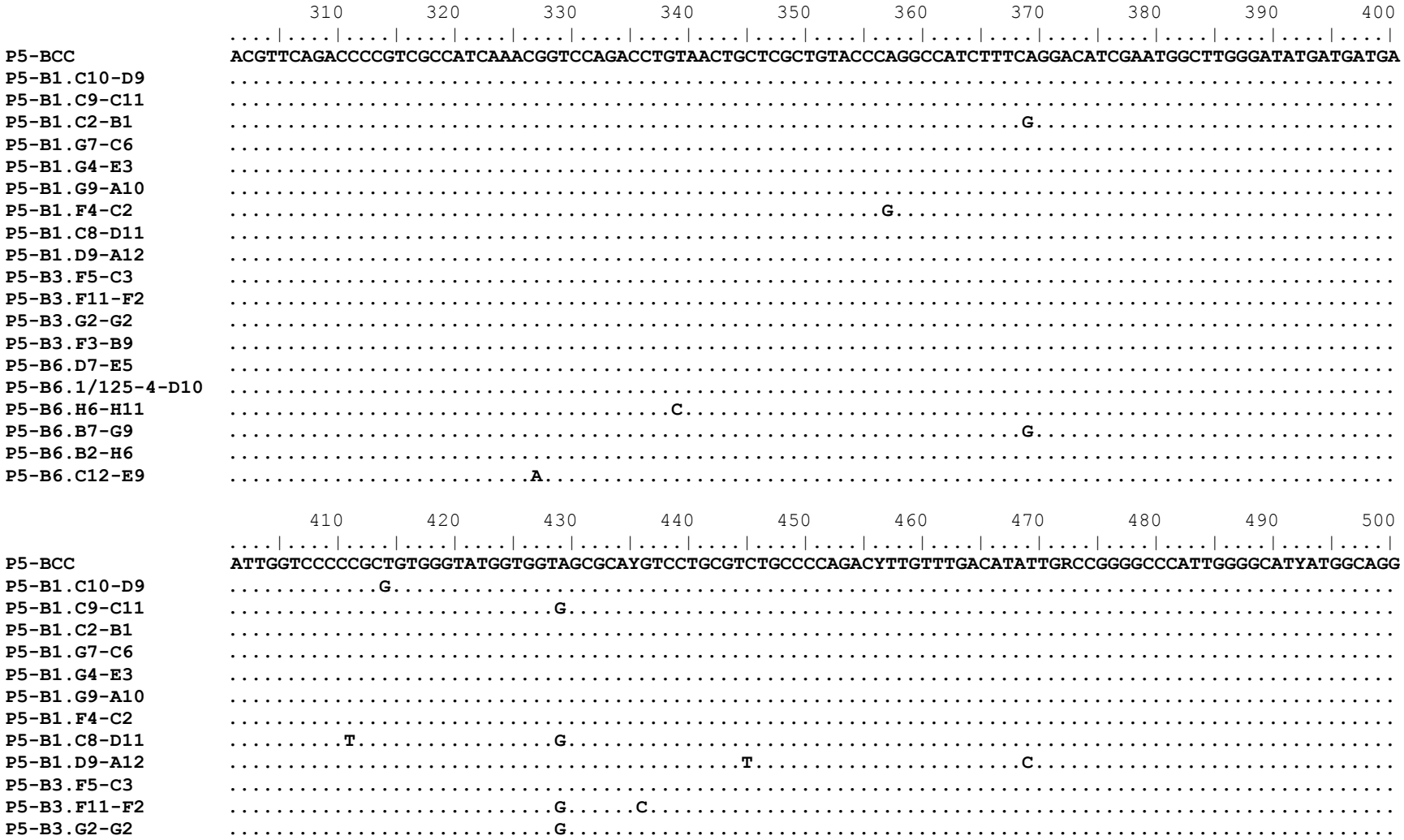

P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12
1810
1820
1830
1840
1850
1860
1870
P4-BCC
CCGCTGGTGACCTACGGCATGACGGGTCTTTGGTCCCTAGCTTTGCTCGTCCTCTTGCTCCCCCAGCGAGCGTAYGCT
P4-B1.B11-A4
P4-B1.C2-A6
P4-B1.C4-G11
P4-B1.C2-B7
P4-B1.F5-D2
P4-B1.E8-G2
P4-B3.E9-B2
P4-B3.E11-C3
P4-B7.F3-E3
P4-B7.F9-H6
P4-B7.G7-E3
P4-B7.H3-C12
P4-B7.B6-C8
P4-Plas.G12-A9
P4-Plas.C3-F3
P4-Plas.G7-C8
P4-Plas.D7-E2
P4-Plas.E6-G5
P4-Plas.H8-H10
P4-Plas.D10-B12

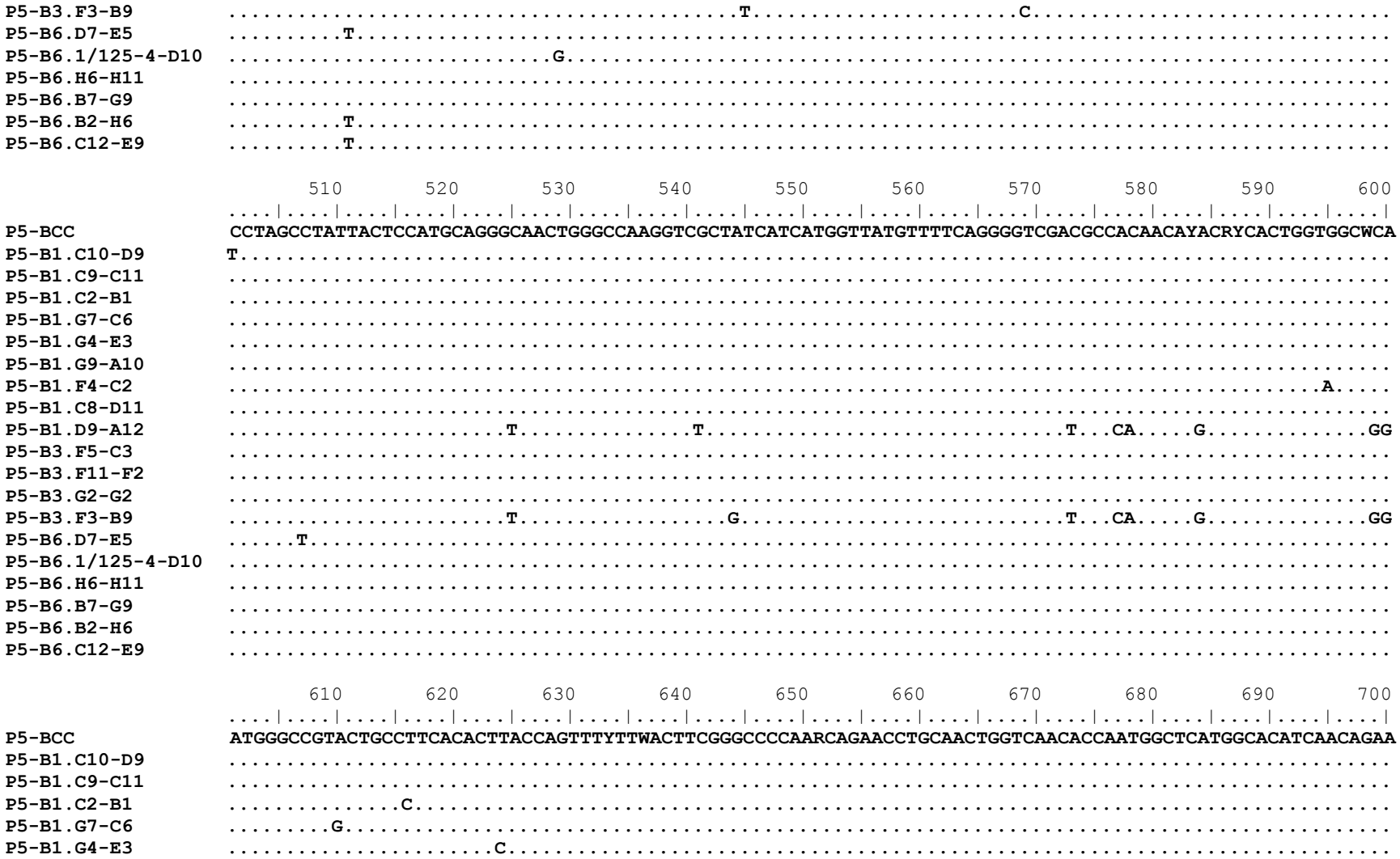
10.1.4 Patient 5

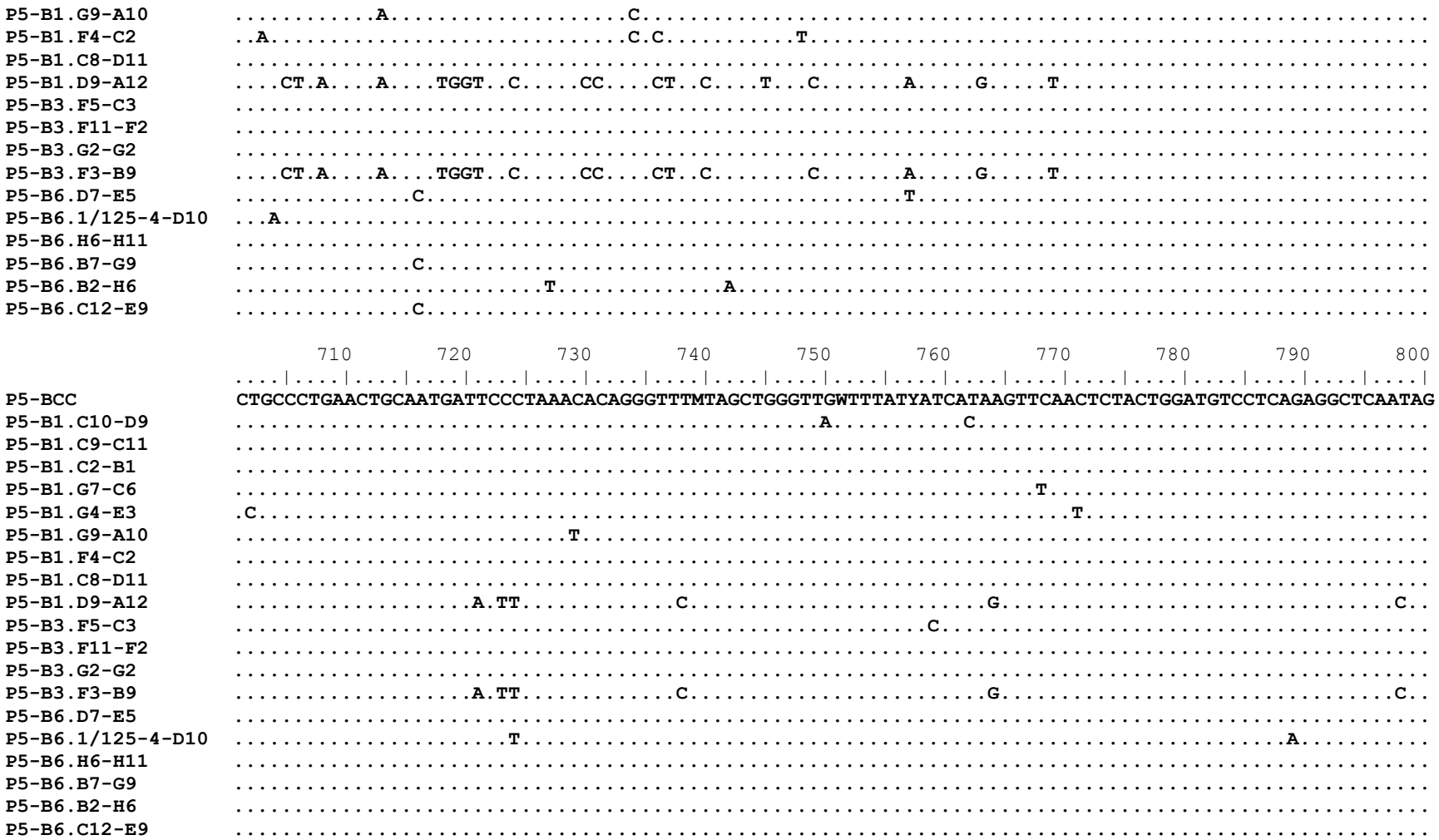
Gene position: E1: 1 - 576nt, E2: 577 - 1693nt, p7: 1683-1872nt, HVR1: 577 – 657nt, HVR494: 913 - 933nt, HVR595: 1153 - 1179nt.

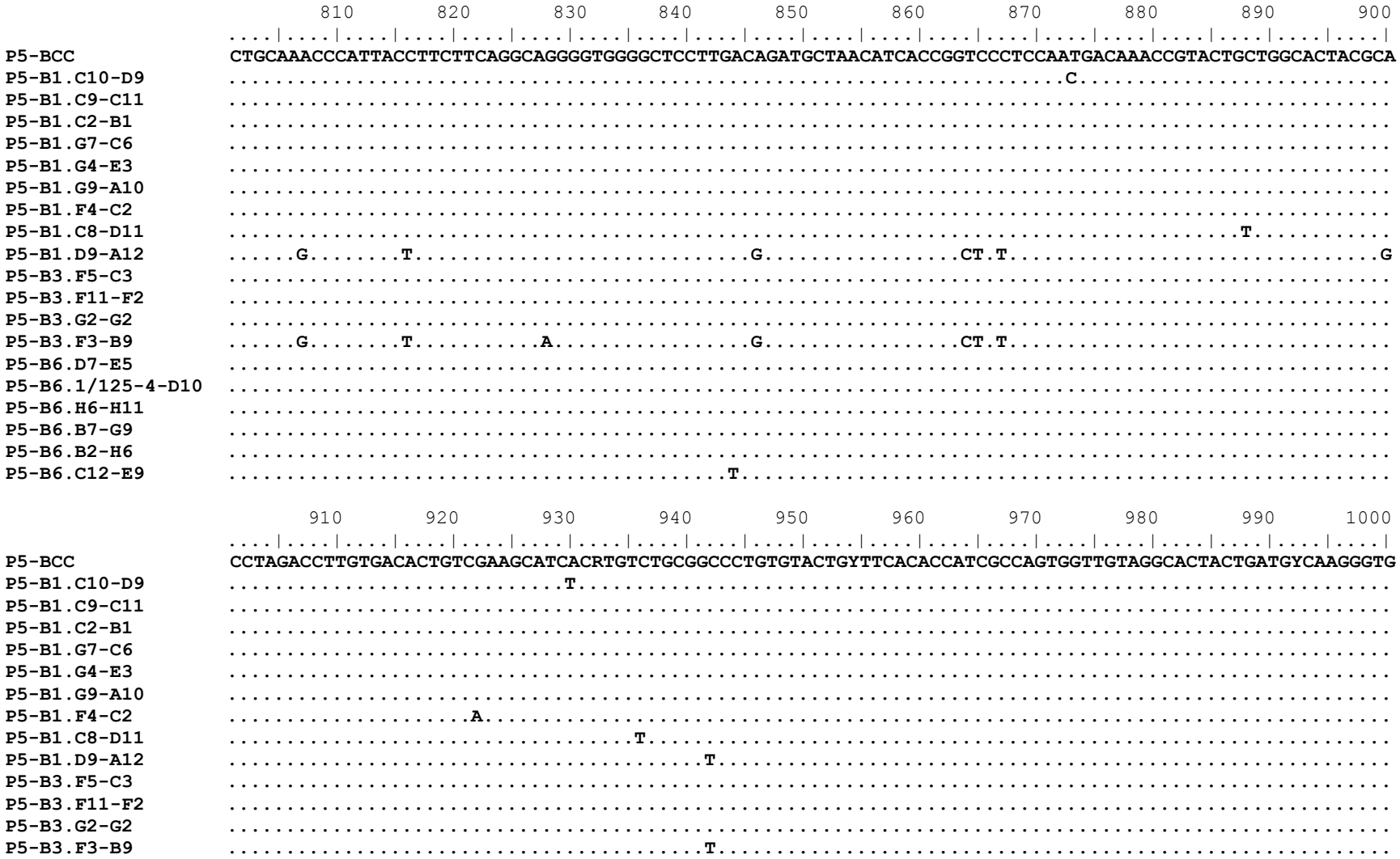
```
                          10        20        30        40        50        60        70        80        90        100
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P5-BCC              YTAGAGTGGCGGAATACGTCTGGCCTYTACATTCTTACCAACGACTGTTCTAATAGCAGTATTGTGTAYGAGGCCGACGATGTTATTCTGCACACACCYG
P5-B1.C10-D9        ....................................C.....................................................................
P5-B1.C9-C11        ....................................................C.....................................................
P5-B1.C2-B1         ..G.....................................................................................................
P5-B1.G7-C6         ..G..A..................................................................................................
P5-B1.G4-E3         .....................................................................................................
P5-B1.G9-A10        .....................................................................................................
P5-B1.F4-C2         ....................C....................................................................................
P5-B1.C8-D11        ....................................................T........C.....................................
P5-B1.D9-A12        ...................................................................................T.................
P5-B3.F5-C3         .....................................................................................................
P5-B3.F11-F2        ....................................................C....................................................
P5-B3.G2-G2         ....................................................C....................................................
P5-B3.F3-B9         ...................................................................................T.................
P5-B6.D7-E5         .....................................................................................................
P5-B6.1/125-4-D10   .....................................................................................................
P5-B6.H6-H11        .....................................................................................................
P5-B6.B7-G9         ..G..................................................................................................
P5-B6.B2-H6         .....................................................................................................
P5-B6.C12-E9        .....................................................................................................

                          110       120       130       140       150       160       170       180       190       200
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P5-BCC              GCTGCGTACCTTGTGTTCAGGACGGCAATAAATCYARGTGCTGGACCCCAGTGACACCTACAGTGGCAGTTAGGTACGTCGGAGCAACCACTGCTTCGAT
P5-B1.C10-D9        ...........................................................................G.........................
P5-B1.C9-C11        .....................................................................................................
P5-B1.C2-B1         .....................................................................................................
```

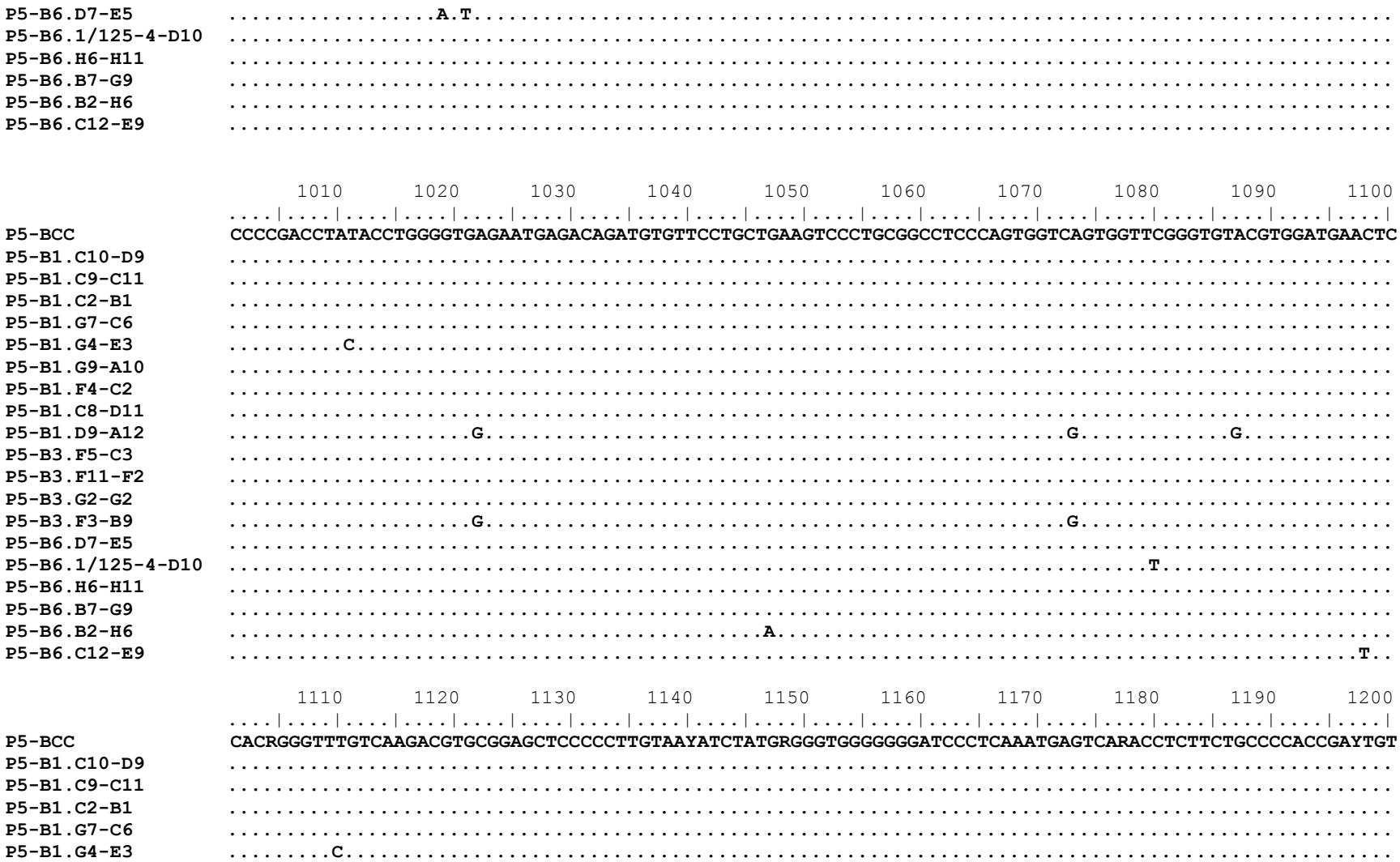

P5-B1.G7-C6
P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2
P5-B1.C8-D11
P5-B1.D9-A12
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9
P5-B6.D7-E5
P5-B6.1/125-4-D10
P5-B6.H6-H11
P5-B6.B7-G9
P5-B6.B2-H6
P5-B6.C12-E9
210
220
230
240
250
260
270
280
290
300
P5-BCC
ACGCAGTCATGTGGACCTGCTAGTGGGCGCAGCCACTATGTGTTYTGCGCTYTACGTGGGTGATATGTGTGGGGCCGTCTTCCTCGTGGGACAAGCCTTC
P5-B1.C10-D9
P5-B1.C9-C11
P5-B1.C2-B1
P5-B1.G7-C6
P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2
P5-B1.C8-D11
P5-B1.D9-A12
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9
P5-B6.D7-E5
P5-B6.1/125-4-D10
P5-B6.H6-H11
P5-B6.B7-G9
P5-B6.B2-H6
P5-B6.C12-E9

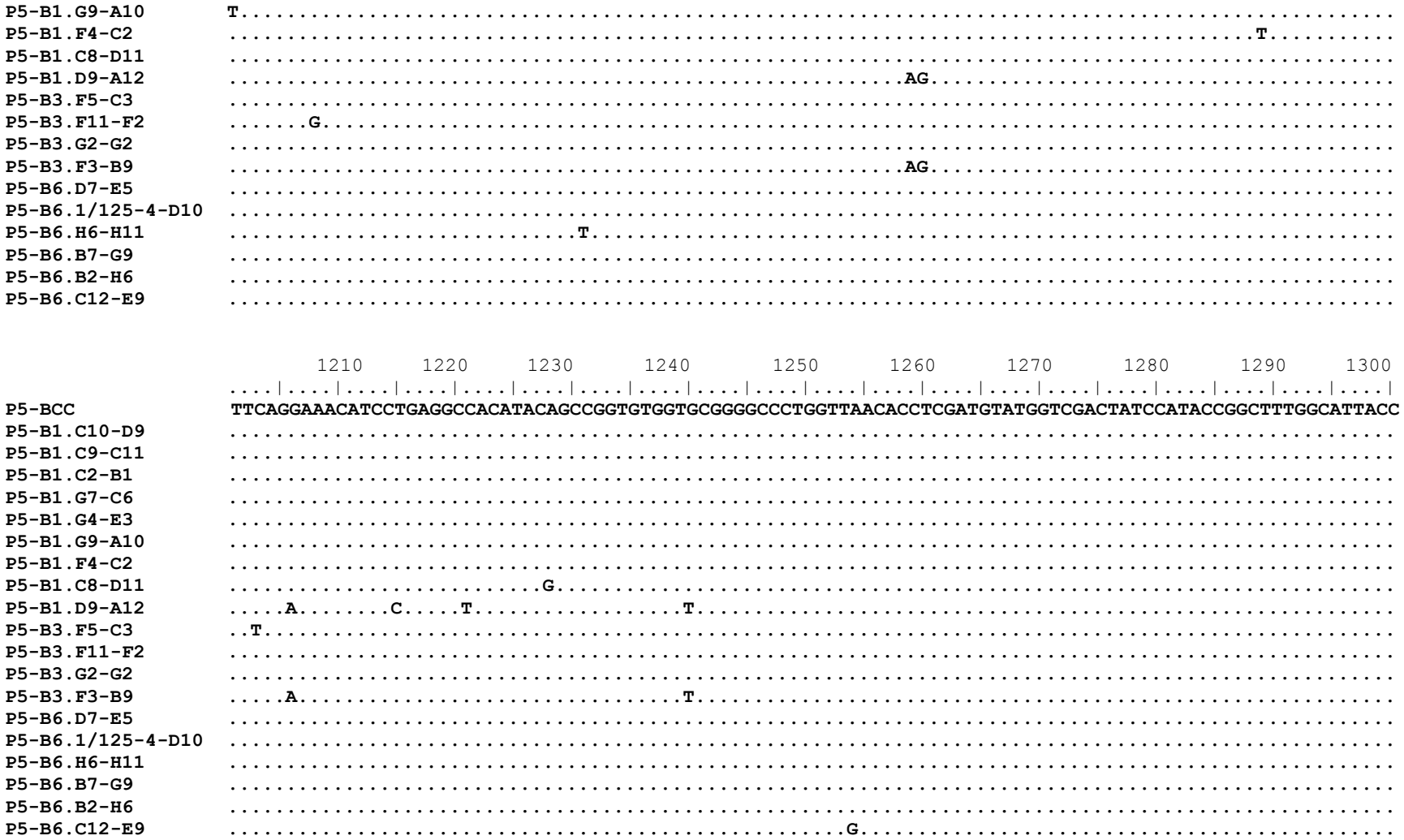

310 320 330 340 350 360 370 380 390 400
P5-BCC ACGTTCAGACCCCGTCGCCATCAAACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCCATCTTTCAGGACATCGAATGGCTTGGGATATGATGATGA
P5-B1.C10-D9
P5-B1.C9-C11
P5-B1.C2-B1 G
P5-B1.G7-C6
P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2 G
P5-B1.C8-D11
P5-B1.D9-A12
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9
P5-B6.D7-E5
P5-B6.1/125-4-D10
P5-B6.H6-H11 C
P5-B6.B7-G9 G
P5-B6.B2-H6
P5-B6.C12-E9 A
410 420 430 440 450 460 470 480 490 500
P5-BCC ATTGGTCCCCCGCTGTGGGTATGGTGGTAGCGCAYGTCCTGCGTCTGCCCCAGACYTTGTTTGACATATTGRCCGGGGCCCATTGGGGCATYATGGCAGG
P5-B1.C10-D9 G
P5-B1.C9-C11 G
P5-B1.C2-B1
P5-B1.G7-C6
P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2
P5-B1.C8-D11 T G
P5-B1.D9-A12 T C
P5-B3.F5-C3
P5-B3.F11-F2 G C
P5-B3.G2-G2 G

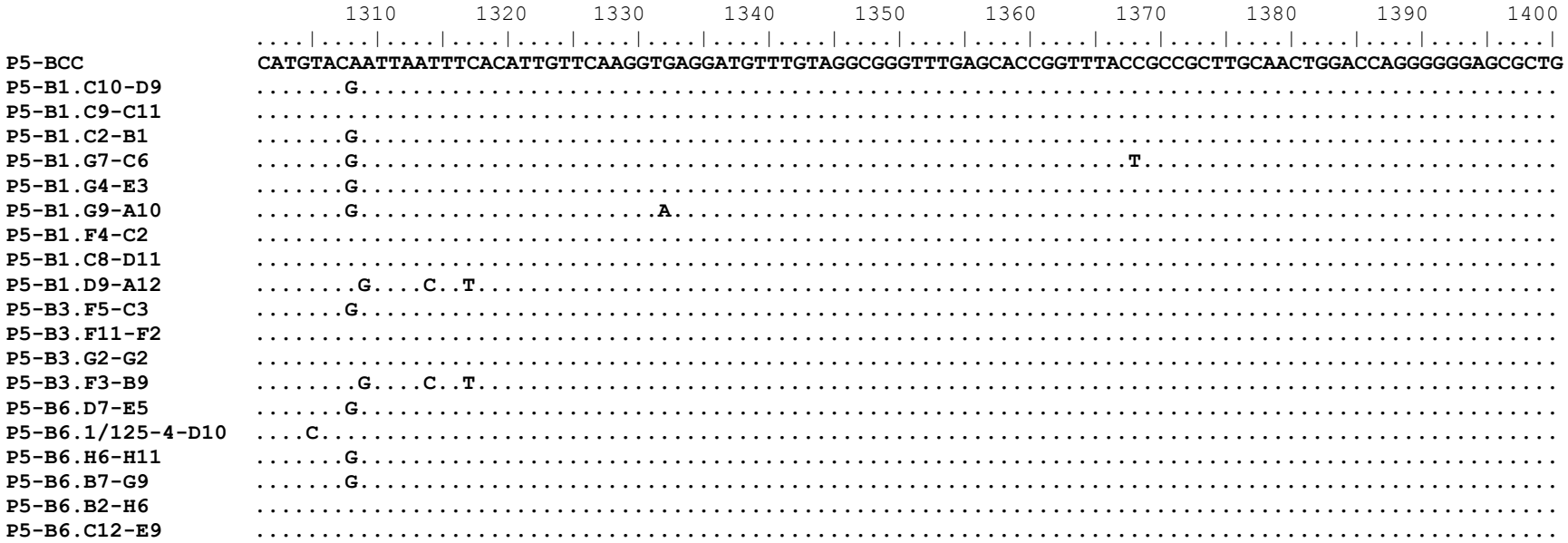

P5-B3.F3-B9 .....T.....C.....
P5-B6.D7-E5 .....T.....
P5-B6.1/125-4-D10 .....G.....
P5-B6.H6-H11 .....
P5-B6.B7-G9 .....
P5-B6.B2-H6 .....T.....
P5-B6.C12-E9 .....T.....
510 520 530 540 550 560 570 580 590 600
P5-BCC CCTAGCCTATTACTCCATGCAGGGCAACTGGGCCAAGGTCGCTATCATCATGGTTATGTTTTCAGGGGTCGACGCCACAACAYACRYCACTGGTGGCWCA
P5-B1.C10-D9 T.....
P5-B1.C9-C11 .....
P5-B1.C2-B1 .....
P5-B1.G7-C6 .....
P5-B1.G4-E3 .....
P5-B1.G9-A10 .....
P5-B1.F4-C2 .....A.....
P5-B1.C8-D11 .....
P5-B1.D9-A12 .....T.....T.....T...CA.....G.....GG
P5-B3.F5-C3 .....
P5-B3.F11-F2 .....
P5-B3.G2-G2 .....
P5-B3.F3-B9 .....T.....G.....T...CA.....G.....GG
P5-B6.D7-E5 .....T.....
P5-B6.1/125-4-D10 .....
P5-B6.H6-H11 .....
P5-B6.B7-G9 .....
P5-B6.B2-H6 .....
P5-B6.C12-E9 .....
610 620 630 640 650 660 670 680 690 700
P5-BCC ATGGGCCGTACTGCCTTCACACTTACCAGTTTYTTWACTTCGGGCCCCAARCAGAACCTGCAACTGGTCAACACCAATGGCTCATGGCACATCAACAGAA
P5-B1.C10-D9 .....
P5-B1.C9-C11 .....
P5-B1.C2-B1 .....C.....
P5-B1.G7-C6 .....G.....
P5-B1.G4-E3 .....C.....

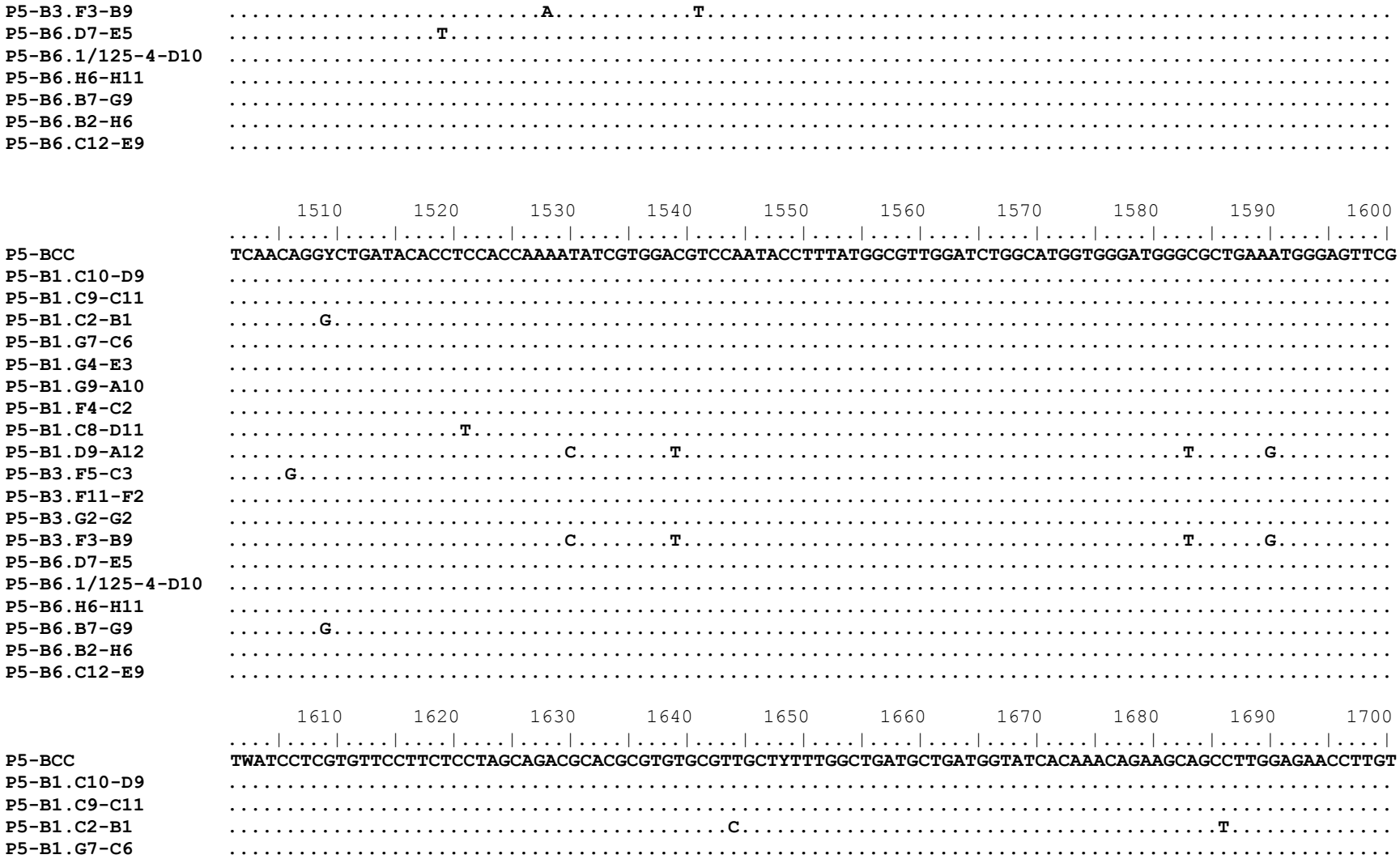

P5-B1.G9-A10 ..........A.....C.....
P5-B1.F4-C2 ..A.....C.C.....T.....
P5-B1.C8-D11 .....
P5-B1.D9-A12 ....CT.A...A....TGGT..C.....CC....CT..C....T...C.....A.....G.....T.....
P5-B3.F5-C3 .....
P5-B3.F11-F2 .....
P5-B3.G2-G2 .....
P5-B3.F3-B9 ....CT.A...A....TGGT..C.....CC....CT..C.....C.....A.....G.....T.....
P5-B6.D7-E5 .............C.....T.....
P5-B6.1/125-4-D10 ...A.....
P5-B6.H6-H11 .....
P5-B6.B7-G9 ....C.....
P5-B6.B2-H6 .....T.....A.....
P5-B6.C12-E9 ....C.....
710 720 730 740 750 760 770 780 790 800
P5-BCC CTGCCCTGAACTGCAATGATTCCCTAAACACAGGGTTTMTAGCTGGGTTGWTTTATYATCATAAGTTCAACTCTACTGGATGTCCTCAGAGGCTCAATAG
P5-B1.C10-D9 .....A.....C.....
P5-B1.C9-C11 .....
P5-B1.C2-B1 .....
P5-B1.G7-C6 .....T.....
P5-B1.G4-E3 .C.....T.....
P5-B1.G9-A10 .....T.....
P5-B1.F4-C2 .....
P5-B1.C8-D11 .....
P5-B1.D9-A12 .....A.TT.....C.....G.....C..
P5-B3.F5-C3 .....C.....
P5-B3.F11-F2 .....
P5-B3.G2-G2 .....
P5-B3.F3-B9 .....A.TT.....C.....G.....C..
P5-B6.D7-E5 .....
P5-B6.1/125-4-D10 .....T.....A.....
P5-B6.H6-H11 .....
P5-B6.B7-G9 .....
P5-B6.B2-H6 .....
P5-B6.C12-E9 .....

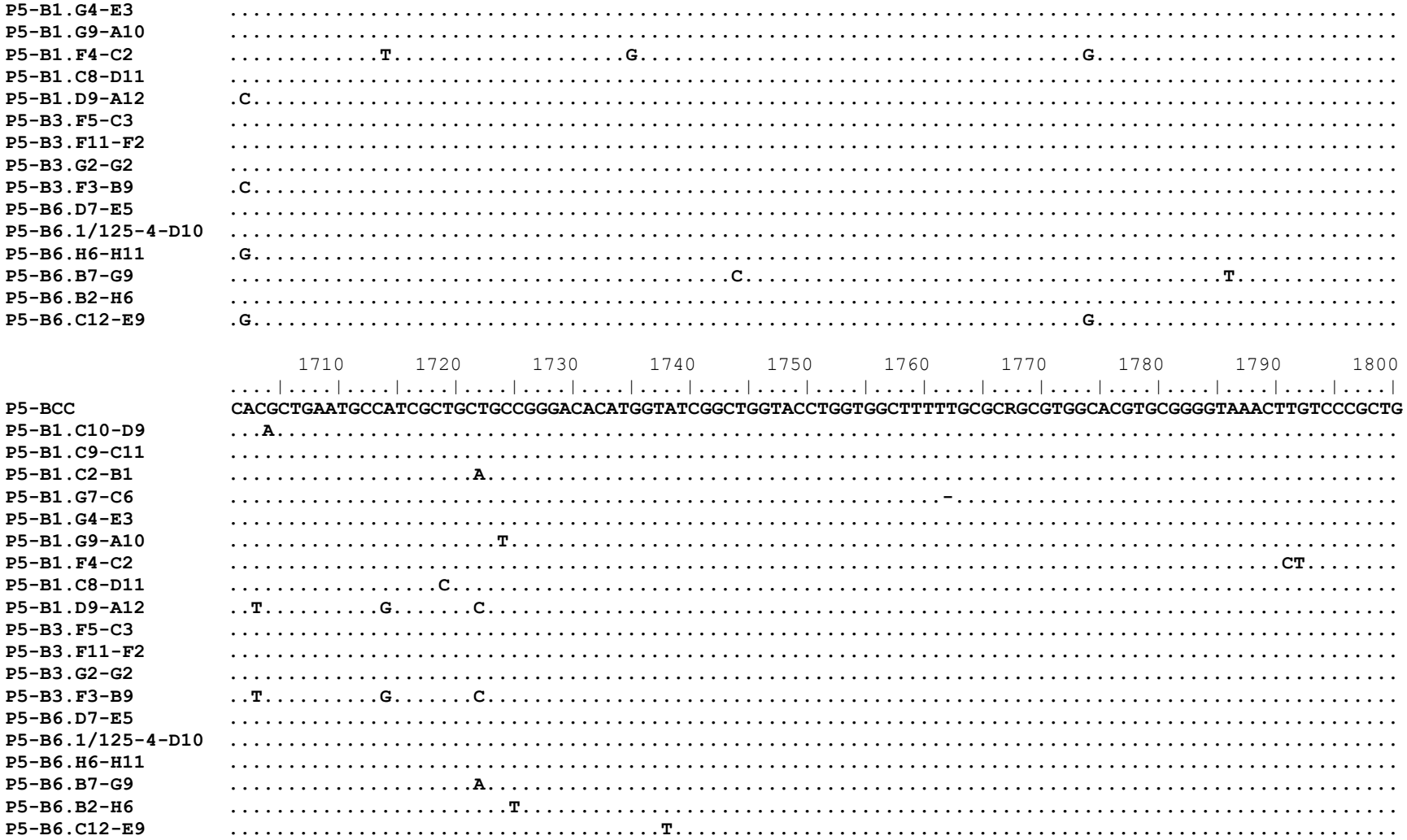

810 820 830 840 850 860 870 880 890 900
P5-BCC CTGCAAACCCATTACCTTCTTCAGGCAGGGGTGGGGCTCCTTGACAGATGCTAACATCACCGGTCCCTCCAATGACAAACCGTACTGCTGGCACTACGCA
P5-B1.C10-D9 C
P5-B1.C9-C11
P5-B1.C2-B1
P5-B1.G7-C6
P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2
P5-B1.C8-D11 T
P5-B1.D9-A12 G T G CT.T G
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9 G T A G CT.T
P5-B6.D7-E5
P5-B6.1/125-4-D10
P5-B6.H6-H11
P5-B6.B7-G9
P5-B6.B2-H6
P5-B6.C12-E9 T
910 920 930 940 950 960 970 980 990 1000
P5-BCC CCTAGACCTTGTGACACTGTCGAAGCATCACRTGTCTGCGGCCCTGTGTACTGYTTCACACCATCGCCAGTGGTTGTAGGCACTACTGATGYCAAGGGTG
P5-B1.C10-D9 T
P5-B1.C9-C11
P5-B1.C2-B1
P5-B1.G7-C6
P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2 A
P5-B1.C8-D11 T
P5-B1.D9-A12 T
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9 T

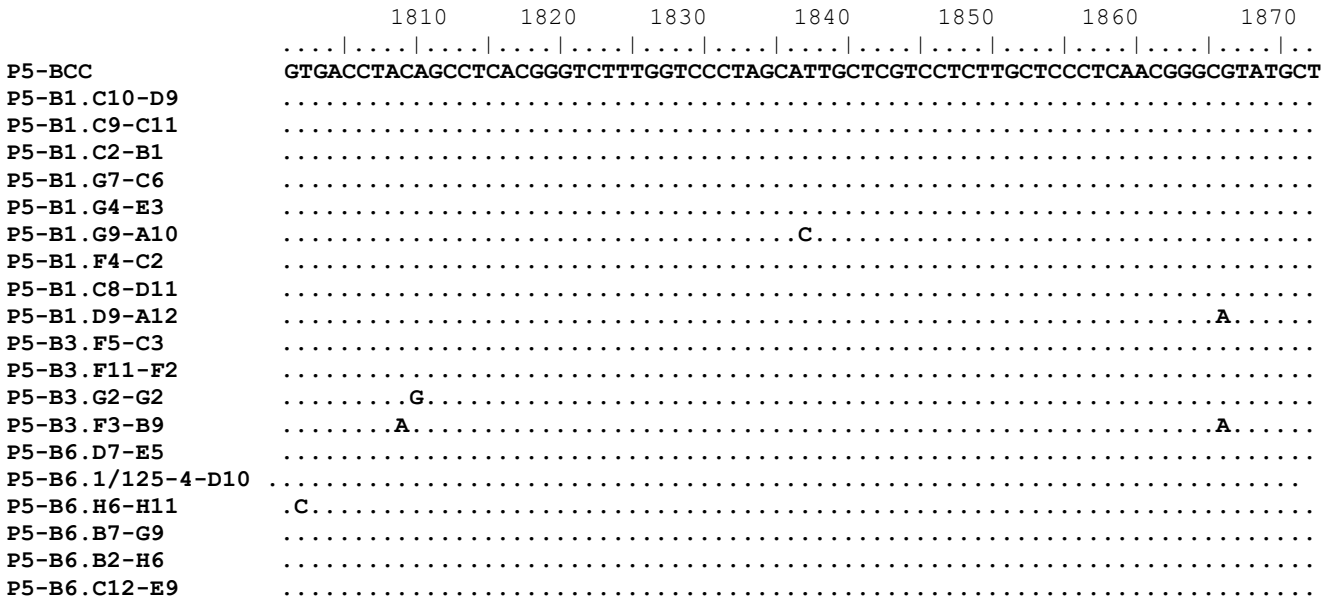

P5-B6.D7-E5 ..................A.T.....
P5-B6.1/125-4-D10 .....
P5-B6.H6-H11 .....
P5-B6.B7-G9 .....
P5-B6.B2-H6 .....
P5-B6.C12-E9 .....
1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
P5-BCC CCCCGACCTATACCTGGGGTGAGAATGAGACAGATGTGTTCCTGCTGAAGTCCCTGCGGCCTCCCAGTGGTCAGTGGTTCGGGTGTACGTGGATGAACTC
P5-B1.C10-D9 .....
P5-B1.C9-C11 .....
P5-B1.C2-B1 .....
P5-B1.G7-C6 .....
P5-B1.G4-E3 ....C.....
P5-B1.G9-A10 .....
P5-B1.F4-C2 .....
P5-B1.C8-D11 .....
P5-B1.D9-A12 ....................G.....G.....G.....
P5-B3.F5-C3 .....
P5-B3.F11-F2 .....
P5-B3.G2-G2 .....
P5-B3.F3-B9 ....................G.....G.....
P5-B6.D7-E5 .....
P5-B6.1/125-4-D10 .....T.....
P5-B6.H6-H11 .....
P5-B6.B7-G9 .....
P5-B6.B2-H6 .....A.....
P5-B6.C12-E9 .....T..
1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
P5-BCC CACRGGGTTTGTCAAGACGTGCGGAGCTCCCCCTTGTAAYATCTATGRGGGTGGGGGGGGATCCCTCAAATGAGTCARACCTCTTCTGCCCCACCGAYTGT
P5-B1.C10-D9 .....
P5-B1.C9-C11 .....
P5-B1.C2-B1 .....
P5-B1.G7-C6 .....
P5-B1.G4-E3 ....C.....


P5-B1.G9-A10
P5-B1.F4-C2
P5-B1.C8-D11
P5-B1.D9-A12
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9
P5-B6.D7-E5
P5-B6.1/125-4-D10
P5-B6.H6-H11
P5-B6.B7-G9
P5-B6.B2-H6
P5-B6.C12-E9
1210
1220
1230
1240
1250
1260
1270
1280
1290
1300
P5-BCC
TTCAGGAAACATCCTGAGGCCACATACAGCCGGTGTGGTGCGGGGCCCTGGTTAACACCTCGATGTATGGTCGACTATCCATACCGGCTTTGGCATTACC


1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
P5-BCC CATGTACAATTAATTTCACATTGTTCAAGGTGAGGATGTTTGTAGGCGGGTTTGAGCACCGGTTTACCGCCGCTTGCAACTGGACCAGGGGGGAGCGCTG
P5-B1.C10-D9 ....G.....
P5-B1.C9-C11 .....
P5-B1.C2-B1 ....G.....
P5-B1.G7-C6 .......G.....T.....
P5-B1.G4-E3 ....G.....
P5-B1.G9-A10 .......G.....A.....
P5-B1.F4-C2 .....
P5-B1.C8-D11 .....
P5-B1.D9-A12 .....G....C..T.....
P5-B3.F5-C3 ....G.....
P5-B3.F11-F2 .....
P5-B3.G2-G2 .....
P5-B3.F3-B9 .....G....C..T.....
P5-B6.D7-E5 ....G.....
P5-B6.1/125-4-D10 ....C.....
P5-B6.H6-H11 ....G.....
P5-B6.B7-G9 ....G.....
P5-B6.B2-H6 .....
P5-B6.C12-E9 .....
1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
P5-BCC CGATATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCAACAACTGAGYTTGCTATACTGCCTTGCTCCTTCACGCCCATGCCYGCATTG
P5-B1.C10-D9 .....
P5-B1.C9-C11 .....
P5-B1.C2-B1 .....
P5-B1.G7-C6 .....
P5-B1.G4-E3 .....
P5-B1.G9-A10 .....
P5-B1.F4-C2 .....T.....A.....
P5-B1.C8-D11 .....
P5-B1.D9-A12 .....C.....A.....T.....
P5-B3.F5-C3 .....
P5-B3.F11-F2 .....
P5-B3.G2-G2 ......T.....


P5-B3.F3-B9 ..........A..........T..........
P5-B6.D7-E5 ....T..........
P5-B6.1/125-4-D10 ....
P5-B6.H6-H11 ....
P5-B6.B7-G9 ....
P5-B6.B2-H6 ....
P5-B6.C12-E9 ....
1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
P5-BCC TCAACAGGYCTGATACACCTCCACCAAAATATCGTGGACGTCCAATACCTTTATGGCGTTGGATCTGGCATGGTGGGATGGGCGCTGAAATGGGAGTTCG
P5-B1.C10-D9 ....
P5-B1.C9-C11 ....
P5-B1.C2-B1 ....G..........
P5-B1.G7-C6 ....
P5-B1.G4-E3 ....
P5-B1.G9-A10 ....
P5-B1.F4-C2 ....
P5-B1.C8-D11 ....T..........
P5-B1.D9-A12 ..........C........T..........T......G..........
P5-B3.F5-C3 .....G..........
P5-B3.F11-F2 ....
P5-B3.G2-G2 ....
P5-B3.F3-B9 ..........C........T..........T......G..........
P5-B6.D7-E5 ....
P5-B6.1/125-4-D10 ....
P5-B6.H6-H11 ....
P5-B6.B7-G9 ....G..........
P5-B6.B2-H6 ....
P5-B6.C12-E9 ....
1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
P5-BCC TWATCCTCGTGTTCCTTCTCCTAGCAGACGCACGCGTGTGCGTTGCTYTTTGGCTGATGCTGATGGTATCACAAACAGAAGCAGCCTTGGAGAACCTTGT
P5-B1.C10-D9 ....
P5-B1.C9-C11 ....
P5-B1.C2-B1 ..........C..........T..........
P5-B1.G7-C6 ....

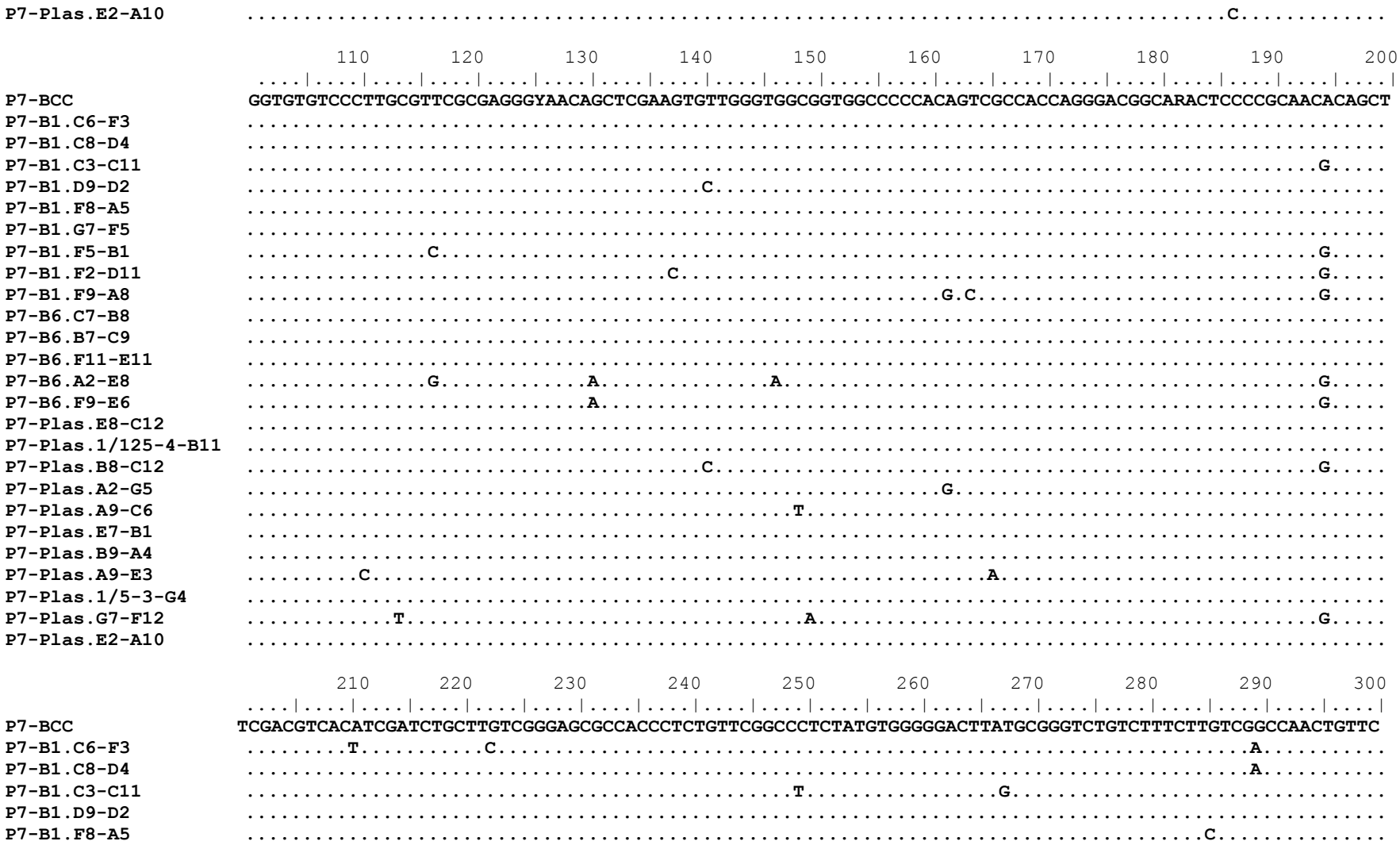

P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2
P5-B1.C8-D11
P5-B1.D9-A12
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9
P5-B6.D7-E5
P5-B6.1/125-4-D10
P5-B6.H6-H11
P5-B6.B7-G9
P5-B6.B2-H6
P5-B6.C12-E9
1710
1720
1730
1740
1750
1760
1770
1780
1790
1800
P5-BCC
CACGCTGAATGCCATCGCTGCTGCCGGGACACATGGTATCGGCTGGTACCTGGTGGCTTTTTGCGCRGCGTGGCACGTGCGGGGTAAACTTGTCCCGCTG
P5-B1.C10-D9
P5-B1.C9-C11
P5-B1.C2-B1
P5-B1.G7-C6
P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2
P5-B1.C8-D11
P5-B1.D9-A12
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9
P5-B6.D7-E5
P5-B6.1/125-4-D10
P5-B6.H6-H11
P5-B6.B7-G9
P5-B6.B2-H6
P5-B6.C12-E9

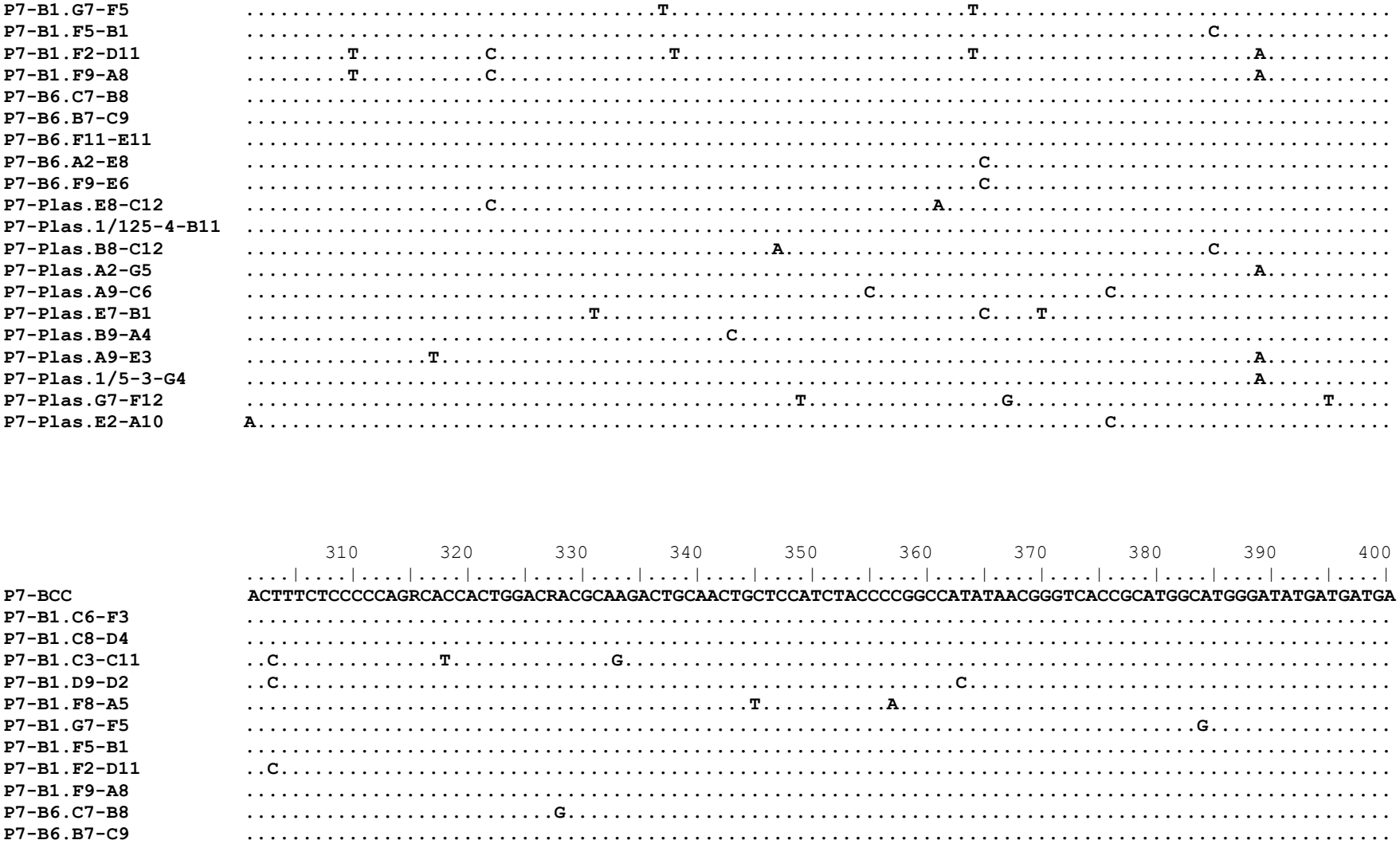

1810 1820 1830 1840 1850 1860 1870
P5-BCC GTGACCTACAGCCTCACGGGTCTTTGGTCCCTAGCATTGCTCGTCCTCTTGCTCCCTCAACGGGCGTATGCT
P5-B1.C10-D9
P5-B1.C9-C11
P5-B1.C2-B1
P5-B1.G7-C6
P5-B1.G4-E3
P5-B1.G9-A10
P5-B1.F4-C2
P5-B1.C8-D11
P5-B1.D9-A12
P5-B3.F5-C3
P5-B3.F11-F2
P5-B3.G2-G2
P5-B3.F3-B9
P5-B6.D7-E5
P5-B6.1/125-4-D10
P5-B6.H6-H11
P5-B6.B7-G9
P5-B6.B2-H6
P5-B6.C12-E9

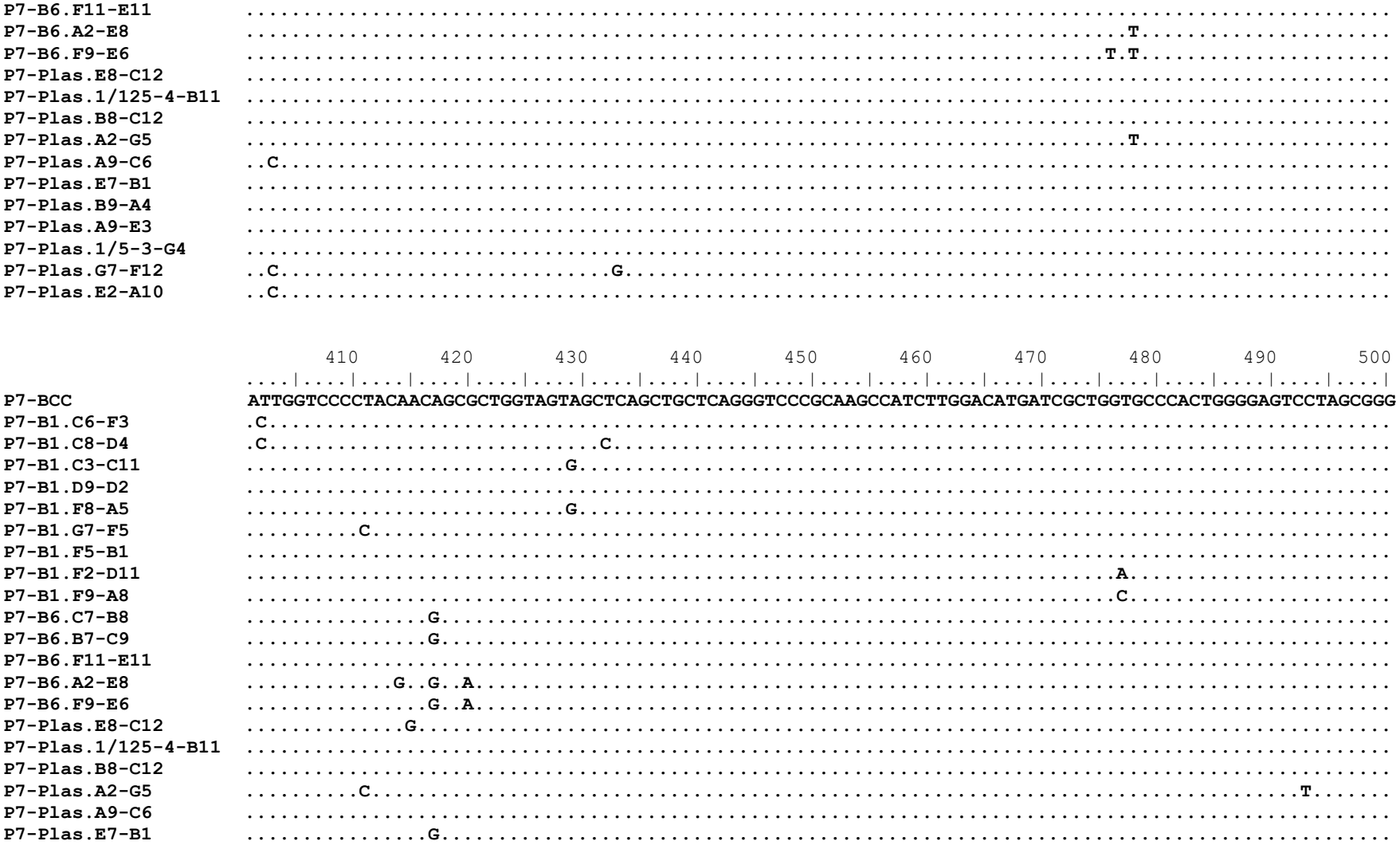
10.1.5 Patient 7

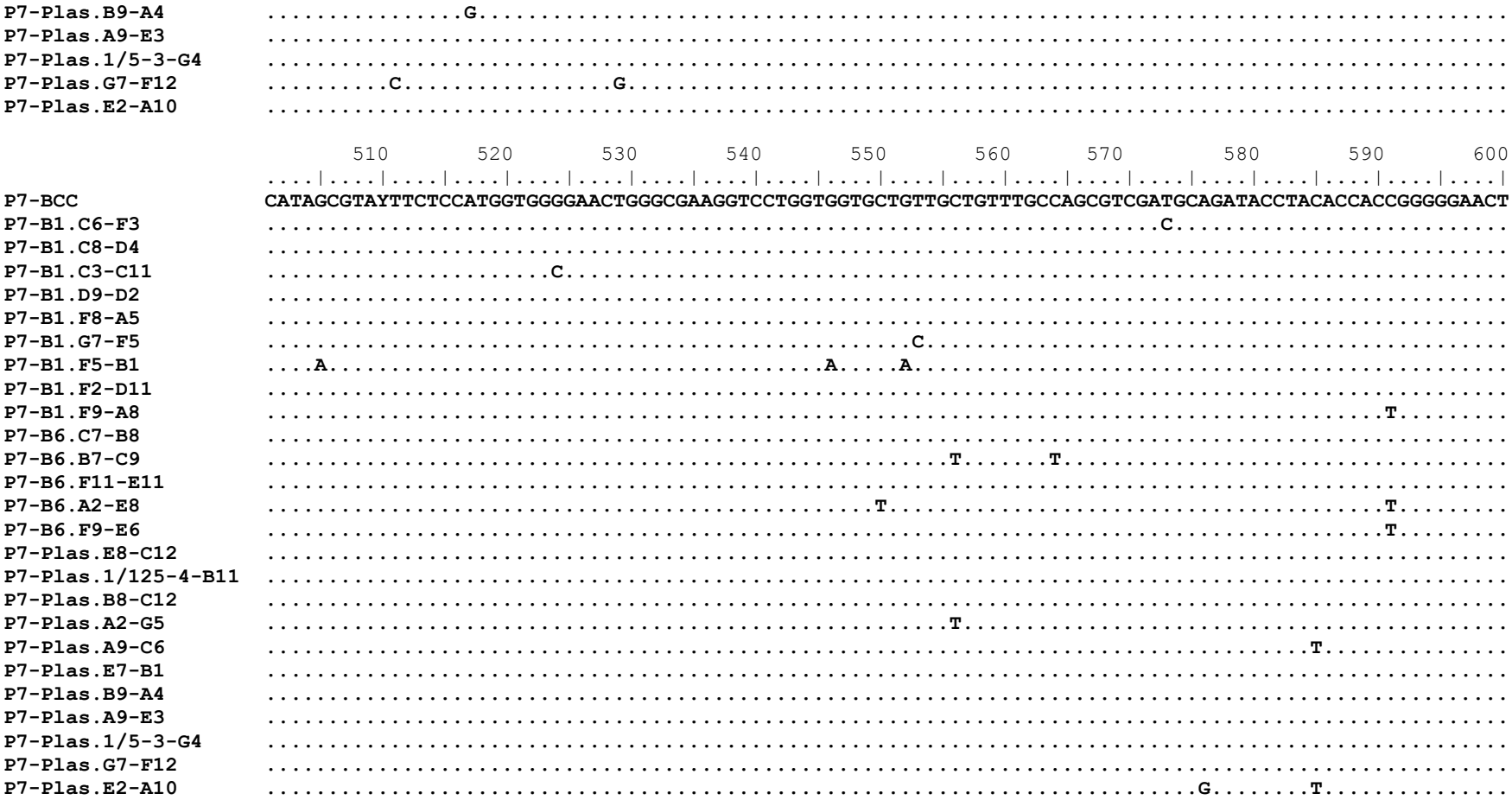
Gene position: E1: 1 - 576nt, E2: 577 - 1665nt, HVR1: 577 – 657nt.

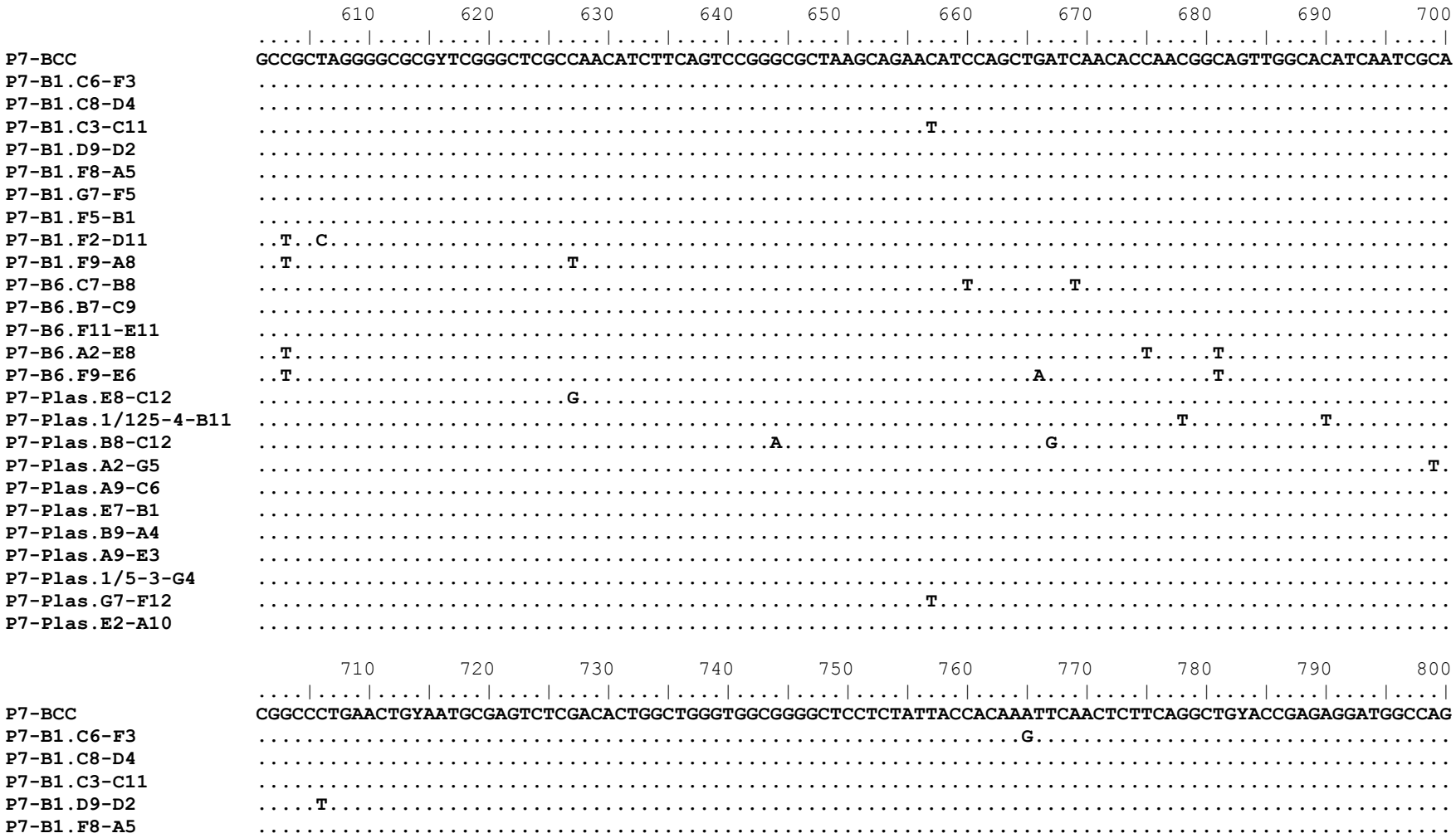
```
                          10        20        30        40        50        60        70        80        90       100
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P7-BCC            TACCAAGTACGCAACTCCTCGGGCATCTACCATGTCACCAATGATTGCCCTAACTCGAGTATTGTGTACGAGACGGCCGATGCCATTCTRCACTCTCCGG
P7-B1.C6-F3       ..........................................................................................................
P7-B1.C8-D4       ..........................................................................................................
P7-B1.C3-C11      ..........................................................................................................
P7-B1.D9-D2       ..........................................................................................................
P7-B1.F8-A5       ..................................................C.......................................................
P7-B1.G7-F5       ..........................................................................................................
P7-B1.F5-B1       ................T.....................................C.......................................................
P7-B1.F2-D11      ........G.................................................................................................
P7-B1.F9-A8       ..........................................C...............................................................
P7-B6.C7-B8       ..........................................................................................................A.
P7-B6.B7-C9       ..........................................................................................................A.
P7-B6.F11-E11     ..................................................................................C.....................
P7-B6.A2-E8       ..........................................................................................................
P7-B6.F9-E6       ..........................................................................................................
P7-Plas.E8-C12    ..........................................................................................................
P7-Plas.1/125-4-B11 ......................................................................................................
P7-Plas.B8-C12    ..........................................................................................................
P7-Plas.A2-G5     ..........................................................................................................
P7-Plas.A9-C6     ..........................................................................................................
P7-Plas.E7-B1     ..........................................................................................................
P7-Plas.B9-A4     ..........................................................................................................
P7-Plas.A9-E3     ..............................................C...........................................................
P7-Plas.1/5-3-G4  ..........................................................................................................
P7-Plas.G7-F12    ..........................................................................................................
```

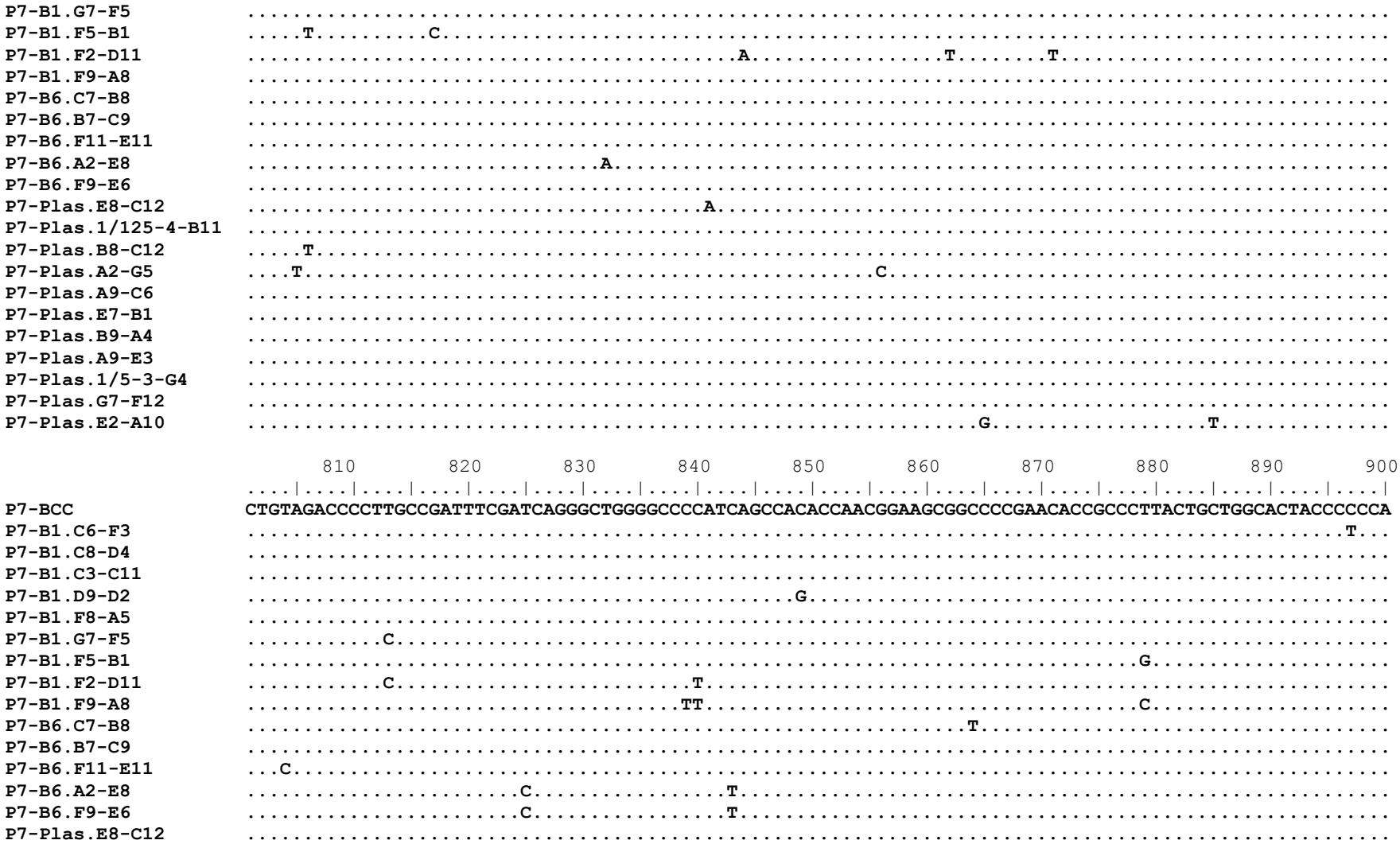

P7-Plas.E2-A10
110 120 130 140 150 160 170 180 190 200
P7-BCC GGTGTGTCCCTTGCGTTCGCGAGGGYAACAGCTCGAAGTGTTGGGTGGCGGTGGCCCCCACAGTCGCCACCAGGGACGGCARACTCCCCGCAACACAGCT
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
210 220 230 240 250 260 270 280 290 300
P7-BCC TCGACGTCACATCGATCTGCTTGTCGGGAGCGCCACCCTCTGTTCGGCCCTCTATGTGGGGGACTTATGCGGGTCTGTCTTTCTTGTCGGCCAACTGTTC
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5

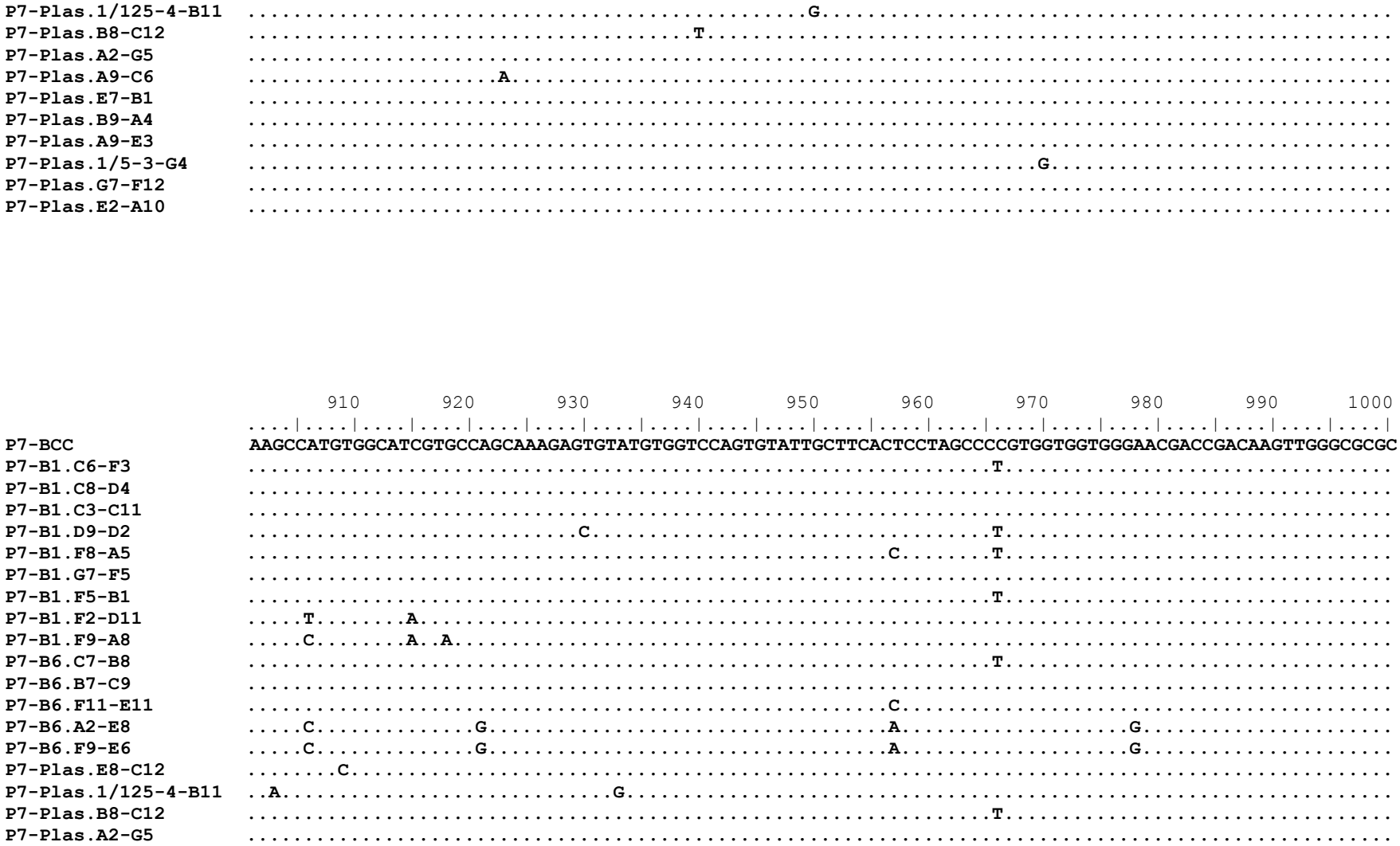

P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
310
320
330
340
350
360
370
380
390
400
P7-BCC
ACTTTCTCCCCCAGRCACCACTGGACRACGCAAGACTGCAACTGCTCCATCTACCCCGGCCATATAACGGGTCACCGCATGGCATGGGATATGATGATGA
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9

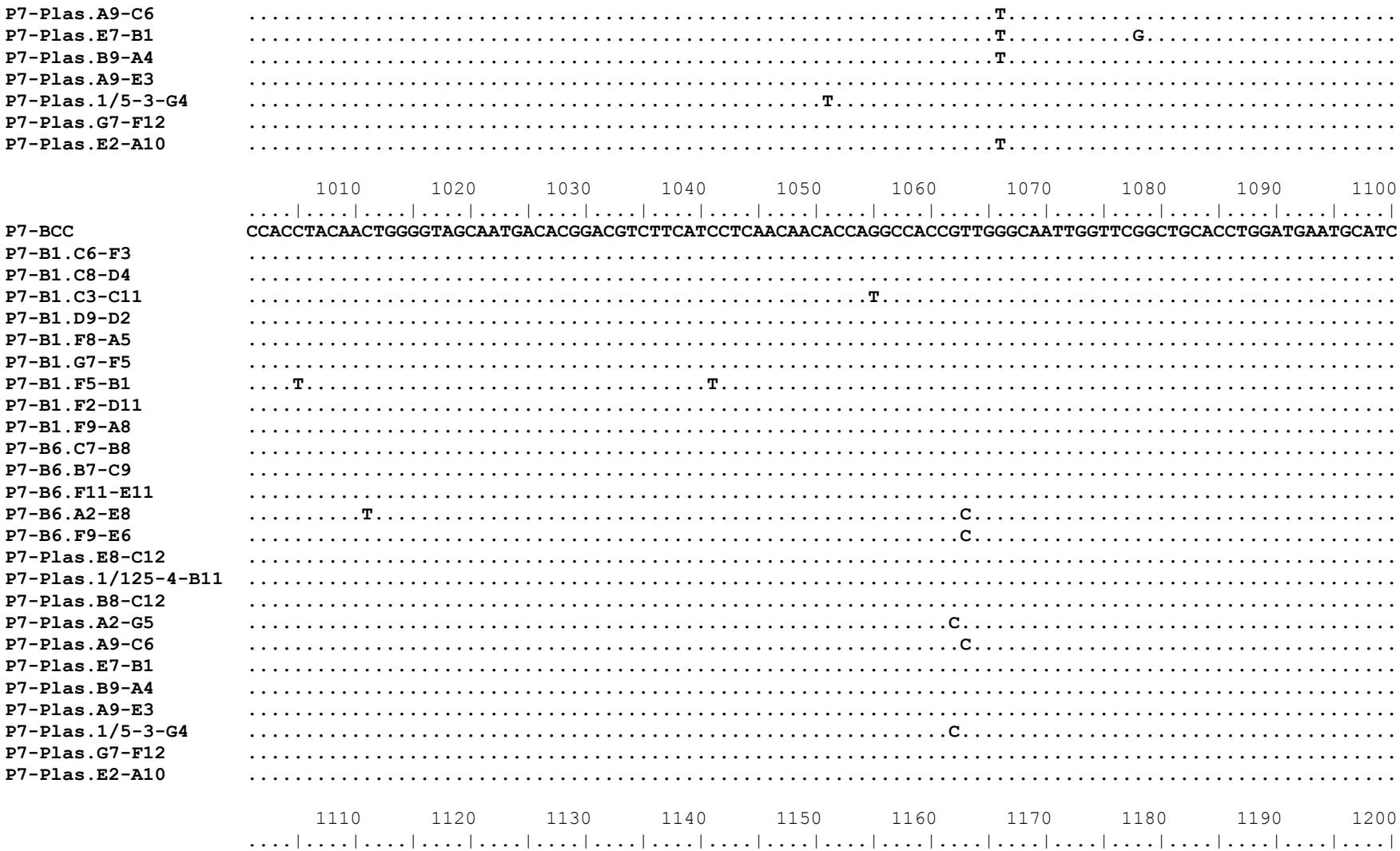

P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
410
420
430
440
450
460
470
480
490
500
P7-BCC
ATTGGTCCCCTACAACAGCGCTGGTAGTAGCTCAGCTGCTCAGGGTCCCGCAAGCCATCTTGGACATGATCGCTGGTGCCCACTGGGGAGTCCTAGCGGG
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1

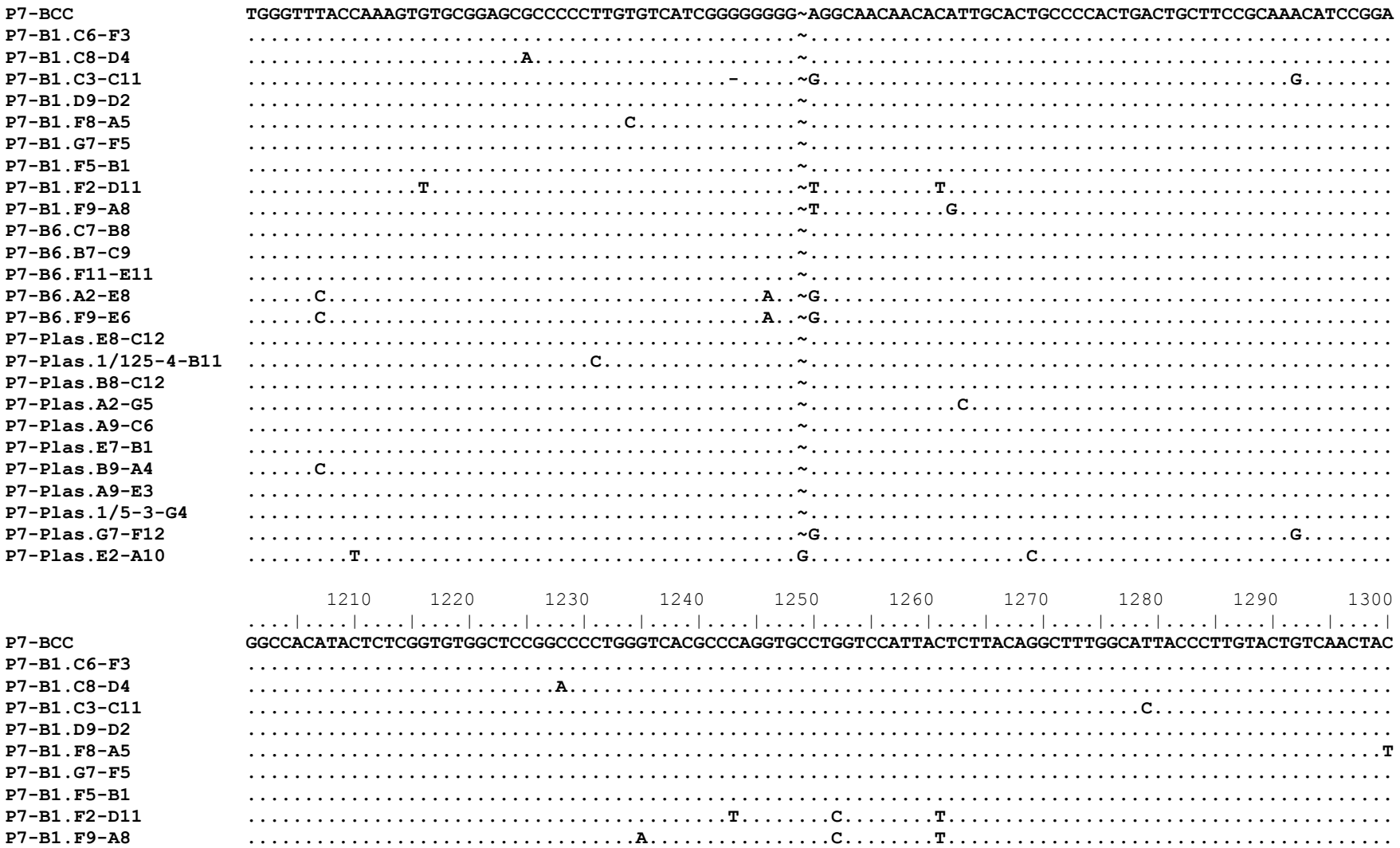

P7-Plas.B9-A4 .....G.....
P7-Plas.A9-E3 .....
P7-Plas.1/5-3-G4 .....
P7-Plas.G7-F12 .....C.....G.....
P7-Plas.E2-A10 .....
510 520 530 540 550 560 570 580 590 600
P7-BCC CATAGCGTAYTTCTCCATGGTGGGGAACTGGGCGAAGGTCCTGGTGGTGCTGTTGCTGTTTGCCAGCGTCGATGCAGATACCTACACCACCGGGGGAACT
P7-B1.C6-F3 .....C.....
P7-B1.C8-D4 .....
P7-B1.C3-C11 .....C.....
P7-B1.D9-D2 .....
P7-B1.F8-A5 .....
P7-B1.G7-F5 .....C.....
P7-B1.F5-B1 ....A.....A.....A.....
P7-B1.F2-D11 .....
P7-B1.F9-A8 .....T.....
P7-B6.C7-B8 .....
P7-B6.B7-C9 .....T.....T.....
P7-B6.F11-E11 .....
P7-B6.A2-E8 .....T.....T.....
P7-B6.F9-E6 .....T.....
P7-Plas.E8-C12 .....
P7-Plas.1/125-4-B11 .....
P7-Plas.B8-C12 .....
P7-Plas.A2-G5 .....T.....
P7-Plas.A9-C6 .....T.....
P7-Plas.E7-B1 .....
P7-Plas.B9-A4 .....
P7-Plas.A9-E3 .....
P7-Plas.1/5-3-G4 .....
P7-Plas.G7-F12 .....
P7-Plas.E2-A10 .....G.....T.....

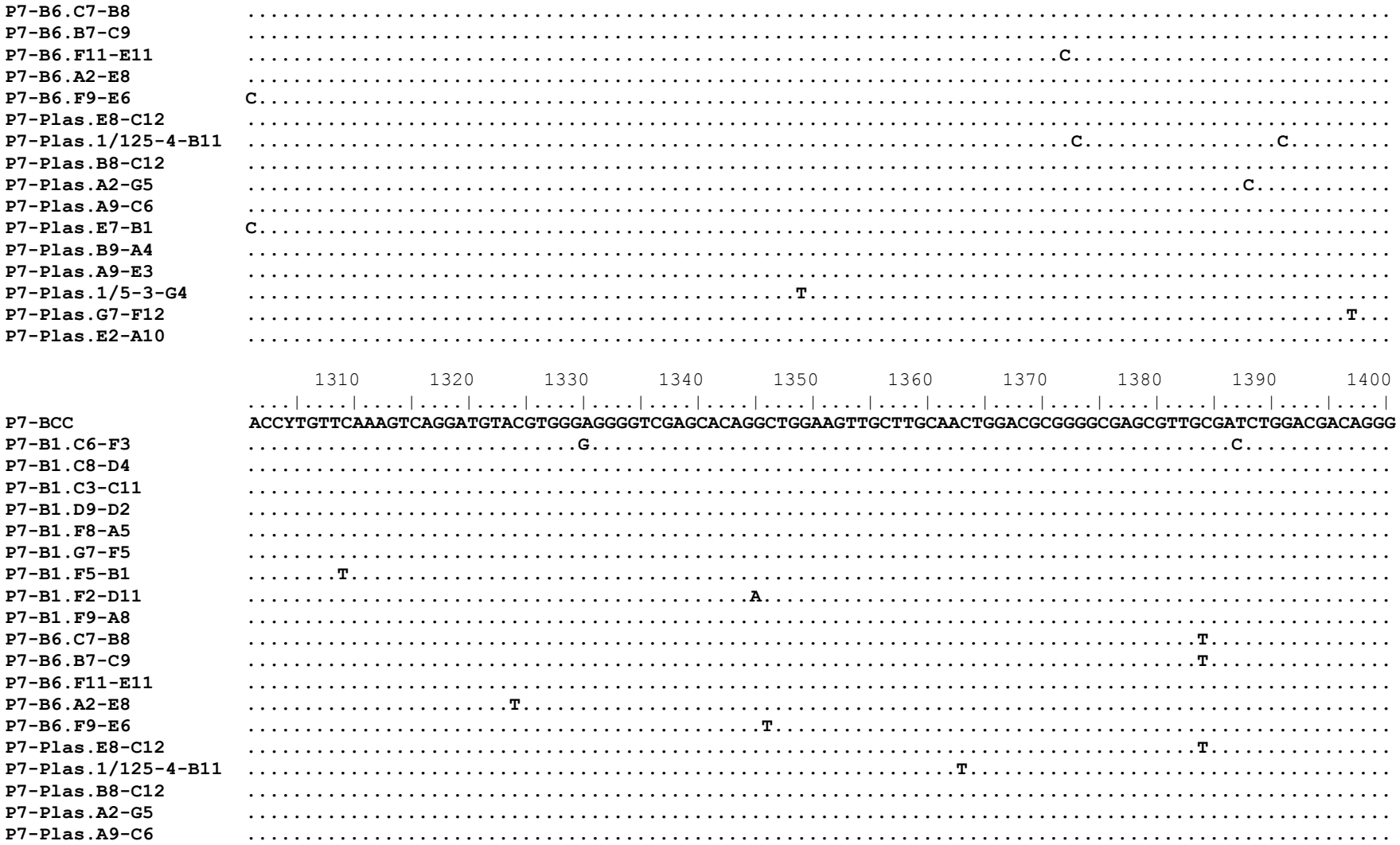

610 620 630 640 650 660 670 680 690 700
P7-BCC GCCGCTAGGGGCGCGYTCGGGCTCGCCAACATCTTCAGTCCGGGCGCTAAGCAGAACATCCAGCTGATCAACACCAACGGCAGTTGGCACATCAATCGCA
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
710 720 730 740 750 760 770 780 790 800
P7-BCC CGGCCCTGAACTGYAATGCGAGTCTCGACACTGGCTGGGTGGCGGGGCTCCTCTATTACCACAAATTCAACTCTTCAGGCTGYACCGAGAGGATGGCCAG
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5

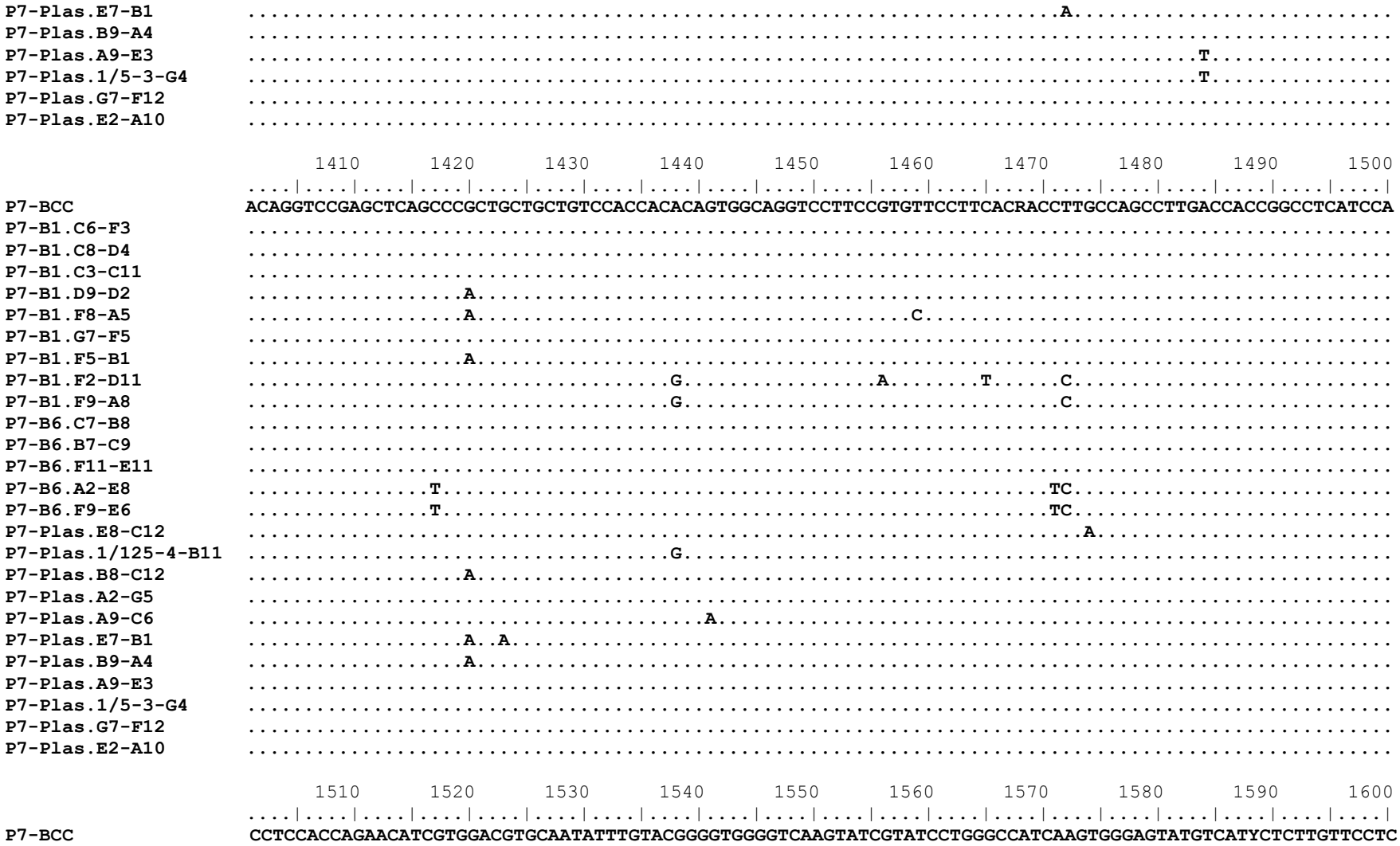

P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
810
820
830
840
850
860
870
880
890
900
P7-BCC
CTGTAGACCCCTTGCCGATTTCGATCAGGGCTGGGGCCCCATCAGCCACACCAACGGAAGCGGCCCCGAACACCGCCCTTACTGCTGGCACTACCCCCCA
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12

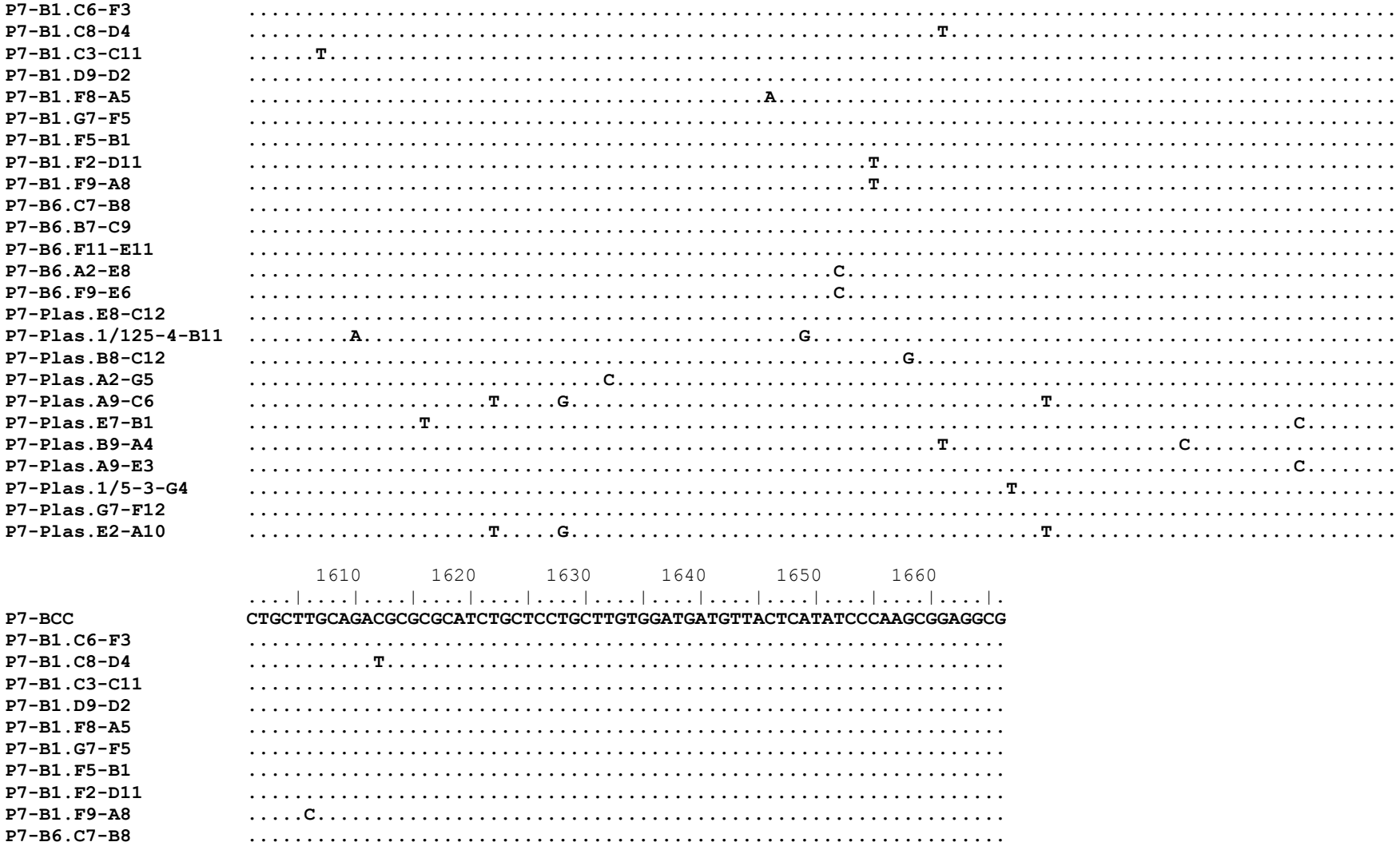

P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
910
920
930
940
950
960
970
980
990
1000
P7-BCC
AAGCCATGTGGCATCGTGCCAGCAAAGAGTGTATGTGGTCCAGTGTATTGCTTCACTCCTAGCCCCGTGGTGGTGGGAACGACCGACAAGTTGGGCGCGC
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5


P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
P7-BCC
CCACCTACAACTGGGGTAGCAATGACACGGACGTCTTCATCCTCAACAACACCAGGCCACCGTTGGGCAATTGGTTCGGCTGCACCTGGATGAATGCATC
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
1110 1120 1130 1140 1150 1160 1170 1180 1190 1200


P7-BCC TGGGTTTACCAAAGTGTGCGGAGCGCCCCCTTGTGTCATCGGGGGGGG~AGGCAACAACACATTGCACTGCCCCACTGACTGCTTCCGCAAACATCCGGA
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
P7-BCC GGCCACATACTCTCGGTGTGGCTCCGGCCCCTGGGTCACGCCCAGGTGCCTGGTCCATTACTCTTACAGGCTTTGGCATTACCCTTGTACTGTCAACTAC
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8


P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
P7-BCC ACCYTGTTCAAAGTCAGGATGTACGTGGGAGGGGTCGAGCACAGGCTGGAAGTTGCTTGCAACTGGACGCGGGGCGAGCGTTGCGATCTGGACGACAGGG
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6


P7-Plas.E7-B1 ....A.....
P7-Plas.B9-A4 ....
P7-Plas.A9-E3 ....T.....
P7-Plas.1/5-3-G4 ....T.....
P7-Plas.G7-F12 ....
P7-Plas.E2-A10 ....
1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
P7-BCC ACAGGTCCGAGCTCAGCCCGCTGCTGCTGTCCACCACACAGTGGCAGGTCCTTCCGTGTTCCTTCACRACCTTGCCAGCCTTGACCACCGGCCTCATCCA
P7-B1.C6-F3 ....
P7-B1.C8-D4 ....
P7-B1.C3-C11 ....
P7-B1.D9-D2 ....A.....
P7-B1.F8-A5 ..........A.....C.....
P7-B1.G7-F5 ....
P7-B1.F5-B1 ....A.....
P7-B1.F2-D11 ..........G.....A.....T.....C.....
P7-B1.F9-A8 ..........G.....C.....
P7-B6.C7-B8 ....
P7-B6.B7-C9 ....
P7-B6.F11-E11 ....
P7-B6.A2-E8 ..........T.....TC.....
P7-B6.F9-E6 ..........T.....TC.....
P7-Plas.E8-C12 ....A.....
P7-Plas.1/125-4-B11 ....G.....
P7-Plas.B8-C12 ....A.....
P7-Plas.A2-G5 ....
P7-Plas.A9-C6 ....A.....
P7-Plas.E7-B1 ..........A..A.....
P7-Plas.B9-A4 ....A.....
P7-Plas.A9-E3 ....
P7-Plas.1/5-3-G4 ....
P7-Plas.G7-F12 ....
P7-Plas.E2-A10 ....
1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
P7-BCC CCTCCACCAGAACATCGTGGACGTGCAATATTTGTACGGGGTGGGGTCAAGTATCGTATCCTGGGCCATCAAGTGGGAGTATGTCATYCTCTTGTTCCTC

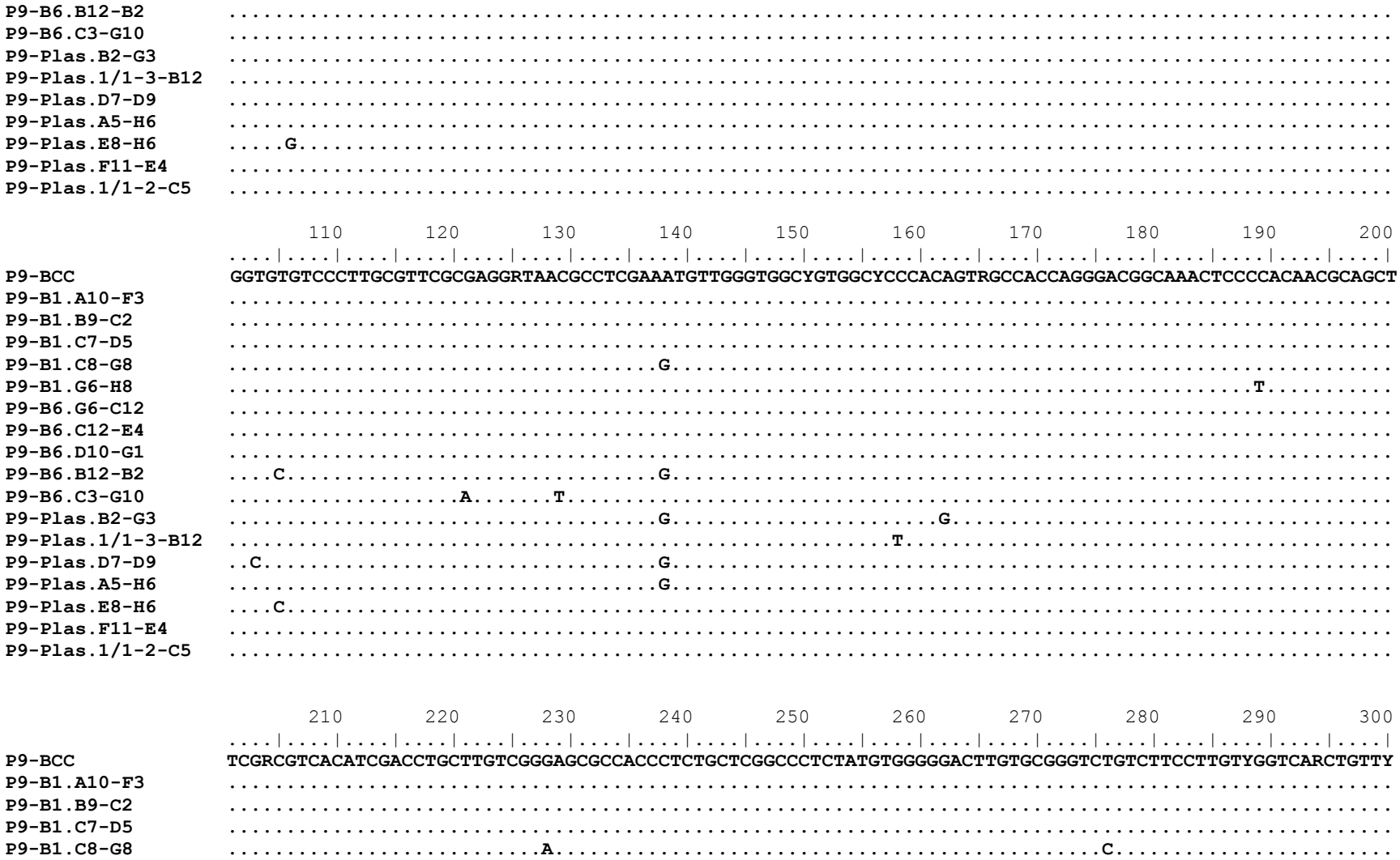

P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8
P7-B6.B7-C9
P7-B6.F11-E11
P7-B6.A2-E8
P7-B6.F9-E6
P7-Plas.E8-C12
P7-Plas.1/125-4-B11
P7-Plas.B8-C12
P7-Plas.A2-G5
P7-Plas.A9-C6
P7-Plas.E7-B1
P7-Plas.B9-A4
P7-Plas.A9-E3
P7-Plas.1/5-3-G4
P7-Plas.G7-F12
P7-Plas.E2-A10
1610
1620
1630
1640
1650
1660
P7-BCC
CTGCTTGCAGACGCGCGCATCTGCTCCTGCTTGTGGATGATGTTACTCATATCCCAAGCGGAGGCG
P7-B1.C6-F3
P7-B1.C8-D4
P7-B1.C3-C11
P7-B1.D9-D2
P7-B1.F8-A5
P7-B1.G7-F5
P7-B1.F5-B1
P7-B1.F2-D11
P7-B1.F9-A8
P7-B6.C7-B8

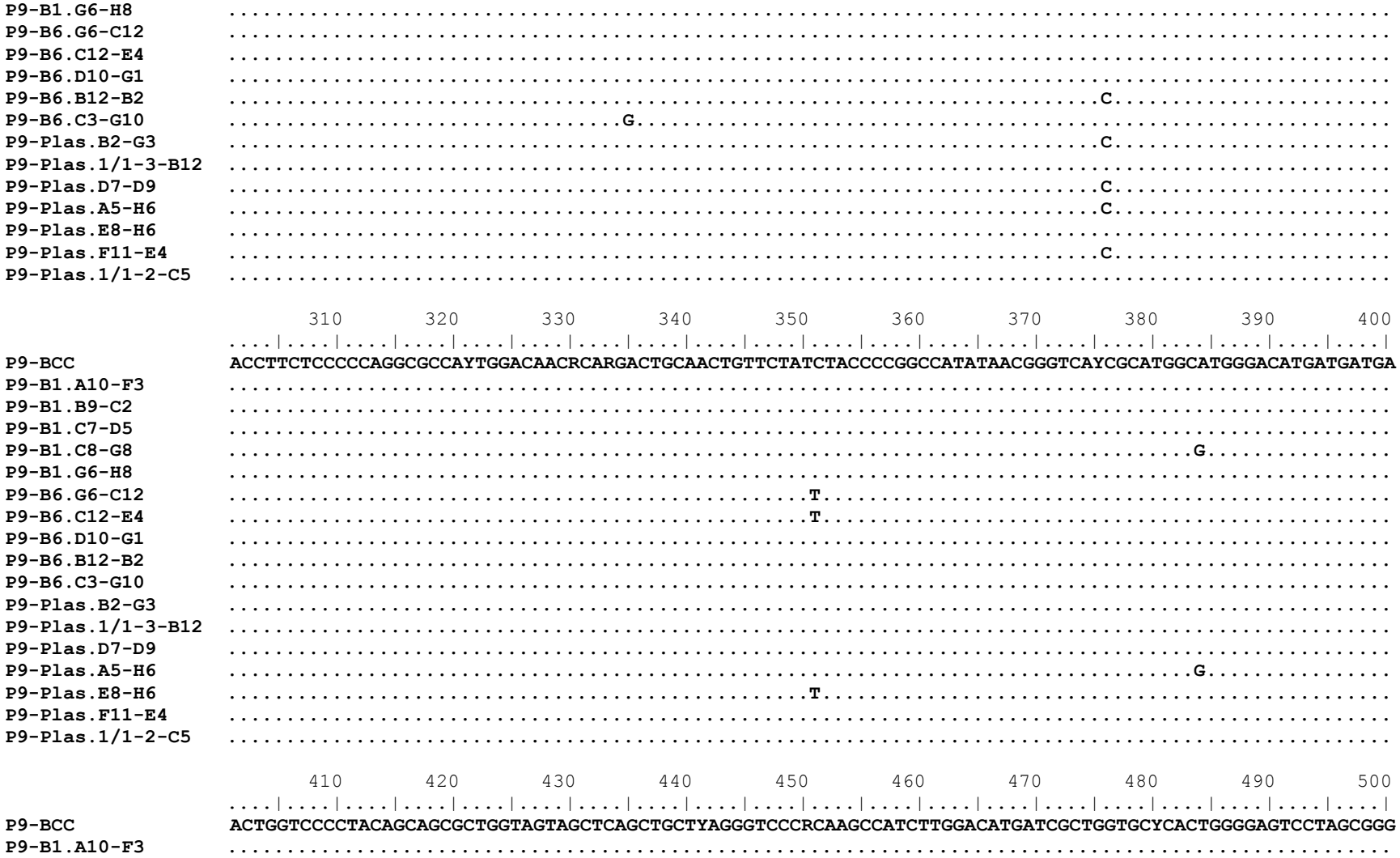
```
P7-B6.B7-C9           ............................................................
P7-B6.F11-E11         ............................................................AA
P7-B6.A2-E8           ............................................................
P7-B6.F9-E6           ............................................................
P7-Plas.E8-C12        ............................................................AA.
P7-Plas.1/125-4-B11   ............................................................
P7-Plas.B8-C12        ............................................................
P7-Plas.A2-G5         ............................................................
P7-Plas.A9-C6         ............................................................
P7-Plas.E7-B1         ............................................................
P7-Plas.B9-A4         ...........................A................................
P7-Plas.A9-E3         ............................................................
P7-Plas.1/5-3-G4      ............................................................
P7-Plas.G7-F12        ............................................................
P7-Plas.E2-A10        ............................................................
```

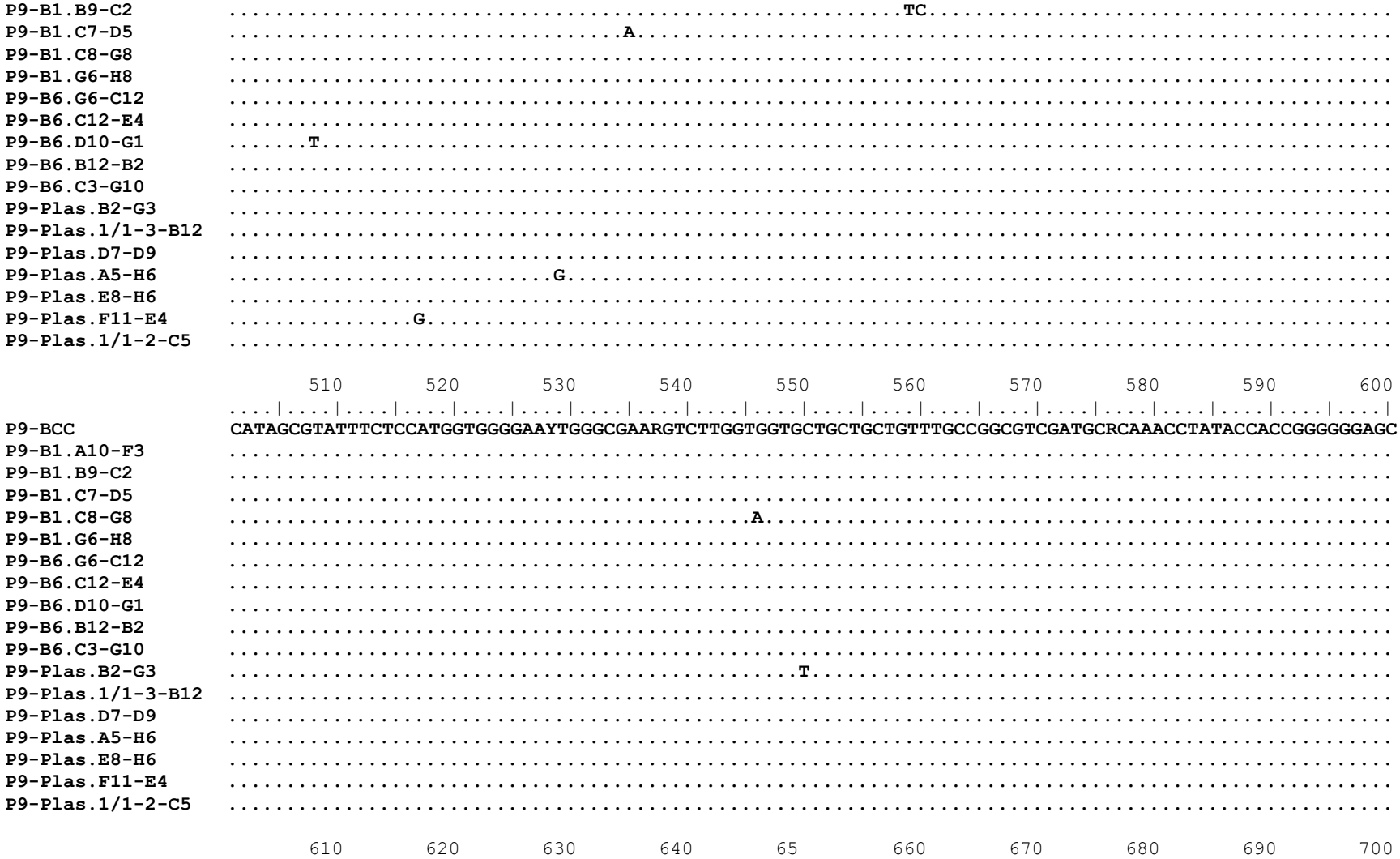
10.1.6 Patient 9

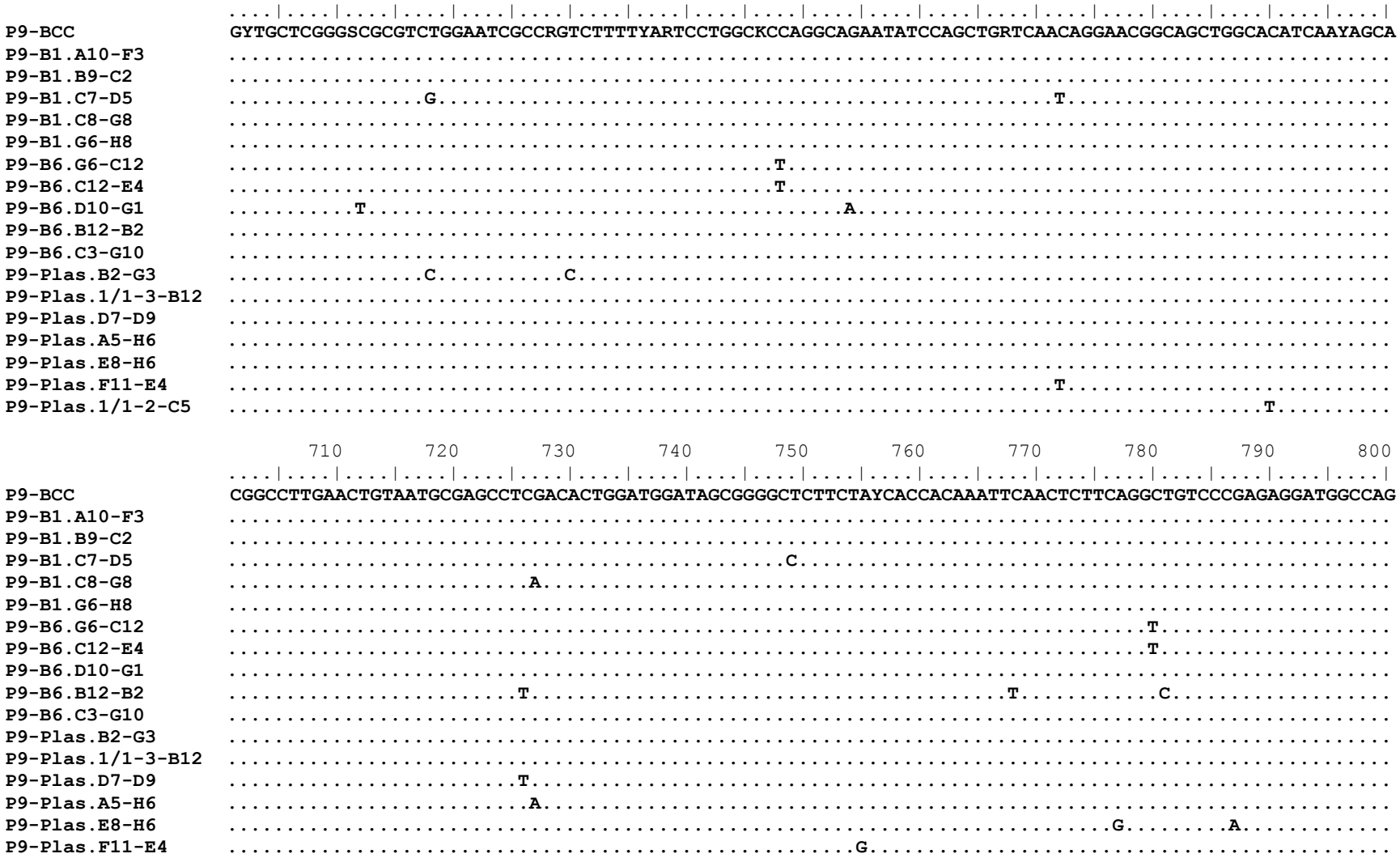
Gene position: E1: 1 - 576nt, E2: 577 - 1665nt, HVR1: 577 – 657nt.

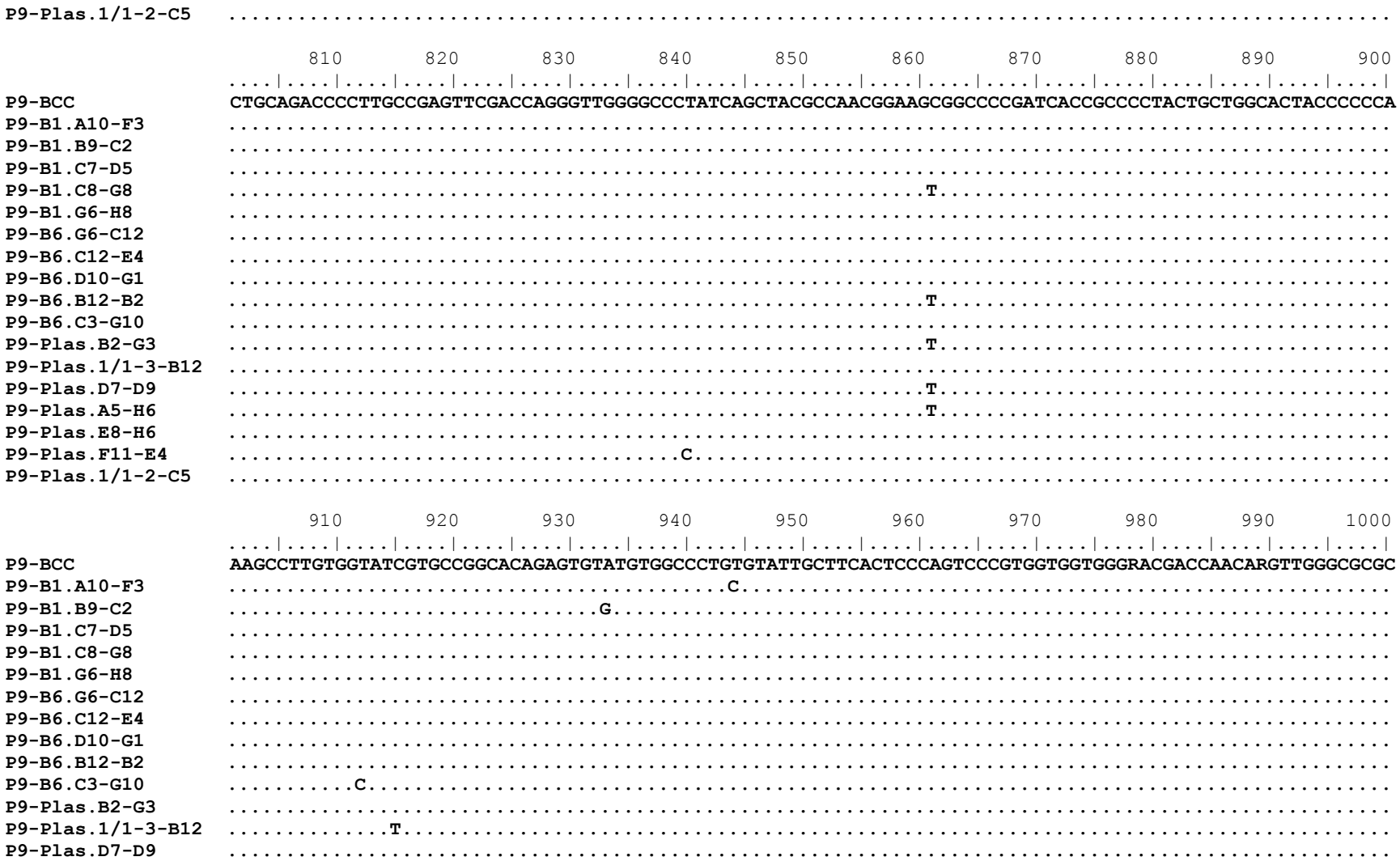
```
                              10        20        30        40        50        60        70        80        90        100
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
P9-BCC                TATCAAGTACGCAACTCCTCGGGCCTCTACCATGTCACCAATGATTGCCCTAACTCGAGTATTGTGTACGAGACGGCCGACACCATCCTACACTTTCCGG
P9-B1.A10-F3          ....................................................................................................
P9-B1.B9-C2           ....................................................................................................
P9-B1.C7-D5           ....................................................................................................
P9-B1.C8-G8           ..........................................................C.........................................
P9-B1.G6-H8           ..........................................................G.........................................
P9-B6.G6-C12          ....................................................................................................
P9-B6.C12-E4          ....................................................................................................
P9-B6.D10-G1          ..........................................A.........................................................
```

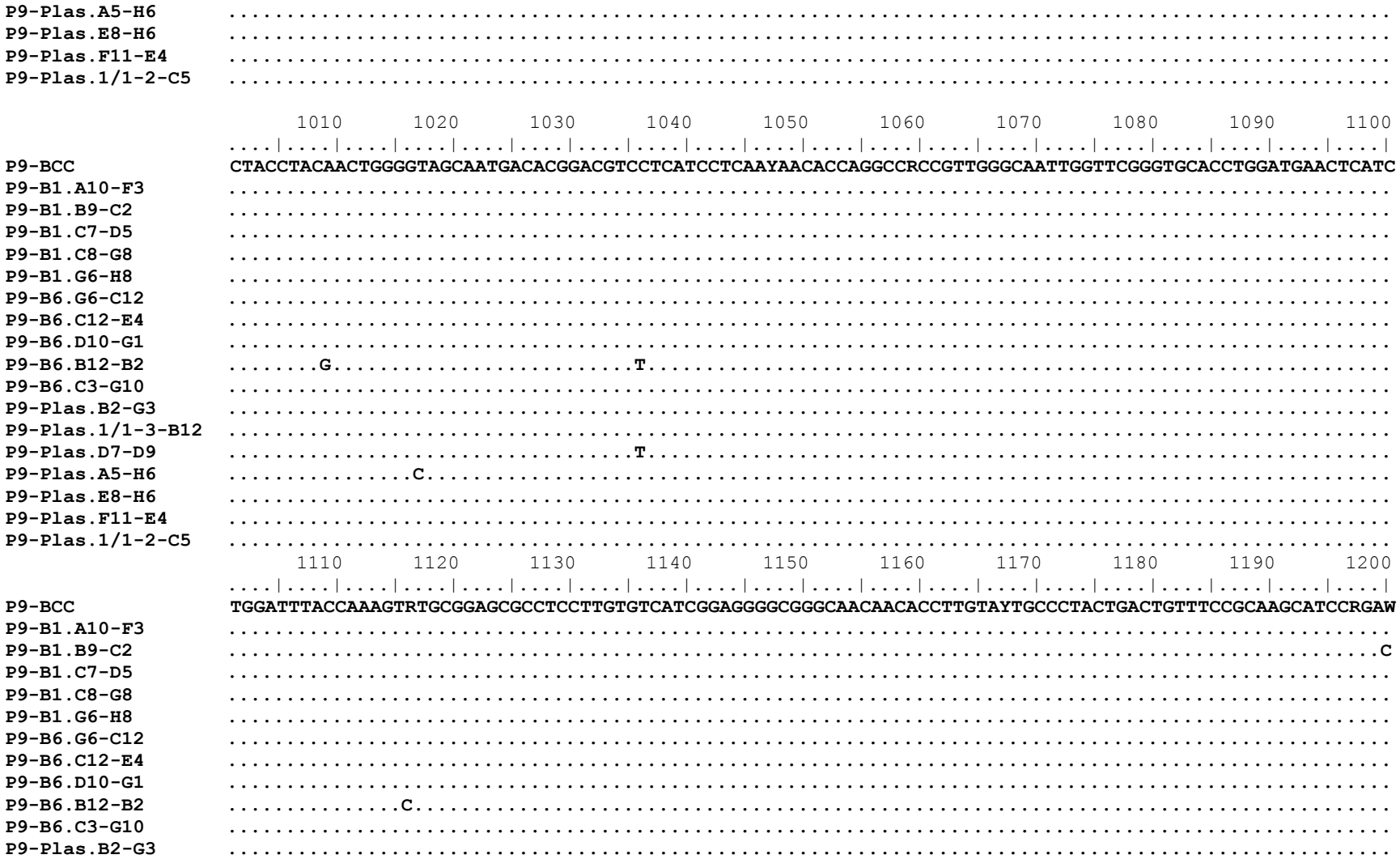

P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
110 120 130 140 150 160 170 180 190 200
P9-BCC GGTGTGTCCCTTGCGTTCGCGAGGRTAACGCCTCGAAATGTTGGGTGGCYGTGGCYCCCACAGTRGCCACCAGGGACGGCAAACTCCCCACAACGCAGCT
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
210 220 230 240 250 260 270 280 290 300
P9-BCC TCGRCGTCACATCGACCTGCTTGTCGGGAGCGCCACCCTCTGCTCGGCCCTCTATGTGGGGGACTTGTGCGGGTCTGTCTTCCTTGTYGGTCARCTGTTY
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8

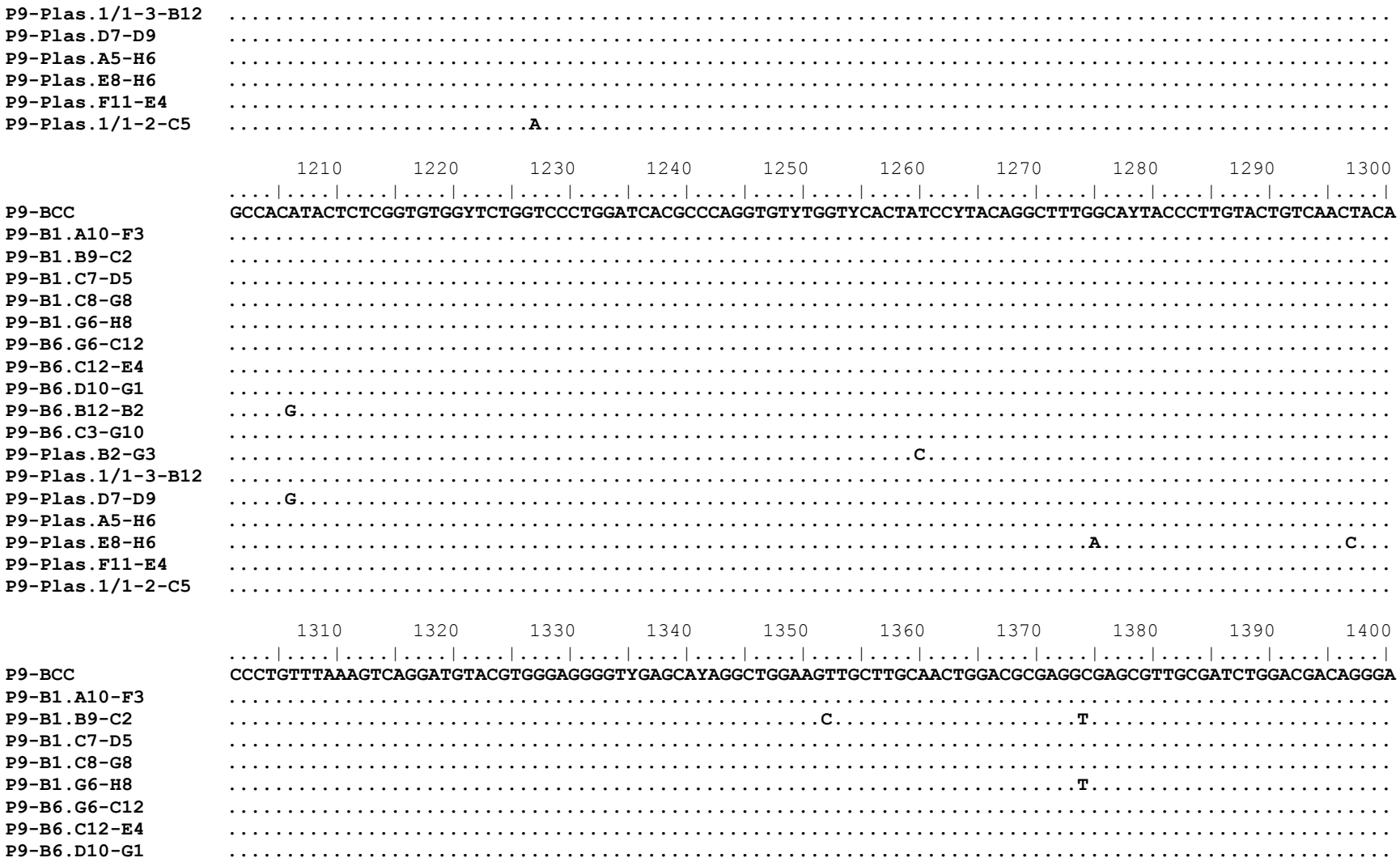

P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
310 320 330 340 350 360 370 380 390 400
P9-BCC ACCTTCTCCCCCAGGCGCCAYTGGACAACRCARGACTGCAACTGTTCTATCTACCCCGGCCATATAACGGGTCAYCGCATGGCATGGGACATGATGATGA
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
410 420 430 440 450 460 470 480 490 500
P9-BCC ACTGGTCCCCTACAGCAGCGCTGGTAGTAGCTCAGCTGCTYAGGGTCCCRCAAGCCATCTTGGACATGATCGCTGGTGCYCACTGGGGAGTCCTAGCGGG
P9-B1.A10-F3

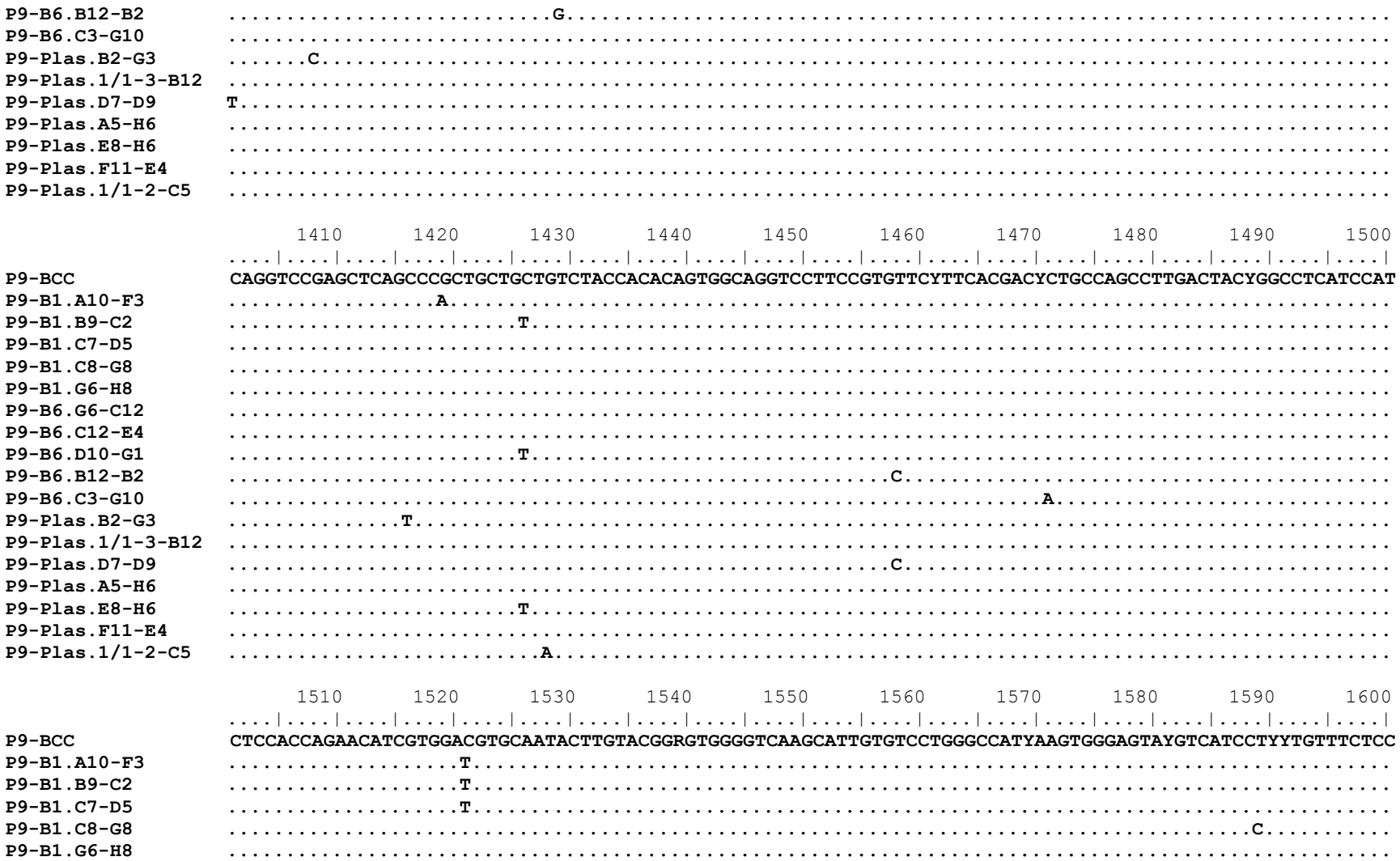

P9-B1.B9-C2 ....TC..........
P9-B1.C7-D5 ....A..........
P9-B1.C8-G8 ....
P9-B1.G6-H8 ....
P9-B6.G6-C12 ....
P9-B6.C12-E4 ....
P9-B6.D10-G1 ....T..........
P9-B6.B12-B2 ....
P9-B6.C3-G10 ....
P9-Plas.B2-G3 ....
P9-Plas.1/1-3-B12 ....
P9-Plas.D7-D9 ....
P9-Plas.A5-H6 ....G..........
P9-Plas.E8-H6 ....
P9-Plas.F11-E4 ....G..........
P9-Plas.1/1-2-C5 ....
510 520 530 540 550 560 570 580 590 600
P9-BCC CATAGCGTATTTCTCCATGGTGGGGAAYTGGGCGAARGTCTTGGTGGTGCTGCTGCTGTTTGCCGGCGTCGATGCRCAAACCTATACCACCGGGGGGAGC
P9-B1.A10-F3 ....
P9-B1.B9-C2 ....
P9-B1.C7-D5 ....
P9-B1.C8-G8 ....A..........
P9-B1.G6-H8 ....
P9-B6.G6-C12 ....
P9-B6.C12-E4 ....
P9-B6.D10-G1 ....
P9-B6.B12-B2 ....
P9-B6.C3-G10 ....
P9-Plas.B2-G3 ....T..........
P9-Plas.1/1-3-B12 ....
P9-Plas.D7-D9 ....
P9-Plas.A5-H6 ....
P9-Plas.E8-H6 ....
P9-Plas.F11-E4 ....
P9-Plas.1/1-2-C5 ....
610 620 630 640 65 660 670 680 690 700

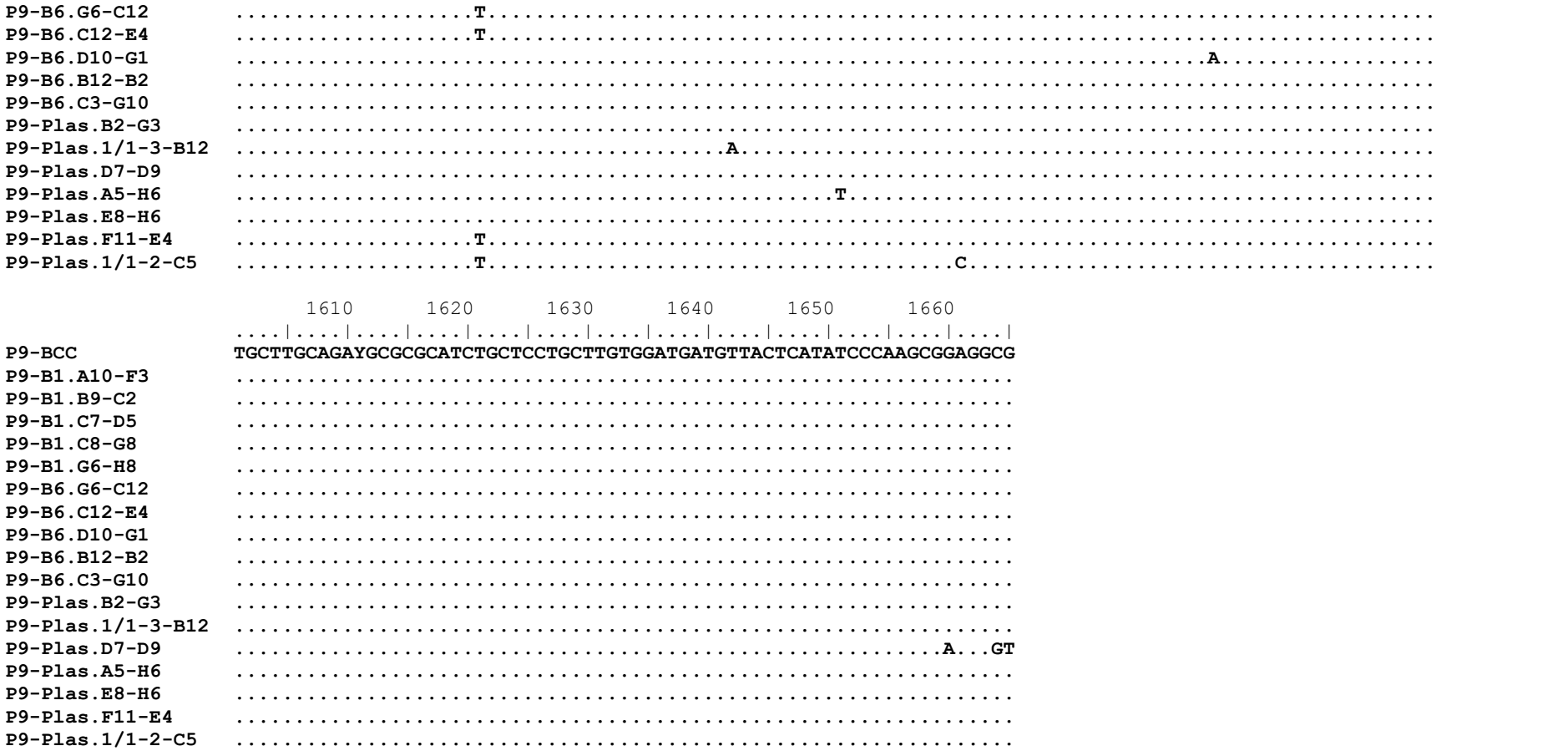

P9-BCC GYTGCTCGGGSCGCGTCTGGAATCGCCRGTCTTTTYARTCCTGGCKCCAGGCAGAATATCCAGCTGRTCAACAGGAACGGCAGCTGGCACATCAAYAGCA
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
710 720 730 740 750 760 770 780 790 800
P9-BCC CGGCCTTGAACTGTAATGCGAGCCTCGACACTGGATGGATAGCGGGGCTCTTCTAYCACCACAAATTCAACTCTTCAGGCTGTCCCGAGAGGATGGCCAG
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4


P9-Plas.1/1-2-C5
810 820 830 840 850 860 870 880 890 900
P9-BCC CTGCAGACCCCTTGCCGAGTTCGACCAGGGTTGGGGCCCTATCAGCTACGCCAACGGAAGCGGCCCCGATCACCGCCCCTACTGCTGGCACTACCCCCCA
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
910 920 930 940 950 960 970 980 990 1000
P9-BCC AAGCCTTGTGGTATCGTGCCGGCACAGAGTGTATGTGGCCCTGTGTATTGCTTCACTCCCAGTCCCGTGGTGGTGGGRACGACCAACARGTTGGGCGCGC
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9


P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
P9-BCC CTACCTACAACTGGGGTAGCAATGACACGGACGTCCTCATCCTCAAYAACACCAGGCCRCCGTTGGGCAATTGGTTCGGGTGCACCTGGATGAACTCATC
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2 .....G.....T.....
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9 .....T.....
P9-Plas.A5-H6 .....C.....
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
P9-BCC TGGATTTACCAAAGTRTGCGGAGCGCCTCCTTGTGTCATCGGAGGGGCGGGCAACAACACCTTGTAYTGCCCTACTGACTGTTTCCGCAAGCATCCRGAW
P9-B1.A10-F3
P9-B1.B9-C2 C
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2 .....C.....
P9-B6.C3-G10
P9-Plas.B2-G3


P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
P9-BCC GCCACATACTCTCGGTGTGGYTCTGGTCCCTGGATCACGCCCAGGTGTYTGGTYCACTATCCYTACAGGCTTTGGCAYTACCCTTGTACTGTCAACTACA
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1
P9-B6.B12-B2
P9-B6.C3-G10
P9-Plas.B2-G3
P9-Plas.1/1-3-B12
P9-Plas.D7-D9
P9-Plas.A5-H6
P9-Plas.E8-H6
P9-Plas.F11-E4
P9-Plas.1/1-2-C5
1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
P9-BCC CCCTGTTTAAAGTCAGGATGTACGTGGGAGGGGTYGAGCAYAGGCTGGAAGTTGCTTGCAACTGGACGCGAGGCGAGCGTTGCGATCTGGACGACAGGGA
P9-B1.A10-F3
P9-B1.B9-C2
P9-B1.C7-D5
P9-B1.C8-G8
P9-B1.G6-H8
P9-B6.G6-C12
P9-B6.C12-E4
P9-B6.D10-G1


P9-B6.B12-B2 ....G..........
P9-B6.C3-G10 ....
P9-Plas.B2-G3 ....C..........
P9-Plas.1/1-3-B12 ....
P9-Plas.D7-D9 T.........
P9-Plas.A5-H6 ....
P9-Plas.E8-H6 ....
P9-Plas.F11-E4 ....
P9-Plas.1/1-2-C5 ....
1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
P9-BCC CAGGTCCGAGCTCAGCCCGCTGCTGCTGTCTACCACACAGTGGCAGGTCCTTCCGTGTTCYTTCACGACYCTGCCAGCCTTGACTACYGGCCTCATCCAT
P9-B1.A10-F3 ....A..........
P9-B1.B9-C2 ....T..........
P9-B1.C7-D5 ....
P9-B1.C8-G8 ....
P9-B1.G6-H8 ....
P9-B6.G6-C12 ....
P9-B6.C12-E4 ....
P9-B6.D10-G1 ....T..........
P9-B6.B12-B2 ....C..........
P9-B6.C3-G10 ....A..........
P9-Plas.B2-G3 ....T..........
P9-Plas.1/1-3-B12 ....
P9-Plas.D7-D9 ....C..........
P9-Plas.A5-H6 ....
P9-Plas.E8-H6 ....T..........
P9-Plas.F11-E4 ....
P9-Plas.1/1-2-C5 ....A..........
1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
P9-BCC CTCCACCAGAACATCGTGGACGTGCAATACTTGTACGGRGTGGGGTCAAGCATTGTGTCCTGGGCCATYAAGTGGGAGTAYGTCATCCTYYTGTTTCTCC
P9-B1.A10-F3 ....T..........
P9-B1.B9-C2 ....T..........
P9-B1.C7-D5 ....T..........
P9-B1.C8-G8 ....C..........
P9-B1.G6-H8 ....


P9-B6.G6-C12 ....T.....
P9-B6.C12-E4 ....T.....
P9-B6.D10-G1 ....A.....
P9-B6.B12-B2 ....
P9-B6.C3-G10 ....
P9-Plas.B2-G3 ....
P9-Plas.1/1-3-B12 ....A.....
P9-Plas.D7-D9 ....
P9-Plas.A5-H6 ....T.....
P9-Plas.E8-H6 ....
P9-Plas.F11-E4 ....T.....
P9-Plas.1/1-2-C5 ..................T.....C.....
1610 1620 1630 1640 1650 1660
P9-BCC TGCTTGCAGAYGCGCGCATCTGCTCCTGCTTGTGGATGATGTTACTCATATCCCAAGCGGAGGCG
P9-B1.A10-F3 .....
P9-B1.B9-C2 .....
P9-B1.C7-D5 .....
P9-B1.C8-G8 .....
P9-B1.G6-H8 .....
P9-B6.G6-C12 .....
P9-B6.C12-E4 .....
P9-B6.D10-G1 .....
P9-B6.B12-B2 .....
P9-B6.C3-G10 .....
P9-Plas.B2-G3 .....
P9-Plas.1/1-3-B12 .....
P9-Plas.D7-D9 .....A...GT
P9-Plas.A5-H6 .....
P9-Plas.E8-H6 .....
P9-Plas.F11-E4 .....
P9-Plas.1/1-2-C5 .....

10.2 Appendix 2

Mean genetic distance between SGA-derived HCV sequences isolated from plasma and liver.

STDEV: Standard Deviation. ND: Not Done, due to limited data.

P-values are calculated by two-way unpaired t-test between the genetic distances of individual sequences.

NS: non-statistical significant, *: $p<0.05$, **: $p<0.005$, ***: $p<0.001$

Table 10.1 : Mean genetic distance (patient 1)

Patient 1 (STDEV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	0.0	1×10^{-2} ($\pm 6 \times 10^{-3}$) ND	3×10^{-3} ($\pm 2 \times 10^{-3}$) ND	3×10^{-3} ($\pm 2 \times 10^{-3}$) ND
Biopsy 1	1×10^{-2} ($\pm 6 \times 10^{-3}$) ND	2×10^{-2} ($\pm 8 \times 10^{-3}$)	2×10^{-2} ($\pm 7 \times 10^{-3}$)*	2×10^{-2} ($\pm 8 \times 10^{-3}$) *
Biopsy 6	3×10^{-3} ($\pm 2 \times 10^{-3}$) ND	2×10^{-2} ($\pm 7 \times 10^{-3}$)*	8×10^{-3} ($\pm 4 \times 10^{-3}$)	1×10^{-2} ($\pm 7 \times 10^{-3}$) **
Plasma	3×10^{-3} ($\pm 2 \times 10^{-3}$) ND	2×10^{-2} ($\pm 8 \times 10^{-3}$) *	1×10^{-2} ($\pm 7 \times 10^{-3}$) **	2×10^{-2} ($\pm 8 \times 10^{-3}$)

Table 10.2: Mean genetic distance (patient 2)

Patient 2 (STDEV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	2×10^{-4} ($\pm 3 \times 10^{-4}$)	3×10^{-2} ($\pm 2 \times 10^{-2}$)***	9×10^{-3} ($\pm 4 \times 10^{-3}$)***	1×10^{-2} ($\pm 6 \times 10^{-3}$)***
Biopsy 1	3×10^{-2} ($\pm 2 \times 10^{-2}$)***	1×10^{-2} ($\pm 5 \times 10^{-3}$)	1×10^{-2} ($\pm 6 \times 10^{-3}$)***	1×10^{-2} ($\pm 2 \times 10^{-2}$)NS
Biopsy 6	9×10^{-3} ($\pm 4 \times 10^{-3}$)***	1×10^{-2} ($\pm 6 \times 10^{-3}$)***	3×10^{-2} ($\pm 2 \times 10^{-2}$)	3×10^{-2} ($\pm 7 \times 10^{-3}$)***
Plasma	1×10^{-2} ($\pm 6 \times 10^{-3}$)***	1×10^{-2} ($\pm 2 \times 10^{-2}$)NS	3×10^{-2} ($\pm 7 \times 10^{-3}$)***	1×10^{-2} ($\pm 5 \times 10^{-3}$)

Table 10.3: Mean genetic distance (patient 4)					
Patient 4 (STDEV)	Consensus	Biopsy 1	Biopsy 3	Biopsy 7	Plasma
Consensus	0.0	2×10^{-3} ($\pm 1 \times 10^{-3}$)^{NS}	1×10^{-3} ($\pm 5 \times 10^{-4}$) ND	3×10^{-3} ($\pm 2 \times 10^{-3}$) ND	2×10^{-3} ($\pm 4 \times 10^{-3}$)ND
Biopsy 1	2×10^{-3} ($\pm 1 \times 10^{-3}$) ^{NS}	1×10^{-2} ($\pm 6 \times 10^{-3}$)	9×10^{-3} ($\pm 7 \times 10^{-3}$) ND	1×10^{-2} ($\pm 6 \times 10^{-3}$)^{NS}	1×10^{-2} ($\pm 7 \times 10^{-3}$)^{NS}
Biopsy 3	1×10^{-3} ($\pm 5 \times 10^{-4}$) ND	9×10^{-3} ($\pm 7 \times 10^{-3}$) ND	4×10^{-3}	1×10^{-2} ($\pm \times 10^{-3}$)	1×10^{-2} ($\pm 5 \times 10^{-3}$)^{NS}
Biopsy 7	3×10^{-3} ($\pm 2 \times 10^{-3}$) ND	1×10^{-2} ($\pm 6 \times 10^{-3}$)^{NS}	1×10^{-2} ($\pm \times 10^{-3}$) ND	1×10^{-2} ($\pm 8 \times 10^{-3}$)	1×10^{-2} ($\pm 8 \times 10^{-3}$)^{NS}
Plasma	2×10^{-3} ($\pm 4 \times 10^{-3}$)ND	1×10^{-2} ($\pm 7 \times 10^{-3}$)^{NS}	1×10^{-2} ($\pm 5 \times 10^{-3}$)^{NS}	1×10^{-2} ($\pm 8 \times 10^{-3}$)^{NS}	1×10^{-2} ($\pm 7 \times 10^{-3}$)^{NS}

Table 10.4: Mean genetic distance (patient 5)				
Patient 5 (STDEV)	Consensus	Biopsy 1	Biopsy 3	Biopsy 6
Consensus	2×10^{-4} ($\pm 3 \times 10^{-4}$)	9×10^{-3} ($\pm 1 \times 10^{-2}$)***	1×10^{-2} ($\pm 2 \times 10^{-2}$)***	5×10^{-3} ($\pm 3 \times 10^{-3}$)***
Biopsy 1	9×10^{-3} ($\pm 1 \times 10^{-2}$)***	3×10^{-2} ($\pm 2 \times 10^{-2}$)	3×10^{-2} ($\pm 2 \times 10^{-2}$) ^{NS}	2×10^{-2} ($\pm 2 \times 10^{-2}$) *
Biopsy 3	1×10^{-2} ($\pm 2 \times 10^{-2}$)***	3×10^{-2} ($\pm 2 \times 10^{-2}$)^{NS}	3×10^{-2} ($\pm 2 \times 10^{-2}$)	2×10^{-2} ($\pm 2 \times 10^{-2}$)**
Biopsy 6	5×10^{-3} ($\pm 3 \times 10^{-3}$)***	2×10^{-2} ($\pm 2 \times 10^{-2}$)*	2×10^{-2} ($\pm 2 \times 10^{-2}$)**	1×10^{-2} ($\pm 5 \times 10^{-3}$)

Table 10.5: Mean genetic distance (patient 7)				
Patient 7 (STDEV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	5×10^{-4} ($\pm 8 \times 10^{-4}$)	1×10^{-2} ($\pm 5 \times 10^{-3}$) ***	1×10^{-2} ($\pm 7 \times 10^{-3}$) ***	8×10^{-3} ($\pm 7 \times 10^{-3}$)***
Biopsy 1	1×10^{-2} ($\pm 5 \times 10^{-3}$)***	2×10^{-2} ($\pm 7 \times 10^{-3}$)	2×10^{-2} ($\pm 9 \times 10^{-3}$)^{NS}	2×10^{-2} ($\pm 9 \times 10^{-3}$)*
Biopsy 6	1×10^{-2} ($\pm 7 \times 10^{-3}$) ***	2×10^{-2} ($\pm 9 \times 10^{-3}$)^{NS}	2×10^{-2} ($\pm 8 \times 10^{-3}$)	2×10^{-2} ($\pm 4 \times 10^{-3}$)**
Plasma	8×10^{-3} ($\pm 7 \times 10^{-3}$)***	2×10^{-2} ($\pm 9 \times 10^{-3}$)*	2×10^{-2} ($\pm 4 \times 10^{-3}$)**	2×10^{-2} ($\pm 5 \times 10^{-3}$)

Table 10.5: Mean genetic distance (patient 9)				
Patient 9 (STDEV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	0.0	4×10^{-2} ($\pm 3 \times 10^{-2}$) ND	4×10^{-2} ($\pm 3 \times 10^{-3}$) ND	2×10^{-2} ($\pm 2 \times 10^{-3}$)ND
Biopsy 1	4×10^{-2} ($\pm 3 \times 10^{-2}$) ND	2×10^{-2} ($\pm 1 \times 10^{-2}$)	2×10^{-2} ($\pm 1 \times 10^{-2}$)NS	2×10^{-2} ($\pm 1 \times 10^{-2}$)NS
Biopsy 6	4×10^{-2} ($\pm 3 \times 10^{-3}$) ND	2×10^{-2} ($\pm 1 \times 10^{-2}$)NS	3×10^{-2} ($\pm 1 \times 10^{-2}$)	4×10^{-3} ($\pm 5 \times 10^{-3}$)NS
Plasma	8×10^{-3} ($\pm 7 \times 10^{-3}$)***	2×10^{-2} ($\pm 9 \times 10^{-3}$)*	2×10^{-2} ($\pm 4 \times 10^{-3}$)**	3×10^{-2} ($\pm 1 \times 10^{-2}$)

10.3 Appendix 3

Shannon entropy calculated for SGA-deriving sequences deriving from the same anatomical site and between sites.

Two-tailed paired t-test: *: $p < 0.05$, **: $p < 0.01$ ***: $p < 0.001$, NS: Not Significant. r^2: Pearson's correlation coefficient. All correlation coefficients are significant ($p > 0.001$)

Table 10.7: Shannon entropy (patient 1)

Patient 1 (r^2)	Consensus	Biopsy 1	Biopsy 6
Consensus	0.024	0.033* (0.24)	0.034*** (0.50)
Biopsy 1	0.033* (0.24)	0.018	0.045*** (0.24)
Biopsy 6	0.034*** (0.50)	0.045*** (0.24)	0.037

Table 10.8: Shannon entropy (patient 2)

Patient 2 (r^2)	Consensus	Biopsy 1	Biopsy 6
Consensus	0.014	0.017*** (0.63)	0.019* (0.21)
Biopsy 1	0.017*** (0.63)	0.022	0.021*** (0.24)
Biopsy 6	0.019* (0.21)	0.021*** (0.24)	0.009

Table 10.9: Shannon entropy (patient 4)				
Patient 4 (r^2)	Consensus	Biopsy 1	Biopsy 6	Biopsy 7
Consensus	0.007	0.016** (0.19)	0.008** (0.22)	0.018*** (0.27)
Biopsy 1	0.016** (0.19)	0.013	0.013*** (0.23)	0.015NS (0.59)
Biopsy 6	0.008** (0.22)	0.013*** (0.23)	0.003	0.017*** (0.18)
Biopsy 7	0.018*** (0.27)	0.015NS (0.59)	0.017*** (0.18)	0.017

Table 10.10: Shannon entropy (patient 5)				
Patient 5 (r^2)	Consensus	Biopsy 1	Biopsy 3	Biopsy 6
Consensus	0.016	0.031*** (0.38)	0.029*** (0.45)	0.028^{NS} (0.09)
Biopsy 1	0.031*** (0.38)	0.035	0.024NS (0.78)	0.035*** (0.44)
Biopsy 3	0.029*** (0.45)	0.024NS (0.78)	0.034	0.036*** (0.32)
Biopsy 6	0.028NS (0.09)	0.035*** (0.44)	0.036*** (0.32)	0.015

Table 10.11: Shannon entropy (patient 7)			
Patient 5 (r^2)	Consensus	Biopsy 1	Biopsy 6
Consensus	0.050	0.061*** (0.21)	0.055*** (0.23)
Biopsy 1	0.061*** (0.21)	0.027	0.038* (0.21)
Biopsy 6	0.055*** (0.23)	0.038* (0.21)	0.020

Table 10.12: Shannon entropy (patient 9)			
Patient 9 (r^2)	Consensus	Biopsy 1	Biopsy 6
Consensus	0.011	0.028*** (0.30)	0.028*** (0.33)
Biopsy 1	0.028*** (0.30)	0.027	0.013^{NS} (0.79)
Biopsy 6	0.028*** (0.33)	0.013^{NS} (0.79)	0.028

10.4 Appendix 4

BLOSUM62 similarity score for inferred HCV amino acid sequences isolated from plasma and liver collected from patient 1, 2, 4, 5, 7 and 9.

STDV: Standard deviation. ND: Not Done, the analysis has not been possible due to limited samples.

Table 10.13: Amino acid identity score by BLOSUM62 (patient 1)

Patient 1 (% (STDEV))	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	ND	97.9 (± 0.70)	98.6 (± 0.15)	97.9 (± 0.72)
Biopsy 1	97.9 (± 0.70)	97.8 (± 0.84)	98.6 (± 1.04)	98.2 (± 1.03)
Biopsy 6	98.6 (± 0.15)	98.6 (± 1.04)	99.6 (± 0.26)	98.6 (± 1.00)
Plasma	97.9 (± 0.72)	98.2 (± 1.03)	98.6 (± 1.00)	98.0 (± 1.14)

Table 10.14: Amino acid identity score by BLOSUM62 (patient 2)

Patient 2 (% (STDEV))	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	ND	95.7 (± 0.18)	95.8 (± 0.15)	95.8 (± 0.14)
Biopsy 1	95.7 (± 0.18)	99.4 (± 0.82)	97.5 (± 2.11)	99.3 (± 0.67)
Biopsy 6	95.8 (± 0.15)	97.5 (± 2.11)	96.7 (± 2.1)	97.7 (± 2.10)
Plasma	95.8 (± 0.14)	99.3 (± 0.67)	97.7 (± 2.10)	99.2 (± 0.70)

Table 10.15: Amino acid identity score by BLOSUM62 (patient 4)					
Patient 4 (% (STDEV))	Consensus	Biopsy 1	Biopsy 3	Biopsy 7	Plasma
Consensus	ND	95.8 (± 0.31)	95.9 (± 0.00)	96.0 (± 0.18)	95.9 (± 0.72)
Biopsy 1	95.8 (± 0.31)	98.0 (± 1.09)	98.7 (± 0.68)	97.4 (± 0.91)	97.9 (± 1.09)
Biopsy 3	95.9 (± 0.00)	98.7 (± 0.68)	99.0 (ND)	97.3 (± 1.27)	97.6 (± 0.85)
Biopsy 7	96.0 (± 0.18)	97.4 (± 0.91)	97.3 (± 1.27)	97.8 (± 1.60)	97.7 (± 1.13)
Plasma	95.9 (± 0.72)	97.9 (± 1.09)	97.6 (± 0.85)	97.7 (± 1.13)	97.7 (± 1.04)

Table 10.16: Amino acid identity score by BLOSUM62 (patient 5)				
Patient 5 (% (STDEV))	Consensus	Biopsy 1	Biopsy 3	Biopsy 6
Consensus	ND	96.7 (± 1.30)	96.4 (± 1.77)	97.2 (± 0.28)
Biopsy 1	96.7 (± 1.30)	97.8 (± 1.77)	97.4 (± 1.94)	98.4 (± 1.48)
Biopsy 3	96.4 (± 1.77)	97.4 (± 1.94)	96.5 (± 2.16)	97.5 (± 1.73)
Biopsy 6	97.2 (± 0.28)	98.4 (± 1.48)	97.5 (± 1.73)	99.18 (± 0.40)

Table 10.17: Amino acid identity score by BLOSUM62 (patient 7)				
Patient 7 (% (STDEV))	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	ND	99.4 (± 0.24)	99.4 (± 0.19)	99.4 (± 0.17)
Biopsy 1	99.4 (± 0.24)	99.4 (± 0.34)	99.3 (± 0.31)	99.3 (± 0.32)
Biopsy 6	99.4 (± 0.19)	99.3 (± 0.31)	99.4 (± 0.26)	99.3 (± 0.27)
Plasma	99.4 (± 0.17)	99.3 (± 0.32)	99.3 (± 0.27)	99.3 (± 0.23)

Table 10.18: Amino acid identity score by BLOSUM62 (patient 9)				
Patient 9 (STDEV)	Consensus	Biopsy 1	Biopsy 6	Plasma
Consensus	ND	98.2 (± 0.08)	98.1 (± 0.27)	98.1 (± 0.29)
Biopsy 1	98.2 (± 0.08)	98.6 (± 0.80)	98.8 (± 0.87)	98.7 (± 0.80)
Biopsy 6	98.1 (± 0.27)	98.8 (± 0.87)	98.6 (± 0.96)	98.6 (± 0.76)
Plasma	98.1 (± 0.29)	98.7 (± 0.80)	98.6 (± 0.76)	98.5 (± 0.80)